Optimal decision-making, plasticity and biases: a modelling study

Juan Camilo Ramírez Idárraga

A thesis submitted for the degree of
Doctor of Philosophy

Supervised by Dr James A. R. Marshall

Department of Computer Science
University of Sheffield
United Kingdom
September, 2015

Abstract

Individuals in nature frequently face decision problems where the information available to them is uncertain and their reproductive success depends on the outcome of their decisions. In these cases natural selection should be expected to favour individuals whose behavioural strategies yield the best reproductive payoffs. It is accepted that decision-makers in nature should evolve to behave *as if* they were Bayesian learners when making decisions on uncertain information (Marshall et al., 2013a,b; McNamara et al., 2006; McNamara and Houston, 1980; Tenenbaum et al., 2006). In this thesis different decision problems from nature are modelled in order to determine the optimal strategies that should be expected to evolve in response to different parameters of the environment, under the assumption that decision-makers should also evolve to make decisions as if they were Bayesian.

One model is proposed to determine the conditions under which *inducible defences*, a type of *phenotypic plasticity*, should be expected to evolve as a defensive mechanism against predators. The model is used to predict when plasticity is the evolutionarily optimal strategy, given the decision-maker's inherited uncertainty regarding predation risk. The model assumes that this inherited uncertainty has been shaped by natural selection in prior generations of the decision-maker's species to reflect the uncertainty exhibited by the predation risk in the environment. It is shown that when this inherited uncertainty is high enough (and thus when the uncertainty exhibited by the predation risk is also high enough) then plasticity is the optimal strategy.

A second model is presented in order to test the hypothesis that decision-makers evolve their Bayesian priors in response to variation in the environment. The results confirm the assumption made in the model described above that the decision-maker's inherited assessment and uncertainty of the predation risk can be shaped by natural selection.

Finally, a third model is presented to determine when decision-makers should be expected to evolve self-deception biases in situations of conflict over resources, such as food, against other decision-makers. This model tests the theory proposed by Trivers (2011), which states that the most evolutionarily successful deceivers are those who self-deceive first. This is because self-deceiving deceivers do not have to pay the physiological costs paid by deceivers who are aware of their deception. In the model presented in this thesis it is shown that self-deception is more likely to be the optimal strategy as the information the decision-maker has access to becomes more uncertain and as the contested resource becomes more valuable.

Declaration

I declare that this thesis is original work, apart from the exceptions indicated by a reference in the text, and that it has not been submitted to any other university for examination.

Acknowledgements

I would like to express enormous gratitude to my supervisor, Dr James Marshall, for his kind guidance and support during the preparation of this thesis. Working with him has been a very enriching and pleasant experience for me.

I would also like to thank Colciencias for their generous financial support for the completion of my Ph.D.

Additional thanks are due to Dr Ángel García and Dr Martha Millán for encouraging me to pursue a Ph.D. and for their valuable guidance and assistance to start this venture.

Finally, I wish to thank all the people that have offered me their inconditional support and encouragement and who have contributed to make my Ph.D. journey such a gratifying experience.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

1.1 The evolution of behaviours as optimal solutions to decision-making problems

By the process of natural selection certain traits become more or less common in a species (Darwin, 1859). These traits can be morphological or behavioural, for instance. The reproductive success of individuals in a species may depend on their solutions to decision-making problems they encounter, given uncertain information from their environments. In this case, natural selection is expected to favour behavioural mechanisms that find the best solutions to these problems, on average (McNamara and Houston, 2009). A behavioural mechanism refers to the set of internal processes (psychological, physiological, molecular, etc) that lead to a behaviour (Fawcett et al., 2015). Optimal solutions may be found analytically for these decision-making problems and individuals should be expected to evolve the behavioural mechanisms that allow them to make decisions in response to these problems *as if* aware of the analytical solutions. This is not to say that individuals should be assumed capable of finding these analytical solutions but rather that their evolved behaviours should simply approximate them, most likely through mechanisms, such as “rules-of-thumb”, other than computing the solutions analytically (McNamara et al., 2006; McNamara and Houston, 1980). Natural environments generally pose great complexity, which makes it very difficult for a single evolved behaviour to be optimal in all possible cases. Thus, behaviours favored by natural selection should be expected to be optimal, on average, even if they are the result of behavioural mechanisms that are not always optimal (McNamara and Houston, 2009).

In the study of evolutionarily-optimal animal behaviour, the solution to a decision-making problem faced by an individual frequently depends on a variable of the environment whose value is unknown to the individual. This variable is part of the state of the environment and can be, for instance, the probability of finding food or encountering a predator in a given location. Despite lacking knowledge about the state of the environment, the individual has access to cues which indicate the value of this variable with a degree of uncertainty. The individual can reduce this uncertainty with repeated gathering of these cues in order to estimate the variable and make an informed decision, thus the evolutionary benefit of the decision depends on the quality of the estimate (Dall et al., 2005). Analytically, it can be shown that using Bayes' theorem, shown in Equation 1.1, is the optimal method to produce these estimates from uncertain data (Marshall et al., 2013b; McNamara et al., 2006; McNamara and Houston, 1980). Thus, it should be expected that individuals of a species evolve the behavioural mechanisms to make decisions as if computing Bayesian estimates, which does not necessarily imply that individuals evolve the means to compute these estimates (Marshall et al., 2013a; McNamara et al., 2006; Tenenbaum et al., 2006). That is to say, the individual should evolve to act as if having a *prior* estimate of the unknown variable of the environment and updating it to a *posterior* estimate with each uncertain cue it gathers (McNamara et al., 2006). An individual's prior estimate represents the individual's innate assessment of the environment variable. It has been hypothesised that this default 'worldview' of the individual is shaped by natural selection and thus is a projection of the environment experienced by the individual's ancestors (McNamara et al., 2006). Often the variable of the environment the individual must estimate is the probability of an event, A, such as the probability of encountering a predator. Thus the individual should evolve to behave as if having a prior estimate, denoted by $P(A)$ in Equation 1.1, and updating it to a posterior estimate, denoted by $P(A|B)$, when the individual encounters new evidence in the form of the observed occurence of a subsequent event B (*e.g.*, a subsequent predator encounter or non-encounter).

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)} \tag{1.1}$$

Foraging in food patches of uncertain quality is an example of a decision problem where a decision-maker should be expected to behave as Bayesian learner (McNamara, 1982). In this scenario the environment consists of a set of food patches. Each type of patch yields an expected reward (*i.e.*, an expected amount of food) when the individual explores it but the individual does not know in advance the magnitude of this expected reward. The individual can spend any amount of time exploring a patch and, when doing so, it has two possible choices: to continue exploring or to give up the current patch and move on to a different one. The individual can explore a patch without knowing for certain that food will be found there and must choose the appropriate time when to move on and explore a different one in order to maximise its food gain and minimise the time invested. The individual can use the information learned from exploring the current patch in order to make its decision. It is assumed that natural selection will favour the foraging strategy that overall maximises the decision-maker's food gain (Krebs et al., 1983; McNamara, 1985; Pyke, 1984; Pyke et al., 1977). McNamara (1982) modeled the scenario faced by an individual in an environment composed of food patches, where the individual updates a prior assessment of each patch to a Bayesian posterior estimate after a number of repeated observations. McNamara (1982) showed that the optimal policy is to change patches when the estimated reward rate at the current patch is below or equal to the reward rate of the whole environment. This relates to the *marginal value theorem*, which predicts that individuals should generally stay longer in the current patch when the resource richness is relatively high and costs of travel to a different patch are high (Charnov, 1976). Thus, this is the behavioural strategy that should be expected to be favoured by natural selection. This is an example of how a mathematical model can be used to determine what behavioural strategy should be expected to evolve in response to a decision problem in a natural environment, given the assumption that decision-makers should be driven by natural selection to behave as if they were Bayesian.

This thesis presents three models to investigate different decision problems in nature. The first model is proposed to investigate when an individual should exhibit *phenotypic plasticity* in the form of *inducible defences.* The second model is presented to show how Bayesian priors can be shaped by natural selection, which is a hypothesis made in the literature of the evolution of animal behaviour. The third aims to show when *self-deception* in a situation of conflict should be expected to evolve. More general details of these models as well as the motivation for each one of these appear below.

1.2 A model of the evolution of plasticity

Phenotypic plasticity is the ability of a genotype to be expressed by different phenotypes, depending on stimuli from the environment (DeWitt and Scheiner, 2004; Price et al., 2003; Whitman and Agrawal, 2009). This is opposed to *fixed phenotypic specialisation* where the phenotype develops independently from environmental stimuli. Chapter 2 investigates *inducible defences*, which constitute a type of phenotypic plasticity observed in some species, such as *Daphnia pulex* (Agrawal et al., 1999; Hammill et al., 2008; Tollrian, 1993). These defences are costly but they improve survival when there is a predation risk. Efforts have been made to predict when this type of plasticity should be expected to evolve instead of specialisation, given the trade-off posed by these defences (Boeing et al., 2005; Hammill et al., 2008; Riessen and Sprules, 1990). In Chapter 2 a model is presented to predict when this type of plasticity should be expected to maximise an individual's expectancy of reaching adulthood when this trade-off is taken into account. The model simulates the lifetime of an individual with inducible defences who must decide when to induce them, under the assumption that the individual has evolved to estimate the predation risk through repeated observations by behaving as a Bayesian learner in order to make its decision.

If the individual chooses induction before making any observation then the individual is said to exhibit phenotypic specialisation, otherwise the individual is said to exhibit phenotypic plasticity. The individual's expectancy of reaching adulthood depends on when the defences are induced since these entail physiological costs if induced when not needed and failing to induce them when they are needed results in an increased death risk by predation.

The model is used to predict, for each environment, whether the individual should choose plasticity or specialisation in order to optimise its expectancy of reaching adulthood. It is shown that this choice depends on the standard deviation of the individual's prior. More specifically, it is shown that for any environment there is a threshold. If the standard deviation of the individual's prior is above or equal to this threshold then the individual maximises its expectancy of reaching adulthood by exhibiting plasticity. Otherwise the individual achieves this by exhibiting specialisation. The model assumes that the individual's prior is inherited and that its standard deviation has been previously shaped by natural selection to reflect the uncertainty of the predation risk in the environment (a hypothesis tested in Chapter 3). Given this assumption and the results obtained from the model, it is concluded that inducible defences maximise the individual's expectancy of reaching adulthood only when the predation risk exhibits a level of uncertainty above a certain minimum, which is consistent with the literature on plasticity. The model is novel in that it takes into account the individual's inherited prior in order to predict when inducible defences should be expected to maximise the individual's expectancy of reaching adulthood.

1.3 A model of the evolution of Bayesian priors

A model is presented in in Chapter 3 to investigate the hypothesis made in Chapter 2, namely that Bayesian priors can be encoded genetically and shaped by natural selection. In the model, which follows an approach similar to that followed by Hinton and Nowlan (1987), each member of a population of Bayesian learners must make repeated observations from the environment in order to estimate the probability that an event occurs (*e.g.*, a predator encounter). This probability is referred to as the environment state and each observation consists of a Bernoulli trial with a probability equal to the environment state. Learners are born with a genetically-encoded prior estimate of the environment state and use the evidence gained from each observation to update their prior estimate to a Bayesian posterior estimate using Equation 1.1. The reproductive success of each individual increases with the accuracy of the estimates made after a fixed number of observations and learners pass on their encoded priors to their offspring but not the information they have learned individually from the environment. The simulated evolution of these individuals shows that their priors are shaped by the *Baldwin effect* (Baldwin, 1896) to allow them to estimate accurately the environment state. Whenever the environment state changes the population evolves priors that approximate the updated environment state since these individuals are subject to a selective pressure to produce the most accurate Bayesian posterior estimates. In addition to this, it is shown that individuals evolve their priors to have high standard deviations when the distribution of environment states also has a high standard deviation. The results obtained and presented in Chapter 3 support the hypothesis made in Chapter 2 that decision-makers can have their priors evolve to have different means and standard deviations, and that the evolved standard deviations reflect the uncertainty of the environment states.

1.4 A model of the evolution of self-deception

Self-deception is a tendency to ignore or deny truthful, objective information in order to convince oneself of a false idea without any awareness of the deception. It is manifest in the apparent biases in decision making observed in animals, including humans, in certain situations. These seem to present an evolutionary puzzle, since one would expect decisions based on biased (unrealistic) information to be suboptimal, unlike unbiased Bayesian estimates. Despite this, self-deception biases are reported frequently in psychology literature (Alicke and Govorun, 2005; McCormick et al., 1986; Pallier et al., 2002; Sharot, 2011a,b; Svenson, 1981) and Johnson and Fowler (2011) have proposed a model to explain how these can evolve, which has received discussion and criticism (Johnson and Fowler, 2013; Marshall et al., 2013a,b). It has been proposed by Trivers (2011) that individuals should evolve to self-deceive first in order to better deceive others. Although the theory proposed by Trivers (2011) has been scrutinised extensively (Bandura, 2011) it has not been formally modelled. Chapter 4 presents the first model designed to investigate Trivers' proposal and is an extension to the one proposed by Johnson and Fowler (2011). In the model presented in Chapter 4 individuals make decisions by taking directly into account the benefits and costs of each outcome and by choosing the course of action that can be estimated as the best with the information available. It is shown that in certain circumstances self-deceiving decision-makers are the most evolutionarily successful, even when there is no deception between these. In a further extension of this model individuals additionally exhibit deception biases and Trivers' premise (that effective deception is less physiologically costly with the aid of self-deception) is incorporated. It is shown that under Trivers' hypothesis natural selection favours individuals that self-deceive as they deceive others. This model shows how a behavioural strategy can be evolutionarily optimal despite being the result of a suboptimal behavioural mechanism (self-deception), which is consistent with the literature on this topic (McNamara and Houston, 2009).

Finally, in Chapter 5 a concluding discussion of the results obtained is presented and future directions to follow are proposed.

Chapter 2

Optimal sampling-specialisation balance in individuals with inducible defences

2.1 Introduction

Some animals are able to induce changes in their phenotypes as a response to stimuli from the environment. These changes can be morphological or behavioural and they result from changes in the environment and not exclusively from the genetic code. Additionally, these changes can be permanent or temporary. This ability to induce changes in phenotype in response to stimuli from the environment is called phenotypic *plasticity* (Price et al., 2003; Scheiner, 1993) and is in contrast to *fixed phenotypic specialisation*, where the phenotype develops independently from stimuli. Plasticity has been broadly defined as “environment-dependent phenotype expression” (DeWitt and Scheiner, 2004) since different plastic individuals from the same species can exhibit different phenotypes if they are exposed to different phenotype-inducing stimuli. If the environment exhibits variations that take place within the lifespan of an individual then these occur more frequently that matching, optimal phenotypes can evolve and individuals face a selective pressure to adjust their phenotypes to make them fit the current environment state (De Jong, 2005; Forsman, 2014). When this occurs, it should be expected that a species evolves phenotypic plasticity, *i.e.*, environment-dependent genotype expression. That is to say, when there is enough instability in the environment, individuals who exhibit phenotypic plasticity should have an evolutionary advantage over those who do not (Price et al., 2003). Specialised strategies may evolve, depending on how environment changes occur. For instance, it has been shown that in spatially-variable environments non-plastic individuals should evolve to 'estimate' favourable environments as more likely than they really are, a form of 'optimistic' behaviour. On the other hand, in temporally-variable environments non-plastic individuals should evolve to be 'pessimistic'. That is to say, they should evolve to 'estimate' favourable environments as less likely than they really are (McNamara et al., 2011).

Phenotypic plasticity is commonly observed in nature, in animals and plants (Belsky et al., 1991; Bradshaw, 1965; DeWitt and Scheiner, 2004; Ellis, 2004; Kuzawa, 2005; Nettle, 2011; Schlichting and Pigliucci, 1998; West-Eberhard, 2003). It is particularly beneficial to plants given their motionless nature, unlike animals, who can more easily escape unhospitable environments (Schlichting, 1986; Sultan, 2000). In plants, phenotypic plasticity commonly consists of changes in thickness and size of leaves and roots. For instance, in soil with low concentration of nutrients roots grow to reach the most amount of nutrients (Sultan, 2000). In animal species one example of phenotypic plasticity is that of *inducible defences*. This is a type of plasticity that consists of morphologies that serve as defensive mechanisms that are developed when a threat is detected. These are observed in *Daphnia pulex*, a species of crustaceans commonly known as water fleas, whose phenotype develops defences when environment cues indicate the presence of phantom midge larvae, *Chaoborus*, who predate on the fleas (Agrawal et al., 1999; Hammill et al., 2008). *Daphnia pulex* individuals detect the *kairomone* (*i.e.*, a chemical substance) released by the larvae and respond by developing neckteeth on their heads as a defensive mechanism (Parejko, 1991; Spitze, 1992; Tollrian, 1995), as illustrated in Figure 2.1. After being induced, the defensive mechanism is kept until maturity, when the *Daphnia pulex* individual has grown to the point that *Chaoborus* is no longer a threat (Walls and Ketola, 1989).

Plasticity has been studied in order to determine the specific conditions that must be met for it to evolve in particular circumstances. Models have been proposed, where plasticity is simulated as a process where a fixed-length period of time is invested by each plastic individual in cue-gathering after which the phenotype develops accordingly (Gabriel, 2006; Jablonka et al., 1995; Levins, 1968; Meyers and Bull, 2002; Moran, 1992; Padilla and Adolph, 1996; Piersma and Drent, 2003) but it has also been modelled as an incremental process in which phenotype development is cumulative, where the individual's development stage is composed of periods of specialisation and cue-detection (Frankenhuis and Panchanathan, 2011). It has been shown how variables of the environment determine the levels of plasticity and specialisation that evolve (Frankenhuis and Panchanathan, 2011). One variable that has been observed to have an important impact in the evolvability of plasticity or specialisation is the uncertainty with which the cues signal the environment state, *i.e.*, the reliability of these cues. It has been shown that an increase in the reliability of environmental cues leads to individuals being more likely to evolve plasticity in order to predict the environment state via cue-detection and adjust their phenotypes accordingly (Frankenhuis and Panchanathan, 2011). If the cues are reliable then the plastic individual becomes more evolutionarily successful because it manages to reduce the probability of developing the wrong phenotype in the current environment (Nepomnaschy and Flinn, 2009). However, this environment-induced development normally requires time and other resources, therefore a more evolutionarily stable strategy should be to have a phenotype specialised for the local conditions of the environment before birth if these are unlikely to change. There is a trade-off between having a specialised phenotype from an early life stage and the risk of developing this early, specialised phenotype in a mismatching environment, because this entails costs (Gluckman et al., 2005) that lead to decreased fitness.

The development of inducible defences in *Daphnia pulex* results in a survival benefit (Havel and Dodson, 1984; Krueger and Dodson, 1981; Parejko and Dodson, 1991) but entails a demographic cost (*i.e.*, a diminished reproductive rate due to a physiological cost) to the species (Black and Dodson, 1990; Havel and Dodson, 1987; Riessen and Sprules, 1990; Walls et al., 1991; Walls and Ketola, 1989) , therefore they pose a trade-off (Riessen, 1992, 1999), and this creates a selective pressure on *Daphnia pulex* to refrain from inducing the defences when they are not needed. If each *Daphnia pulex* individual could 'know' (no conscious act being implied) in advance that predators are abundant in the environment then this individual would maximise its survival by specialising as early as possible by inducing its defences. On the other hand, if the individual could anticipate that predators are absent, then it would maximise its survival by never inducing its defences and saving physiological resources that could be invested in some other way. In nature, however, this individual does not have this information in advance. The uncertainty in the information the individual has access to in order to predict threats results in a trade-off between learning from the environment (plasticity) and early specialisation.

The demographic benefits and costs of this type of defences have been measured by Boeing et al. (2005) by comparing *Daphnia pulex*'s population growth, denoted by r, when these are exposed to different environments, with varying degrees of predator presence. In their experiments, two *Daphnia pulex* clones are used. The first reacts to the presence of

Chaoborus, by inducing its defences, and is referred to as ‘RC’. The other is non-reactive and is referred to as ‘NRC’. These two clones are exposed to three environments. In the first, denoted by ‘C’, no predators are present. In the second denoted by ‘P’, *Chaoborus* swim freely and can predate on *Daphnia*. In the third, denoted by ‘K’, *Chaoborus* are present but they are separated from *Daphnia* by a mesh tube; in this manner their presence is known to *Daphnia* via *Chaoborus*’ kairomone but they do not represent a threat. The demographic cost of *Daphnia*’s defences was measured as $r_{(\mathrm{C,RC})} - r_{(\mathrm{K,RC})}$. That is to say, as the difference in growth observed in populations of predator-responsive *Daphnia* individuals when these do not have to induce their defences (*i.e.*, when they live in a predator-free environment) and when they do induce them without being affected by risk of predation (*i.e.*, when they live in an environment where the presence of predators is manifest even though these pose no threats). The demographic benefit of *Daphnia*’s defences was measured as $r_{(\mathrm{P,RC})} - r_{(\mathrm{P,NRC})}$. That is to say, as the difference in growth observed between a responsive population and and an unresponsive population when they are exposed to environments with predation risk. In summarising their results, Boeing et al. conclude that the cost paid by *Daphnia* for inducing its defences is manifest in a decrease of 32.3% in population growth whereas the benefit is measured as an increase of 68.4% (Boeing et al., 2005). This is consistent with prior efforts to estimate these benefits and costs (Riessen and Sprules, 1990) but no predictions are made about the level of predation risk beyond which induction should be favoured by natural selection to occur (Hammill et al., 2008).

Hammill et al. (2008) investigated how the benefits and costs of inducible defences in *Daphnia pulex* individuals affects their fitness, in order to determine the threshold predator risk at which an *‘undefended’* individual faces selection pressure to transition to *‘defended’*. It should be expected that a defended individual’s fitness increases relative to an undefended’s as the predation risk increases. The threshold for induction is given by the point where the two fitness values intersect (Roff, 2002). Hammill et al. (2008) measured an individual’s fitness as being directly proportional to the individual’s life expectancy and *lifetime reproductive success* (LRS), which is measured as the overall number of offspring of an individual. The individual’s life expectancy and LRS are increased and decreased by induction of the defences, respectively. The life expectancy and LRS of two clones, Cyril and Colin, were measured with and without defences. This was achieved by exposing defended and undefended individuals of each clone to different levels of predation risk and predation cues and measuring the resulting lifespan and number of offspring of each individual. These data were used to calculate the expected fitness of each individual, defended and undefended, for different levels of predator risk. Using this theoretical expected fitness, a theoretical (predator risk) threshold of induction was found for Cyril. The theoretical induction threshold of Colin was found to be much higher than the predator concentrations used experimentally. Hammill et al. (2008) explain that this is because, in the case of Colin, the fitness of defended and undefended individuals is uniformly low and the difference in fitness between the two is very small. Therefore the defences do not offer a significant benefit in terms of fitness and Colin individuals should have evolved low sensitivity to *Chaoborus* in order to avoid the costs incurred into when inducing the defences. With their theoretical model, Hammill et al. (2008) are able to predict the level of predator concentration that makes induction necessary to maximise fitness (in terms of both survival and reproductive success). However this threshold is not in relation to any individual's inherited prior estimate of predation risk.

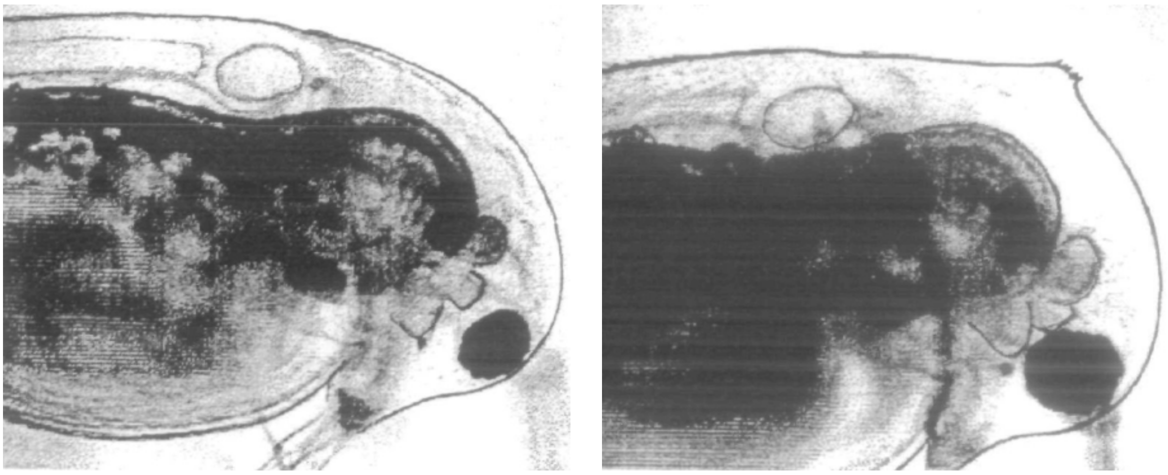
Frankenhuis and Panchanathan (2011) proposed an evolutionary model of the evolutionary trade-off between plasticity and specialisation considering the individual's prior. In the model, a group of individuals live in an environment that can be in one of two permanent states: 0 or 1. An environment in state 1 can be thought of as one where predators (which represent a death risk for the individuals modelled) are encountered with higher probability than an environment in state 0. Similarly, the individuals' phenotypes are numerical and range between 0 and 1. The reproductive success of an individual depends on how closely its phenotype matches the current environment state. At birth individuals immediately fully disperse and settle, after which each individual is permanently located in the same environment whereas the whole population can eventually become scattered across different environments with different states. The lifetime of the individual is composed of 20 discrete steps. In each step one of two mutually-exclusive actions is performed: sampling or specialising. By choosing to specialise the individual shifts its phenotype incrementally towards the optimal phenotype that matches what the individual believes is the environment state. Otherwise, by choosing to sample the individual receives a noisy cue as to the environment state. The probability that a cue is truthful is called the *cue validity* and is constant for all states. With each cue the individual updates its prior regarding the environment state using Bayes' theorem. The individual has a prior estimate of the environment state. The means of the priors considered in the model are 0.5, 0.1, 0.3, 0.7, and 0.9 (Frankenhuis and Panchanathan, 2011).

Frankenhuis and Panchanathan (2011) use dynamic programming to determine the optimal choice (either sample or specialise) an individual should make at every life step, for different environment prior probabilities (*i.e.*, the prior probability of the environment being in one state or the other) and cue validities (*i.e.*, the probability of receiving a correct cue from the environment). The sequence of optimal choices at each life step comprises the lifetime of an individual with the best fitness value and therefore this sequence is the strategy that should be expected to be favoured by natural selection in an evolutionary scenario. The results obtained show that plasticity generally decreases as one environment becomes more likely than the other and increases with cue validity. Three categories of cue validity are considered: low (0.55), intermediate (0.75), and high (0.95). Investment on sampling is observed to be greatest when cue validity is intermediate (0.75) and it decreases as cue validity decreases and increases. This is because as cue validity increases, an individual requires fewer observations to estimate correctly the environment state because these are more reliable. On the other hand, when cues are less reliable an optimal individual needs more sampling in order to produce an estimate of the same quality as with reliable cues. However individuals who invest all their time sampling without specialising do not increase their fitness. Results also show that reliance on sampling decreases as one environment becomes increasingly more likely than the other. This is because if one environment state is much more likely than the other then natural selection will simply favour individuals who sample little or not at all and specialise early always matching the environment state that is most probable.

When cue validity is low, individuals evolve to develop early, devote little time to sampling and their belief (*i.e.*, certainty) about the environment state is also low. In the intermediate category, the most investment on sampling is observed. In addition to this, increased learning paired with increased cue validity results in increased certainty about the environment state. Finally, when cue validity is high, individuals spend less time on sampling than with intermediate cue validities because then individuals require fewer observations to estimate correctly the environment state before developing the optimal phenotype. Certainty about the environment state is maximum when cue validity is high.

In the model proposed by Frankenhuis and Panchanathan (2011), phenotype development is an incremental process and the results achieved coincide with other models of adaptive development by concluding that the evolvability of phenotypic plasticity depends mainly on two variables. The first one is the individuals' priors. As these becomes less informative (*i.e.*, more uncertain or closer to $\frac{1}{2}$) plasticity becomes more likely to evolve. The second variable is the cues' validity. The evolvability of plasticity occurs mainly when cue validity is intermediate.

In Section 2.2 a model is proposed to simulate the phenotypic development of an individual analogous to *Daphnia pulex* fleas. The individual possesses inducible defences that improve its probability of survival against a predator but these imply costs to the individual when these are induced when they are not needed. In this model these costs are in terms of decreased survival. The individual's lifetime consists of a sequence of predator encounters and non-encounters and the probability of surviving each one of these with and without defences is a parameter of the environment state. The individual accumulates the information from past encounters and non-encounters in order to update its prior estimate of predator risk to a Bayesian posterior, which is used by the individual in order to decide whether to induce the defences or postpone them until more evidence is collected. As in the model proposed by Frankenhuis and Panchanathan (2011), there is no fixed period of time devoted to cue-gathering. Instead, the individual can accumulate cues from the environment and induce its defences after it estimates that the predation risk is high. The purpose of the model is to determine when the individual should optimally choose to induce the defences, given the accumulated evidence of predation risk, in order to maximise its expectancy of reaching adulthood. This is used to determine the conditions under which the individual should be expected to exhibit plasticity or early specialisation. A dynamic programming algorithm is provided along with the model, in order to determine the optimal choice the individual must make at any given moment of its lifetime, given the environment state, the evidence collected by the individual and the individual's prior. Repeated experimentation with this algorithm shows that there are two types of thresholds. One is a threshold on the standard deviation of the individual's prior that determines whether the individual should use plasticity or specialisation in order to maximise expectancy of reaching adulthood. When plasticity yields the maximum survival expectancy then the algorithm finds the threshold evidence (i.e., number of observed predator encounters and non-encounters) upon which the individual should optimally choose to induce the defences. That is to say, the thresholds found by the algorithm are not in terms of predator concentration, as in the model proposed by Hammill et al. (2008), but in terms of the individual's prior and the evidence (number of past predator encounters and non-encounters) accumulated until a given moment. It is shown that in realistic environments, the individual maximises its expectancy of reaching adulthood by exhibiting plasticity when the standard deviation of its inherited prior is above or equal to the threshold, which depends on the environment state. Otherwise, it is shown that



(a) Without neckteeth

(b) With neckteeth

Figure 2.1: Difference in head shape between undefended (2.1a) and defended (2.1b) *Daphnia pulex*. Photos reproduced from Tollrian (1993).

the individual maximises its expectancy of reaching adulthood by specialising as early as possible.

The model assumes that the individual's prior has been inherited and shaped by natural selection in previous generations of the individual's species to reflect the uncertainty in the predation risk (a hypothesis tested in Chapter 3). Given this assumption and the results obtained, it is concluded that inducible defences maximise the individual's expectancy of reaching adulthood when the predation risk has been historically uncertain (*i.e.*, when the predation risk has not been constant but has exhibited great variability) in the evolutionary history of the individual's species. This is consistent with the literature on phenotypic plasticity described earlier. However, the model is novel in that it predicts when inducible defences maximise the individual's expectancy of reaching adulthood by taking into account the individual's inherited prior in order to make these predictions.

2.2 The model

The model simulates a scenario similar to that faced by *Daphnia pulex*, described in Section 2.1. Namely, the phenotypic development of an individual with inducible defences who must decide when to induce them, depending on the individual's estimated probability of encountering a predator, given that these defences are costly if induced when they are not needed. The individual's lifetime consists of a sequence of independent, discrete life steps. At any life step t one predator encounter occurs as a Bernoulli event. The probability of such predator encounter is unknown to the individual and non-variable over time. At each step the individual has the choice to induce the defences *before* the encounter or non-encounter has taken place. The individual holds a beta prior, encoded with hyperparameters α and β, that represents its default estimate of the probability of encountering a predator. The beta prior belongs to a type of distributions called *conjugate distributions*.[1] These have the property that application of Bayes' theorem (given in Equation 1.1) to update a conjugate prior results in a posterior of the same family of distributions. That is to say, if the prior is beta-distributed the posterior will also be a beta-distributed. Additionally the calculation will require little mathematical effort. The mean of this prior is the individual's *prior estimate of predation risk*, denoted by μ and given by Equation 2.1. The individual updates it to a Bayesian posterior with each observation (*i.e.*, with each encounter or non-encounter). The mean of the posterior, after s encounters and f non-encounters, is referred to as the *posterior estimate of predation risk* and denoted by $\phi_{s,f}$ and given by Equation 2.2.

$$\mu = \frac{\alpha}{\alpha + \beta} \tag{2.1}$$

$$\phi_{s,f} = \frac{s + \alpha}{s + f + \alpha + \beta} \tag{2.2}$$

The individual's prior is assumed to have been inherited and shaped by natural selection in previous generations to allow members of the individual's species to estimate accurately the predation risk. This assumption is examined in more detail in Chapter 3. As stated above, the individual's default estimate of predation risk is denoted by μ and given by the mean of the individual's inherited prior. The uncertainty the individual has on this default estimate increases as the standard deviation of the prior increases. This uncertainty (the standard deviation of the individual's inherited prior) is denoted by σ and given by Equation 2.3, where α and β are the hyperparameters of the individual's prior.

$$\sigma = \sqrt{\frac{\alpha\beta}{(\alpha + \beta)^2(\alpha + \beta + 1)}} \tag{2.3}$$

An individual's prior with mean μ and a high standard deviation is assumed to have evolved to reflect that in previous evolutionary generations the predation risk is, on average, μ with high variation. On the other hand, an individual's prior with mean μ and a low standard deviation is assumed to have evolved to reflect that in the previous evolutionary history of the individual's species the predation risk is, on average, μ with low variation. It should be expected that the individual's certainty on its own inherited prior estimate

---

[1] A prior distribution and its Bayesian posterior distribution are called *conjugate distributions* if they belong to the same family. In this case the prior is known as a *conjugate prior*.

	Without defence	With defence
No predator encountered	$p_{0,0}$	$p_{0,1}$
Predator encountered	$p_{1,0}$	$p_{1,1}$

Table 2.1: Payoff matrix listing the probability that an individual in the model survives each life step, depending on whether a predator is encountered and whether the individual's phenotypic defences have already been developed.

of predation risk should have an influence on the decisions made by the individual. This will be confirmed later in Section 2.4.

The probability that the individual survives at time t and makes it to $t+1$ depends on whether a predator is encountered and whether defences have been developed or not. The notation $p_{\mathrm{p,d}}$ is used to denote the probability that the individual survives each life step t, where p (predator) and d (defences) are binary (0 or 1). The former indicates whether a predator is encountered and the latter indicates whether the defences have already been developed. By default, an individual has no defences and survives each t with probability $p_{0,0}$ if no predator is encountered, otherwise the individual survives with probability $p_{1,0}$. If the individual has developed defences prior to t then it survives with probability $p_{0,1}$ if a predator is not encountered, and with probability $p_{1,1}$ otherwise. The sequence of life steps continues indefinitely until the individual gets killed by a predator or until it dies from other causes. That is to say, when encountering a predator on a given life step, the individual dies with probability $1 - p_{1,0}$ if no defences have been induced and with probability $1 - p_{1,1}$ otherwise. If a predator is not encountered at any life step, the individual dies with probability $1 - p_{0,0}$, if no defences have been induced, with probability $1 - p_{0,1}$ otherwise. The four variables $p_{0,0}$, $p_{0,1}$, $p_{1,0}$, $p_{1,1}$ are referred to as the *survival probabilities*. The *environment state*, denoted by S, is given by the 4-tuple $S = (p_{0,0}, p_{0,1}, p_{1,0}, p_{1,1})$.

An environment state is considered to be realistic if it satisfies the following three conditions. The first is that $p_{1,1} > p_{1,0}$, because in the presence of a predator, survival must be more probable when defences have been developed. The second is that $p_{0,0} > p_{0,1}$, because development of defences is costly and in the absence of threats it must become a waste of resources that reduces survival. The third is that $p_{0,0} > p_{1,0}$ because the probability of survival without defences must be lower in the presence of predators than in the absence of these. In other words, any environment state $S = (p_{0,0}, p_{0,1}, p_{1,0}, p_{1,1})$ is considered to be realistic as long as $p_{0,0}, p_{0,1}, p_{1,0}, p_{1,1} \in (0, 1)$ and $p_{0,0} > p_{1,0}, p_{0,1}$ and $p_{1,1} > p_{1,0}$. The individual's survival at any life step is given by the payoff matrix shown in Table 2.1.

The model assumes that the individual's interest is to survive its early developmental stages and to reach adulthood. The notation T is used to refer to the life step where the individual reaches adulthood, hence the individual's interest is to maximise its expectancy of surviving until life step T. That is to say, the individual needs to maximise its *expectancy of reaching adulthood.* Thus the *optimal decision* to be made at life step t, when s predator encounters and f non-encounters have occurred in previous life steps (*i.e.*, when $t-1 = s+f$), is the one that maximises the individual's future expectancy of reaching adulthood in environment state S, given s and f. The *optimal strategy* is comprised by the sequence of optimal decisions given a sequence of observations (*i.e.*, encounters and/or non-encounters) made by the individual. The problem is to determine the optimal strategy to be followed by an individual with a beta-distributed prior, encoded as two beta parameters α and β, in an environment state $S = (p_{0,0}, p_{0,1}, p_{1,0}, p_{1,1})$. A solution is proposed in Section 2.3.

2.3 The optimal strategy as the solution to a one-armed bandit problem

The problem of deciding the optimal strategy for an individual (*i.e.*, when to induce defences) in the model introduced in Section 2.2 can be formulated as a *one-armed bandit problem* (Gittins et al., 2011) as follows. In each life step the choice of inducing defences or postponing them is represented as 'arms' A_1 and A_0, respectively. Prior to any life step t the individual has observed s predator encounters and f non-encounters, and therefore $t-1 = s+f$. Then at time t, before making its decision, the individual's updated posterior *estimated* probability of encountering a predator is given by Equation 2.2.

At life step t, the individual chooses whether to induce its defences (by 'pulling' arm A_1) or not (by 'pulling' arm A_0) and each choice yields a reward consisting of future survival time. The individual's future survival time after inducing the defences when s predator encounters and f non-encounters have occurred is a random variable denoted by $V_{s,f}$. Similarly, the future survival time after postponing the defences is denoted by $U_{s,f}$. The individual's *estimated* survival probability after postponing the defences is denoted by $\sigma_{U_{s,f}}$ and the *estimated* survival probability after inducing them is denoted by $\sigma_{V_{s,f}}$. These two estimated probabilities are given by Equation 2.4 and Equation 2.5, respectively.

$$\sigma_{U_{s,f}} = (1 - \phi_{s,f})p_{00} + \phi_{s,f}p_{10} \tag{2.4}$$

$$\sigma_{V_{s,f}} = (1 - \phi_{s,f})\ p_{01} + \phi_{s,f}\ p_{11} \tag{2.5}$$

The individual's future survival time after induction of the defences, denoted by $V_{s,f}$, is a negative-binomially distributed variable. Using Bernoulli terminology, $V_{s,f}$ is the number of 'successes' (*i.e.*, the number of life steps the individual survives), each of which with probability $\sigma_{V_{s,f}}$, before one 'failure' occurs (*i.e.*, before the life step where the individual dies). That is to say, $V_{s,f} \sim \mathrm{NB}(1, \sigma_{V_{s,f}})$. [2] The individual's expectancy of reaching adulthood after inducing the defences is thus given by the mean value of $V_{s,f}$, shown in Equation 2.6.

$$\mathrm{E}[V_{s,f}] = \frac{\sigma_{V_{s,f}}}{1 - \sigma_{V_{s,f}}} \tag{2.6}$$

The notation $R_{s,f}$ is used to refer to the individual's highest expectancy of reaching adulthood at life step t, after s predator encounters and f non-encounters in previous steps (with $t - 1 = s + f$). $R_{s,f}$ is given by Equation 2.7, where $I[s, f]$ denotes the individual's optimal decision, in such a way that $I[s, f] = 1$ if the individual's optimal decision is to induce the defences at time t and $I[s, f] = 0$ otherwise. If $I[s, f] = 1$ then the highest expectancy of reaching adulthood is given by $E[V_{s,f}]$. If $I[s, f] = 0$ then the highest expectancy of reaching adulthood is given by $\sigma_{U_{s,f}}\ (1 + R^*_{t+1})$, where R^*_{t+1} denotes the the highest expectancy of reaching adulthood in the next life step $(t + 1)$ and is given by Equation 2.8.

$$R_{s,f} = I[s, f]\ \mathrm{E}[V_{s,f}] + (1 - I[s, f])\ \sigma_{U_{s,f}}\ (1 + R^*_{t+1}) \tag{2.7}$$

$$R^*_{t+1} = \phi_{s,f}\ R_{s+1,f} + (1 - \phi_{s,f})\ R_{s,f+1} \tag{2.8}$$

The individual's optimal strategy is given by matrix I, where $I[s, f]$ denotes the individual's optimal decision when s predator encounters and f non-encounters have occurred, as described above. Matrix I can be found with dynamic programming by calculating first the highest expectancy of reaching adulthood at future life step T as the base case of the recurrence depicted in Equation 2.7. This expectancy of reaching adulthood can be approximated by assuming that in the base case the choice of not inducing defences is permanent. That is to say, by assuming that if defences are not induced at T then they cannot be induced in the future. Under this assumption, the individual's future survival time after postponing the defences at time T, after s_T predator encounters and f_T non-encounters (with $T - 1 = s_T + f_T$), is a negative-binomially distributed variable denoted by U_{s_T,f_T}. Using Bernoulli terminology, U_{s_T,f_T} is the number of 'successes' (*i.e.*, the number of life steps the individual survives), each of which with probability $\sigma_{U_{s_T,f_T}}$ (defined in

---

[2] The notation $X \sim \mathrm{NB}(r, p)$ is used to express that random variable X follows a negative binomial distribution and represents the number of Bernoulli 'successes' that occur until r 'failures' have occurred, when the probability of each 'success' is p.

Equation 2.4), before one 'failure' occurs (*i.e.*, before the life step where the individual dies). That is to say, $U_{s_T,f_T} \sim \mathrm{NB}(1, \sigma_{U_{s_T,f_T}})$. The expectancy of reaching adulthood of the individual after not inducing the defences at time T is thus given by the mean value of U_{s_T,f_T}, shown in Equation 2.9.

$$E[U_{s_T,f_T}] = \frac{\sigma_{U_{s_T,f_T}}}{1 - \sigma_{U_{s_T,f_T}}} \tag{2.9}$$

Since at time T either decision is permanent, this terminal time serves as a way to model the time where the individual is no longer able to undergo phenotypic development, *i.e.*, the time where the individual reaches full maturity and is no longer able to induce its defences. Thus T is the maximum possible developmental time of the individual in the model. This limit is similar to that in the model of Frankenhuis and Panchanathan (2011), where development takes place over the course of 20 life steps. In the model presented in this chapter, different values of T can be used and thus different developmental periods can be considered.

Algorithm 1 solves the multi-armed bandit problem described above and finds the decision-maker's optimal strategy, denoted by I. The algorithm receives the environment state $S = (p_{00}, p_{01}, p_{10}, p_{11})$ and the individual's prior estimate of predation risk as input (as a beta prior with parameters α and β), which is assumed to have been inherited by the individual and shaped by natural selection, a hypothesis examined in Chapter 3. The algorithm also receives an integer T representing the time until the individual's maturation and after which the defences are assumed not to be able to be induced. This value is used as the base case of the recurrence introduced in Equation 2.7. The highest expectancy of reaching adulthood when defences are not induced in the base case is calculated as shown in Equation 2.9.

The results returned by the algorithm were observed to vary understandably as a function of its arguments. Since there are several of these, a formal prediction of the results is difficult. However, important conclusions were reached after repeated experimentation with different parameters to the algorithm. Algorithm 1 was used to determine the optimal strategy of an individual in environment states $S_1 = (p_{00} = 0.99, p_{01} = 0.2, p_{10} = 0.1, p_{11} = 0.5)$, $S_2 = (p_{00} = 0.95, p_{01} = 0.2, p_{10} = 0.1, p_{11} = 0.7)$, $S_3 = (p_{00} = 0.99, p_{01} = 0.5, p_{10} = 0.1, p_{11} = 0.35)$, and $S_4 = (p_{00} = 0.95, p_{01} = 0.7, p_{10} = 0.1, p_{11} = 0.4)$. These environment states comply with the restrictions between survival probabilities described in Section 2.2 and thus are realistic. Individuals with different prior estimates in the interval $[0, 1]$

were tested. A terminal time $T = 1,000$ was used. The results obtained are presented in Section 2.4. Even though these results are restricted to experimentation with these environments, it can be hypothesised that similar results should be obtained with other realistic environments.

Algorithm 1

```
procedure FindOptimalStrategy($\alpha, \beta, p_{00}, p_{01}, p_{10}, p_{11}, T$)
    $I \leftarrow$ matrix-of-zeros$[T, T]$
    for each $s, f \in \mathbb{Z}^+ \cup \{0\}$ such that $s + f = T - 1$ do
        $\phi_{s,f} \leftarrow \frac{s+\alpha}{s+f+\alpha+\beta}$
        $\sigma_{U_{s_T,f_T}} \leftarrow (1 - \phi_{s,f})p_{00} + \phi_{s,f}p_{10}$
        $R_U \leftarrow \frac{\sigma_{U_{s_T,f_T}}}{1-\sigma_{U_{s_T,f_T}}}$
        $\sigma_{V_{s,f}} \leftarrow (1 - \phi_{s,f})p_{01} + \phi_{s,f}p_{11}$
        $R_V \leftarrow \frac{\sigma_{V_{s,f}}}{1-\sigma_{V_{s,f}}}$
        if $R_V > R_U$ then
            $I[s, f] \leftarrow 1$
            $R_{s,f} \leftarrow R_V$
        else
            $I[s, f] \leftarrow 0$
            $R_{s,f} \leftarrow R_U$
        end if
    end for
    for $t = T - 1 \rightarrow 1$ do
        for each $s, f \in \mathbb{Z}^+ \cup \{0\}$ such that $s + f = t - 1$ do
            $\phi_{s,f} \leftarrow \frac{s+\alpha}{s+f+\alpha+\beta}$
            $\sigma_{U_{s,f}} \leftarrow (1 - \phi_{s,f})p_{00} + \phi_{s,f}p_{10}$
            $R_U \leftarrow \sigma_{U_{s,f}}\ (1 + \phi_{s,f}R_{s+1,f} + (1 - \phi_{s,f})R_{s,f+1})$
            $\sigma_{V_{s,f}} \leftarrow (1 - \phi_{s,f})p_{01} + \phi_{s,f}p_{11}$
            $R_V \leftarrow \frac{\sigma_{V_{s,f}}}{1-\sigma_{V_{s,f}}}$
            if $R_V > R_U$ then
                $I[s, f] \leftarrow 1$
                $R_{s,f} \leftarrow R_V$
            else
                $I[s, f] \leftarrow 0$
                $R_{s,f} \leftarrow R_U$
            end if
        end for
    end for
    return $I$
end procedure
```

2.4 Results

Two broad types of optimal strategies were found. One is plasticity, *i.e.*, postponing induction of the defences until evidence of predation risk is found. The other is to induce the defences at time $t = 1$, *i.e.*, before any evidence is collected. This strategy can be referred to as *earliest specialisation* and is opposed to plasticity, since it consists of ignoring all information from the environment and deciding to specialise as soon as possible.

Repeated experimentation with Algorithm 1 indicates that the optimal strategy depends on the mean and standard deviation of the individual's prior. In Section 2.2 it is explained that the standard deviation of the individual's inherited prior is a measure of the individual's uncertainty on its inherited default estimate of the predation risk. From the experimental tests carried out with Algorithm 1 it was conjectured that for any prior estimate μ (the mean of the individual's inherited beta prior), plasticity is the optimal strategy if and only if the standard deviation of the individual's prior is above or equal to a threshold that depends on the environment state. This threshold is referred to as the *inherited uncertainty threshold*. The notation δ_μ^S is used to refer to the inherited uncertainty threshold in environment state S of an individual whose prior estimate of predation risk is μ. The experiments show that the optimal strategy of any individual in an environment state S is plasticity if and only if $\sigma \geq \delta_\mu^S$, where μ and σ denote the mean and standard deviation, respectively, of the individual's inherited prior. This conjecture was arrived at after a purely empirical method consisting of repeated experimentation with Algorithm 1 (a method described further below) and a formal demonstration is difficult. The values of δ_μ^S for different environment states S and individual priors were found through repeated experimentation. However the predictions made have been observed to be consistent throughout the experiments carried out, the results of which are described as follows.

Experimentation with S_1 has showed, for instance, that $\delta_{0.67}^{S_1} \approx 0.104$. Thus plasticity is the optimal strategy in S_1 of an individual whose prior estimate is $\mu = 0.67$ if and only if $\sigma \geq \delta_{0.67}^{S_1}$, where σ is the standard deviation of the individual's inherited prior. This is illustrated in Figure 2.2a and Figure 2.2b which show graphically the optimal strategy of two individuals whose prior estimates of predation risk are equal to 0.67.[3] The optimal decision, given any s (number of predator encounters) and f (number of non-encounters), is given by the colour of the area where the (s, f) coordinate is located. Black means that the optimal decision is to postpone inducing the defences and continue sampling. Grey means the optimal decision is to induce them. Only the individual whose strategy is depicted in Figure 2.2a satisfies the condition that makes plasticity the optimal strategy. The two figures show the same general pattern except for the $(0, 0)$ coordinate, and this shows the optimal strategy depicted in Figure 2.2a is plasticity whereas the one depicted in Figure 2.2b is specialisation.

---

[3] Details on how the plots in Figure 2.2 were produced appear in in Section B.2.

In order to estimate empirically that $\delta^{S_1}_{0.67} \approx 0.104$ the following method was used. Given S_1 and $\mu = 0.67$, individual priors with different beta hyperparameters α and β in the interval $[0.01, 1,000]$ such that $\mu = \frac{\alpha}{\alpha+\beta}$ were tested. Thus, in all of these priors it holds that $\alpha = \beta$ if and only if $\mu = 0.5$. Additionally it holds that $\alpha > \beta$ when $\mu > 0.5$ and that $\alpha < \beta$ when $\mu < 0.5$. Each combination of these hyperparameters results in a different beta distribution with mean μ but with different standard deviation. It was observed that plasticity was always the optimal strategy in the cases where this standard deviation was above or equal to 0.104 (approximately) whereas earliest specialisation was always the optimal strategy when this standard deviation was below 0.104. The two plots shown in Figure 2.2 illustrate the results obtained with only two of the beta priors (with mean $\mu = 0.67$) considered. The reason for showing only these two is to illustrate the existence of the inherited uncertainty threshold $\delta^{S_1}_{0.67}$ and how the standard deviation of the individual's prior determines the optimal strategy when the mean is $\mu = 0.67$. This empirical method was used with other values of μ in both S_1 and S_2.

Results analogous to the above are shown in Figure 2.3 where S_2 is tested instead.[4] The two plots in Figure 2.3 illustrate that $\delta^{S_2}_{0.67} \approx 0.342$ and that therefore plasticity is the optimal strategy in S_2 of an individual whose prior estimate is $\mu = 0.67$ if and only if the standard deviation of the individual's prior is above or equal to 0.342. The thresholds $\delta^{S_1}_{0.67}$ and $\delta^{S_2}_{0.67}$ differ even though they refer to the same prior mean (0.67) and this indicates they depend on the environment state. Additional tests with different values of μ were carried out and the results were consistent with the conjecture stated earlier that plasticity is the optimal strategy only when the standard deviation of the individual's prior is above the inherited uncertainty threshold. Additional, analogous results can be found in Figure 2.4[5] and Figure 2.5[6] where S_3 and S_4 are tested, respectively.

The conjecture stated earlier indicates that a critical variable that determines whether plasticity is the optimal strategy or not is the individual's inherited uncertainty regarding predation risk (*i.e.*, the standard deviation of the individual's prior). More specifically, the results indicate that plasticity is more likely to be the optimal strategy when the individual holds greater uncertainty regarding its inherited prior estimate of predation risk. On the other hand, earliest specialisation is more likely to be the optimal strategy when the individual holds lower uncertainty regarding its inherited prior estimate of predation risk.

As explained in Section 2.2, this model assumes that the individual's prior is inherited

---

[4] Details on how the plots in Figure 2.3 were produced appear in in Section B.3.

[5] Details on how the plots in Figure 2.4 were produced appear in in Section B.4.

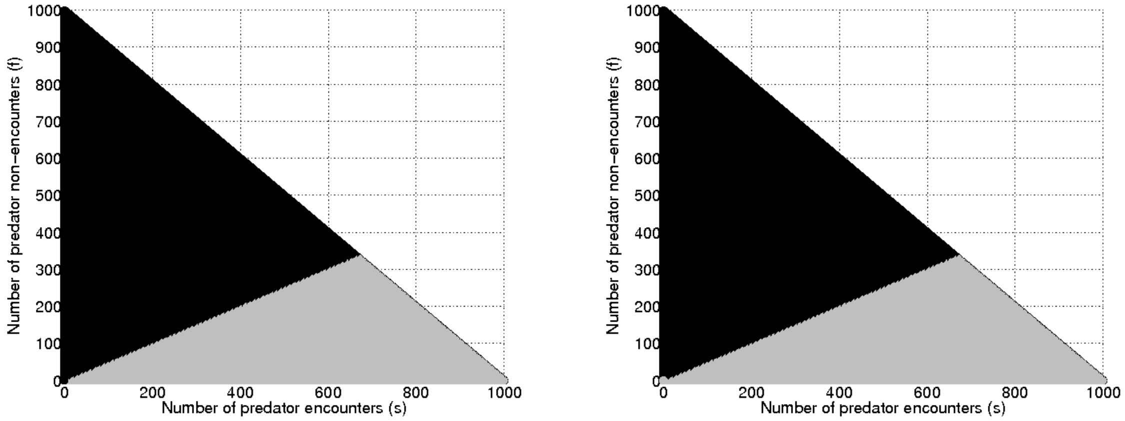
[6] Details on how the plots in Figure 2.5 were produced appear in in Section B.5.

and has been shaped by natural selection in previous generations of the individual's species, an assumption that is explored in more detail in Chapter 3. It can be hypothesised that an individual's inherited prior has evolved to have a high standard deviation to reflect the fact that in previous generations of the individual's species the predation rate has exhibited high variation and, thus, high uncertainty. Under this assumption (tested in Chapter 3), the results described earlier indicate that plasticity should be more likely to the be the optimal strategy when the predation risk has been historically uncertain. On the other hand, when the predation risk has been historically certain (*i.e.*, predictable and/or with little variation) in previous generations of the individual's species, the optimal strategy should be more likely to be earliest specialisation. In other words, plasticity becomes more likely to be the optimal strategy as the individual's evolved prior reflects more uncertainty (manifest in a high standard deviation), which in turn is hypothesised to occur when predation risk has been highly uncertain in the previous evolutionary generations of the individual's species (this hypothesis is tested in Chapter 3). A formal prediction of the uncertainty threshold is difficult, given the complexity of Algorithm 1. However the empirical finding of this threshold is in agreement with the literature, which states that plasticity should be beneficial (and favoured by natural selection) when there is high variability in the environment (De Jong, 2005; DeWitt and Scheiner, 2004; Forsman, 2014; Price et al., 2003; Scheiner, 1993). This is because an inherited (and evolved) prior with a standard deviation high enough to be above the threshold should be indicative of an environment with high variability (a hypothesis shown to be true in Chapter 3) and only in such environment should plasticity be optimal, in agreement with the literature (De Jong, 2005; DeWitt and Scheiner, 2004; Forsman, 2014; Price et al., 2003; Scheiner, 1993).

The conjecture stated earlier was tested with different individual prior means and environment states and the results were always consistent. These experimental tests showed that $\delta_{\mu}^{S_1} = 0$ when $\mu < 0.67$ and that $\delta_{\mu}^{S_1} = +\infty$ when $\mu > 0.94$. In other words, when the mean of the individual's inherited prior is below 0.67 the optimal strategy in S_1 is plasticity, regardless of the standard deviation (which is always above zero in beta distributions). Similarly, when the mean of the individual's inherited prior estimate of predation risk is above 0.94 the optimal strategy in S_1 is earliest specialisation, independently of how low or high the standard deviation is. That is to say, the interval $[0.67, 0.94]$ is where the standard deviation of the individual's prior has been observed experimentally to determine the individual's optimal strategy in environment state S_1. This interval can be referred to as the *critical interval* of the environment state. The notation Λ_S is used to denote the critical interval of environment state S. Therefore the experiments carried out indicate that $\Lambda_{S_1} \approx [0.67, 0.94]$. In the same manner, the experiments with S_2, S_3, and S_4 have shown that $\Lambda_{S_2} \approx [0.56, 0.81]$, $\Lambda_{S_3} \approx [0.67, 0.96]$, and $\Lambda_{S_4} \approx [0.46, 0.87]$, respectively. These findings indicate that each critical interval is specific to each environment state. Thus it can be hypothesised that tests on other environments should reveal different critical intervals. Figure 2.6 plots the inherited uncertainty threshold for each μ in the critical interval of each one of the four environments tested. That is to say, Figure 2.6a shows the values of $\delta_\mu^{S_1}$ for all values of μ in Λ_{S_1} and the other three subplots in Figure 2.6 display analogous information regarding environments S_2, S_3, and S_4.[7] For each individual prior's mean μ on the horizontal axis, the values on the vertical axis above the curve (*i.e.*, the values y in the points (μ, y) above the curve) in Figure 2.6a are the individual prior's standard deviations with which plasticity is the optimal strategy in environment S_1. The other three subplots in Figure 2.6 show analogous information regarding the other three realistic environments considered. The four curves shown in Figure 2.6 are noticeably different and this is most likely because the four environments differ. The curve in Figure 2.6a is very similar to the one depicted in Figure 2.6c, although the latter exhibits a faster decrease with values of μ above 0.90. This difference is probably due to the fact that in S_1 the survival probability of an defended individual when predators are absent (*i.e.*, p_{01}) is higher than the survival probability of a defended individual when predators are present (*i.e.*, p_{11}), whereas the opposite happens in S_3. That is to say, in environment S_1 it holds that $p_{01} < p_{11}$ whereas in S_3 it holds that $p_{01} > p_{11}$. Then it can be hypothesised that the faster decrease of the curve observed in Figure 2.6c occurs as the value of p_{01} increases relative to that of p_{11}. The same occurs with Figure 2.6b and Figure 2.6d. Both bear a resemblance but the latter decreases faster as μ increases above 0.80, and in environment S_2 the environment probabilities are such that $p_{01} < p_{11}$ whereas in S_4 it holds that $p_{01} > p_{11}$. Hence it can be conjectured that the accelerated decrease of the curves in Figure 2.6c and Figure 2.6d compared to the curves in Figure 2.6a and Figure 2.6b, respectively, occurs as p_{01} increases relative to p_{11}. This conjecture is purely experimental, but it could be addressed formally in future work. Additionally, the curve in Figure 2.6a appears to decrease more steeply than that in Figure 2.6b as μ increases. A reasoning similar to the one presented earlier can be followed to reach a conjecture regarding the difference between these curves. The survival probability in the absence of predators and defences (*i.e.*, p_{00}) is higher in S_1 whereas the survival probability in the presence of predators when defences have been in-

---

[7] Details on how the four subplots in Figure 2.6 were produced appear in in Section B.6.



(a) $\mu_1 \approx 0.67$, $\sigma_1 \approx 0.104$

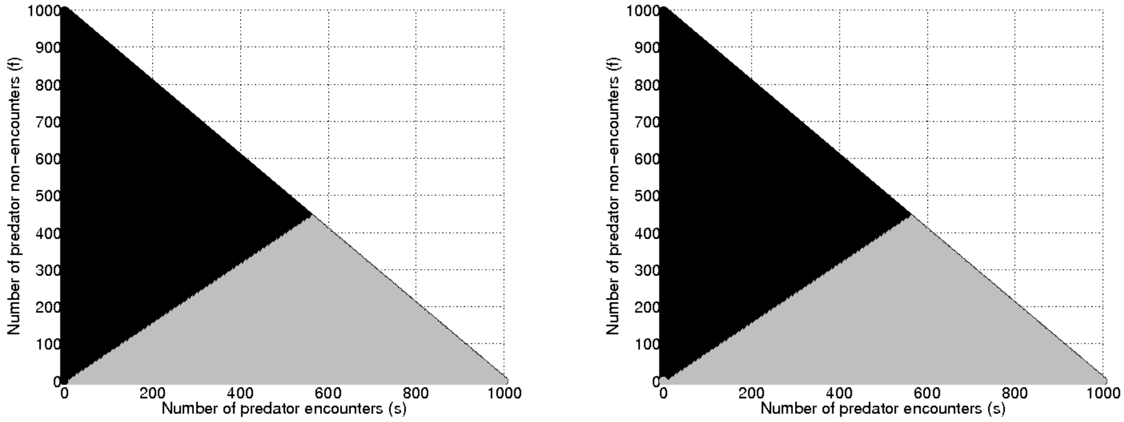
(b) $\mu_2 \approx 0.67$, $\sigma_2 \approx 0.103$

Figure 2.2: Optimal decision $I[s, f]$ in an environment state $S_1 = (p_{00} = 0.99, p_{01} = 0.2, p_{10} = 0.1, p_{11} = 0.5)$ with two different individual priors. Figure 2.2a shows the optimal decision when the individual's prior has a mean μ_1 and standard deviation σ_1 and Figure 2.2b shows the optimal decision when the individual's prior has a mean μ_2 and standard deviation σ_2. This optimal decision is given by the colour of the area where the (s, f) coordinate is located. Black means that the optimal decision is to postpone inducing the defences and continue sampling. Grey means the optimal decision is to induce them. The priors have the same mean but in Figure 2.2a the standard deviation is higher than in Figure 2.2b. Experimentally it has been determined that $\delta^{S_1}_{0.67} \approx 0.104$. Plasticity is the optimal strategy in Figure 2.2a (as indicated by the black colour at coordinate $(0, 0)$) because $\sigma_1 \geq \delta^{S_1}_{0.67}$ and earliest specialisation is the optimal strategy in Figure 2.2b (as indicated by the grey colour at coordinate $(0, 0)$) because $\sigma_2 < \delta^{S_1}_{0.67}$. Details on how to produce these plots appear in Section B.2.

duced (*i.e.*, p_{11}) is higher in S_2. It could be hypothesised that a simultaneous decrease and increase of p_{00} and p_{11}, respectively (as in a transition from S_1 to S_2), should result in a less rapidly decreasing curve. Figure 2.6c and Figure 2.6d exhibit a similar difference and the same hypothesis applies to these figures. As stated earlier, a formal prediction of the results obtained with Algorithm 1 is difficult and so is a formal proof of these conjectures. However, these hypotheses could be explored in more detail in the future, to show how tweaking these parameters could possibly change the size of the space of individual priors' standard deviations with which plasticity is optimal, thus making this strategy more or less likely to be optimal.

2.5 Conclusions

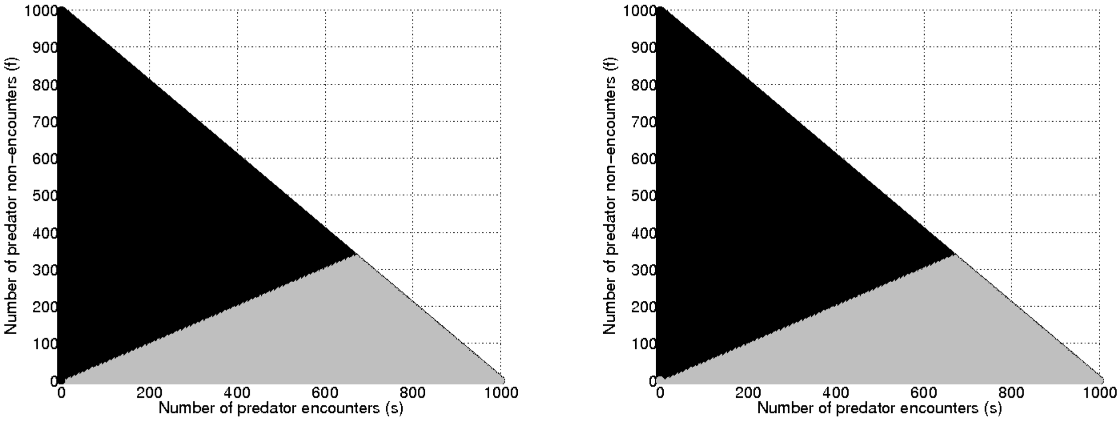
The model presented in Section 2.2 recreates the phenotypic development of an individual similar to *Daphnia pulex* and Algorithm 1 determines *when* this individual should induce its defences in order to maximise its expectancy of reaching adulthood. The time when induction should optimally occur depends on the survival costs of inducing or not, given the



(a) $\mu_1 \approx 0.67$, $\sigma_1 \approx 0.342$

(b) $\mu_2 \approx 0.67$, $\sigma_2 \approx 0.329$

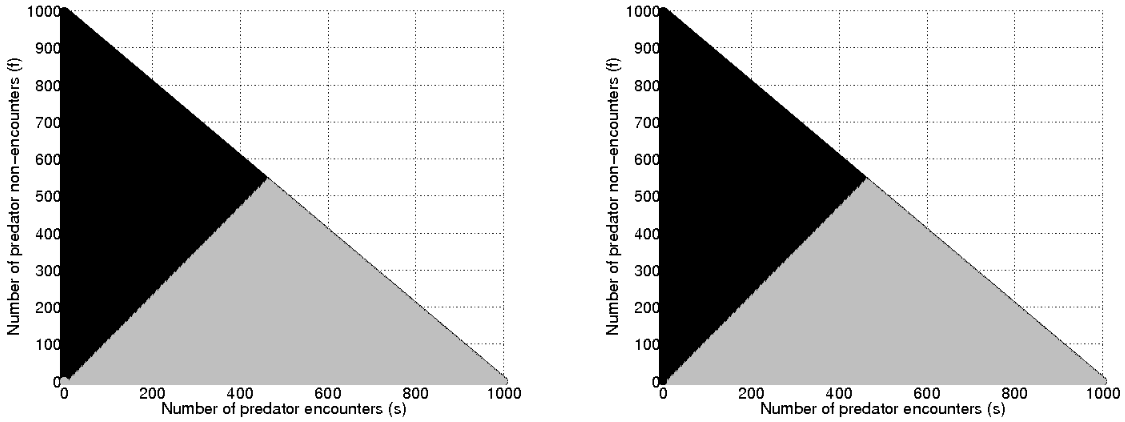
Figure 2.3: Optimal decision $I[s, f]$ in an environment state $S_2 = (p_{00} = 0.95, p_{01} = 0.2, p_{10} = 0.1, p_{11} = 0.7)$ with two different individual priors. Figure 2.3a shows the optimal decision when the individual's prior has a mean μ_1 and standard deviation σ_1 and Figure 2.3b shows the optimal decision when the individual's prior has a mean μ_2 and standard deviation σ_2. This optimal decision is given by the colour of the area where the (s, f) coordinate is located. Black means that the optimal decision is to postpone inducing the defences and continue sampling. Grey means the optimal decision is to induce them. The priors have the same mean but in Figure 2.3a the standard deviation is higher than in Figure 2.3b. Experimentally it has been determined that $\delta_{0.67}^{S_2} \approx 0.342$. Plasticity is the optimal strategy in Figure 2.3a (as indicated by the black colour at coordinate $(0, 0)$) because $\sigma_1 \geq \delta_{0.67}^{S_2}$ and earliest specialisation is the optimal strategy in Figure 2.3b (as indicated by the grey colour at coordinate $(0, 0)$) because $\sigma_2 < \delta_{0.67}^{S_2}$. Details on how to produce these plots appear in Section B.3.



(a) $\mu_1 \approx 0.67$, $\sigma_1 \approx 0.115$

(b) $\mu_2 \approx 0.67$, $\sigma_2 \approx 0.114$

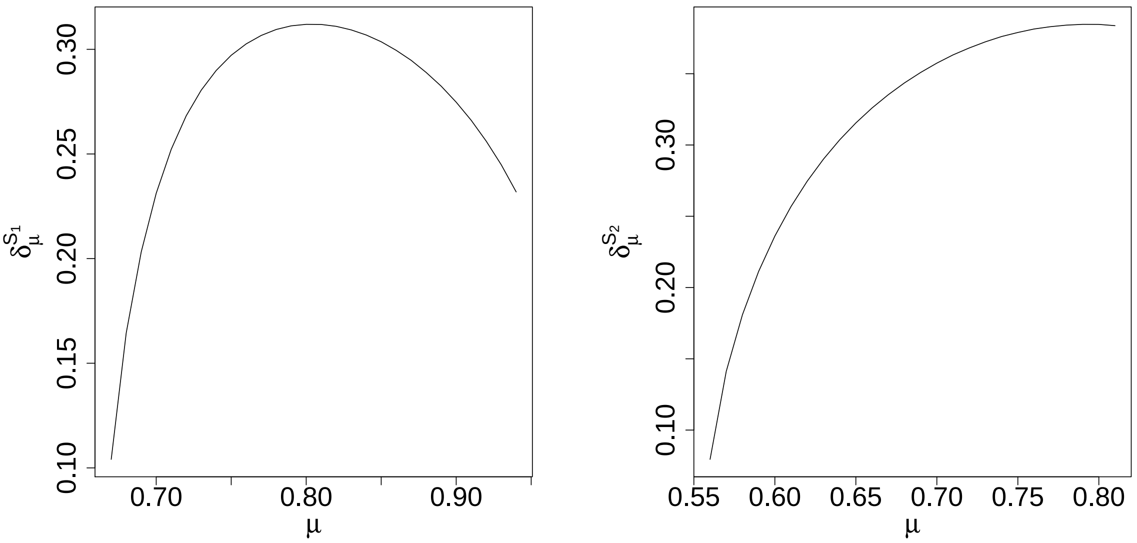
Figure 2.4: Optimal decision $I[s, f]$ in an environment state $S_3 = (p_{00} = 0.99, p_{01} = 0.5, p_{10} = 0.1, p_{11} = 0.35)$ with two different individual priors. Figure 2.4a shows the optimal decision when the individual's prior has a mean μ_1 and standard deviation σ_1 and Figure 2.4b shows the optimal decision when the individual's prior has a mean μ_2 and standard deviation σ_2. This optimal decision is given by the colour of the area where the (s, f) coordinate is located. Black means that the optimal decision is to postpone inducing the defences and continue sampling. Grey means the optimal decision is to induce them. The priors have the same mean but in Figure 2.4a the standard deviation is higher than in Figure 2.4b. Experimentally it has been determined that $\delta_{0.67}^{S_3} \approx 0.115$. Plasticity is the optimal strategy in Figure 2.4a (as indicated by the black colour at coordinate $(0, 0)$) because $\sigma_1 \geq \delta_{0.67}^{S_3}$ and earliest specialisation is the optimal strategy in Figure 2.4b (as indicated by the grey colour at coordinate $(0, 0)$) because $\sigma_2 < \delta_{0.67}^{S_3}$. Details on how to produce these plots appear in Section B.4.



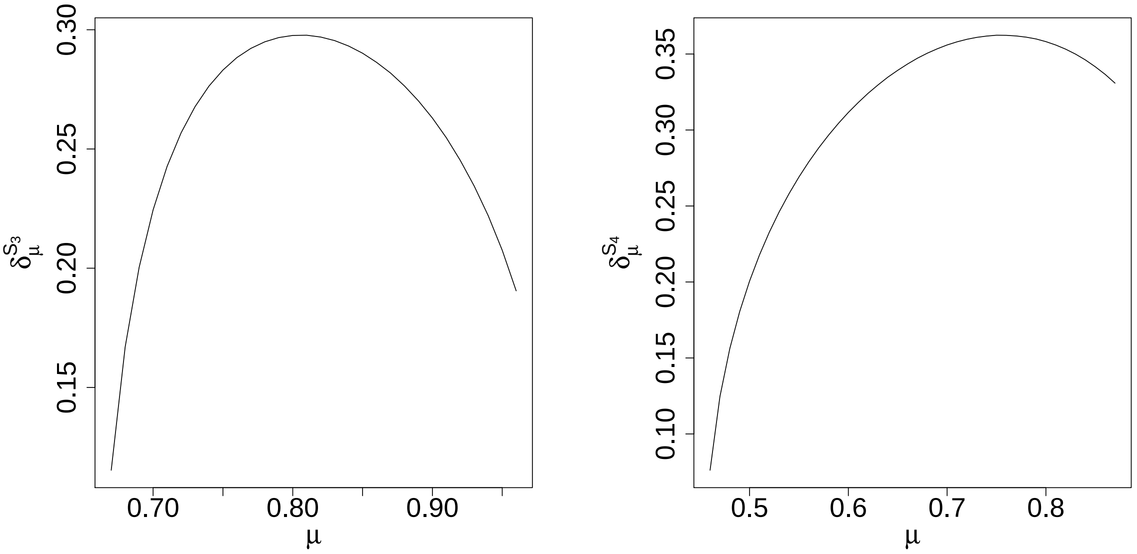
(a) $\mu_1 \approx 0.67$, $\sigma_1 \approx 0.342$

(b) $\mu_2 \approx 0.67$, $\sigma_2 \approx 0.299$

Figure 2.5: Optimal decision $I[s, f]$ in an environment state $S_4 = (p_{00} = 0.95, p_{01} = 0.7, p_{10} = 0.1, p_{11} = 0.4)$ with two different individual priors. Figure 2.5a shows the optimal decision when the individual's prior has a mean μ_1 and standard deviation σ_1 and Figure 2.5b shows the optimal decision when the individual's prior has a mean μ_2 and standard deviation σ_2. This optimal decision is given by the colour of the area where the (s, f) coordinate is located. Black means that the optimal decision is to postpone inducing the defences and continue sampling. Grey means the optimal decision is to induce them. The priors have the same mean but in Figure 2.5a the standard deviation is higher than in Figure 2.5b. Experimentally it has been determined that $\delta^{S_4}_{0.67} \approx 0.342$. Plasticity is the optimal strategy in Figure 2.5a (as indicated by the black colour at coordinate $(0, 0)$) because $\sigma_1 \geq \delta^{S_4}_{0.67}$ and earliest specialisation is the optimal strategy in Figure 2.5b (as indicated by the grey colour at coordinate $(0, 0)$) because $\sigma_2 < \delta^{S_4}_{0.67}$. Details on how to produce these plots appear in Section B.5.



(a) $\delta_\mu^{S_1}$ for all $\mu \in \Lambda_{S_1}$ (b) $\delta_\mu^{S_2}$ for all $\mu \in \Lambda_{S_2}$



(c) $\delta_\mu^{S_3}$ for all $\mu \in \Lambda_{S_3}$ (d) $\delta_\mu^{S_4}$ for all $\mu \in \Lambda_{S_4}$

Figure 2.6: The inherited uncertainty threshold plotted against the individual prior's mean μ in the critical interval of environments S_1, S_2, S_3, and S_4. That is to say, Figure 2.6a shows the values of $\delta_\mu^{S_1}$ for all values of μ in Λ_{S_1} and the other three subplots in Figure 2.6 display analogous information regarding environments S_2, S_3, and S_4. Details on how to produce these plots appear in Section B.6.

environment state, and on the number of past predator encounters and non-encounters experienced by the individual, which comprise the evidence accumulated up to a given moment. The individual uses this evidence to update its prior estimate of predation threat to a Bayesian posterior. The probability that an individual survives each predator encounter or non-encounter with or without defences is given by the environment state. For any given amount of evidence and environment state, the optimal decision (induce the defences or postpone them) is the one that maximises the individual's expectancy of reaching adulthood. The optimal strategy is the sequence of optimal decisions to be made in response to each piece of sequential evidence. Algorithm 1 uses dynamic programming in order to find the optimal strategy for any environment state and individual prior. This algorithm is used to find the optimal strategy of individuals with different priors in realistic environment states, *i.e.*, the optimal strategy under realistic assumptions about survival probabilities.

The results presented in Section 2.4 show that the algorithm normally finds two types of optimal strategies. The first is earliest specialisation, which consists of inducing the defences before making any observations from the environment. The second is plasticity, which consist of postponing induction of the defences until the predation risk assessed by the individual makes induction adaptive.

The results obtained after repeated experimentation with Algorithm 1 indicate that the standard deviation of the individual's inherited prior plays an important role when determining the optimal strategy. As described in Section 2.2, the standard deviation of the individual's inherited prior measures the individual's inherited uncertainty regarding predation risk. More specifically, it was determined that given any environment state S and prior estimate of predation risk μ (which is the mean of the individual's inherited prior), the optimal strategy is plasticity if and only if the standard deviation of the inherited prior is above or equal to a threshold that depends on S and μ, denoted by $\delta_{\mu}^{S_1}$ and referred to as the *inherited uncertainty threshold*. In other words, if the standard deviation of the individual's inherited prior is below the threshold, then the optimal strategy is earliest specialisation, otherwise the optimal strategy is plasticity. In the latter case, the algorithm shows the optimal decision to make at any point in the life history of the individual, given the amount of accumulated evidence (*i.e.*, number of past predator encounters and non-encounters) until then. This is illustrated in Section 2.4 with repeated tests in two realistic environment states. In each case, $\delta_{\mu}^{S_1}$ is determined empirically and it is shown that an individual whose prior estimate of predation risk is μ maximises its expectancy of reaching adulthood in environment state S by exhibiting plasticity if and only if the standard deviation of its inherited prior is above or equal to $\delta_{\mu}^{S_1}$.

It is accepted that natural selection should favour individuals who behave *as if* computing Bayesian estimates (Marshall et al., 2013b; McNamara et al., 2006; McNamara and Houston, 1980) and therefore it should be expected that individuals evolve priors that allow them to make these estimates more accurately. This is not to say that adapted individuals in nature carry out Bayesian computations when processing uncertain data but rather that the cognitive machinery they use should approximate the same behaviour of an optimal Bayesian decision-maker (Marshall et al., 2013a; McNamara et al., 2006; Tenenbaum et al., 2006). In Chapter 3 it is shown how these Bayesian priors can be inherited and shaped by natural selection. It can be hypothesised that the fact that plasiticity maximises the expectancy of reaching adulthood only when the standard deviation of an individual's inherited (and previously evolved) prior is above an *inherited uncertainty threshold* indicates that plasticity is the optimal strategy only when the predation risk has been subject to great uncertainty in the past evolutionary history of the individual's species. If the predation risk is uncertain and unpredictable, then it should be expected that individuals evolve priors with high standard deviations in order to reflect the uncertainty of the predation risk (a hypothesis examined in more detail in Chapter 3). When this uncertainty is high enough and as a result the standard deviation of the evolved prior exceeds the inherited uncertainty threshold, plasticity yields the best expectancy of reaching adulthood. On the other hand, if the predation risk is certain and predictable, then it should be expected that individuals evolve priors with low standard deviations in order to reflect the low uncertainty of the predation risk (a hypothesis examined in more detail in Chapter 3). When this uncertainty is low enough and as a result the standard deviation of the evolved prior fails to exceed the inherited uncertainty threshold, earliest specialisation yields the best expectancy of reaching adulthood. This observed correlation between the uncertainty of the environment state and the survival benefit provided by the inducible defences in the model is consistent with the literature on plasticity cited in Section 2.1. One key assumption in this model, is that as the uncertainty in the environment increases then the individual's inherited prior should have a higher standard deviation, as evolved by the individual's ancestors to reflect the environmental uncertainty (an assumption shown to be true in Chapter 3). Therefore the model shows analytically that as the environment becomes uncertain then plasticity becomes more adaptive, as a consequence of the individual's inherited prior, shaped by natural selection in previous generations of the individual's species to have a higher standard deviation. This result complements the literature on plasticity cited in Section 2.1 and recent research on the prediction of plasticity as a consequence of inherited priors (Frankenhuis and Panchanathan, 2011; Stamps and Krishnan, 2014a,b).

One difference between the model presented in Section 2.2 and that of Frankenhuis and Panchanathan (2011) is that in the former cues have maximum validity, *i.e.*, they are always truthful. In the model proposed by Frankenhuis and Panchanathan (2011), each observation is an indicator of predator presence that can have varying degrees of certainty (called 'cue validity') whereas in the model presented in Section 2.2 each observation consists of an actual predator encounter or non-encounter that has a direct effect on the individual's survival, although an encounter does not necessarily result in the individual's death. Since the cues are valid in the model presented in Section 2.2 then individuals are expected to rely on these in order to develop the phenotype that maximises their expected survival; this in accordance to the predictions made by Frankenhuis and Panchanathan (2011) and Nepomnaschy and Flinn (2009). In addition to this, in the model proposed by Frankenhuis and Panchanathan (2011) the developmental stage of an individual is modelled as 20 time periods, whereas in the model presented in Section 2.2 the developmental history of the individual is limited by the terminal time T. Any terminal time T can be used as a parameter to Algorithm 1, although in the experiments carried out in Section 2.4 this parameter is set to $1,000$. The model presented in Section 2.2 differs from that of Hammill et al. (2008) in what is optimised. The former is concerned with optimising expectancy of reaching adulthood (which may not be the only fitness factor) whereas the latter optimises fitness, which is directly proportional to both the individual's expectancy of reaching adulthood (which is increased by the defences) and the individual's reproductive rate (which is decreased by the defences). In addition to this, Hammill et al. (2008) determine an induction threshold in terms of predator concentration whereas Algorithm 1 determines two thresholds. The first is a threshold on the standard deviation of the individual's inherited prior that depends directly on the environment state and determines whether the optimal strategy is plasticity or earliest specialisation. Thus the model presented in Section 2.2 associates the optimality of plasticity to the individual's prior, which can be inherited and shaped by natural selection, a hypothesis examined in Chapter 3. When plasticity is found to be the optimal strategy, the second threshold is the number of accumulated past predator encounters and non-encounters that make induction the optimal decision.

The results obtained in Section 2.4 are purely theoretical but provide an insight into the conditions that must be met for plasticity to evolve in nature. In the model presented in Section 2.3 these conditions are defined by the individual's probability of surviving each predator encounter or non-encounter, with and without defences. If the parallels of these four survival probabilites in nature could be measured in a species with permanent, inducible defences, (*Daphnia pulex* for instance) in a way similar to the work of Hammill et al. (2008), then the threshold observed experimentally in this species should be consistent with the results presented in Section 2.4. It would of of interest to examine this hypothesis in the future using realistic information from nature.

Chapter 3

Can natural selection encode Bayesian priors?

3.1 Introduction

Animals normally exhibit behaviours that allow them to survive and reproduce. These may entail a level of risk, given that the individuals can not always control or predict the outcome of their actions. For instance, foraging may imply exposure to predators but should nonetheless evolve if the benefit of finding food outweighs *on average* the long-term cost of taking the risk (Frankenhuis and Del Giudice, 2012). Some behaviours are phenotypic expressions under purely genetic control, therefore they are inherited, not acquired, and that is why they can evolve. These inherited behaviours evolve as adaptations that overall should be expected to enhance the chances of survival and reproduction of the individual (Williams, 1966). In addition to this, the evolutionary advantage provided by an animal's behaviour depends partially on the environment the animal occupies. In a hypothetically unchanging environment, it should be expected that natural selection favours individuals with behaviours that optimise their chances of survival in such a habitat. If these behaviours are optimal only in a specific configuration of the environment and if this configuration changes then new optimal behaviours are likely to evolve in order to reflect this variation. If each one of these changes occurs more frequently than a new matching optimal behaviour can evolve natural selection is likely to favour a learning ability that allows individuals to infer the configuration of the environment and adjust their behaviours accordingly instead of having these encoded genetically. The environment may provide cues that indicate its current state and these may be accompanied by some degree of uncertainty. An individual with this ability gathers these cues (Baldwin, 1896; Dall et al., 2005; DonaldsonMatasci et al., 2010; Levins, 1968; Meyers and Bull, 2002; Piersma and Drent, 2003) in order to anticipate the conditions of the environment and implement a strategy that maximises its chances of reaching a goal, such as finding food or avoiding a predator, for instance. The individual reduces uncertainty by collecting more information and the accuracy of the predictions made by the individual increases with the time spent gathering data from the environment. With a learning ability the individual may be able to infer the environment state but the information learned by the individual during its lifetime is not communicated to the genotype. However the learning ability itself may be passed on to offspring. In this case those who are more efficient at learning (*i.e.*, those who require less training to learn or those who do so more precisely) have an advantage and should be favoured by natural selection. This phenomenon was originally known as the *Baldwin effect* (Baldwin, 1896), which is a particular case of *phenotypic plasticity* (Bradshaw, 1965; DeWitt and Scheiner, 2004; Schlichting and Pigliucci, 1998; West-Eberhard, 2003), described in Chapter 2. The Baldwin effect occurs when an individual is able to search locally the space of phenotypic strategies in order to find the one that yields the best reproductive rate. Thus individuals with an inherited ability to perform this search faster will have an evolutionary advantage (Dennett, 2003). As a result of this, the ability to gain or learn information at the phenotypic level eventually has an evolutionary repercussion at the genetic level even though information learned from the environment is not transferred to the genetic code nor passed on to offspring through genetic inheritance. In addition to this, a possible, though not necessary, consequence of the Baldwin effect is that individuals later evolve the adaptations learned as fixed traits rather than acquiring them through learning (Scheiner, 2014). Even though the Baldwin effect results in an apparently advantageous phenotypic flexibility, it also implies costs because in order to learn the individual must invest time and other resources that would otherwise be used to satisfy other physiological requirements.

The Baldwin effect was modelled by Hinton and Nowlan (1987) in order to show how selection for a learning ability allows evolution to reach a space of alternatives that would not normally be reachable in practice by a traditional evolutionary search. The model simulates the evolution within a group of learning individuals, each one of these with a genome that encodes a neural network in such a way that each gene expresses whether a connection in the network is present or not. In the initial population each genome is initialised with a random genome of size 20 and alleles 0 (connection absent), 1 (connection present), and ‘?’ (connection undecided). A set of connections is arbitrarily chosen as the ‘correct’ one. Thereafter every individual undergoes a ‘learning’ process consisting of replacing the ‘?’ genes by either 0 or 1 in random trials until the correct set of connections is produced or until a limit of $1,000$ trials has been exceeded. An individual who never manages to learn the correct network receives minimum fitness whereas another one who does receives a fitness score that increases with the number of unused trials. The fittest individuals are selected to reproduce through one-point crossover and the offspring of these form a new population that replaces the previous one. After enough generations the population becomes dominated by individuals who require the fewest trials to learn the good network. Even though these individuals may not have inherited by chance the correct set of connections their genomes are in enough proximity to the objective network so that they can learn the remaining connections with little training. In similar simulations run without learning, the evolutionary search cannot find the correct network.

It has been shown that learners should be optimised by natural selection to behave *as if* they are computing Bayes-optimal estimates when trying to infer information from uncertain data (Marshall et al., 2013b; McNamara et al., 2006; McNamara and Houston, 1980). Even though real animals are unlikely to perform Bayesian computations, natural selection should favour individuals whose behaviour resembles that of Bayesian learners. This refers to cases where an individual is frequently faced with a decision problem whose solution depends on a variable of the environment unknown to the individual. This variable can be, for instance, the predation risk and/or the probability of finding food in an unfamiliar location. The individual faces a selective pressure to make the decision that maximises its survival and reproductive rate (*e.g.*, avoiding danger when the predation risk is too high or foraging in a different place if the probability of finding food in the currrent location is low). Despite lacking access to the actual state of the environment, the individual may have access to noisy cues that reveal this variable with some degree of error. Populations of individuals facing repeatedly these situations should evolve to behave as if having a *prior* and updating it to a Bayesian *posterior* with each noisy cue collected from the environment.

Trimmer et al. (2011) illustrate this by considering the scenario where an individual must decide whether to perform an action or not. If the individual acts it can succeed or fail. In the former case the individual receives a benefit, b, and in the latter it pays a penalty cost, c. The benefit has a positive impact on the individual’s fitness by increasing it whereas the cost decreases it. The event where the individual succeeds is denoted by A and the success probability is $p_A = P(A)$. Trimmer et al. (2011) show formally that the individual maximises its gain in the long term by choosing to act if and only if $p_A > \frac{c}{b+c}$. In other words, if the individual finds itself facing the same decision in repeated ocassions, its overall payoff in the long term will be positive if and only if it chooses to act only when the inequality holds true. It is assumed that the value of the benefit and the cost involved are known to the individual at the moment of making the decision but not p_A. However, the individual can estimate this probability through repeated trials or experimental observations of A. Two types of learners are considered: *frequentists* and *Bayesians*. A frequentist learner estimates p_A as $\frac{k}{n}$ after observing k occurrences of A (*i.e.*, k successes) during n trials. On the other hand a Bayesian learner starts off with a default estimate of an underlying distribution of the possible values of $p_A = P(A)$, called the *prior* estimate, before making any observation. The prior is the Bayesian learner's default belief about p_A. When new evidence becomes available in the form of an observation of an event B (*e.g.*, a subsequent observation trial) then the Bayesian learner updates its belief to a *posterior* estimate of p_A given that B has occurred. By Bayes' theorem this updated estimate is given by Equation 1.1. That is to say, at any moment the current estimate is a distribution of the possible values of p_A, and this estimate is updated with each repeated observation. If the original prior is a uniform distribution then it is called an *uninformative prior*, meaning that the Bayesian learner starts with an unbiased default belief. The updated posterior becomes more accurate as the number of observations increases.

Since decision makers are face a selective pressure to behave as Bayesians when estimating an uncertain variable of the environment, a decision maker's prior represents its default belief regarding this variable. In other words, the individual's prior represents its default 'worldview' and it has been hypothesised that this prior is inherited and subject to natural selection (McNamara et al., 2006). In other words, it has been hypothesised that populations of decision makers should evolve different priors in response to different environments. The main objective of this chapter is to show how priors can effectively be inherited and shaped by natural selection. Section 3.2 introduces an evolutionary model where a population of Bayesian learners face the scenario presented by Trimmer et al. (2011). Every individual must estimate the probability p_A of an event A and has a genetically-encoded beta prior that serves as its default estimate of p_A. Every learner makes repeated independent observation trials in order to update its estimate of p_A with the evidence gained. Fitness is measured in terms of the ability to quickly produce an estimate that approximates accurately the real probability, therefore accurate learners are selected for even though the information gained by them through observation trials is not passed on to their offspring. Results show how Bayesian populations experience the Baldwin effect as individuals whose priors are closer to p_A are favoured by natural selection and therefore the value of p_A drives the evolution of these. With any change in the value of p_A the population's priors change as natural selection favours individuals who can learn and estimate the new value faster. When p_A values are beta-distributed with hyperparameters α and β, the Bayesian population is shown to evolve genomes that reflect these parameters as if these were being estimated by Bayesian learners. The model is extended in Section 3.8 to incorporate frequentist learners and have these compete with Bayesians in order to determine the circumstances in which one group is more evolutionarily successful than the other. Repeated competition between Bayesians and frequentists shows that the former are favoured by natural selection when the environment poses the most uncertainty whereas the latter are more evolutionarily successful when the environment poses the least, a result that is in agreement with previous literature on optimal decision-making (Marshall et al., 2013b; McNamara et al., 2006; McNamara and Houston, 1980).

3.2 The Bayesian-Baldwin model

The model simulates the evolution of a population of 100 individuals who are required to estimate the probability, p_A, of an event A. The probability p_A is a parameter of the environment unknown to the population and is referred to as the *environment state*. Every population member makes n repeated observations of a Bernoulli trial with probability p_A, denoted by Θ_{p_A}, in order to estimate the environment state based on the number of occurences and non-occurences of A. Every individual has a beta-distributed *prior estimate* of p_A encoded as a binary genome $[\alpha, \beta]$. The mean of the prior of any individual with genome $[\alpha, \beta]$ is denoted by $\mu_{\alpha,\beta}$ and given by Equation 3.1. The individual uses the information accumulated from the repeated Bernoulli trials in order to update its prior to a Bayesian posterior. The individual's posterior estimate of p_A is denoted by $\phi_{\alpha,\beta}(p_A)$ and is given by Equation 3.2, where k is the number of observed occurrences of A after the n repetitions of Θ_{p_A}. The number of observations n is also a parameter of the model called the *learning length* and is constant for all individuals. The genes α and β are always restricted to the semi-closed interval $(0, 100]$ and in the initial population they are normally-distributed with mean 1.0 and variance 0.01. This is done to ensure that all the individual's priors are around $\frac{1}{2}$, by Equation 3.1, and therefore uninformative.

$$\mu_{\alpha,\beta} = \frac{\alpha}{\alpha + \beta} \tag{3.1}$$

$$\phi_{\alpha,\beta}(p_A) = \frac{k + \alpha}{n + \alpha + \beta} \tag{3.2}$$

The fitness of every individual is measured in terms of how accurately this individual estimates a non-empty multiset of m beta-distributed environment states, $S = \{p_1, p_2, p_3, ..., p_m\}$, called the *environment.* The beta distribution the m values in an environment S originate from is called the *environment distribution.* The *environment hyperparameters* are denoted by α_S and β_S. The fitness of an individual with genome $[\alpha, \beta]$ in environment S is denoted by $f_S(\alpha, \beta)$ and given by Equation 3.3.

$$f_S(\alpha, \beta) = \frac{1}{1 + \sum_{p_i \in S} [p_i - \phi_{\alpha,\beta}(p_i)]^2} \tag{3.3}$$

A new population is created by performing 100 matings with arithmetic crossover between members of the previous population. The parents in each mating are selected randomly from the previous population and each parent's probability of being selected is proportional to its fitness. The offspring are mutated by adding Gaussian-noise, with mean and variance equal to 0.00 and 0.01 respectively. The mutation rate is set to 0.01 in order to facilitate convergence, given that the population size is 100. Higher rates were found to produce less stability of the evolved genomes. The offspring make up a new population, which in turn replaces the previous one. Evolution runs for $10,000$ generations. The evolutionary model is briefly summarised in Figure 3.1.

Two types of evolutionary simulations are considered. In the first the environment S is replaced every $1,000$ generations with a new one containing the same number of beta-distributed environment states, with hyperparameters α_S and β_S. In the second, environment changes occur every $2,000$ generations. It can be expected that if the environment changes periodically then it is harder for individuals to keep up with these changes with evolved priors. An environment is more *stable* than another if the former changes less frequently than the latter.

The model, as described above, simulates a situation in which an individual needs to estimate the probability of an event before taking any action. For instance, A could refer to the event a predator is nearby, and the indivividuals would need to determine (not necessarily as a conscious act) how probable this event is in order to follow an appropriate

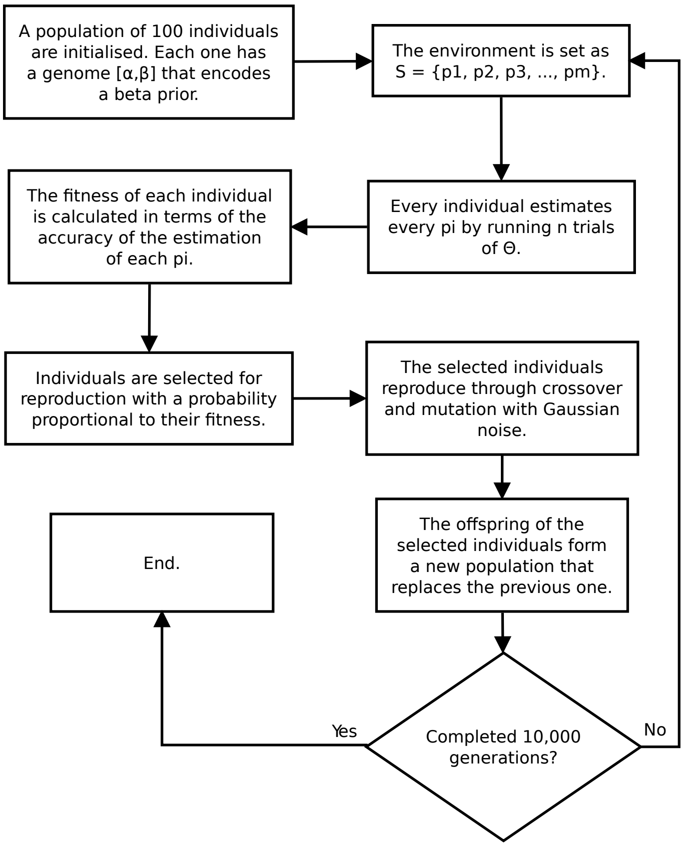


Figure 3.1: Summary of the Baldwinian evolution model of Bayesian learners.

course of action, such as developing defences (Agrawal et al., 1999; Hammill et al., 2008), as discussed in Chapter 2. The individual is more likely to use its resources efficiently if it makes the right decisions and this is more likely to occur with accurate estimates of p_A. A learner in nature would normally use the information gained from the environment in order to make a decision or take an action, such as *Daphnia pulex* individuals developing defences when chemical cues indicate the presence of predators, and this decision would likely have a repercussion in the survival or reproductive success of the decision-maker. For instance, if predators are frequent and a *Daphnia pulex* individual accurately deduces this then it will grow the defences required and this investment will be beneficial. Otherwise, if the individual fails to predict the presence of predators, its survival and reproduction will be compromised. In the model described in this section, however, every individuals assesses p_A but the decision this individual would make using this information is not simulated and neither are the consequences of the actions that would be taken by the learner. Instead fitness is calculated directly in terms of the accuracy of the estimates produced. This simplification is reasonable because in a real-world scenario it should be expected that the survival and reproductive success of the decision-maker is associated directly to the accuracy of its assessments.

3.3 The learning length as a selective pressure

Given an environment S, the *default error* of an individual x, with genome $[\alpha_x, \beta_x]$, is a measure of the discrepancy between x's prior and the environment states in S. It is denoted by $\delta_S(x)$ and given by Equation 3.4. On the other hand, x's *posterior error* is a measure of the inaccuracy of the estimations made by x. It is denoted by $\Delta_S(x)$ and given by Equation 3.5, where k_i is the number of occurrences of A observed by x after n repetitions of Θ_{p_i}, for each $p_i \in S$.

$$\delta_S(x) = \frac{1}{m} \sum_{p_i \in S} \left| \frac{\alpha_x}{\alpha_x + \beta_x} - p_i \right| \tag{3.4}$$

$$\Delta_S(x) = \frac{1}{m} \sum_{p_i \in S} \left| \frac{k_i + \alpha_x}{n + \alpha_x + \beta_x} - p_i \right| \tag{3.5}$$

The *accuracy* of an individual x when estimating S is inversely proportional to the posterior error of x in this environment, therefore an individual with the minimum posterior error in a population of Bayesians should have the maximum accuracy, relative to that of other members of the population. Every time x runs an observation trial of Θ_{p_i} the discrepancy between p_i and this individual's updated posterior estimate is likely to be reduced. For this reason the x's accuracy should be expected to increase with the learning length (*i.e.*, the number of allowed observation trials of Θ_{p_i} per individual), which is denoted by n. Therefore the fitness score of the learner should also increase with n, which is to be expected since with each observation the individual reduces the uncertainty about the environment state and its estimate approaches the correct value.

If an environment state is irrational then the minimum posterior error by any Bayesian in the genetic pool is understandably limited by n and the individual's estimate may never match the actual objective value. For instance, if $p_i = \frac{\pi}{10}$ then a Bayesian can only approximate this value and the accuracy of such estimate will improve with higher values of n without ever reaching the actual probability. But if the environment state is rational then a Bayesian might be able to estimate it with perfect accuracy after n observations of Θ_{p_i} if this individual's prior is close enough to p_i. For instance, if $p_i = 0.3$ then a Bayesian x with genome $[\alpha_x = 27, \beta_x = 63]$ could estimate this value with perfect accuracy after $n = 10$ trials of Θ_{p_i} during which $k = 3$ occurrences of A are observed. In any case, the minimum posterior error achievable in the population is limited by n and individuals who achieve it are the ones who receive the maximum fitness relative to others' and should be the ones most likely to reproduce. As n increases, selective pressure over

Bayesians is reduced because then more of these individuals are likely to estimate p_i with the maximum achievable accuracy. On the other hand, with lower values of n selective pressure is stronger and it should be expected that an increasingly limited number of Bayesians reproduce, whereas with higher values of n the selective presion is weaker and individuals with greater posterior errors should also be able to reproduce and pass on their genetically-encoded priors. In the evolutionary simulation presented in later sections (Section 3.5, Section 3.6, and Section 3.7) the learning length is set to $n = 2$, for speed and simplicity, in order to test a high selective pressure, and it can be predicted that a higher value would result simply in less selective pressure and less exclusivity of the priors evolved.

3.4 Types of environments

The uncertainty of the Bernoulli trial Θ_{p_i} increases with its variance and is maximum when $p_i = 0.5$ because then Θ_{p_i} is equivalent to the flipping of a fair coin. Similarly, the certainty posed by an environment S increases with the standard deviation of the environment distribution because then the environment states in S should be expected to spread more evenly in $[0, 1]$. An environment whose values tend to be concentrated in a small subinterval of $[0, 1]$ is said to be less uncertain than another one whose values are more evenly spread over this interval. Different choices of environment hyperparameters α_S and β_S should result in certain values p_i appearing more frequently in S than others from interval $[0, 1]$, thus different environment hyperparameters should result in environments with varying degrees of uncertainty. Four types of environment distributions are considered in the model and the probability density functions of these are plotted in Figure 3.2.

An environment S with beta hyperparameters α_S and β_S such that $\alpha_S, \beta_S > 1$ can be of one of three types, depending on the size of α_S relative to β_S. If $\alpha_S > \beta_S$ then S is said to be of *Type I*. On the other hand, S is said to be of *Type II* if $\alpha_S < \beta_S$. Otherwise, if $\alpha_S = \beta_S$, then S is said to be of *Type III*. The density functions of these three types of beta distributions are plotted in Figure 3.2 in blue, black, and green, respectively. With any of these three environment distributions, the fitness of any Bayesian learner x increases with its inherited prior's proximity to the mean of the environment distribution because then its default and posterior errors decrease.

If the hyperparameters α_S and β are such that $\alpha_S, \beta_S < 1$ and $\alpha_S = \beta_S$ then the environment distribution is said to be of *Type IV* and has a mean equal to $\frac{1}{2}$ and two modes, one equal to 0 and another one equal to 1, as illustrated in Figure 3.2 in red. Thus

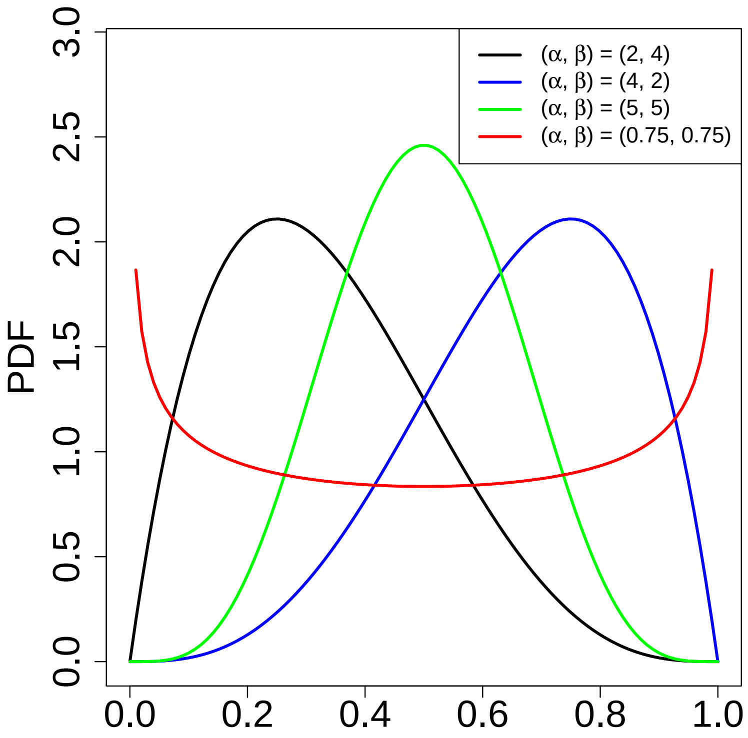


Figure 3.2: Shape of the density functions of the four types of environment distributions considered in the model. Each one of these is a beta distribution with hyperparameters α_S and β_S. An environment S is said to be of Type I when $\alpha_S, \beta_S > 1$ and $\alpha_S > \beta_S$; of type II when $\alpha_S, \beta_S > 1$ and $\alpha_S < \beta_S$; of type III when $\alpha_S, \beta_S > 0$ and $\alpha_S = \beta_S$; and of type IV when $\alpha_S, \beta_S < 1$ and $\alpha_S = \beta_S$.

environment states in S should be expected to be either very high (near 1.0) or low (near 0.0). Thus in the same environment S it should be expected that some environment states p_i make A occur very frequently and others make A occur very rarely. If a Bayesian learner x has a prior near the mean of the environment distribution (*i.e.*, near 0.5) its default error will still tend to be large when estimating every p_i. As the type IV hyperparameters decrease the environment states in S should be expected to be closer to either 0.0 or 1.0. On the other hand, as the type IV hyperparameters increase the environment states in S become more uniformly distributed in the interval $[0, 1]$. Given the above, the uncertainty posed by an environments of type IV is greater than that any of the other three types and it increases as α_S and β_S increase.

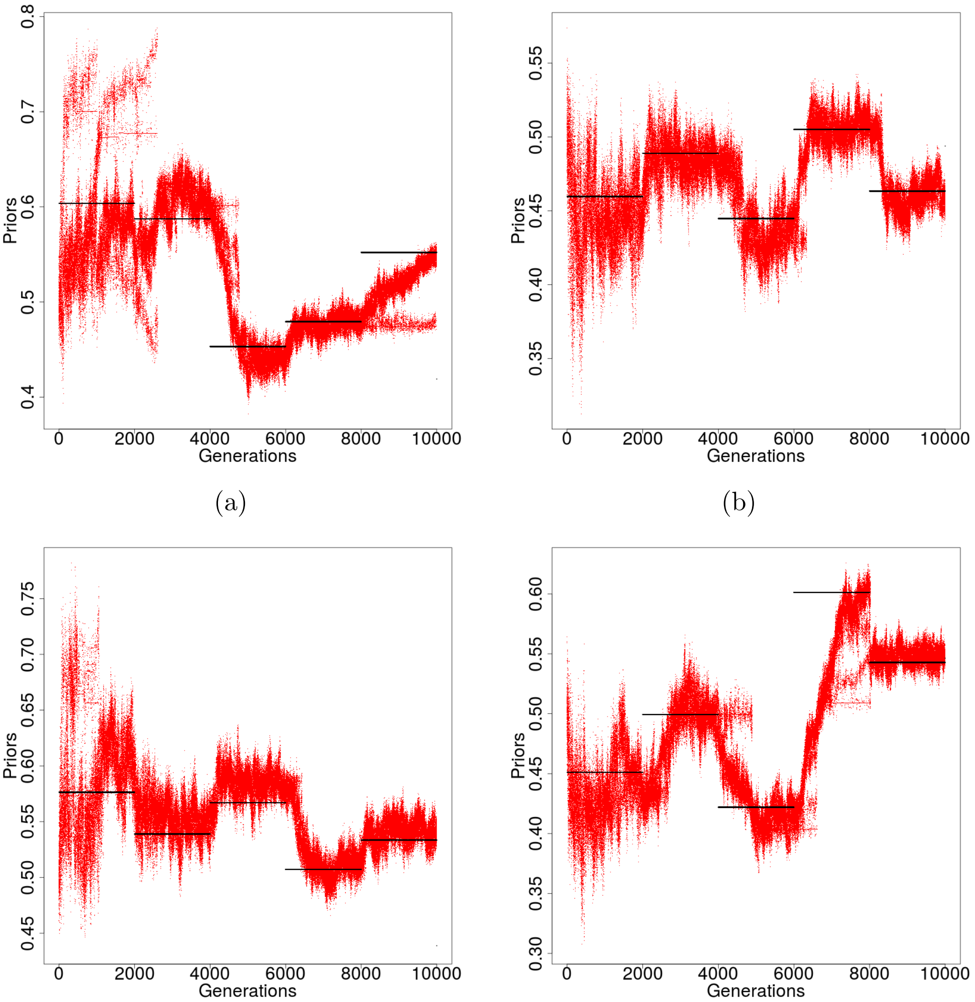
3.5 Evolutionary simulations to illustrate the selective pressure to estimate the environment state

Four evolutionary simulations were run with a population of Bayesians in an environment S of size 1 with hyperparameters $\alpha_S = 50$ and $\beta_S = 50$ (*i.e.*, an environment of type I) that changes every $2,000$ generations. The priors of the individuals as these evolve in the four simulations are plotted in Figure 3.3 as red dots whereas horizontal lines display the single environment state $p_1 \in S$ as it changes periodically.[1] The plots show that the population becomes rapidly dominated by learners whose priors are closer to the environment state p_1 (whose values are displayed as horizontal black lines) each generation. As evolution progresses and S changes every $2,000$ generations, the priors of the population change accordingly moving towards the current environment state becoming increasingly closer to p_1 in later generations. The Baldwin effect drives the evolution of the Bayesian learners because natural selection favours those who have the ability to 'learn' the correct p_1 by updating their genetically-encoded priors to a posterior that approximates the environment state with only $n = 2$ observations. Every individual's ability to learn is directly associated to its default error. Given any two individuals, x and y, if the former's default error is smaller than the latter's then y requires more training (*i.e.*, a greater learning length n) in order to estimate p_1 with the same accuracy as x.

The individuals that are favoured by natural selection in the simulations depicted in Figure 3.3 are those whose priors are close enough to p_1 in order for them to be able to estimate this environment state with the best accuracy in the whole population with only two observations of Θ_{p_1}. Even though these individuals eventually learn (*i.e.*, estimate approximately) the actual value of p_1, this 'knowledge' is not communicated to their offspring via genetic inheritance. In subsequent generations, these offspring must learn the environment state by themselves but they will be likely to have inherited from their parents the ability to learn and estimate the environment accurately with little training (*i.e.*, a prior that results in a default error that permits the estimation of p_1 with only two observations of Θ_{p_1}) and they are likely to pass on this ability to their own offspring. On the other hand, individuals whose priors are too far from the actual p_1 require more training in order to estimate the environment state and therefore are selected against until becoming extinct. As n increases, the selective pressure becomes weaker and individuals with priors further away from the environment state (*i.e.*, individuals with higher default errors) should be able to survive and reproduce.

---

[1] Technical details in Section C.3.



(a) (b)

(c) (d)

Figure 3.3: The learners' evolved Bayesian priors plotted as red dots each generation in an environment S of size 1 with hyperparameters $\alpha_S = 50$ and $\beta_S = 50$, as the environment state $p_1 \in S$ (horizontal black lines) changes periodically. Technical details in Section C.3.

3.6 Evolutionary simulations to illustrate the selective pressure to estimate the environment hyperparameters

Four evolutionary simulations were run with a Bayesian population in an environment S of size 1 with hyperparameters $\alpha_S = 25$ and $\beta_S = 75$ (*i.e.*, an environment of type II) that changes every $1,000$ generations. The α_x and β_x genes of every individual x in the population are plotted in Figure 3.4 as red and blue dots, respectively.[2] Each plot shows that from early generations the population evolves in such a manner that every learner x has a genome $[\alpha_x, \beta_x]$ such that the $\alpha_x < \beta_x$. This shows that natural selection drives the population to reflect the parameters of the environment distribution, which in this case are $\alpha_S = 25$ and $\beta_S = 75$. The genomes in the population do not necessarily match accurately these parameters but they do reflect the inequality relationship between α_S and β_S, *i.e.*, natural selection drives the population to estimate that $\alpha_S < \beta_S$. That is to say, the evolutionary search is able to discover that the environment distribution is of type II.

Analogous results are observed when equivalent sets of four simulations are run with environment hyperparameters $[\alpha_S = 95, \beta_S = 15]$ (type I), and $[\alpha_S = 75, \beta_S = 75]$ (type III), as shown in Figure 3.5, and Figure 3.6, respectively.[3] The plots show that natural selection favours individuals x whose genomes $[\alpha_x, \beta_x]$ exhibit the same equality or inequality relationship as the environment hyperparameters α_S and β_S. That is to say, when the environment hyperparameters satisfy $\alpha_S > \beta_S$ natural selection favours individual x if its genome is such that $\alpha_x > \beta_x$. Analogously, the same relationship between environment hyperparameters and evolved genomes is observed when $\alpha_S < \beta_S$ and when $\alpha_S = \beta_S$.

The results described above occur because the expected value of any $p_i \in S$ is given by $\overline{p_i} = \frac{\alpha_S}{\alpha_S + \beta_S}$ and the default error of any individual x with genome $[\alpha_x, \beta_x]$ tends to decrease as β_x approaches $\left(\frac{1-\overline{p_i}}{\overline{p_i}}\right)\alpha_x$ because then the individual's prior approaches $\overline{p_i}$ and the default error is smaller. More generally it can be predicted that:

1. When $\frac{\alpha_S}{\alpha_S + \beta_S} > 0.5$ (*i.e.*, when the environment is of type I) natural selection should favour every individual x whose genome $[\alpha_x, \beta_x]$ is such that $\alpha_x > \beta_x$.
2. When $\frac{\alpha_S}{\alpha_S + \beta_S} < 0.5$ (*i.e.*, when the environment is of type II) natural selection should favour every individual x whose genome $[\alpha_x, \beta_x]$ is such that $\alpha_x < \beta_x$.

---

[2] Technical details in Section C.4.

[3] Technical details to produce Figure 3.5 and Figure 3.6 appear in Section C.5 and Section C.6, respectively.

3. When $\frac{\alpha_S}{\alpha_S+\beta_S} = 0.5$ (*i.e.*, when the environment is of type III) natural selection should favour every individual x whose genome $[\alpha_x, \beta_x]$ is such that $\alpha_x \approx \beta_x$.

Four separate populations of Bayesians were evolved in a environment S of size 1 with hyperparameters $\alpha_S = 0.001$ and $\beta_S = 0.001$ (*i.e.*, an environment of type IV) that changes every $1,000$ generations. The priors of the learners as these evolve are plotted in Figure 3.7.[4] Figure 3.8 plots the α and β genes of the population as red and red blue, respectively.[5] The genomes in the population reflect the changes in the environment as they occur every $1,000$ generations. When the environment state $p_1 \in S$ is temporarily near 1.0 then the population becomes dominated with genomes $[\alpha_x, \beta_x]$ such that $\alpha_x > \beta_x$ whereas when the environment state is near 0.0 then the genomes of most individuals are such that $\alpha_x < \beta_x$. The individuals that are favoured by natural selection are those whose default error allows them to estimate the environment state p_1 with only two observations of Θ_{p_1}. When p_1 changes from being near 1.0 to being near 0.0 and vice versa new individuals with appropriate priors and default errors are selected. No single prior or genome type is observed to be prevalent over the course of the whole simulation because of the instability of the environment. Individuals are evolutionarily fit only for very specific and temporary environment states. In Section 3.7, evolutionary simulations are run with type IV environments of size greater than 1. It is shown that in these simulations natural selection favours individuals whose genomes reflect, to some extent, the environment hyperparameters.

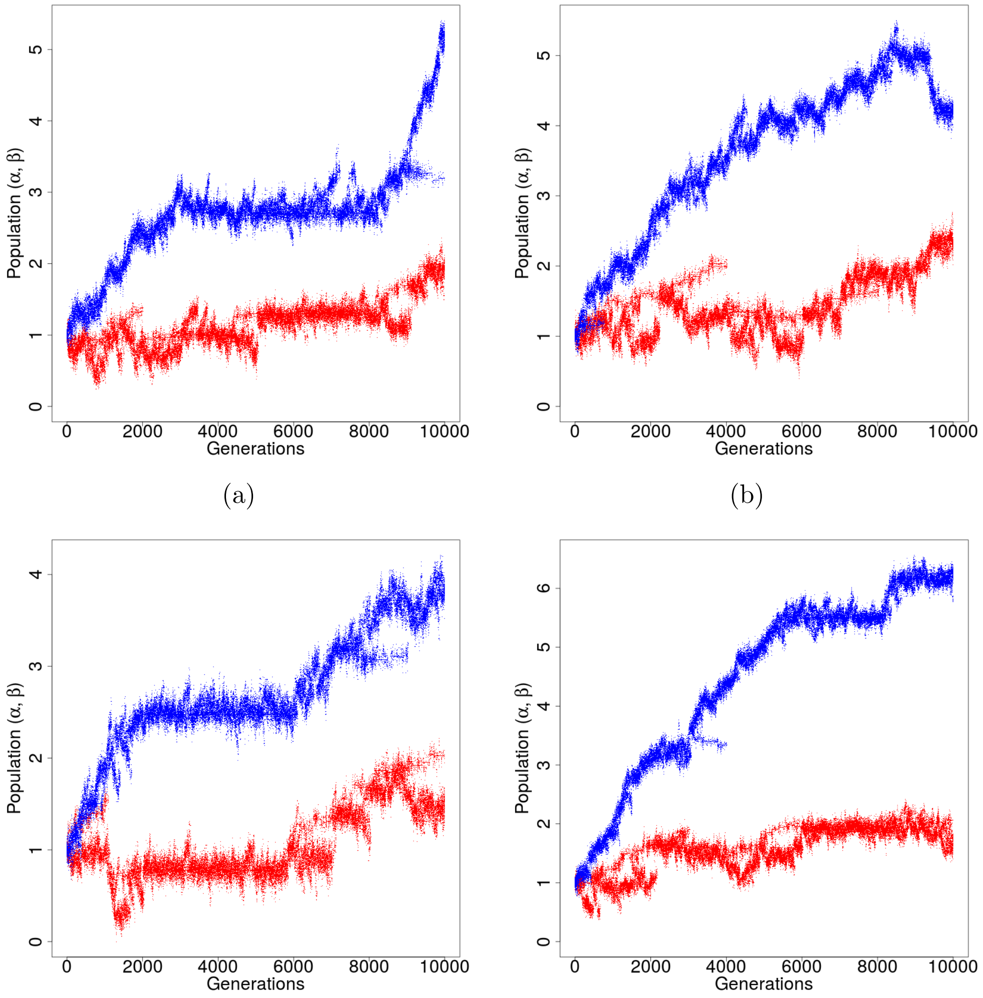
3.7 Evolutionary simulations to illustrate the selective pressure to estimate the hyperparameters of a type IV environment

Four evolutionary simulations were run with a population of Bayesians in a type IV environment of size 10 with environment hyperparameters $\alpha_S = 0.25$ and $\beta_S = 0.25$ that changes every generation, *i.e.*, the set of ten beta-distributed environment states each learner must estimate is different every generation. This simulates a scenario where a variable of the environment may exhibit some variability observed by an individual during its lifetime, as it may occur in a natural scenario, *e.g.*, foraging in different food patches of variable quality (McNamara, 1982). Figure 3.9 shows the population genomes

---

[4] Technical details to produce Figure 3.7 appear in Section C.7.

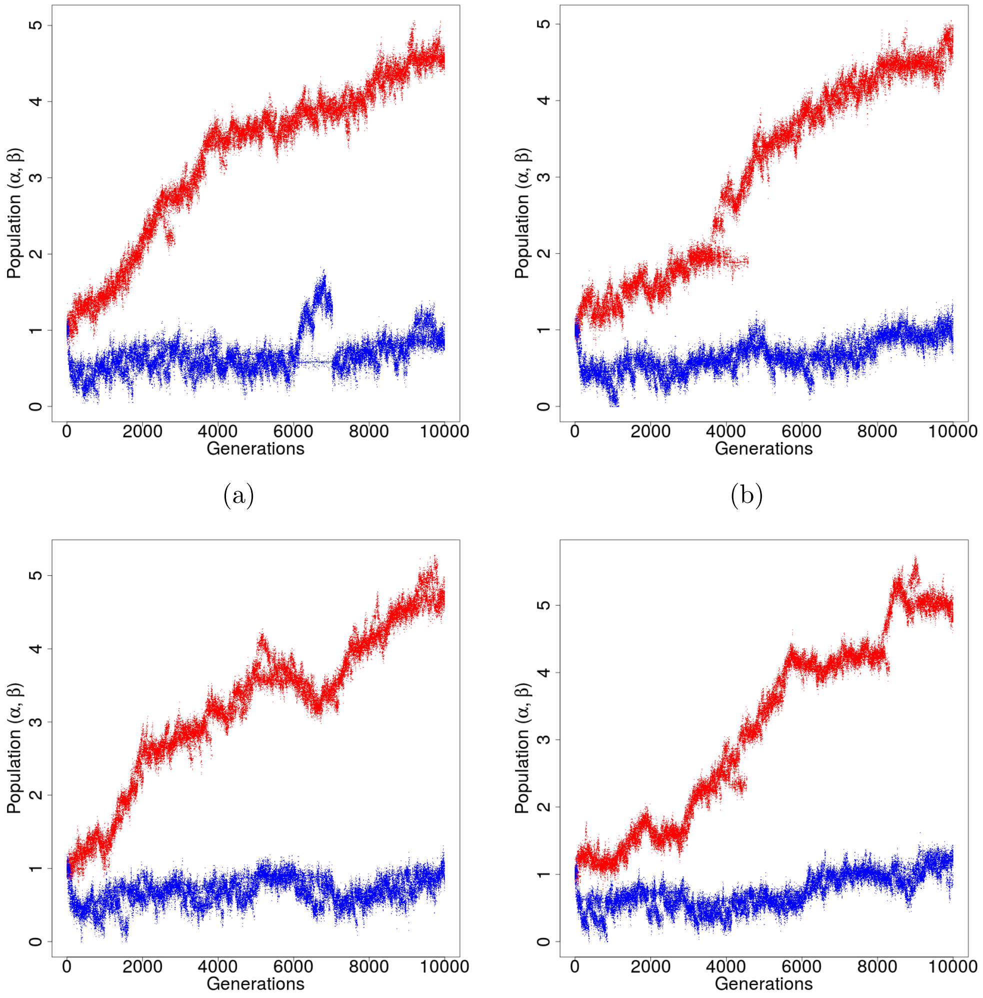
[5] Technical details to produce Figure 3.8 appear in Section C.7



(a) (b)

(c) (d)

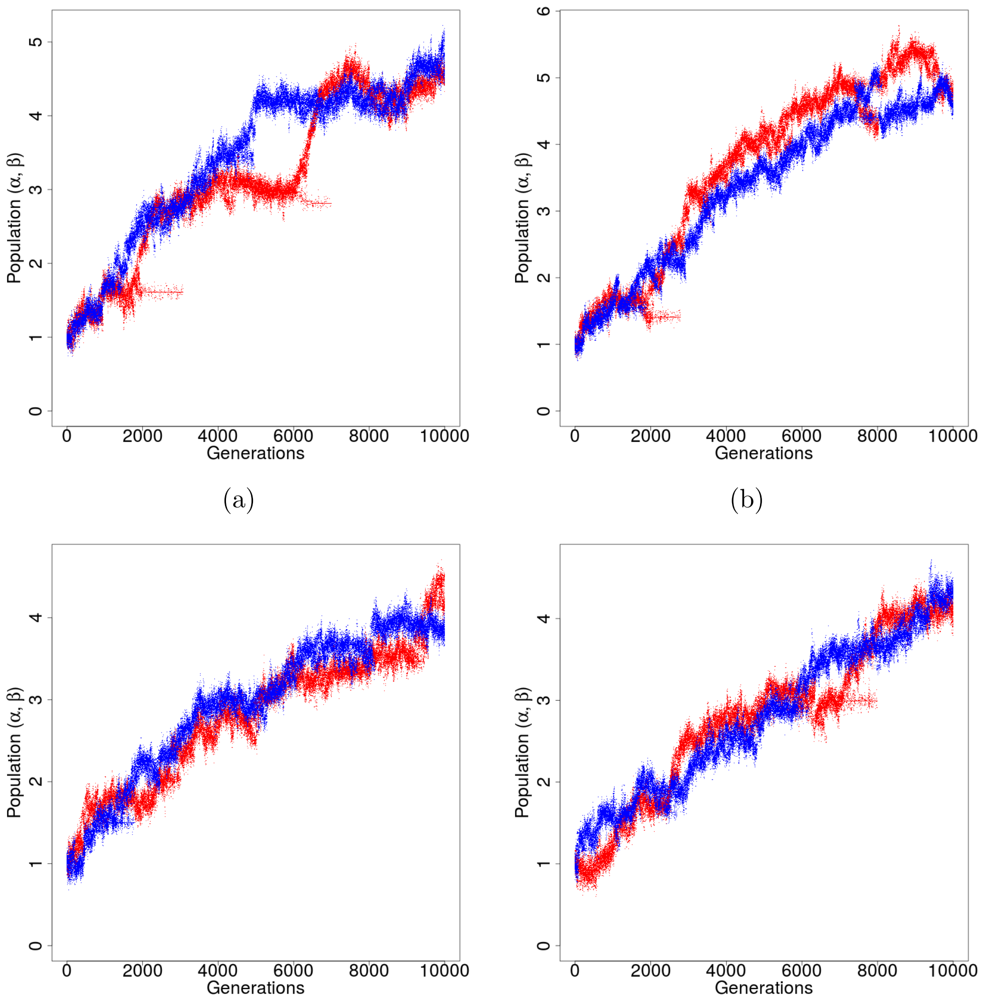
Figure 3.4: Evolved α (red dots) and β (blue dots) genes of a population of $N = 100$ Bayesian learners evolving over the course of $10,000$ generations in an environment S of size 1 with hyperparameter $\alpha_S = 25$ and $\beta_S = 75$ when the environment state $p_1 \in S$ changes once every $1,000$ generations. Details in Section C.4.



(a) (b)

(c) (d)

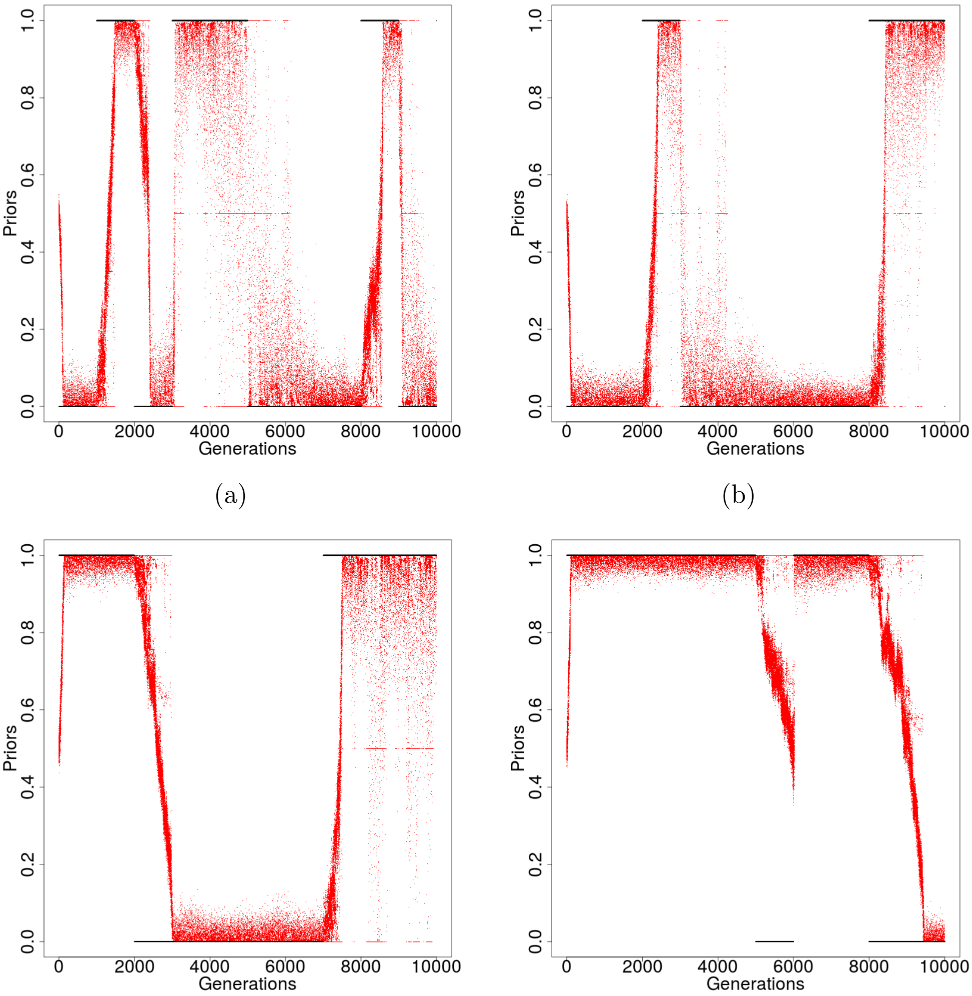
Figure 3.5: Evolved α (red dots) and β (blue dots) genes of a population of $N = 100$ Bayesian learners evolving over the course of $10,000$ generations in an environment S of size 1 with hyperparameters $\alpha_S = 95$ and $\beta_S = 15$ when the environment state $p_1 \in S$ changes once every $1,000$ generations. Details in Section C.5.



(a) (b)

(c) (d)

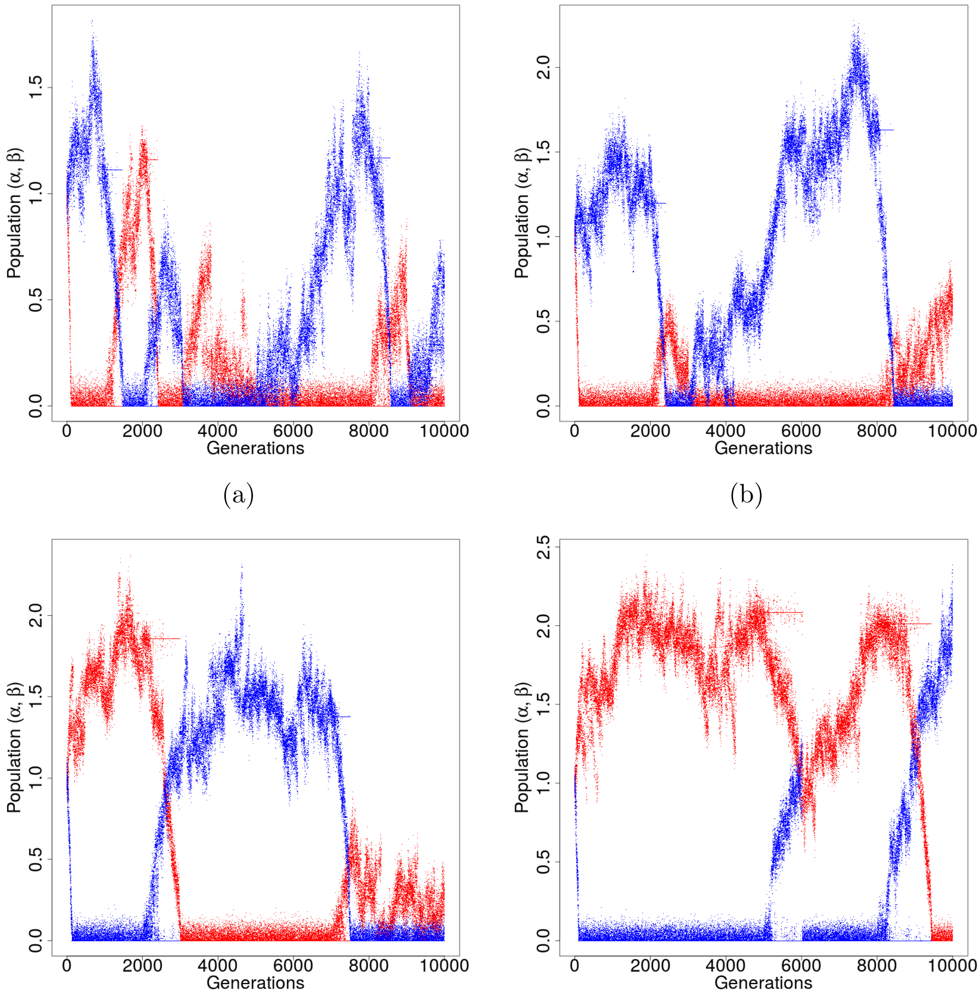
Figure 3.6: Evolved α (red dots) and β (blue dots) genes of a population of $N = 100$ Bayesian learners evolving over the course of 10,000 generations in an environment S of size 1 with hyperparameters $\alpha_S = 75$ and $\beta_S = 75$ when the environment state $p_1 \in S$ changes once every 1,000 generations. Details in Section C.6.



(a) (b)

(c) (d)

Figure 3.7: Evolved priors of a population of $N = 100$ Bayesian learners over the course of $10,000$ generations in an environment S of size 1 with hyperparameters $\alpha_S = 0.001$ and $\beta_S = 0.001$ when the environment state $p_1 \in S$ changes once every $1,000$ generations. The environment state is displayed as black lines. Details in Section C.7.



(a) (b)

(c) (d)

Figure 3.8: Evolved α and β genes in the population for the same four runs of the experiment displayed in Figure 3.7. These values appear to switch when p changes temporarily from values near 0.0 to 1.0 and vice versa. Distribution of the α (red dots) and β (blue dots) genes in the same four simulations displayed in Figure 3.7. Details in Section C.7.

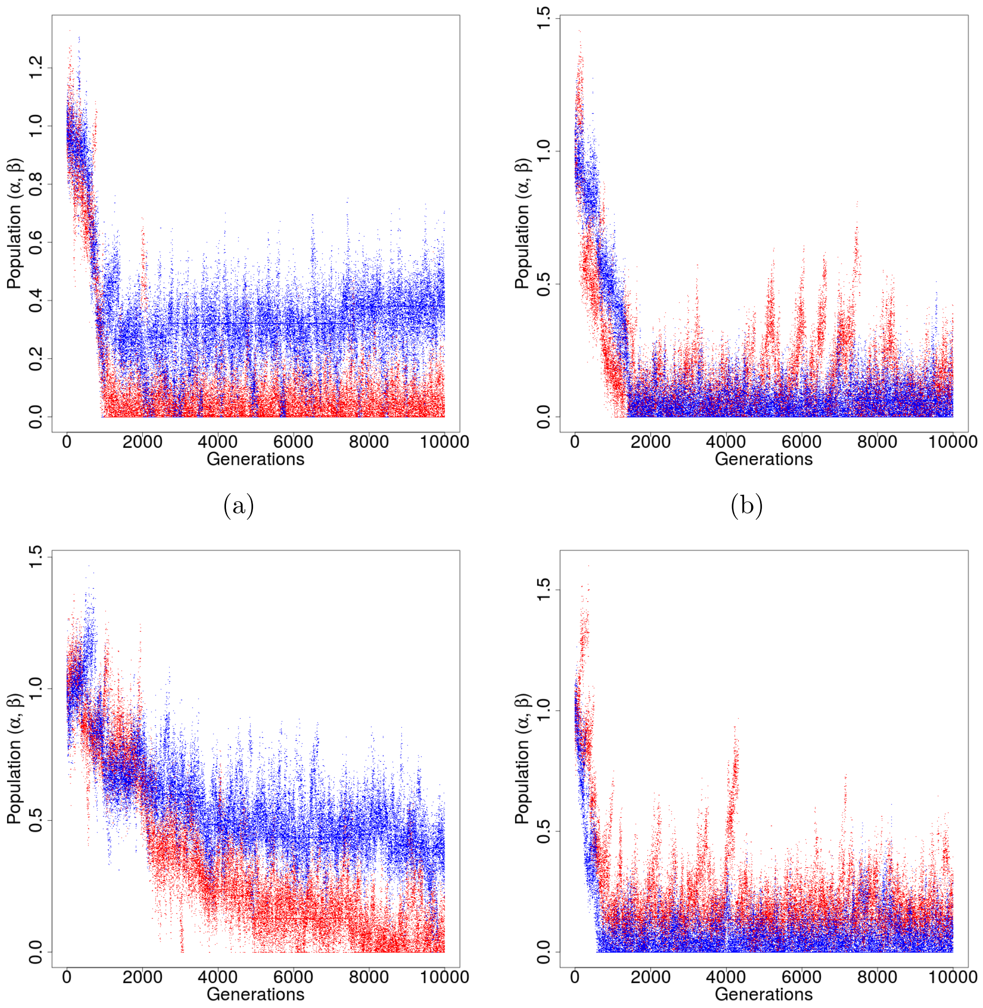
over the course of $10,000$ generations.[6] Figure 3.10 and Figure 3.11 show analogous results from simulations with environment hyperparameters $[\alpha_S = 0.25, \beta_S = 0.75]$ and $[\alpha_S = 0.75, \beta_S = 0.25]$, respectively.[7] The evolutionary history of Bayesians is not necessarily the same in any two simulations with the same environment but a common pattern can be identified. These results show experimentally that the components α and β of the population's genomes generally evolve to be below 1.0, or at least not far above 1.0. In this regard the evolved Bayesian genomes agree with the parameters of the environment distribution. The genes of the population do not necessarily match the actual parameters of the environment distribution but they do not drift indefinitely toward values increasingly greater than 1.0, as in the simulations run in the previous section with environments of types other than IV. This is because the selective pressure imposed by the learning length (n) makes individuals evolve genomes with the least default errors in the average case. In addition to this, by evolving lower hyperparameters α and β, the individuals are also evolving priors with higher standard deviations that match more closely the standard deviation (*i.e.*, the uncertainty) of the environment distribution, which is higher in type IV environments. The results of the four evolutionary searches indicate that the minimal default error in the average case is obtained by a Bayesian x if this individual's genome resembles the parameters of a type IV environment distribution. Therefore, through an evolutionary search for individuals who are able to learn a set of objective environment states, natural selection ends up favouring genomes that resemble the environment parameters, although with a degree of error.

The results described above occur because the standard deviation (and thus the uncertainty) of an environment of type IV is higher than that of any of the other types of environments considered in the model, as explained in Section 3.4. In the evolutionary simulations described above decision-makers evolve their beta hyperparameters to be below 1.0, and in many cases very near zero, because in this manner the resulting high standard deviations of their evolved priors reflect more closely the high uncertainty (*i.e.*, the high standard deviation) of the type IV environment. The evolutionary simulations show that by doing this the individuals are more capable to estimate accurately, on average, all the environment states. This result supports the hypothesis made in Chapter 2, that Bayesian decision-makers evolve their priors to have high standard deviations to match the uncertainty or unpredictability of the environment, *i.e.*, to match the high standard deviation

---

[6] Technical details in Section C.8.

[7] Technical details to produce Figure 3.10 and Figure 3.11 appear in Section C.9 and Section C.10, respectively.



(a) (b)

(c) (d)

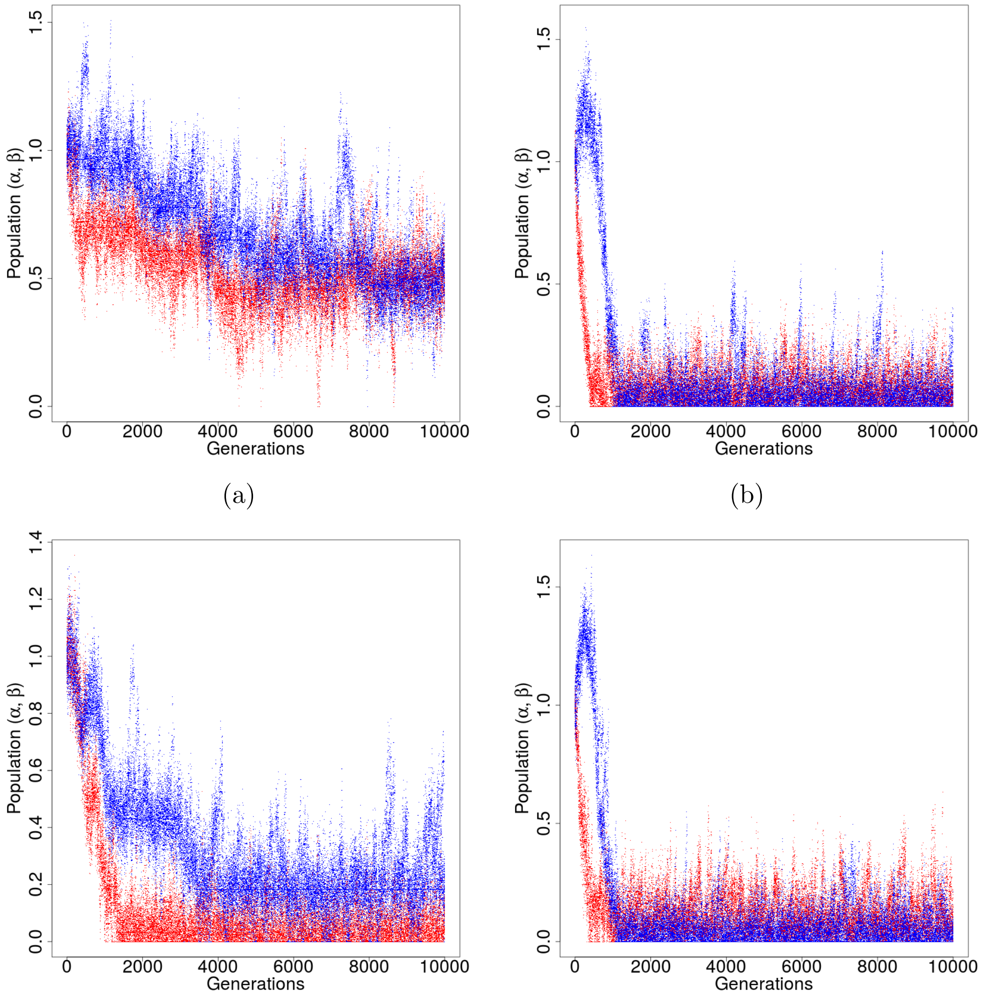
Figure 3.9: Evolved α (red dots) and β (blue dots) genes of a population of $N = 100$ Bayesian learners evolving over the course of $10,000$ generations in an environment S of size 10 with hyperparameter $\alpha_S = 0.25$ and $\beta_S = 0.25$ when the environment S changes every generation. Technical details in Section C.8.

of the observations (predation risk in this case) these decision-makers must estimate.

3.8 The Bayesian-Baldwin model with Bayesian and frequentist learners

In this section the model introduced in Section 3.2 is extended to include frequentist learners. The fitness of every Bayesian x is evaluated as shown in Equation 3.3. On the other hand, a frequentist y estimates each $p_i \in S$ as shown in Equation 3.6, where s is the number of ocurrences of A observed by the frequentist after n repetitions of Θ_{p_i}. Its fitness is calculated as shown in Equation 3.7.

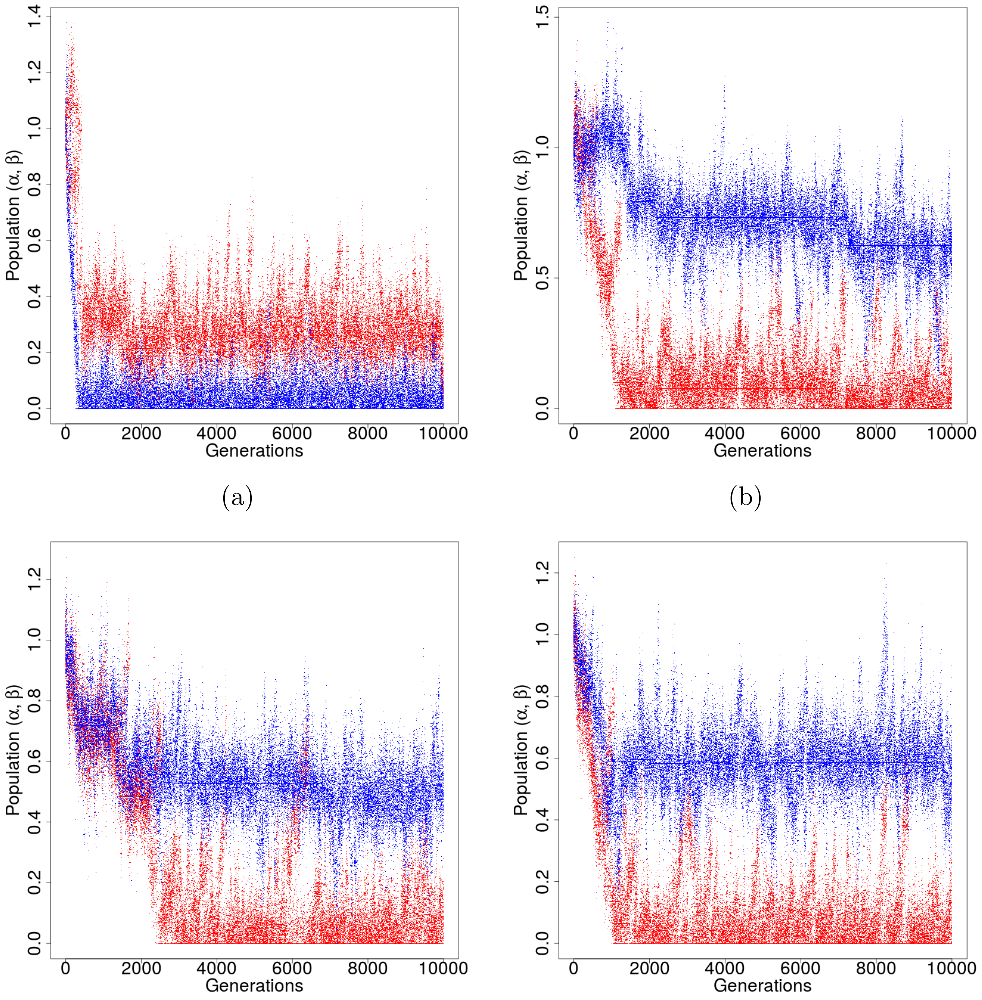
$$\phi_y(p_i) = \frac{s}{n} \tag{3.6}$$



(a) (b)

(c) (d)

Figure 3.10: Evolved α (red dots) and β (blue dots) genes of a population of $N = 100$ Bayesian learners evolving over the course of $10,000$ generations in an environment S of size 10 with hyperparameter $\alpha_S = 0.25$ and $\beta_S = 0.75$ when the environment S changes every generation. Technical details in Section C.9.



(a) (b)

(c) (d)

Figure 3.11: Evolved α (red dots) and β (blue dots) genes of a population of $N = 100$ Bayesian learners evolving over the course of $10,000$ generations in an environment S of size 10 with hyperparameter $\alpha_S = 0.75$ and $\beta_S = 0.25$ when the environment S changes every generation. Technical details in Section C.10.

$$f_S(y) = \frac{1}{1 + \sum_{p_i \in S}[p_i - \phi_y(p_i)]^2} \tag{3.7}$$

Selection, reproduction and other details of the evolutionary simulation are as described in Section 3.2. In this version of the model, each individual carries a third gene indicating whether the individual is Bayesian or frequentist. Thus, individuals of both types are mixed at reproduction and, as a consequence of this, crossover and mutation can sometimes result in Bayesian parents producing frequentist offspring and *vice versa*. A population is said to be *balanced* if it contains even proportions of individuals belonging to each group, Bayesians or frequentists. Otherwise it is said to be *unbalanced*.

Evolutionary simulations were run with different environment hyperparameters in order to determine in which of these environments Bayesians peform better or worse than frequentists. However, it was observed that in two or more evolutionary competitions run independently with the same environment the populations may not evolve in the same manner. For instance in two different runs it can occur that up to the $1,000$-th generation the population is balanced whereas in other repetitions of the same simulation one group is outnumbering the other. Therefore one single run of a simulation is insufficient evidence to reach a decisive conclusion about the evolutionary advantage or disadvantage of Bayesians against frequentists with a given environment. In order to gather enough statistics to reach a compelling conjecture regarding the likelihood that the Bayesian group outperforms its frequentist counterpart, each simulation with each environment was repeated a total of $R = 100$ times in order to perform a statistical test. The average Bayesian population in the r-th simulation run $(1 \leq r \leq R)$ with environment S is denoted by $\overline{b_r}(S)$ and calculated as shown in Equation 3.8, where $b_g(S)$ is the number of Bayesian individuals in the population at the g-th generation with environment S and $G = 1,000$ is the number of generations per run.

$$\overline{b_r}(S) = \frac{1}{G}\sum_{g=1}^{G} b_g(S) \tag{3.8}$$

Bayesians are said to have dominated the r-th run with environment S if $\overline{b_r}(S) > \frac{1}{2}N$, where $N = 100$ is the population size. The value $\overline{b_r}(S)$ is an acceptable measure for comparison of evolutionary success because the total population size is kept constant at $N = 100$.[8] After the $R = 100$ runs of the evolutionary competition are completed,

---

[8] Another possible statistic to take from each simulation would have been the final proportion of Bayesians and frequentists at the $1,000$-th generation in each run. This idea was rejected because it neglects Bayesian population figures that could have occurred in previous generations. In some cases, for example, it was observed that even though one group had had the largest population most generations, a sudden (perhaps temporary) change would take place towards the end of the simulation (relatively few generations before the $1,000$-th) switching the relative amounts of both groups even by a small margin. In such cases merely a measure of the final proportions would not have been very descriptive of the events

the *experimental Bayesian population* with each environment S is denoted by $\overline{B_S}$ and calculated as shown in Equation 3.9.

$$\overline{B_S} = \frac{1}{R}\sum_{r=1}^{R}\overline{b_r}(S) \tag{3.9}$$

Since the total population size is kept constant at $N = 100$ in every run then it follows that $\overline{B_S} \in [0, N]$, for any environment S. The experimental Bayesian population $\overline{B_S}$ serves as a measure of performance of Bayesians against frequentist in any environment S. If $\overline{B_S}$ is observed to be near N then it can be concluded that Bayesians are more evolutionarily successful than frequentists when both groups compete in environment S. On the other hand, if $\overline{B_S}$ is observed to be near zero then it can be concluded that Bayesians are at a disadvantage against frequentists in environment S. Finally, if $\overline{B_S} \approx \frac{1}{2}N$ then it can be concluded that neither group outperforms the other in environment S.

Section 3.9 presents the evolutionary simulations with competition between Bayesians and frequentists with different environments and the results show in which of these environments Bayesians dominate.

3.9 Competition between frequentist and Bayesians in a type IV environment

Repeated evolutionary competitions between Bayesians and frequentists were simulated in type IV environments of size 10 that change every generation. Bayesians were observed to have dominated in 74 of 100 runs with environment hyperparameters $[\alpha_S = 0.99, \beta_S = 0.99]$ and frequentists were observed to dominate in the remaining 26 runs. A binomial test was used to assess the statistical significance of this result by first assuming no prevalence of either Bayesians of frequentists as the default hypothesis, *i.e.*, by assuming the probability of each simulation resulting in Bayesian dominance is $p_0 = \frac{1}{2}$. Tests were performed with a significance level of 5%. Let X be the binomially-distributed variable representing the number of simulations dominated by Bayesians after $R = 100$ runs in an environment S of size 10 with hyperparameters $[\alpha_S = 0.99, \beta_S = 0.99]$, when S changes every generation. The probability that $X \geq 74$ after $R = 100$ simulations is given by $P(X \geq 74 | p_0 = \frac{1}{2}) = P(X \leq 26 | p_0 = \frac{1}{2}) = 8.33 \times 10^{-7}$. It can be concluded that there is strong evidence that Bayesians exhibit an evolutionary advantage over frequentists in environment S because $8.33 \times 10^{-7} < 0.05$.

---

that actually took place up to that point.

More generally speaking, when $R = 100$ simulations are run, the results are statistically-significant at the 5% level if $X \geq 59$ or $X \leq 41$, given that $P(X \geq 59|p_0 = \frac{1}{2}) = P(X \leq 41|p_0 = \frac{1}{2}) \approx 0.04$. When simulation were repeated $R = 100$ times in environments with hyperparameters $[\alpha = 0.90, \beta = 0.90]$, $[\alpha = 0.80, \beta = 0.80]$, $[\alpha = 0.70, \beta = 0.70]$ and $[\alpha = 0.60, \beta = 0.60]$, Bayesian dominance in the $G = 1,000$ generations was observed in 79, 70, 68 and 60 runs, respectively. Since these observations comply with the criteria for statistical significance mentioned above it can be concluded that there is strong evidence that with these environments Bayesians are more likely to overpopulate than their frequentist competitors. On the other hand, when simulations were run in environments with hyperparameters $[\alpha_S = 0.30, \beta_S = 0.30]$, $[\alpha_S = 0.20, \beta_S = 0.20]$, and $[\alpha_S = 0.10, \beta_S = 0.10]$ it was observed that $X < 41$. Therefore it can be generally expected that Bayesians outnumber frequentists when the environment is of type IV and the hyperparameters are above 0.60. On the other hand frequentists outnumber Bayesians when the hyperparameters are below 0.30.

Figure 3.12 plots $\overline{B_S}$ in repeated simulations run with different environment hyperparameters α_S and β_S in the interval $(0, 1]$, when the environment is of size 10 and changes every generation and every individual makes $n = 2$ observations when estimating each environment state.[9] The darker red areas in Figure 3.12 correspond to simulations where the environment is most uncertain, as described in Section 3.4. It is precisely in these environments where Bayesians are observed to have an advantage over frequentists as the experimental Bayesian population is above half the total population size. On the other side, the darker blue areas in Figure 3.12 correspond to simulations with environments that exhibit the least uncertainty, as described in Section 3.4. It is in these environments where frequentists are observed to have an advantage over Bayesians as the experimental Bayesian population is below half the total population size.

Figure 3.13 shows analogous results when $\alpha_S, \beta_S \in [10, 100]$ (*i.e.*, in environments I, II, and III).[10] The darker red areas in the plot pinpoint the regions where Bayesian superiority is greatest and these coincide with environments that exhibit the greatest uncertainty, as described in Section 3.4. On the other hand, the darker blue areas in the plot indicate regions where frequentists are more successful and these coincide with environments that exhibit the least uncertainty, as described in Section 3.4. This is in agreement with the known result that Bayesian learners should be favoured by natural selection when information is uncertain (Marshall et al., 2013b; McNamara et al., 2006;

---

[9] Technical details in Section C.11.

[10] Technical details in Section C.12.

McNamara and Houston, 1980).

The advantage Bayesians have over frequentists, as shown in Figure 3.12 and Figure 3.13, is partially due to the limited number of observations ($n = 2$) each individual is allowed to make when estimating each p_i in the environment. Any frequentist would be likely to infer each p_i with better accuracy if it were allowed to make a high enough number of observations of Θ_{p_i}. For instance, $n = 1,000$ tosses of a coin should yield enough information to make an approximate estimation of the probability of the coin landing heads. But even in this case a Bayesian is likely to have the advantage if its prior is close enough to p_i (which is likely to occur if the prior approximates the expected value of the environment distribution, when the environment is of type I, II, or III) because then the default error of this individual will tend to be small and its estimate of every p_i after the same number of observations should generally be more accurate than that computed by the frequentist. If natural selection favours accuracy in the estimates made by learners then for any number of allowed observations it should be expected that at some point a Bayesian would emerge (through mutation, for instance) with a prior close enough to p_i and who could compute p_i with better accuracy than a frequentist would. This is likely to occur with Bayesians whose priors are near the expected value of the environment distribution and these would be favoured by natural selection over frequentists.

Figure 3.14[11] and Figure 3.15[12] plot $\overline{B_S}$ in simulations with 10-sized environments S where individuals make $n = 5$ observations of Θ_{p_i} when estimating each $p_i \in S$. Figure 3.14 and Figure 3.15 show a pattern of relative Bayesian dominance that resembles that displayed in Figure 3.12 and Figure 3.13. That is to say, from both plots it can be appreciated that Bayesians generally see their evolutionary performance increased as the environment parameters approach each other and the environment becomes more uncertain. Conversely, Bayesians display a relatively lower performance as the distance between the environment parameters increases and the environment becomes less uncertain. However, the increased number of allowed observations reduces the selective pressure against frequentists because then frequentist estimates approach Bayesian estimates, thus the evolutionary search in the simulations depicted in Figure 3.14 and Figure 3.15 is unable to find Bayesians with better accuracy than in the simulations depicted in Figure 3.12 and Figure 3.13. This results in an increase in frequentist population figures between the two simulations, even though both are run with the same environments. For this reason, the experimental Bayesian populations ($\overline{B_S}$) observed in Figure 3.14 and Figure 3.15 are lower

---

[11] Technical details in Section C.13.

[12] Technical details in Section C.14.

than those in Figure 3.12 and Figure 3.13. It can then be generally expected that a higher number of allowed observations results in reduced selective pressure against frequentists but also that the overall pattern remains, namely that Bayesians exhibit an increased advantage as the parameters of the environment distribution increase and and approach each other. This is further confirmed in Figure 3.16[13] and Figure 3.17[14], which plots $\overline{B_S}$ in simulations where individuals make $n = 10$ observations of Θ_{p_i} when estimating each $p_i \in S$. In the two plots it can be observed that the experimental Bayesian populations are lower than those in Figure 3.14 and Figure 3.15, even though the same pattern depicted in the previous plots remains, namely that Bayesian population figures are higher in the environments that pose the most uncertainty.

The performance of a Bayesian individual relative to that of a frequentist when estimating an environment state p_i depends on how far the former's inherited prior estimate is from p_i. If the Bayesian's prior estimate and p_i are approximately equal or differ little then this individual is likely to estimate p_i accurately since the prior estimate acts as a head start. In this case, the prior provides an advantage the frequentist individual lacks. On the other hand, if the Bayesian's prior estimate and p_i differ greatly then the Bayesian may not succeed in updating this prior estimate to a posterior estimate matching the accuracy a frequentist would achieve in the same number of observations ($n = 2$). For this reason, the accuracy of the Bayesian individual increases as the difference between the prior estimate and p_i decreases. Since greater fitness is awarded to individuals that achieve greater accuracy then it should be expected that a Bayesian is at an advantage over frequentist opponents if the environment states these individuals face tend to differ little from the Bayesian's prior estimate. More generally, it should be expected that the fitness of a Bayesian individual increases relative to that of a frequentist opponent as the difference between the Bayesian's prior estimate and the expected value of the environment distribution decreases.

Figure 3.12 and Figure 3.13 show that the evolutionary performance of Bayesians against frequentists increases as the environment hyperparameters converge and decreases as these diverge. It can be theorised that one contributing factor to this result is the difference between the Bayesians' prior estimates at the beginning of each simulation and the mean of the environment distribution, since this results in decreased performance for Bayesians, as explained above. At the beginning of each simulation, Bayesians are approximately unbiased, *i.e.*, their prior estimates are equal to or very near 0.5, since the priors

---

[13] Technical details in Section C.15.

[14] Technical details in Section C.16.

are set to be uninformative. Greater convergence of the environment hyperparameters results in an environment distribution with a mean closer to 0.5 whereas greater divergence results in an environment distribution with a mean farther from 0.5. Because of this and the reasons explained above, at the beginning of each simulation Bayesians should be expected to perform worse when the environment hyperparameters diverge than when these hyperparameters converge. It is likely that the initial Bayesian underperformance occurred when hyperparameters diverge can create an early advantage to frequentists that persists until the end of the simulation, which is what is observed in Figure 3.12 and Figure 3.13.

In addition to the above, a Bayesian individual x's genome components are always positive (*i.e.*, $\alpha_x, \beta_x > 0$) therefore x's inherited prior estimate can never be zero or one. This is because x's prior estimate, given by $\frac{\alpha_x}{\alpha_x+\beta_x}$, is always restricted to the interval $(0, 1)$ when $\alpha_x, \beta_x > 0$. Therefore, a Bayesian individual is likely to have a default error when the environment states p_i tend to be always very near zero or always very near one, which occurs when there is great divergence between the environment hyperparameters. When there is great divergence between the environment hyperparameters, the environment poses minimum uncertainty because the Bernoulli events Θ_{p_i} observed by the individuals are very predictable since they almost always result in Bernoulli 'failures' or almost always result Bernoulli 'successes'. A frequentist that observes only Bernoulli 'failures' or only Bernoulli 'successes' estimates p_i as 0 and 1, respectively. On the other hand, a Bayesian that observes only 'failures' or only 'successes' uses this information to steer its non-zero and non-one prior towards a posterior that approximates 0 or 1, respectively, while never matching any of these values. As the divergence between the environment hyperparameters increases, p_i should generally approach either 0 or 1 more closely and in these cases a frequentist estimate equal to 0 or 1, respectively, becomes more accurate than a Bayesian estimate. In other words, as the divergence between the environment hyperparameters increases, the precision of frequentist estimates should increase compared to that of Bayesian estimates. It can be hypothesised that this phenomenon contributes to the increase in the evolutionary performance of frequentists when there is great divergence between the environment hyperparameters, in detriment of the performance of Bayesians, as shown in Figure 3.12 and Figure 3.13. In general, the results obtained indicate that two variables that impact the evolutionary dominance of Bayesians or frequentists are the environment uncertainty (*i.e.*, the unpredictability of environment states) and the learning length (*i.e.*, n, the number of observations made by each individual).

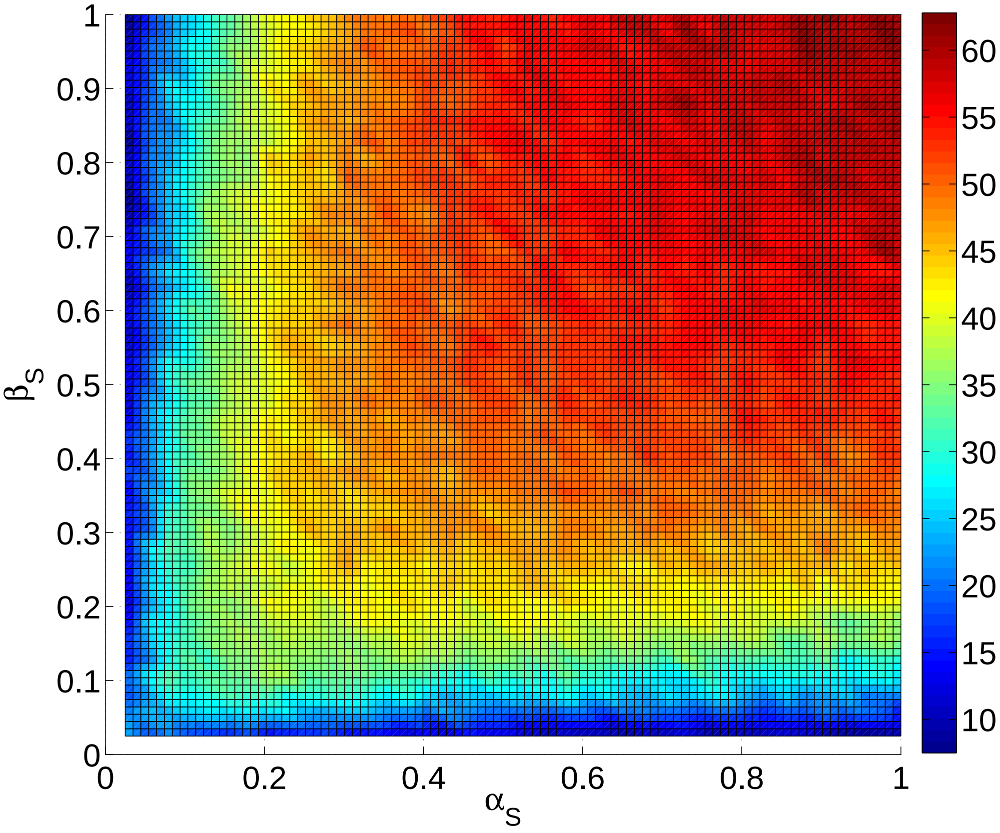


Figure 3.12: Experimental Bayesian populations observed in evolutionary competitions against frequentists in environments of size 10 with different hyperparameters $\alpha_S, \beta_S \in (0, 1]$ when the environment changes every generation. Every individual makes $n = 2$ observations when estimating each environment state. The axes indicate the environment parameters α_S and β_S and the colour indicates the experimental Bayesian population, $\overline{B_S}$ (where $R = 100$ is the number of runs with each environment and $G = 10,000$ is the number of generations in each run), in evolutionary equlibrium in the environment with hyperparameters α_S and β_S. The evolutionary advantage of Bayesians against frequentists is expressed by the experimental Bayesian population in equlibrium. If this measure is above half the population size then Bayesians are more evolutionarily successful. In environments where this value is below Bayesians are outperformed by frequentists. Technical details in Section C.11.

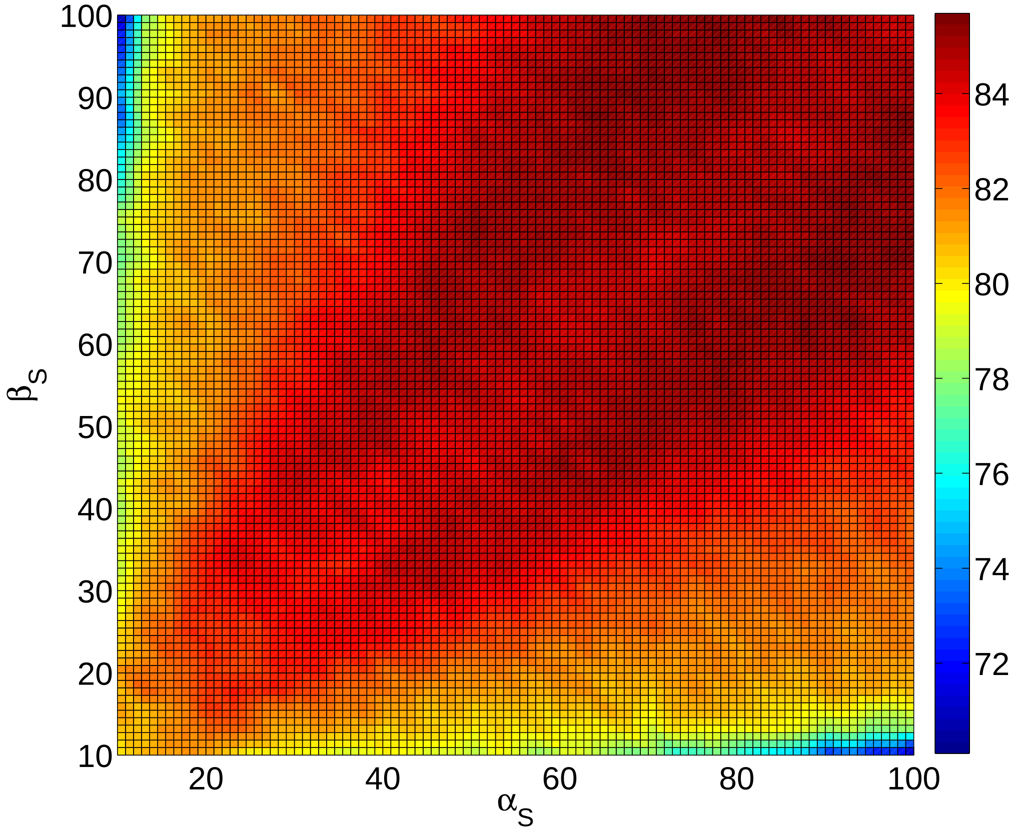


Figure 3.13: Experimental Bayesian populations observed in evolutionary competitions against frequentists in environments of size 10 with different hyperparameters $\alpha_S, \beta_S \in [10, 100]$ when the environment changes every generation. Every individual makes $n = 2$ observations when estimating each environment state. The axes indicate the environment parameters α_S and β_S and the colour indicates the experimental Bayesian population, $\overline{B_S}$ (where $R = 100$ is the number of runs with each environment and $G = 10,000$ is the number of generations in each run), in evolutionary equlibrium in the environment with hyperparameters α_S and β_S. The evolutionary advantage of Bayesians against frequentists is expressed by the experimental Bayesian population in equlibrium. If this measure is above half the population size then Bayesians are more evolutionarily successful. In environments where this value is below, Bayesians are outperformed by frequentists. Details in Section C.12.

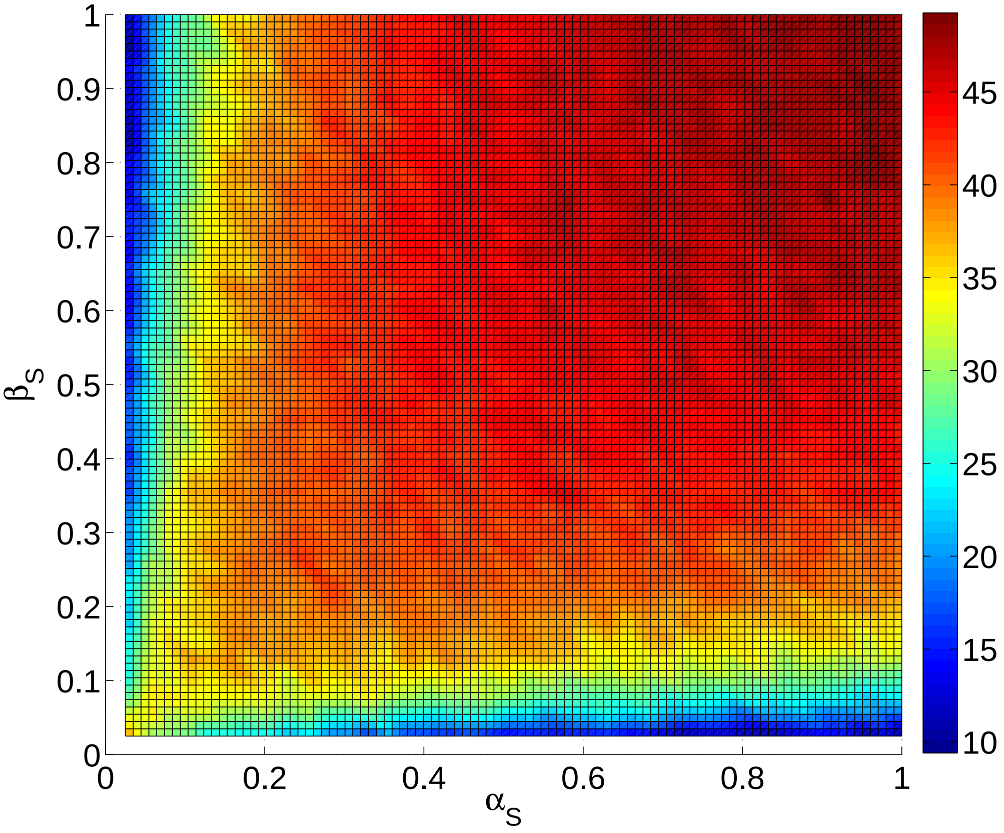


Figure 3.14: Experimental Bayesian populations observed in evolutionary competitions against frequentists in environments of size 10 with different hyperparameters $\alpha_S, \beta_S \in (0, 1]$ when the environment changes every generation. Every individual makes $n = 5$ observations when estimating each environment state. The axes indicate the environment parameters α_S and β_S and the colour indicates the experimental Bayesian population, $\overline{B_S}$ (where $R = 100$ is the number of runs with each environment and $G = 10,000$ is the number of generations in each run), in evolutionary equlibrium in the environment with hyperparameters α_S and β_S. The evolutionary advantage of Bayesians against frequentists is expressed by the experimental Bayesian population in equlibrium. If this measure is above half the population size then Bayesians are more evolutionarily successful. In environments where this value is below, Bayesians are outperformed by frequentists. Technical details in Section C.13.

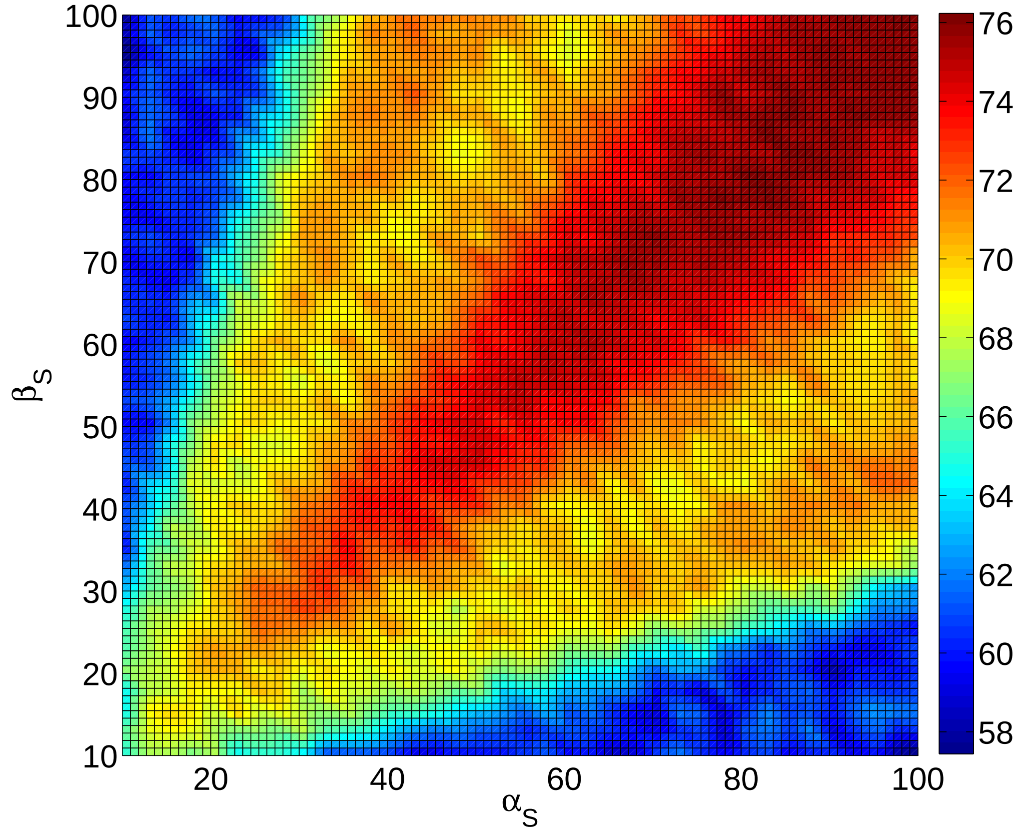


Figure 3.15: Experimental Bayesian populations observed in evolutionary competitions against frequentists in environments of size 10 with different hyperparameters $\alpha_S, \beta_S \in [10, 100]$ when the environment changes every generation. Every individual makes $n = 5$ observations when estimating each environment state. The axes indicate the environment parameters α_S and β_S and the colour indicates the experimental Bayesian population, $\overline{B_S}$ (where $R = 100$ is the number of runs with each environment and $G = 10,000$ is the number of generations in each run), in evolutionary equlibrium in the environment with hyperparameters α_S and β_S. The evolutionary advantage of Bayesians against frequentists is expressed by the experimental Bayesian population in equlibrium. If this measure is above half the population size then Bayesians are more evolutionarily successful. In environments where this value is below, Bayesians are outperformed by frequentists. Details in Section C.14.

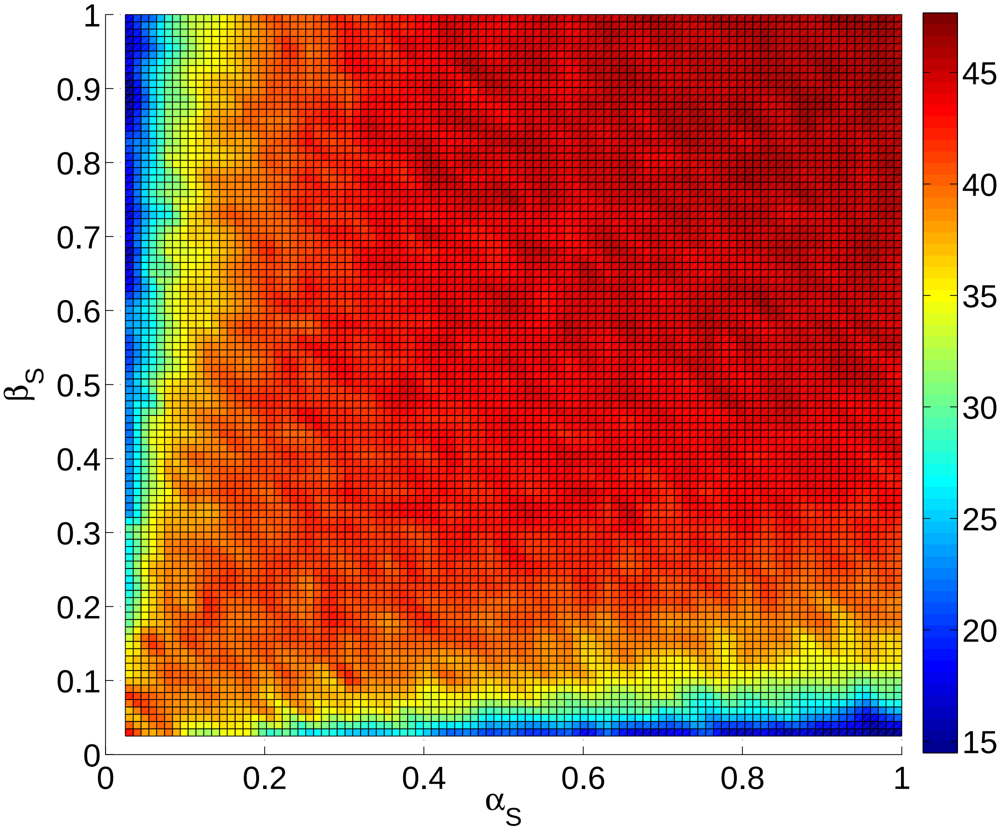


Figure 3.16: Experimental Bayesian populations observed in evolutionary competitions against frequentists in environments of size 10 with different hyperparameters $\alpha_S, \beta_S \in (0, 1]$ when the environment changes every generation. Every individual makes $n = 10$ observations when estimating each environment state. The axes indicate the environment parameters α_S and β_S and the colour indicates the experimental Bayesian population, $\overline{B_S}$ (where $R = 100$ is the number of runs with each environment and $G = 10,000$ is the number of generations in each run), in evolutionary equlibrium in the environment with hyperparameters α_S and β_S. The evolutionary advantage of Bayesians against frequentists is expressed by the experimental Bayesian population in equlibrium. If this measure is above half the population size then Bayesians are more evolutionarily successful. In environments where this value is below, Bayesians are outperformed by frequentists. Details in Section C.15.

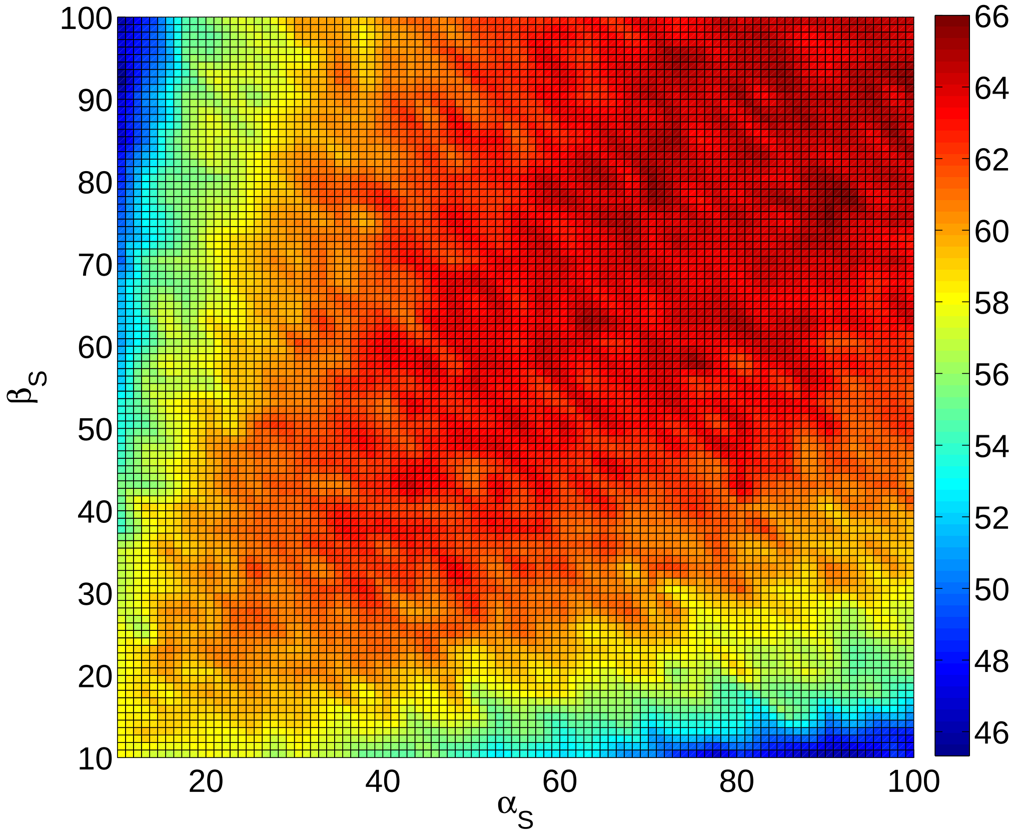


Figure 3.17: Experimental Bayesian populations observed in evolutionary competitions against frequentists in environments of size 10 with different hyperparameters $\alpha_S, \beta_S \in [10, 100]$ when the environment changes every generation. Every individual makes $n = 10$ observations when estimating each environment state. The axes indicate the environment parameters α_S and β_S and the colour indicates the experimental Bayesian population, $\overline{B_S}$ (where $R = 100$ is the number of runs with each environment and $G = 10,000$ is the number of generations in each run), in evolutionary equlibrium in the environment with hyperparameters α_S and β_S. The evolutionary advantage of Bayesians against frequentists is expressed by the experimental Bayesian population in equlibrium. If this measure is above half the population size then Bayesians are more evolutionarily successful. In environments where this value is below, Bayesians are outperformed by frequentists. Details in Section C.16.

3.10 Distribution of Bayesian populations in evolutionary equilibrium against frequentists

Evolutionary competitions between Bayesians and frequentists were repeated independently with different environments S of size 10 that change every generation. Each individual makes $n = 2$ observations of each environment state. Different environment hyperparameters α_S and β_S such that $\alpha_S = \beta_S$ were considered since these provide the environments with the highest uncertainty, as explained in Section 3.4. That is to say, environments with these hyperparameters should be expected to exhibit the greatest variability and thus should be less predictable by the individuals. With each pair of the environment hyperparameters considered, a total of $R = 50,000$ simulations were run independently. Each run was set to last $G = 10,000$ generations. The average Bayesian population measured at the end of each run, denoted by $\overline{b_r}(S)$ $(1 \leq r \leq R)$, was calculated and the frequency of each observation was plotted. Figure 3.18 shows the plotted frequencies when environment hyperparameters are below 1.0 and Figure 3.19 shows these data when the environment hyperparameters are above or equal to 1.0.

The subplots on Figure 3.18 and Figure 3.19 overall show that Bayesian dominance increases with the environment hyperparameters when these are equal. That is to say, if $\alpha_S = \beta_S$ then Bayesians become more evolutionarily successful than frequentists as α_S and β_S increase. This means that in environments of maximum uncertainty, Bayesian dominance increases as the uncertainty of the environment increases, *i.e.*, Bayesians become more evolutionarily successful than frequentists as the environment states become less predictable. These results are in agreement to those presented in Section 3.9, specifically those depicted in Figure 3.12. In Figure 3.18a most observations of the average Bayesian population gather just above 40, which is the approximate value shown in Figure 3.12 when $\alpha_S = \beta_S = 0.25$. The other subplots in Figure 3.18 and Figure 3.19 show that this mode increases with higher environment parameters, which is in agreement with the overall pattern observed in Figure 3.12 and Figure 3.13.

Evolutionary stability at the end of each simulation run was approximate only, since oscillation in the number of evolved Bayesians and frequentists was still observed towards the end of the $G = 10,000$ generations. However, this oscillation did not affect the overall dominance of one group over the other. This is shown graphically in Figure 3.20[15] and Figure 3.21.[16] These figures plot the number of evolved Bayesians and frequentists over

---

[15] Technical details in Section C.19.

[16] Technical details in Section C.20.

the course of $G = 10,000$ generations with different environment paramaters α_S and β_S. The number of Bayesian and frequentist individuals observed each generation are plotted in red and blue, respectively. Each plot depicts the population figures observed in only one run with each hyperparameter pair, but repeated runs exhibited similar results. Figure 3.20 plots the results using the same environment hyperparameters used in Figure 3.18 and Figure 3.21 plots the results using the same environment hyperparameters used in Figure 3.19.

Figure 3.20a shows that generally frequentists evolve to be more numerous than Bayesians when $\alpha_S = \beta_S = 0.25$. Oscillation is observed in the population figures of both individuals, however, frequentist dominance is observable. From the subplot it can be seen that Bayesian population figures oscillate around 40, which is in agreement with the results presented in Figure 3.18a. The subsequent subplots in Figure 3.20 and in Figure 3.21 show that the number of Bayesians increases as higher environment hyperparameters are used, despite the oscillation exhibited by these evolved population figures.

It was observed that this oscillation occurred even with longer simulation runs (*i.e.*, with more than $G = 10,000$ generations). The evolutionary stability of the Bayesian population after $G = 10,000$ generations was measured as follows. During the last $2,000$ generations of each simulation with the hyperparameters used in Figure 3.18 and Figure 3.19 the number of evolved Bayesians was measured each generation in order to calculate the average of these observations and their standard deviation.[17] This process was repeated ten times with each pair of environment hyperparameters. With each pair of environment hyperparameters the average number of evolved Bayesians was found to be approximately equal to the mode depicted graphically in the corresponding subplot in Figure 3.18 and Figure 3.19. That is to say, when $\alpha_S = \beta_S = 0.25$ the average Bayesian population during the last $2,000$ generations of each simulation was around 45 and the minimum observed of all the observations made. This average approximately agrees with the mode depicted in Figure 3.18a. On the other hand, When $\alpha_S = \beta_S = 4$ the average was around 72 and the maximum observed. This average approximately agrees with the mode depicted in Figure 3.19d. The standard deviation of these observations was observed to decrease as the environment hyperparameters increased. When $\alpha_S = \beta_S = 0.25$ this standard deviation was around 11 and the maximum observed. On the other hand, when $\alpha_S = \beta_S = 4$ this standard deviation was around 5 and the minimum observed. These measurements show that even though oscillation in the number of evolved Bayesians always occurs, the over-

---

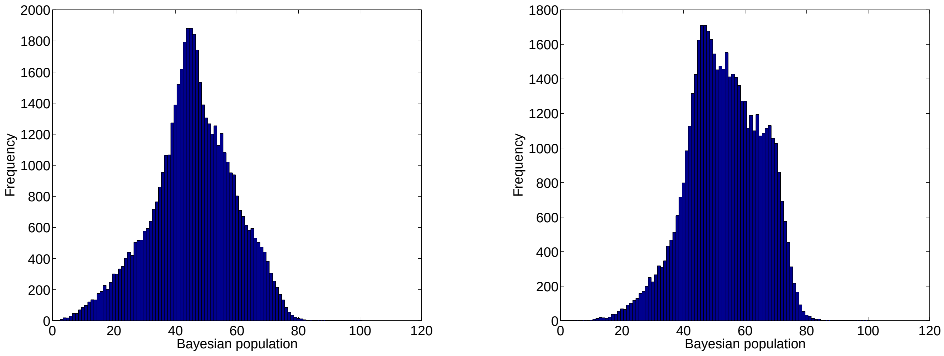
[17] Technical details in Section C.21.

all pattern described earlier is always observed. Bayesian dominance increases with the environment hyperparameters when these are equal. With smaller hyperparameters frequentists gain advantage and the Bayesian population figures in the last $2,000$ generations of each run exhibit less stability. With higher hyperparameters this stability increases and Bayesians become increasingly more dominant. Since a pattern of dominance by one group or the other was generally observable despite oscillating figures, the maximum number of generations was set to $G = 10,000$ in order to simplify the computation required to obtain the results presented in Section 3.9, as well as those shown in Figure 3.18 and Figure 3.19.

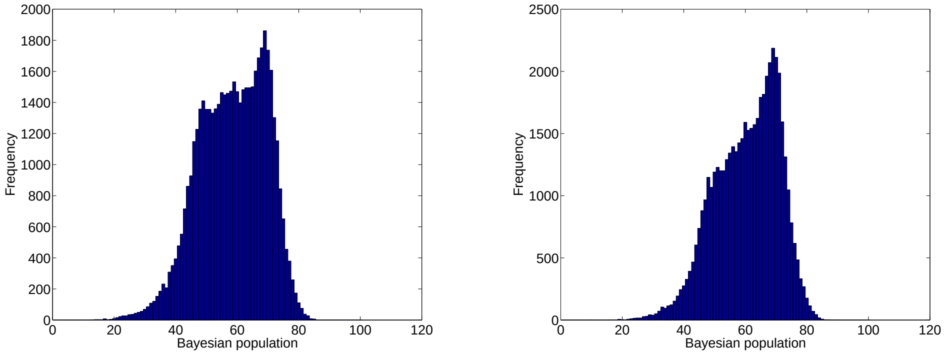
These results complement those presented in Section 3.9, that evolutionary equilibrium becomes more frequent with Bayesian dominance as the environment hyperparameters and the uncertainty posed by the environment grow. The results are also in agreement with the accepted proposal that natural selection should favour decision-makers that behave as if they were using Bayesian inference when information is uncertain (Marshall et al., 2013b; McNamara et al., 2006; McNamara and Houston, 1980).

3.11 Conclusions

The evolutionary model presented in Section 3.2 is proposed to show how Bayesian priors can be shaped by natural selection. In the model, individuals with genetically-encoded priors undergo a training process consisting of repeated Bernoulli observations in order to estimate the probability of an event. This probability is called the environment state. Each individual's genome is binary and consists of the hyperparameters of a beta distribution whose mean is the individual's prior estimate of the environment state. Every individual updates its prior to a posterior with each Bernoulli observation and fitness is measured in terms of the accuracy of the individual's posterior estimate after a limited number of observations. Evolutionary simulations are run where individuals in the initial population have unbiased priors and environment states are beta-distributed. These simulations show that the individuals evolve to have priors that approximate the environment state. This occurs because with these priors they can estimate the environment state with the best accuracy in the number of observations allowed. The individuals achieve this by evolving genomes whose two hyperparameters exhibit the same ratio as that of the environment hyperparameters. The Baldwin effect (Baldwin, 1896) occurs when the priors evolve as adaptations to the environment. The results presented in Section 3.5 and Section 3.6 support the hypothesis made in Chapter 2, that a decision-maker's prior can be shaped by natural selection to have any mean and standard deviation in response to the certainty

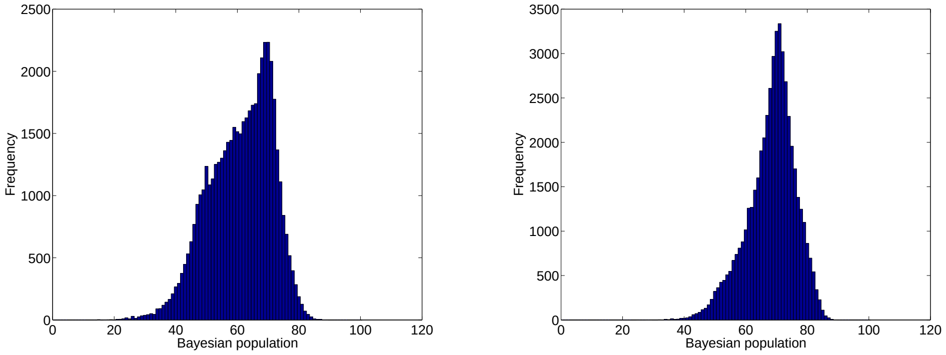


(a) $\alpha_S = 0.25$, $\beta_S = 0.25$ (b) $\alpha_S = 0.50$, $\beta_S = 0.50$

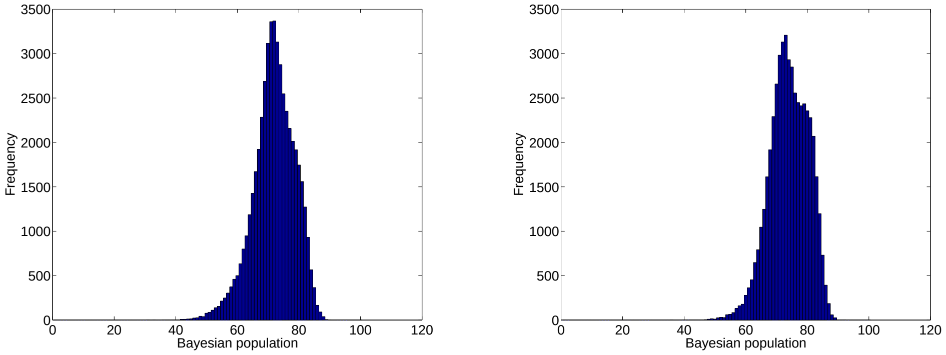


(c) $\alpha_S = 0.80$, $\beta_S = 0.80$ (d) $\alpha_S = 0.95$, $\beta_S = 0.95$

Figure 3.18: Frequency of observed values of the average Bayesian population, denoted by $\overline{b_r}(S)$, with four pairs of environment hyperparameters α_S and β_S such that $\alpha_S = \beta_S$ and $0 < \alpha_S, \beta_S < 1$. With each hyperparameter pair a total of $R = 50,000$ evolutionary competitions between Bayesians and frequentists are run independently. Each simulation runs for $G = 10,000$ generations. The environment S is of size 10 and changes every generation. Each individual makes $n = 2$ observations of each environment state. In each competition the total population size (including Bayesians and frequentists) is kept constant at 100. The four subplots show that Bayesian dominance becomes more frequent as the environment hyperparameters increase. Technical details appear in Section C.17.

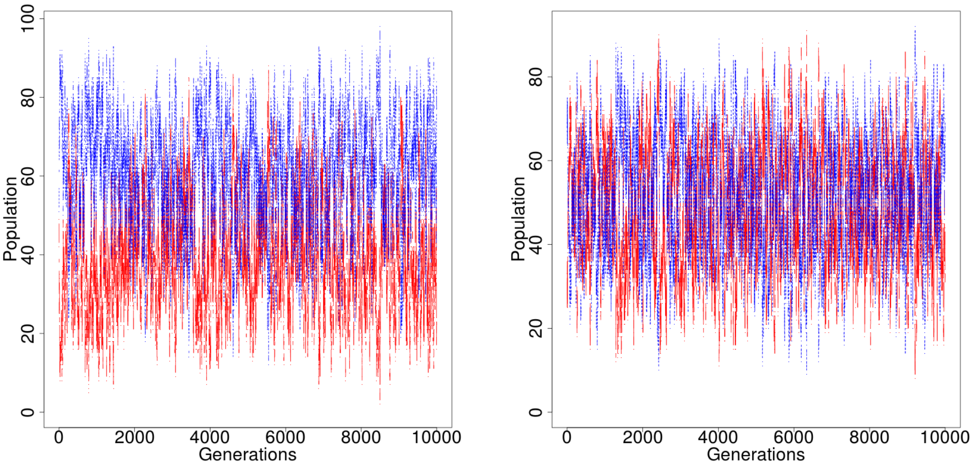


(a) $\alpha_S = 1$, $\beta_S = 1$ (b) $\alpha_S = 2$, $\beta_S = 2$

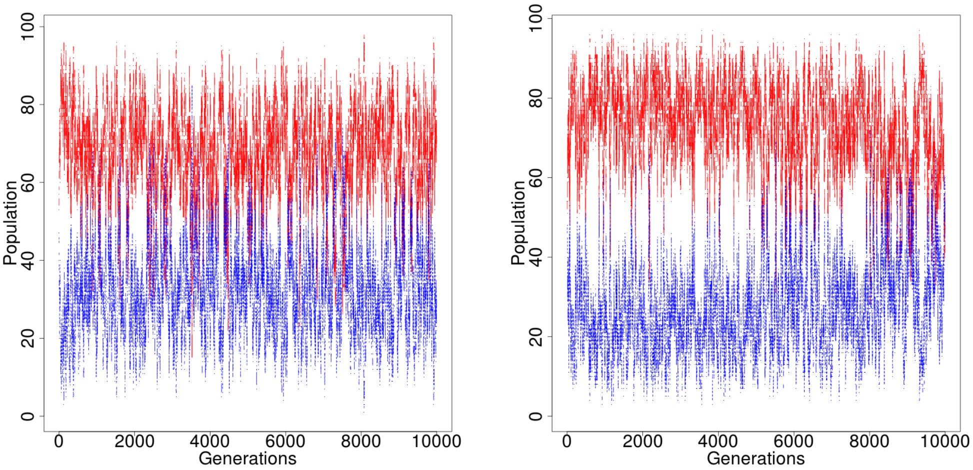


(c) $\alpha_S = 3$, $\beta_S = 3$ (d) $\alpha_S = 4$, $\beta_S = 4$

Figure 3.19: Frequency of observed values of the average Bayesian population, denoted by $\overline{b_r}(S)$, with four pairs of environment hyperparameters α_S and β_S such that $\alpha_S = \beta_S$ and $\alpha_S, \beta_S \geq 1$. With each hyperparameter pair a total of $R = 50,000$ evolutionary competitions between Bayesians and frequentists are run independently. Each simulation runs for $G = 10,000$ generations. The environment S is of size 10 and changes every generation. Each individual makes $n = 2$ observations of each environment state. In each competition the total population size (including Bayesians and frequentists) is kept constant at 100. The four subplots show that Bayesian dominance becomes more frequent as the environment hyperparameters increase. Technical details appear in Section C.18.

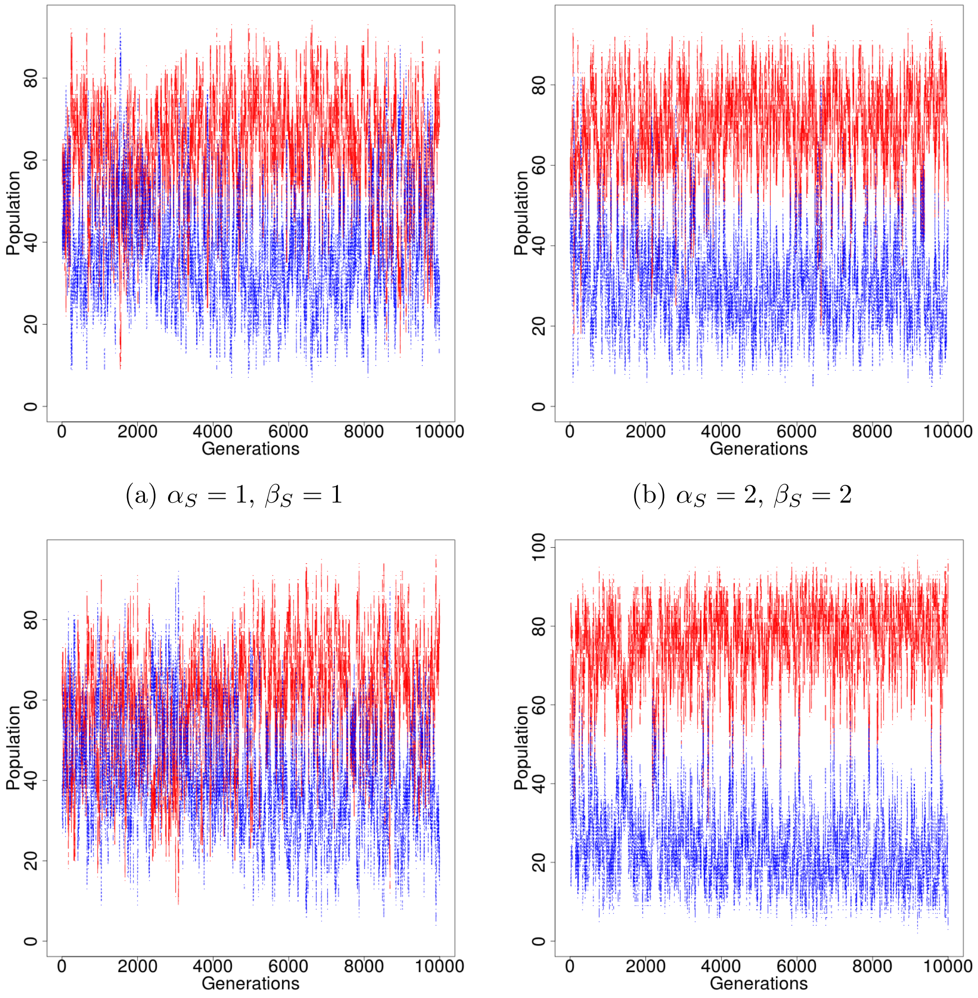


(a) $\alpha_S = 0.25$, $\beta_S = 0.25$ (b) $\alpha_S = 0.50$, $\beta_S = 0.50$



(c) $\alpha_S = 0.80$, $\beta_S = 0.80$ (d) $\alpha_S = 0.95$, $\beta_S = 0.95$

Figure 3.20: Number of Bayesians (red) and frequentists (blue) in evolutionary competitions between the two groups with different environment parameters α_S and β_S such that $\alpha_S = \beta_S$ and $0 < \alpha_S, \beta_S < 1$. Each simulation runs for $G = 10,000$ generations. The environment S is of size 10 and changes every generation. Each individual makes $n = 2$ observations of each environment state. In each competition the total population size (including Bayesians and frequentists) is kept constant at 100. Technical details appear in Section C.19.



(a) $\alpha_S = 1$, $\beta_S = 1$

(b) $\alpha_S = 2$, $\beta_S = 2$

(c) $\alpha_S = 3$, $\beta_S = 3$

(d) $\alpha_S = 4$, $\beta_S = 4$

Figure 3.21: Number of Bayesians (red) and frequentists (blue) in evolutionary competitions between the two groups with different environment parameters α_S and β_S such that $\alpha_S = \beta_S$ and $\alpha_S, \beta_S \geq 1$. Each simulation runs for $G = 10,000$ generations. The environment S is of size 10 and changes every generation. Each individual makes $n = 2$ observations of each environment state. In each competition the total population size (including Bayesians and frequentists) is kept constant at 100. Technical details appear in Section C.20.

of the environment.

The model presented in Section 3.8 extends the one introduced in Section 3.2 by including frequentist learners in the population and allowing an evolutionary competition between these and Bayesians learners when environment states are beta-distributed. For any environment state p_i, any Bayesian whose inherited prior is in close enough proximity to p_i may produce an estimate that is more accurate than the one a frequentist can produce in the number of Bernoulli observations all individuals are allowed to make, whatever this number is set to be. This close proximity between the prior and the objective p_i serves as a learning advantage that can be inherited because then the Bayesian will be able to make an accurate enough estimate with limited training. The Bayesian individual will therefore be more likely to reproduce than the frequentist and will also be likely to pass on its successful prior to its offspring. Since frequentists lack this advantage, they should tend to perform worse than Bayesians, especially if training is limited.

Evolutionary competitions between Bayesians and frequentists show that the former indeed perform better on average when the number of Bernoulli observations is limited and uncertainty is high, whereas the latter are more evolutionarily successful when uncertainty is minimal. In Section 3.4 it is shown that the uncertainty posed by the environment is given by the beta distribution of the environment states and increases as these become more evenly spread over the interval $[0, 1]$. Figure 3.12 and Figure 3.13 show that Bayesians outnumber frequentists when uncertainty is highest. These figures also show that frequentists are favoured by natural selection when uncertainty is minimal.

Scientific literature shows that decision-makers should evolve to behave in the same manner as Bayesian decision-makers if information is uncertain (Marshall et al., 2013b; McNamara et al., 2006; McNamara and Houston, 1980). This does not necessarily imply that individuals favoured by natural selection in nature are truly computing Bayesian estimates (Marshall et al., 2013a; McNamara et al., 2006; Tenenbaum et al., 2006) but rather that their decision-making machinery should lead them to make the same decisions as those made by a theoretical Bayesian. The results presented in this chapter provide evidence supporting this literature. In the evolutionary simulations, each individual's default 'worldview' (McNamara et al., 2006) is given by its inherited prior, and thus is a reflection of the environment experienced by its ancestors. The model shows how estimating the environment simply by being (or by behaving as) a Bayesian (*i.e.*, by updating a prior to a posterior) may not be enough to be evolutionarily successful if there is a selective pressure to produce accurate estimates efficiently (*i.e.*, with little training).

The model shows that in this case evolutionary success depends on how accurately the Bayesian individual's inherited 'worldview' reflects the environment. Competition between Bayesians and frequentists shows how Bayesians can use their evolved priors to outperform frequentists in an evolutionary race, as they are expected to when information is uncertain (Marshall et al., 2013b; McNamara et al., 2006; McNamara and Houston, 1980). This is evidence that the unpredictability of the environment should determine the decision-making strategy of an individual in nature, since both animals and humans are sensitive to the variability (i.e., the unpredictability) of each known choice when making decisions (Hayden and Platt, 2009; Kacelnik and Bateson, 1996).

In the model, individuals estimate the environment state, however, the model does not take into account what decision is made by the individual with the information learned, nor its consequences, and only awards fitness in terms of the accuracy of the computed estimate. However, in a natural scenario, an individual would normally make a choice or an assumption based on this information and its survival and/or reproduction rate would depend on the outcome of the choice/assumption made. For instance, successful foraging in patches could depend on the individual being able to infer the quality of the current patch (McNamara, 1982) and successful development into adulthood could depend on the individual being able to estimate the predation risk (*e.g.*, an individual facing a scenario similar to that of *Daphnia pulex*, described in Chapter 2). The results presented in this chapter show how a Bayesian individual should, in theory, evolve its biases in these situations and it should be expected that in an identical natural scenario decision-makers should evolve to make decisions as if they were using the same priors. Nonetheless, the model could be extended in the future to make more realistic assumptions by simulating full environment conditions and awarding fitness to individuals based on the actual outcome of the decisions made using the information learned.

Chapter 4

Self-deception can evolve under appropriate costs

This chapter has been published on *Current Zoology*, issue 61(2), pages 382 – 396.

4.1 Introduction

Deception in animals (no conscious intention being implied) refers to the signaling of false information from one individual to another and is normally beneficial to the signaller and detrimental to the receiver (Semple and McComb, 1996). For this reason some animals are observed to evolve strategies to deceive others, although natural selection is expected to also favour individuals who are able to 'see' through the deception. In addition to this, signaling may also be costly in order to be considered reliable, for example in mating situations (Zahavi, 1975). Arms races may occur between deceivers (no conscious intention being implied) and deception-uncovering species, with each group under selection to outsmart the other. Unlike deception, self-deception can be sensibly hypothesised not to be evolutionarily stable by itself, because animals who make decisions on false information seem more likely to make bad choices that could lead to negative consequences, such as injury and death. Especially in situations where conflict is likely it is sensible to expect that self-deceiving individuals tend to make suboptimal decisions, for instance risking injury through fighting a stronger opponent, and that in the long term they end up being less evolutionarily successful than others who use truthful information (Marshall et al., 2013b). Despite this, self-deception biases are claimed to occur frequently. For instance, animals and humans sometimes behave as if their subjective confidence in their skills in a given moment is above the objective measure of such attributes (Alicke and Govorun, 2005; McCormick et al., 1986; Pallier et al., 2002; Svenson, 1981). Surveys have shown most drivers rate their own skills as above average (McCormick et al., 1986; Svenson, 1981) and most students regard themselves as above-average leaders (Alicke and Govorun, 2005). It has also been documented that people who are unskilled for a task often fail to recognise their lack of competence, a phenomenon known as the Dunning—Kruger effect (Kruger and Dunning, 1999). Psychological tests have also shown that people tend to overestimate the probability of positive events (*e.g.*, career success) and to underestimate the probability of negative events (*e.g.*, onset of a serious illness) (Sharot, 2011b). Additional studies have shown that these optimistic expectations are not necessarily deterred by knowledge of past, realistic information. For instance, newly married couples tend to overestimate the likelihood of having long marriages despite reported divorce rates of around 50% (Sharot, 2011a). Some stroke patients, who are aware of their condition, have been observed to deceive themselves into thinking that their paralysis is due to factors other than their illness (Ramachandran, 1996). Similar studies have found that surveyed students also rate others as above average (Klar and Giladi, 1997). These results appear to show that individuals are generally unable to estimate correctly the average capability in a group (Brooks and Swann, 2011; Chambers and Windschitl, 2004) and that they have a tendency to overestimate the skills of others.

Self-deception has been defined as a misrepresentation of reality (Trivers, 2000). At its simplest, this would correspond to using a biased estimate of the probability of an event in decision-making. It generally comes in the form of a bias, which is a tendency to act prejudicedly or behave in a way that apparently does not conform to rationality. Biases can be of one of two types: cognitive biases, which are perceptual biases in the subjective experience of an individual, and behavioural biases, which are manifest in behaviours that depart from the optimal fitness-maximizing strategy (Marshall et al., 2013b). Behavioural biases can be diagnosed ignoring the internal decision-making machinery of the individual (McKay and Efferson, 2010). Cognitive biases, on the other hand, are generally hard to diagnose, and attempts to classify or explain them are often controversial (Dougherty et al., 1999; Marshall et al., 2013b). The apparent overconfidence exhibited by students and drivers in the surveys mentioned above are usually diagnosed by psychologists as an example of a cognitive bias (Alicke and Govorun, 2005; McCormick et al., 1986; Pallier et al., 2002; Svenson, 1981). Even though a bias may result in seemingly unreasonable behaviour it could evolve if, for instance, the bias is part or the by-product of a larger behavioural trait that overall proves to be individually advantageous. Thus cognitive biases may evolve given appropriate decision machinery, whereas behavioural biases should be expected not to.

It has been proposed by Trivers (Trivers, 2011; von Hippel and Trivers, 2011b) that the most evolutionarily successful deceivers in nature are those that self-deceive first. That is to say, *unconscious deceivers* (*i.e.*, those who unwittingly 'lie' to themselves just as they lie to others) are favoured by natural selection over *conscious deceivers* (*i.e.*, those who intentionally attempt to be deceitful while acting on truthful information). Trivers hypothesises that this is because conscious deceivers have to pay a considerable cognitive cost in order to avoid exhibiting involuntary responses that would allow others to see through the deception. Studies have shown, for instance, that humans tend to decrease their movement (in feet, legs, or hands) (Davis and Hadiks, 1995; DePaulo, 1992; Ekman, 1989; Ekman et al., 1991; Vrij, 1995) and to exhibit slow or confused speech as well as a higher pitched voice (DePaulo et al., 1985; Zuckerman et al., 1981, 1985) when lying consciously. Humans who act as conscious deceivers exhibit these involuntary responses, however, studies have shown that in some cases it is the opposite reactions (*e.g.*, increased movement in hands) that are often perceived as signs of deception (Vrij and Semin, 1996; Vrij et al., 1996). Trivers' hypothesis is that unconscious deceivers do not exhibit the same involuntary responses, since they believe the lie, and thus tend to be more successful cheaters and do not have to pay the same cognitive cost as conscious deceivers. Deceivers are likely to be penalised if their deception is discovered (Moller, 1987; Rohwer, 1977; Rohwer and Rohwer, 1978; Trivers, 1991), thus deceivers who hide the involuntary signs that reveal their deception should be more successful than those who do not. According to Trivers' theory a tendency towards self-deception evolves as a supportive by-product of the ability to deceive others, and the cost of lying to oneself is outweighed by the benefit brought by the ability to lie convincingly to adversaries. The theory proposed by Trivers has received extensive discussion from different commentators (Bandura, 2011) and Trivers has addressed these criticisms (von Hippel and Trivers, 2011a). One point that has been raised is that in situations of conflict, a deceiver may succeed in discouraging competitors from fighting (*e.g.*, by feigning a strength higher than the actual one) but it is likely that at some point the deception may be uncovered by others and that then the deceiver will face serious consequences, such as injury or death, as pointed out earlier (Frey and Voland, 2011; Funder, 2011; Marshall et al., 2013b). In such case the eventual cost of being discovered may be higher than the advantage posed by deceiving others, and self-deception should not evolve. This point has not been addressed by Trivers (Marshall et al., 2013b).

Section 4.2 introduces an extension to the model of the evolution of overconfidence proposed by Johnson and Fowler (2011) in order to investigate the evolution of self-deception given statistically-optimal behavioural machinery. With this it is shown analytically and computationally that under certain circumstances overconfidence evolves even when decision-makers use a theoretically optimal decision rule as suggested by Marshall et al. (Marshall et al., 2013b). In this case overconfidence or underconfidence are cognitive biases assuming a particular decision machinery, since they lead to optimal behaviour, rather than a sub-optimal behavioural bias (Marshall et al., 2013b). The new model is extended in Section 4.3 to incorporate deception biases in order to test Trivers' theory (Trivers, 2011) by showing that deception is favoured by natural selection when self-deception reduces cognitive or other costs. These self-deception biases are shown to be evolutionarily stable in a situation of conflict, one scenario not addressed by Trivers when replying to their critics (Marshall et al., 2013b; von Hippel and Trivers, 2011a).

In the model presented in Section 4.2 individuals do not attempt to deceive others because the purpose is to compare the self-deception biases with (model in Section 4.3) and without (model in Section 4.2) deception between individuals. Analysis and results of the models introduced in Section 4.2 and Section 4.3 are presented in Section 4.4 and Section 4.5, respectively. Supplementary information is presented in Appendix A.

4.2 A simplified owner-challenger model with internal biases

This section introduces an extension to Johnson and Fowler (J&F)'s model (Johnson and Fowler, 2011) of the evolution of overconfidence. In the extension, individuals self-deceive but do not deceive others. This extension, called the *simplified owner-challenger model*, is further extended in Section 4.3 to allow individuals to both self-deceive and deceive. The purpose of having the two models is to compare the level of self-deception that evolves in the absence of selective pressure to deceive others (in the simplified model presented in this section), and compare it with the level of self-deception that evolves when this selective pressure is present (in the generalised model presented in Section 4.3).

The definition of the simplified owner-challenger model is similar to that proposed by Johnson and Fowler (2011) and can be formulated as follows. Each individual has a fighting capability, denoted by θ. Given any two individuals, i and j, with capabilities

θ_i and θ_j respectively, the former would defeat the latter if $\theta_i > \theta_j$ should a conflict between them occur. The capability advantage i has over j is defined as $A = \theta_i - \theta_j$ and i defeats j if $A > 0$. As in the model originally proposed by Johnson and Fowler (2011), A is modelled as a standard normal random variable (*i.e.*, $A \sim \mathcal{N}(0, 1)$). This is achieved by making individual capabilities (θ_i and θ_j) normally distributed with mean zero and standard deviation $\sqrt{\frac{1}{2}}$.[1] The marginal probability that i defeats j is given by $p_W = P(A > 0) = \frac{1}{2}$.

Every individual i also has an *internal bias* (*i.e.*, a self-deception bias), denoted by k_i, that distorts its perception of its own capability in such a manner that i always acts as if its capability is $\theta_i + k_i$. In addition to this, i's *perception* of j's capability, denoted by $\widehat{\theta_j}$, is normally-distributed with mean θ_j and standard deviation σ_ϵ (*i.e.*, $\widehat{\theta_j} \sim \mathcal{N}(\theta_j, \sigma_\epsilon)$). The *perception error size*, σ_ϵ, is a non-negative parameter of the model. In this manner the model simulates perception errors as they occur in nature, which are due in part to environmental factors beyond the control of each individual, as well as being due to sensory noise. In this manner the advantage i *perceives* it has over an opponent j is given by $\widehat{A} = \theta_i + k_i - \widehat{\theta_j}$.

A conflict between two individuals over a resource occurs in an *owner-challenger encounter* as shown in Figure 4.1, where r is the value of the contested resource and c is a cost both individuals pay if they fight. The encounter involves the *owner* of the resource, who arrives at it first, and a *challenger*, who arrives subsequently and decides whether to claim the resource. If the challenger claims the owner decides whether to fight for the resource or abandon it to the challenger. Both r and c are constant and positive, and each individual decides in sequence whether to fight or not.

In a more realistic scenario, r, c, and σ_ϵ would likely vary from individual to individual and from encounter to encounter. For instance, an individual who has collected many resources will value a newly encountered resource less than an individual who has collected none. Similarly, the cost of a fight will probably be higher for an individual who has been injured badly from previously lost fights than for an individual who has lost none. In addition to this, in a natural scenario perceptual capabilities as well as conditions of the local environment (*e.g.*, low visibility that affects the ability to visualise the opponent) would be likely to vary from encounter to encounter, resulting in different perception errors between individuals. However, r, c, and σ_ϵ have been kept constant among all individuals for simplicity because in this manner the formal analysis of the model (Section 4.3) and

---

[1] If X and Y are normally distributed random variables, *i.e.*, $X \sim \mathcal{N}(\mu_X, \sigma_X)$ and $Y \sim \mathcal{N}(\mu_Y, \sigma_Y)$, then $Z = X + Y$ is also normally distributed with mean $\mu_X + \mu_Y$ and standard deviation $\sqrt{\sigma_X{}^2 + \sigma_Y{}^2}$.

a further extension (Section 4.4) of it are much more manageable and by doing so it is easier to identify the minimal conditions that are sufficient for the evolution of the biases of interest.

There are two differences between the owner-challenger model and the one proposed by J&F (Johnson and Fowler, 2011). The first is that in J&F's model the two competing individuals decide synchronously whether to fight or not whereas in the owner-challenger model the two decisions are made asynchronously and in sequence. The second is that in J&F's model individuals make their decisions based only on their perceived advantage over their respective opponents while ignoring the benefits (r) and costs (c) of each decision, whereas in the owner-challenger model these variables are taken into account by each individual. In this manner, the owner-challenger model addresses two main criticisms of J&F's (Johnson and Fowler, 2013; Marshall et al., 2013a,b). The first one is that since in J&F's model contests over resources are synchronous they can lead to valuable resources remaining unclaimed if no individual chooses to contest them, while the second is that in J&F's model individuals use an arbitrary and unrealistic decision rule in deciding whether to contest (Marshall et al., 2013a,b). The first criticism is addressed by allowing individuals to use whether they arrived at a resource first or second to determine their strategy, thereby creating a uncorrelated asymmetry (Maynard Smith, 1982) and allowing low value resources to be claimed by one individual. The second criticism is addressed by enabling individuals to use the estimated payoffs associated with different outcomes, and an estimate of the probability of winning, to determine whether to contest a resource.

A realistic scenario in the owner-challenger model is that where both individuals use all the relevant information when making their respective decisions. However the mathematical analysis of the model becomes difficult when both decision-makers behave in this manner. For this reason a simplified version of the model is analysed first, where the decision of the challenger j is always to claim the resource and fight whereas the owner i makes its decision (after having been challenged by j) by using the following reasoning; first i estimates its own probability of winning as $\widehat{p_W} = P(\widehat{\theta_j} < \theta_i + k_i)$. Then i estimates its expected payoff from the hypothetical fight as $\widehat{F} = \widehat{p_W}(r-c) + (1-\widehat{p_W})(-c) = \widehat{p_W}r - c$. This individual then decides to defend the resource if and only if this estimated payoff is higher than zero. This in turn occurs if and only if $\widehat{p_W} > \frac{c}{r}$ (Marshall et al., 2013b). This decision rule is rational from the perspective of an owner because it uses all the relevant information available to estimate the expected payoff from a fight and the final decision is made if and only if the evidence suggests that this estimate is positive. In the long run

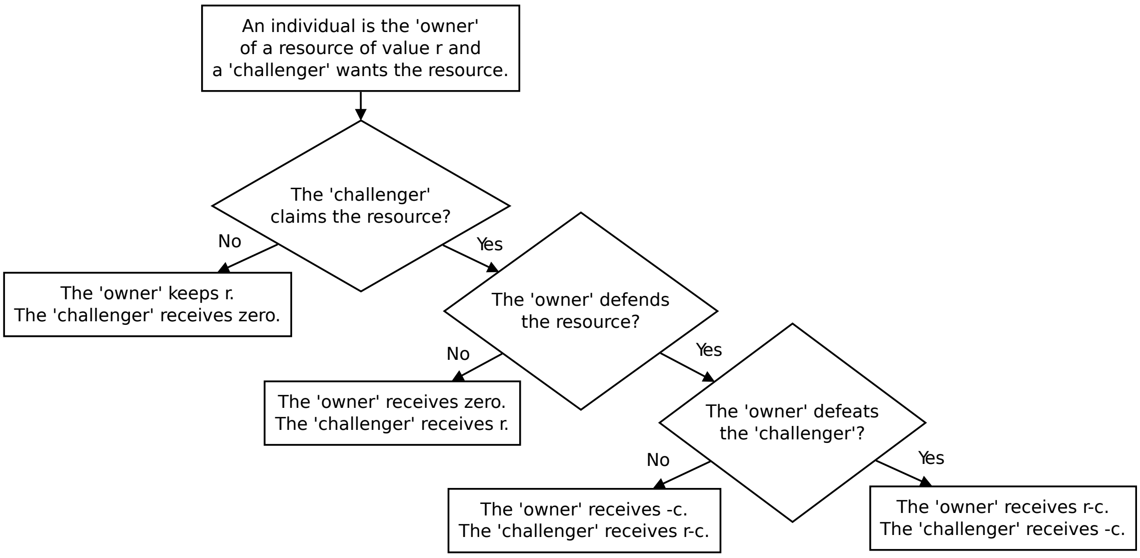


Figure 4.1: An owner-challenger encounter occurs when one individual is the *owner* of a resource and then a *challenger* arrives with the intention of claiming the resource. Both parties decide asynchronously whether to fight over the resource or surrender it to the other individual. If both fight then both pay a cost $-c$ but the winner additionally receives the resource value r. The strongest individual wins the fight. If both have the same capability then the winner is decided randomly with each individual having equal probability of winning.

this rule should yield a positive payoff to an owner on average after repeated encounters with random challengers.

The estimate $\widehat{F}$ does not include the weighted payoff received by an individual when the opponent withdraws from conflict, therefore every owner works under the assumption that the opponent is always intent to fight. This assumption is clearly correct from the perspective of the owner because its decision-making takes place only after having been challenged. However a rational challenger should not always claim, since this ignores the probability that the owner will defend the resource rather than abandon it uncontested. Therefore a challenger that always claims should be expected to perform worse in the long term (*i.e.*, after repeated encounters against random owners) than an owner. The simplified model is proposed in this manner, with *always-aggressive* challengers, in order to determine analytically what values of r, c, and σ_ϵ make internal biases necessary for owners to receive the highest long-term payoffs, even when these individuals use the rational decision rule stated above and when no deception between individuals is present. The analysis of this model and the results are presented in Section 4.4.

4.3 The generalised owner-challenger model with role-dependent internal and external biases

This section introduces a generalised version of the simplified model presented in Section 4.2 in order to simulate the scenario where every decision-maker additionally has an *external bias* (*i.e.*, a deception bias) that alters the capability this individual signals to any opponent. The larger the external bias the greater the baseline capability signalled to competitors. Given any two individuals, x and y, what x perceives is y's projected capability, distorted first by y's external bias and then by x's own perception error. The actual attribute remains unchanged but y may be able to deceive x into thinking that y's capability is greater (or lower) than it actually is, thus making x less (or more) willing to fight. The model aims to test the theory proposed by Trivers (2011). By incorporating the premises of the theory (namely costs paid for conscious deception of others) computational simulations are run to determine in what circumstances, if any, self-deception evolves in order to facilitate the deception of opponents.

An individual with a non-zero external bias exerts a form of deception, or dishonest signaling. The use of a positive external bias is similar to *deimatic behaviour*, in which an animal, feeling in danger, makes a physical display, possibly involving changes in shape, position, and/or colour, in order to appear threatening (probably more than the animal actually is) and to dissuade an opponent from attacking. Examples of deimatic individuals include some species of frog, who, in the presence of a threat, inflate themselves with air and raise their hind legs in order to appear larger (Martins, 1989). The dishonest signal sent by an individual with a positive external bias could also be compared to *Batesian mimicry*, where a harmless individual imitates the signals of a harmful one, in order to discourage attacks from predators. Examples of Batesian species include *Lampropeltis elapsoides*, a nonvenomous snake who exhibits the colour pattern of the venomous *Micrurus fulvius* (Kikuchi and Pfennig, 2010).

In the generalised model each decision-maker has two types of bias. An *internal bias*, denoted by k, that influences the perception the individual has of itself (as in the simplified model of the previous section), and an *external bias*, denoted by s, that distorts the capability it displays to opponents. Both biases comprise together a *deception pair* denoted by $[k, s]$. Any individual x with internal bias k_x and external bias s_x believes that its own capability is $\theta_x + k_x$ and attempts to deceive any potential opponent y into believing that x's capability is $\theta_x + s_x$. Then what y perceives is a normal deviate of the projected capability with standard deviation σ_ϵ, the perception error size.

Each decision-maker in the model holds two deception pairs, $[k_o, s_o]$ and $[k_c, s_c]$. The first one of these is expressed when the individual is playing the role of an owner and the second one when the individual is a challenger. The first pair can be referred to as the individual's *owner biases* and the latter as the *challenger biases.* Alternatively the first pair can be referred to as the *owner strategy* of the individual while the second pair is the *challenger strategy.* A strategy is *symmetrically biased* if its internal and external biases are equal otherwise it is *asymmetrically biased.* Asymmetrically-biased individuals represent organisms in nature that exercise 'conscious' deception because they attempt to project an image of themselves that differs from their true self-perception. On the other hand, symmetrically-biased individuals represent organisms that do not deceive or deceive 'unconsciously', because if they spread false information it is only because they 'believe' it as well.

In each encounter every individual expresses only the pair of biases that match the role (owner or challenger) the individual is playing at that moment. Given an owner x and a challenger y, x estimates its probability of winning as $\widehat{p_{W,x}} = P(\widehat{\theta_y} < \theta_x + k_{o,x})$, where $\widehat{\theta_y} \sim \mathcal{N}(\theta_y + s_{c,y}, \sigma_\epsilon)$, given that in the encounter this individual estimates its capability as $\theta_x + k_{o,x}$ (with x's internal owner bias) and that of y as a normal deviate of $\theta_y + s_{c,y}$ (with y's external challenger bias) with perception error size σ_ϵ. On the other hand, y estimates its probability of winning as $\widehat{p_{W,y}} = P(\widehat{\theta_x} < \theta_y + k_{c,y})$, where $\widehat{\theta_x} \sim \mathcal{N}(\theta_x + s_{o,x}, \sigma_\epsilon)$, given that in the encounter this individual estimates its capability as $\theta_y + k_{c,y}$ (with y's internal challenger bias) and that of x as a normal deviate of $\theta_x + s_{o,x}$ (with x's external owner bias) with perception error size σ_ϵ. The owner x estimates its expected payoff as $\widehat{F_x} = \widehat{p_{W,x}}(r - c) + (1 - \widehat{p_{W,x}})(-c) = \widehat{p_{W,x}}r - c$ and fights if and only if this estimate is positive. Similarly, the challenger y estimates its expected payoff as $\widehat{F_y} = \widehat{p_{W,y}}r - c$ and decides to fight if and only if $\widehat{F_y} > 0$. As explained in Section 4.2, this decision rule (*i.e.*, fighting if and only if the estimated payoff is positive) is rational from the perspective of an owner but not necessarily so from the perspective of a challenger. This is because for the challenger a rationally estimated payoff would necessarily include an estimate of the probability of the owner contesting the resource. However the model becomes difficult to analyse if the challenger is set to estimate this probability. For this reason the model considers the simplified scenario where both challengers and owners use the same rule because the former can be realistically assumed to be conservative when forced to work with imperfect information. In this manner the challenger only challenges when it estimates that it can win the resource even if the owner fights back. Despite not being rigourously rational this challenger behaviour is sensible and realistic.

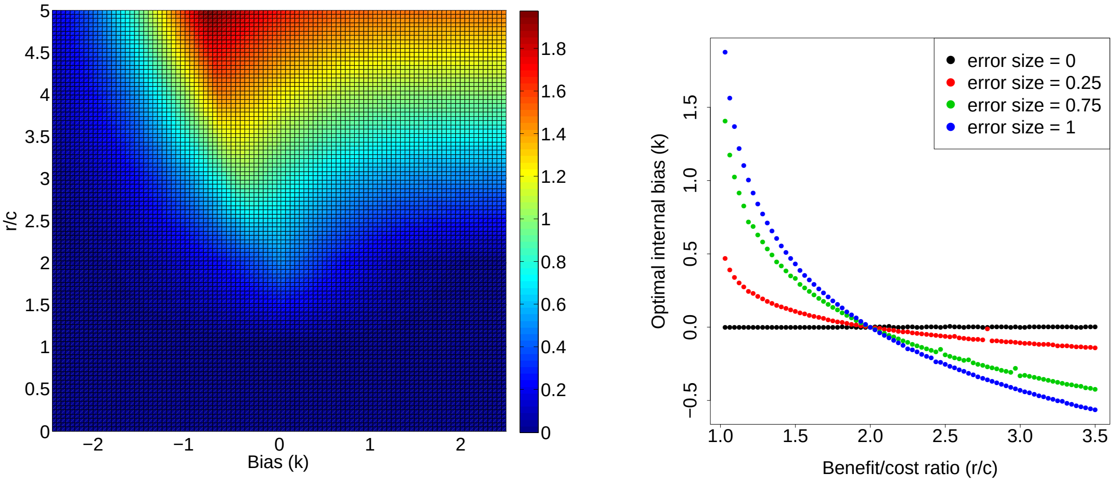
Trivers' premise that deception is more costly in the absence of self-deception (Trivers, 2011) is incorporated into the model by having each individual pay a *conscious deception cost* that penalizes asymmetrical strategies, regardless of whether a fight actually takes place or not. The cost paid by an individual is proportional to the discrepancy between the internal and the external biases in the strategy exercised by this individual in the encounter, with a proportionality constant $\lambda \in [0, 1]$. That is to say, the cost paid by the owner x increases with the difference between this individual's owner biases and is given by $\lambda|k_{o,x} - s_{o,x}|$, whereas the cost paid by the challenger y increases with the difference between this individual's challenger biases and is given by $\lambda|k_{c,y} - s_{c,y}|$.

Dishonest signaling may serve as a way to avoid the cost derived from a physical conflict by discouraging an opponent from fighting (an individual with a high external bias may dissuade an opponent from fighting); however, in nature such signaling is also costly, even though the cost paid in exchange for the ability to cheat opponents (*e.g.*, through having to invest in ornamentation) may be less than the one paid for taking part in a fight (*e.g.*, through sustaining an injury) (Backwell et al., 2000; Zahavi, 1975). This premise is incorporated into the model by having each individual pay a *dishonest signaling cost* proportional to the square of the external bias in the strategy played by the individual in an encounter against an opponent. The proportionality constant is denoted $\omega \in [0, 1]$ and the cost paid by an owner x is thus given by $\omega {s_{o,x}}^2$ whereas the cost paid a challenger y is given by $\omega {s_{c,y}}^2$. The conscious deception cost and the dishonest signaling cost paid by an individual are subtracted from the payoff received by this decision-maker from the encounter. For instance, if an owner x wins a fight against a challenger y then x's final payoff is $F_x = r - c - \lambda|k_{o,x} - s_{o,x}| - \omega {s_{o,x}}^2$ whereas y's is $F_y = -c - \lambda|k_{c,y} - s_{c,y}| - \omega {s_{c,y}}^2$. Clearly positive factors λ and ω together impose a selective pressure on decision-makers driving them towards becoming less deceptive and more symmetrically biased. It can be hypothesised that without the former parameter individuals should evolve to be asymmetrically biased whereas without the latter individuals should evolve to be maximally deceptive. Given any two individuals, x and y, y's internal bias can evolve so that y disregards the uninformative signal originated from x's external bias. In the absence of the dishonest signaling cost this would escalate. For this reason, this cost was set to be proportional to the square of the external bias in order to prevent signallers' external biases and receivers' internal biases from increasing indefinitely in an evolutionary arms race.

A set of evolutionary simulations were run with the role-dependent owner-challenger model under Triver's premise (*i.e.*, with large enough values of λ and ω) as follows. Firstly a population of decision-makers is initialised randomly with standard normal biases, then each generation every individual x is paired at random with exactly one adversary y in the population. The fitness of x is calculated as the average of its payoff when playing owner and challenger against y, and it increases with the resources (each one of these with value r) x manages to protect (as an owner) and/or usurp from y (as a challenger) and decreases with the number of fights x involves itself in (because each fight comes with a cost c). The fitness of x depends on the decisions this individual makes and how advantageously it influences the decisions of y, who is also trying to maximise its own gain. Fitness proportional selection (Baker, 1987) is used to determine which individuals reproduce, with normally-distributed mutations. Evolution runs until no considerable changes are observed and the population is assumed to be in equilibrium. Full details of the model are presented in Section A.3. Results from the model are presented in Section 4.5.

4.4 Results with the simplified owner-challenger model with internal biases

The expected payoff $F(k)$ of an owner with internal bias k depends on the value of the resources contested (*i.e.*, r), the cost of each fight (*i.e.*, c), and the error made when estimating the capability of an opponent (*i.e.*, σ_ϵ). This is approximated with a numerical method described in Section A.1. The expected payoffs for different values of $\frac{r}{c}$ and k when $\sigma_\epsilon = 1$ are plotted in Figure 4.2a. The plot shows that the highest payoff is obtained by owners with biases near zero when $\frac{r}{c} = 2$. But as this ratio increases it is owners with negative biases who receive the highest expected payoffs. Therefore owners that underestimate their own strength are the ones that in the long term perform the best against always-aggressive challengers when $\sigma_\epsilon = 1$ and the value of the contested resource (r) outweighs the cost of a confrontation (c). In Section A.2 it is formally shown that when $\frac{r}{c} \in (0, 1]$ owners never retaliate against always-aggressive challengers and end up receiving the same payoff (zero), regardless of owners' bias; this is because the value of the resource is offset by the cost of the inevitable fight. Therefore it can be concluded that as long as $\frac{r}{c} \leq 1$ no owner should perform better than the other and no bias can be considered optimal. On the other hand, the biases that maximise $F(k)$ when $\frac{r}{c} \in (1, +\infty)$ were found



(a) (b)

Figure 4.2: Figure 4.2a shows the owner's expected payoff (F) in the simplified model against random always-aggressive challengers as a function of confidence biases (k) and benefit cost ratios ($\frac{r}{c}$) when $\sigma_\epsilon = 1$. Figure 4.2b shows the self-deception biases (k) that yield the highest payoffs to an owner in the long term when facing random always-aggressive challengers along different benefit/cost ratios ($\frac{r}{c}$) and perception error sizes (σ_ϵ). Details on how to produce these plots appear in Section D.2.

numerically and plotted in Figure 4.2b.[2] In Section A.2 it is shown that any owner i's fighting probability is given by $p_F = P(A > z - k_i)$, where z is an advantage threshold for conflict that depends on $\frac{r}{c}$ and σ_ϵ. It is also shown that if i is an optimal-decision maker then $p_F = \frac{1}{2}$. Therefore after a large enough number of repeated encounters with random challengers an optimal owner should fight back in half of these encounters, because $P(A > 0) = \frac{1}{2}$. This means that given any $\frac{r}{c}$ and σ_ϵ, only owners with biases equal to z may be optimal decision-makers because only these have fighting probabilities equal to $\frac{1}{2}$. This is confirmed by the numerical results displayed in Figure 4.2b. If $\sigma_\epsilon = 0$ then $z = 0$ (Technical details in Section A.2) and the highest payoffs are received by owners with approximately zero bias. Unbiased individuals also get the best payoffs when $\sigma_\epsilon > 0$ and $\frac{r}{c} = 2$ because then $z = 0$. If $\sigma_\epsilon > 0$ then z decreases below zero as $\frac{r}{c}$ increases above 2 and therefore negative owner biases yield the highest payoffs. Additionally if $\sigma_\epsilon > 0$ then z increases above zero as $\frac{r}{c}$ decreases below 2 and owners achieve the maximum gain with positive biases. It can be concluded that owners require biases to optimise their payoffs if and only if the perception error is present (*i.e.*, if $\sigma_\epsilon > 0$). That is to say, internal biases serve as a means to compensate for information noise, given the assumed decision rule.

---

[2] Details on how to produce the plots in Figure 4.2 appear in Section D.2.

4.5 Results with the generalised owner-challenger model with role-dependent biases

The difference between deception and self-deception biases were measured when populations were in evolutionary equilibrium in different simulations with different values of $\frac{r}{c}$ and σ_ϵ. If these biases tend to evolve to have the same value when it is costly to have them differ, then this would support the theory proposed by Trivers (Trivers, 2011). Figure 4.3 shows the average owner bias differences ($\overline{|k_o - s_o|}$) and the average challenger bias differences ($\overline{|k_c - s_c|}$) when the population is in evolutionary equilibrium in simulations run with parameters $\frac{r}{c} \in [1, 5]$, $\sigma_\epsilon \in [0, 4]$, $\lambda \in \{0, 0.5, 1\}$ and $\omega = 1$.[3] That is to say, Figure 4.3 shows the level of symmetry of owner and challenger strategies in equilibrium with ($\lambda = 0.5$ and $\lambda = 1$) and without ($\lambda = 0$) Trivers' premise that deception is more costly to the deceiver when it is unconscious. The plots show that when $\lambda > 0$ the symmetry in owner and challenger strategies generally increases as $\frac{r}{c}$ and σ_ϵ increase together. With high enough values of these two parameters the internal and external biases evolve to be approximately equal, which is consistent with Trivers' theory because as the parameters increase, natural selection increasingly favours individuals that self-deceive just as much as they deceive others because they avoid the cognitive cost of conscious deception, and the effects of acting based on false information are more than offset by the effects of manipulating others' perceptions. Figure 4.3 also shows that asymmetry can be stable as the perception error decreases and $\frac{r}{c}$ increases. That is to say, as higher costs are paid for the ability of being consciously deceitful, it pays off more to be an unconscious deceiver, unless the perception errors are low (allowing the owner to make decisions on more certain information) and the value of the contested resource greatly outweighs the costs of a fight.

Figure 4.3a shows that if $\lambda = 0$ then owner bias differences are generally lowest when $\frac{r}{c} \in [2, 2.5]$ and they increase as $\frac{r}{c}$ increases and decreases away from this interval. That is to say, if $\frac{r}{c} \in [2, 2.5]$ then evolutionary equilibrium generally occurs when the population exercise owner strategies that are symmetrically biased (*i.e.*, when owners are unconscious deceivers), otherwise equilibrium generally occurs when the population exercise owner strategies that are symmetrically unbiased (*i.e.*, when owners are conscious deceivers). The figure also shows that owner bias differences increase and decrease with σ_ϵ. This means that as the information available becomes noisier then it pays off more to be an asymmetrically-biased owner (*i.e.*, a consciously-deceiving owner). A similar

---

[3] Details on how to produce these plots appear in Section D.3.

pattern occurs in challenger strategies, as shown in Figure 4.3b, although the bias differences observed in these strategies when the population is in evolutionary equilibrium are generally higher. That is to say, differences in challenger strategies increase with σ_ϵ and as $\frac{r}{c}$ increases and decreases away from $[2, 3.5]$. All the above is similar to what is observed in Figure 4.2b where the magnitude of optimal internal bias exhibit a similar relationship with $\frac{r}{c}$ and σ_ϵ, *i.e.*, higher magnitudes as $\frac{r}{c}$ increases and decreases from 2 and as σ_ϵ increases. It is reasonable to assume that this similarity is due to the same causes (*i.e.*, information noise) although a formal demonstration of this (such as the one provided for the simplified model in Section 4.4) is difficult in the generalised model with role-dependent biases.

Additional evolutionary simulations were run with the same parameters with external biases absent. The purpose of this was to compare the evolved internal biases in owner and challenger strategies in the absence and presence of external biases. The difference in magnitude of internal biases in equilibrium when these evolve in the presence and absence of external biases was measured by running separate evolutionary simulations with $(s_o, s_c \neq 0)$ and then without $(s_o, s_c = 0)$ deception biases. Individuals pay a conscious deception cost $(\lambda = 1)$ and a dishonest signaling cost $(\omega = 1)$ only in simulations where external biases are present. The average internal bias in owner strategies when the population is in evolutionary equilibrium in simulations with external biases is denoted by $\overline{k_o^{\text{d}}}|_{\lambda=1}$ and the average internal bias in challenger strategies is denoted by $\overline{k_c^{\text{d}}}|_{\lambda=1}$. The average internal bias in owner strategies when individuals are in evolutionary equilibrium in simulations without external biases is denoted by $\overline{k_o^{\text{nd}}}|_{\lambda=0}$ and the average internal bias in challenger strategies is denoted by $\overline{k_c^{\text{nd}}}|_{\lambda=0}$. The difference between evolved internal biases with and without external biases ($\overline{k_o^{\text{d}}}|_{\lambda=1} - \overline{k_o^{\text{nd}}}|_{\lambda=0}$ and $\overline{k_c^{\text{d}}}|_{\lambda=1} - \overline{k_c^{\text{nd}}}|_{\lambda=0}$) for each choice of $\frac{r}{c}$ and σ_ϵ are plotted in Figure 4.4.[4] The two plots show that there are values of $\frac{r}{c}$ and σ_ϵ for which these differences are generally positive and that these differences tend to increase as $\frac{r}{c}$ and σ_ϵ increase. This implies that with high enough values of $\frac{r}{c}$ and σ_ϵ the magnitude of the evolved internal biases increases in the presence of external biases, which means that the ability to deceive others requires an increase in self-deception in order to be evolutionarily stable. It can be hypothesised that the internal biases that evolve without external biases do so for a reason similar to the one explained in Section 4.4 for the simplified model with always-aggressive challengers (*i.e.*, noise in the information available), although a formal demonstration of this is harder in the generalised model with

---

[4] Details on how to produce these plots appear in Section D.4.

role-dependent biases.

In evolutionary simulations with both internal and external biases where individuals are forced to pay a cost for the ability of conscious deception (*i.e.*, for exercising asymmetrical strategies) there is a difference in the magnitude of the internal biases that evolve compared to the internal biases evolved in the absence of this cost, as shown in Figure 4.5.[5] The average internal bias in the equilibrium population is measured when individuals evolve with ($\lambda = 1$) and without ($\lambda = 0$) paying a cost for conscious deception and then the differences between these averages is calculated for each choice of $\frac{r}{c}$ and σ_ϵ. The two plots in Figure 4.5 show that these differences tend to increase with $\frac{r}{c}$ and σ_ϵ and that they are generally above zero with high enough values of these two parameters. This implies that the magnitude of the self-deception that evolves under the influence of the conscious-deception cost tends to become larger than the self-deception that evolves without this cost as $\frac{r}{c}$ and σ_ϵ increase. That is to say, with high enough values of $\frac{r}{c}$ and σ_ϵ, self-deception is effectively higher under Trivers' premise that conscious deception is costly (Trivers, 2011). In addition to this, self-deception evolved under Trivers' premise increases as the information decision-makers use becomes noisier (*i.e.*, as σ_ϵ increases) and as the value of the contested resource grows relative to the cost of a fight (*i.e.*, as $\frac{r}{c}$ increases).

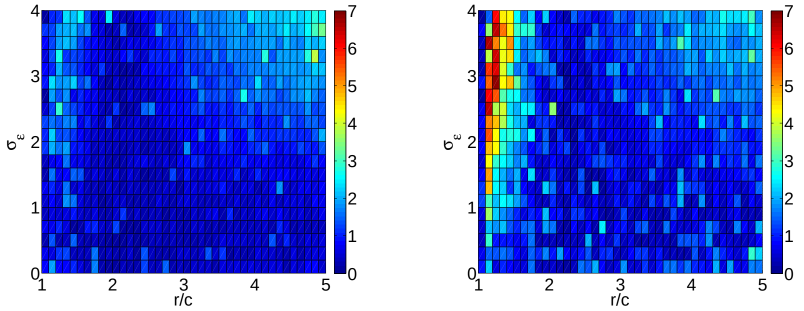
4.6 Conclusions

The owner-challenger model extends the one proposed by Johnson and Fowler (2011) and offers two improvements over the original model. The first is that resources are never left unclaimed and the second is that individuals use a rational decision rule by taking into account the costs and benefits of each decision. Two versions of the model are considered: (1) the model with internal biases, introduced in Section 4.2, and (2) the model with role-dependent internal and external biases, introduced in Section 4.3. The model with internal biases aims to determine the evolvability of self-deceptive cognitive biases, given that decision-makers use a rational decision rule. The model with role-dependent internal and external biases introduces dishonest signaling and aims to test the theory proposed by Trivers, which states that self-deception should evolve if individuals face a selective pressure to deceive each other and that self-deceiving deceivers have an evolutionary advantage over other deceitful individuals because the former do not have to pay the cognitive costs of concealing a lie consciously.

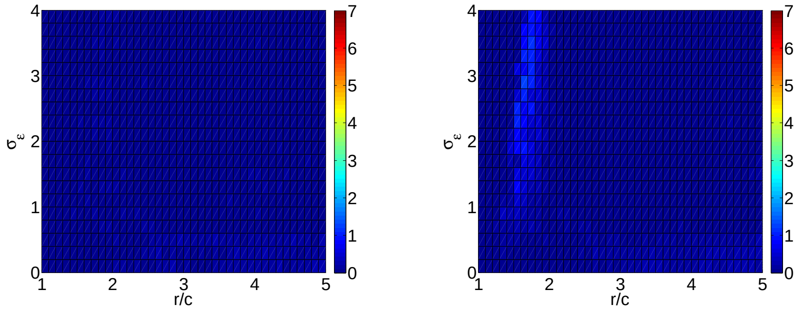
The baseline results with the owner-challenger model with internal biases introduced

---

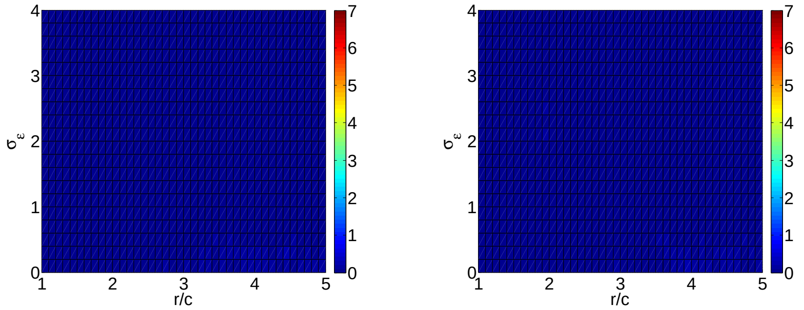
[5] Details on how to produce these plots appear in Section D.5.



(a) $\overline{|k_o - s_o|}$, $\lambda = 0$. (b) $\overline{|k_c - s_c|}$, $\lambda = 0$.

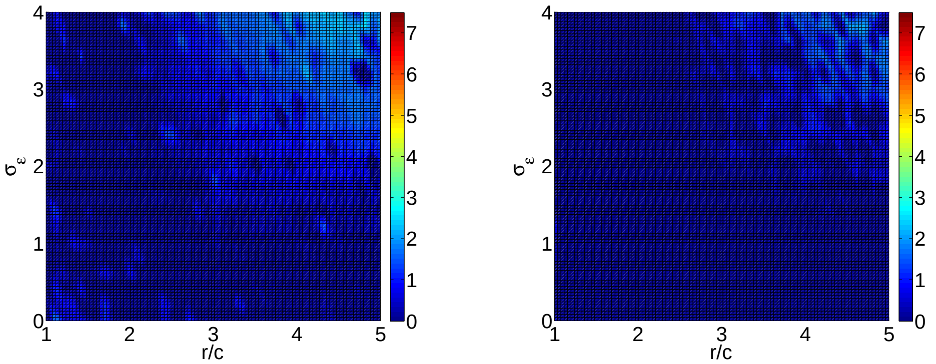


(c) $\overline{|k_o - s_o|}$, $\lambda = 0.5$. (d) $\overline{|k_c - s_c|}$, $\lambda = 0.5$.



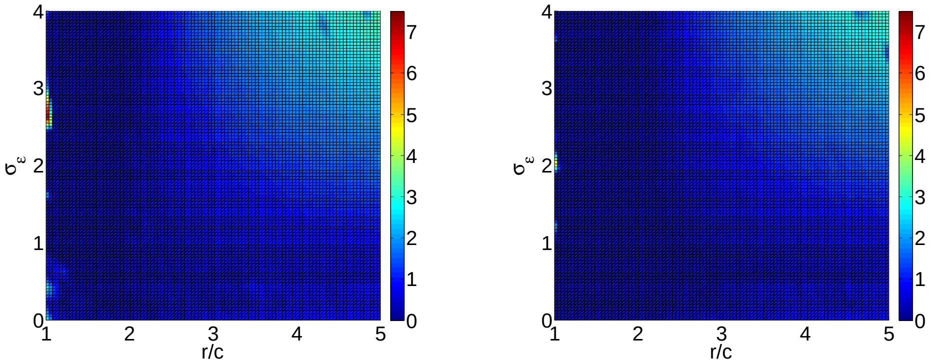
(e) $\overline{|k_o - s_o|}$, $\lambda = 1$. (f) $\overline{|k_c - s_c|}$, $\lambda = 1$.

Figure 4.3: Differences between internal (self-deception) and external (deception) biases evolved with different benefit/cost ratios ($\frac{r}{c}$), perception errors (σ_ϵ), with fixed conscious deception costs (λ) and a fixed dishonest signaling cost ($\omega = 1$). With each combination of these parameters, evolutionary simulations are run with populations composed of individuals with role-dependent biases. When equilibrium is reached the average difference of owner biases ($\overline{|k_o - s_o|}$) and the average difference of challenger biases ($\overline{|k_c - s_c|}$) in the population are calculated and plotted. These plots show values of $\frac{r}{c}$ and σ_ϵ with which symmetrical (darker blue) and asymmetrical (darker red) strategies are stable. As conscious deception costs increase differences between internal and external biases tend to decrease towards zero. Details on how to produce these plots appear in Section D.3.



(a) $\overline{k_o^{\mathrm{d}}}|_{\lambda=1} - \overline{k_o^{\mathrm{nd}}}|_{\lambda=0}$ (b) $\overline{k_c^{\mathrm{d}}}|_{\lambda=1} - \overline{k_c^{\mathrm{nd}}}|_{\lambda=0}$

Figure 4.4: Difference between internal (self-deception) biases evolved in the presence and absence of external biases. When external biases are present, individuals pay a conscious deception cost ($\lambda = 1$) and a dishonest signaling cost ($\omega = 1$). The notations $\overline{k_o^{\mathrm{nd}}}|_{\lambda=0}$ and $\overline{k_c^{\mathrm{nd}}}|_{\lambda=0}$ are used to refer to the average internal bias in owner and challenger strategies, respectively, in populations in evolutionary equilibrium when individuals evolve with no external biases (*i.e.*, when $s_{o,x}, s_{c,x} = 0$ for every individual x). The notations $\overline{k_o^{\mathrm{d}}}|_{\lambda=1}$ and $\overline{k_c^{\mathrm{d}}}|_{\lambda=1}$ are used to refer to the average internal bias in owner and challenger strategies, respectively, in populations in evolutionary equilibrium when individuals evolve with external biases. Figure 4.4a shows the owner difference $\overline{k_o^{\mathrm{d}}}|_{\lambda=1} - \overline{k_o^{\mathrm{nd}}}|_{\lambda=0}$ for different benefit/cost ratios ($\frac{r}{c}$) and perception errors (σ_ϵ), whereas Figure 4.4b shows the challenger difference $\overline{k_c^{\mathrm{d}}}|_{\lambda=1} - \overline{k_c^{\mathrm{nd}}}|_{\lambda=0}$. The plots show that these differences tend to increase with $\frac{r}{c}$ and σ_ϵ. This implies that as these parameters increase, an increase in the ability to deceive others (from zero external bias to nonzero external bias) requires an increase in the magnitude of self-deception in order to be stable. Details on how to produce these plots appear in Section D.4.



(a) $\overline{k_o^{\mathrm{d}}}|_{\lambda=1} - \overline{k_o^{\mathrm{d}}}|_{\lambda=0}$ (b) $\overline{k_c^{\mathrm{d}}}|_{\lambda=1} - \overline{k_c^{\mathrm{d}}}|_{\lambda=0}$

Figure 4.5: Differences in the magnitude of internal biases evolved with ($\lambda = 1$) and without ($\lambda = 0$) conscious deception costs. In both plots the dishonest signaling cost is $\omega = 1$. The average owner internal bias evolved in the presence of external biases when $\lambda = 1$ is denoted by $\overline{k_o^{\mathrm{d}}}|_{\lambda=1}$ and the same average when $\lambda = 0$ is denoted by $\overline{k_o^{\mathrm{d}}}|_{\lambda=0}$. The average challenger internal bias evolved in the presence of external biases when $\lambda = 1$ is denoted by $\overline{k_c^{\mathrm{d}}}|_{\lambda=1}$ and the same average when $\lambda = 0$ is denoted by $\overline{k_c^{\mathrm{d}}}|_{\lambda=0}$. The plots show that the differences $\overline{k_o^{\mathrm{d}}}|_{\lambda=1} - \overline{k_o^{\mathrm{d}}}|_{\lambda=0}$ and $\overline{k_c^{\mathrm{d}}}|_{\lambda=1} - \overline{k_c^{\mathrm{d}}}|_{\lambda=0}$ are generally nonzero and that they increase with $\frac{r}{c}$ and σ_ϵ. This implies that internal biases evolved with conscious deception costs (*i.e.*, with $\lambda = 1$) tend to become larger than internal biases evolved without this assumption (*i.e.*, when $\lambda = 0$) as $\frac{r}{c}$ and σ_ϵ increase. Details on how to produce these plots appear in Section D.5.

in Section 4.2 show that, given an assumed optimal decision rule taking proper account of probabilities costs and benefits of outcomes, suggested by Marshall et al. (Marshall et al., 2013b), biases provide a way for owners in the model to compensate for perception errors when their opponents are certain to fight. This is illustrated in Figure 4.2b, where it is shown that if there are perception errors ($\sigma_\epsilon > 0$) then optimal behaviour requires non-zero biases, the sole exception being when $\frac{r}{c} = 2$. If errors are not present ($\sigma_\epsilon = 0$) then owners do not require any biases to behave optimally. But if errors are present then self-deception biases are required to gain the best payoffs in the long term. These results provided a baseline optimal level of self-deception to compare the extended model against.

The extended owner-challenger model with role-dependent internal and external biases was introduced in Section 4.3 with the purpose of examining the evolution of self-deception as a supporting mechanism of deception. Symmetrically-biased individuals are those who self-deceive just as much as they attempt to deceive others, and can be considered to be ‘unconscious’ deceivers. On the other hand asymmetrically-biased individuals are those who project an image of themselves that differs from their self-perception, and are analogous to ‘conscious’ deceivers. The premise of Trivers’ theory was incorporated into the model by having each decision-maker pay a dishonest signaling cost (for having the ability to deceive others through external biases) and a conscious deception cost (for exercising asymmetrically-biased strategies, *i.e.*, for being consciously deceitful), emulating the physiological costs that deceivers in nature have to pay, according to Trivers’ proposal (Trivers, 2011). Evolutionary simulations with the model show that when these costs are present then symmetrically-biased, self-deceiving individuals are more evolutionarily successful than those who attempt to deceive others while attempting to act on truthful information, when the benefit/cost ratio and the perception error are high enough (Figure 4.3). In other words, self-deceiving deceivers are more likely to evolve as the benefit/cost ratio and the perception error increase, when the conscious deception and dishonest signaling costs are present. The internal biases evolved when individuals attempt to deceive others generally exceed those that are evolved when individuals cannot deceive others (Figure 4.4), as information becomes noisier (*i.e.*, as σ_ϵ increases) and the benefit/cost ratio becomes larger. That is to say, in order to be evolutionarily stable, an increase in deceitful behaviour requires an increase in self-deceiving behaviour when there are physiological costs for exercising deception and for doing so consciously. Further simulations show that internal biases are also generally higher when there is a conscious deception cost (*i.e.*, when $\lambda = 1$) than when this cost is absent (*i.e.*, when $\lambda = 0$), and also that the difference between the internal biases evolved with and without this cost generally increases with $\frac{r}{c}$ and σ_ϵ (Figure 4.5). From these numerical experiments it can be concluded that Trivers' theory generally holds true in situations of conflict if two conditions are met. First, the value of the contested resource must sufficiently exceed the cost of the fight required to claim the resource. Second, there must be a high enough degree of uncertainty in the information the decision-maker uses to assess its chances of winning the fight. As the value of the resource and the uncertainty increase, from the perspective of an individual it tends to payoff more in the long term to self-deceive as much as to attempt to deceive others, when conscious deception and dishonest signaling are physiologically costly. Then it should be expected that when these conditions are met, self-deceiving fighters evolve, as predicted by Trivers' theory (Marshall et al., 2013b; von Hippel and Trivers, 2011a).

The theory by Trivers has received considerable discussion, and it is possible that it will continue to be debated whether this theory correctly explains the apparent self-deception biases observed in humans, such as the ones presented in Section 4.1. The model presented in this article aims to test this theory in the particular case of a situation of conflict. The motivation for proposing this model is the point raised by commentators that the risk of injury faced by a self-deceiving, deceitful fighter is likely to be higher than the benefit received from discouraging an opponent from fighting by means of a deceitful exhibition of strength, and that therefore self-deceiving deceivers should not evolve. The model presented in this article simulates a situation of conflict where it is shown that, under the premises of the theory, self-deceiving, deceitful fighters do evolve. Given this, the generalised model proposed in this chapter serves as a first attempt to formally address the evolution of self-deception in situations of conflict, and the results obtained complement Trivers' proposal (Marshall et al., 2013b; von Hippel and Trivers, 2011a). The model could be extended in the future. For instance, it would be of interest to examine how biases evolve when the two opponents share the contested resource if both estimate that their capabilities are approximately equal. It would also be interesting to determine the impact of more biologically-realistic assumptions on the results presented in this chapter, such as population variation in perception errors, costs of conflict and so on.

Chapter 5

Conclusions

This thesis presents different mathematical models that recreate decision problems in natural scenarios in order to determine the behavioural strategy that should be expected to evolve in response to these problems. These models are based on the accepted hypothesis that decision-makers in nature are compelled by natural selection to exhibit the same behaviour that should be expected to be displayed by Bayesian decision-makers when facing the same problem with uncertain information (Marshall et al., 2013b; McNamara et al., 2006; McNamara and Houston, 1980). In each model the optimal behavioural strategy is found analytically or using evolutionary simulations. The model presented in Chapter 2 investigates when individuals in nature should display phenotypic plasticity in the form of inducible defences and Chapter 4 investigates when individuals should self-deceive in situations of conflict. In the model presented in Chapter 2 it is assumed that decision-makers can evolve the machinery to behave as if having priors of different means and standard deviations. This assumption is tested in Chapter 3 and found to be true. The results obtained with these models can be summarised as follows.

The model presented in Chapter 2 simulates the developmental history of an individual with inducible defences. These pose a trade-off to the individual since they are costly and only beneficial when there is a predation risk. During its lifetime the individual measures the predation risk by making repeated observations of the environment, each time updating its inherited prior estimate of the predation risk to a Bayesian posterior estimate, using Bayes' theorem, given in Equation 1.1. In each observation, a predator encounter occurs as a Bernoulli trial. The environment state consists of the probability that the individual survives each predator encounter or non-encounter with and without defences. The individual may induce the defences at any point, when any amount of evidence of predation risk has been collected.

Given any amount of evidence, the optimal decision (induce or postpone) is the one that should be expected to maximise the individual's future expectancy of reaching adulthood. The optimal strategy is the sequence of optimal decisions to make in response to a sequence of observations. A dynamic programming algorithm is used to determine the optimal decision the individual should make, given the individual's prior, the amount of accumulated evidence of predation risk and the environment state. Realistic environments were tested and two types of optimal strategies were found. The first is earliest specialisation, which consists of inducing the defences prior to making any observation, thus discarding any evidence of predation risk that may be collected. The second is plasticity, which consists of inducing the defences only after accumulating evidence of predation risk. The results obtained and described in Chapter 2 show that there is a threshold on the standard deviation of the individual's prior that determines the individual's optimal strategy. The optimal strategy is earliest specialisation when the standard deviation of the individual's prior is below this threshold whereas plasticity is the optimal strategy when the individual's prior is above or equal. When plasticity is the optimal strategy, the algorithm determines the amount of accumulated evidence (*i.e.*, number of past predator encounters and non-encounters) that makes induction the optimal decision. It should be expected that the the standard deviation of the individual's inherited prior evolves to reflect the uncertainty of the predation risk (a hypothesis confirmed in Chapter 3), thus a high standard deviation should be indicative of high uncertainty. Therefore an individual in the model maximises its expectancy of reaching adulthood by exhibiting plasticity only when the predation risk in the environment exhibits great uncertainty. This is in agreement with previous literature that states that plasticity should be expected to provide an evolutionary benefit only when the environment poses a high enough instability (Fischer et al., 2014; Ord et al., 2010; Price et al., 2003; Scheiner, 1993; Stephens, 1991). The model presented in Chapter 2 provides an analytical demonstration of this, taking into account the individual's inherited prior in order to predict when plasticity maximises the individual's expectancy of reaching adulthood.

Given the above, the model presented in Chapter 2 serves as a theoretical complement to the work of Hammill et al. (2008), which has been carried out with a similar purpose. However the predictions made by the model presented in Chapter 2 are made in terms of the individual's prior and refer to the individual's expectancy of reaching adulthood, not to its fitness as in the model of Hammill et al. (2008). The model could be improved in the future by modelling the individual's reproductive rate in order to estimate its fitness.

This could be done in order to predict properly the conditions (*e.g.*, environment state, individual's prior, etc) that must be met for inducible defences to be favoured by natural selection. This would offer a full theoretical complement to the work of Hammill et al. (2008). With survival probabilities (described in Chapter 2) measured experimentally this theoretical model should return results in agreement with the predictions made by Hammill et al. (2008). In addition to this, it may be possible to demonstrate formally the empirical conclusions achieved with the model in Chapter 2. This chapter was chronologically the last one of this thesis to be written and, due to time restrictions, attempts to produce a more formal explanation of the experimental observations described in Section 2.4 can only be suggested as future work.

Chapter 3 introduces a model to show how Bayesian priors can be encoded genetically and shaped by natural selection. This is in support of the assumption made in the model introduced in Chapter 2, that individuals can inherit different priors and that depending on these the individual should maximise its expectancy of reaching adulthood by exhibiting plasticity or specialisation. The model presented in Chapter 3 simulates a group of individuals whose evolutionary fitness depends on their ability to estimate accurately the probability of an event, referred to as the environment state. Each individual achieves this by making repeated Bernoulli trials that occur with a probability equal to the environment state. The individual has an inherited beta prior and updates it to a Bayesian posterior with each Bernoulli observation. The mean of the posterior after a certain number of observations is the individual's estimate of the environment state. Individuals that produce the most accurate estimates are the most likely to reproduce. During reproduction individuals pass on their inherited priors, with mutation, to their offspring but not the information learned from the observations made.

Evolutionary simulations are run where the environment state changes periodically and it is shown that the individuals' priors evolve to be within a range of each updated environment state after enough generations. That is to say, individuals evolve their priors to approximate the current environment state, whatever this may be. It is also shown that when the environment states are beta-distributed, the individuals evolve the hyperparameters of their priors to resemble the hyperparameters of the environment beta distribution. By doing this the Bayesian individuals evolve their priors to reflect the standard deviation of the environment distribution and thus the uncertainty of the environment.

The model is extended to include frequentist individuals, *i.e.*, those who use frequentist inference when estimating the environment state. Evolutionary simulations with competition between Bayesians and frequentists are run. The results show that Bayesians become more evolutionarily stable than frequentists as the environment states become more uncertain. This is consistent with the accepted proposal that natural selection should favour individuals that behave *as if* Bayesian learners when processing uncertain information (Marshall et al., 2013b; McNamara et al., 2006; McNamara and Houston, 1980). Thus the results show that Bayesian individuals evolve their priors in order to be able to estimate the environment state accurately and that they outperform frequentists, as they are expected to, when both types of learners are faced with uncertain information.

Chapter 4 addresses the theory of the evolution of self-deception proposed by Trivers (2011) and aims to test it in the particular case of a situation of conflict, a scenario raised by commentators of the proposal (Frey and Voland, 2011; Funder, 2011; Marshall et al., 2013b) and not addressed by Trivers (Marshall et al., 2013b). In order to do this, a situation of conflict for resources between two individuals is modelled. Each individual faces a selective pressure to maximise its gains, in terms of collected resources, and minimise its losses, derived from engaging in fights. In order to achieve this, each individual is able to measure the fighting capability of the opponent in order to decide whether or not to fight. Each individual may 'deceive' its opponent (by feigning greater or lower capability than the actual one) and 'self-deceive' (by over- or underestimating its capability advantage over the opponent). The individual pays a cost proportional to the difference between the level of deception and self-deception it exercises. This is done in order to incorporate the hypothesis stated by Trivers (2011) that deception is more costly to the deceiver if this individual is aware of the deception. The results show that self-deception does evolve in situations of conflict under the premises of Trivers (2011) when two conditions are met. The first is that the cost of each fight must be sufficiently exceeded by the value of the resource in contest. The second is that each individual must face a high enough level of uncertainty when assessing its own capability and that of the opponent. The results show that, given the conditions stated, a suboptimal behavioural mechanism (self-deception) can lead to an evolutionarily optimal behaviour, which is consistent with the literature (McNamara and Houston, 2009). These results support the theory proposed by Trivers (2011) by showing how the predictions made by the proposal apply in situations of conflict. The model presented in Chapter 4 is the first to achieve this and address the point raised by commentators of the proposal that self-deception should not evolve in situations of conflict (Bandura, 2011). For simplicity, numerous assumptions are made in model. Improvements can be made by replacing these assumptions with more realistic ones. This could be achieved by having each individual assess the value of the contested resouce and the cost of the potential fight depending on their individual previous histories. For instance, given any two individuals, the same resource should appear more appealing to the one individual who has previously gained the least amount of resources. Similarly, the individual who has recently lost the most fights should be more inclined to avoid a conflict in order to prevent further injury. In addition to this, perception errors could vary from individual to individual and, possibly, from encounter to encounter, depending on environmental conditions.

In general, individuals in nature should reasonably be hypothesised to evolve the machinery to process optimally the information they use when making crucial decisions (e.g., when deciding whether or not to compete over a valuable resource, as in the model presented in Chapter 4). In this regard, the existence of biases may appear counterintuitive for this purpose, since these simply distort this information processing. For instance, even though the literature indicates that individuals should behave as Bayesians (Marshall et al., 2013a,b; McNamara et al., 2006; McNamara and Houston, 1980; Tenenbaum et al., 2006), people are reported to be overconfident regarding the outcome of future events despite known evidence to the contrary (Sharot, 2011a,b). However, the theory proposed by Trivers and the results presented in Chapter 4 show that a bias can be part of a behavioral mechanism that overall is beneficial, even though in other circumstances the bias alone could lead to detrimental behaviour. Then it could be sensibly hypothesised that the biases reported in the literature (Alicke and Govorun, 2005; McCormick et al., 1986; Pallier et al., 2002; Sharot, 2011a,b; Svenson, 1981) could potentially have similar explanations.

The models presented in this thesis and summarised above are purely theoretical and make no use of experimental data collected from nature. However the variables that are taken into account are realistic and thus the predictions made by the models should be consistent with equivalent scenarios in nature. This is exemplified, for instance, in the model introduced in Chapter 3 where Bayesian learners are shown to win an evolutionary race against frequentists exactly as they should be expected to in nature. The models presented in this thesis make some assumptions in order to make each problem more tractable. However, it is reasonable to hypothesise that these assumptions do not compromise greatly the generality of the results obtained. Nonetheless the models could be improved in the future by incorporating more realistic assumptions, such as variable resource values and perception errors in the model introduced in Chapter 4 and experimental data in the model introduced in Chapter 2.

Bibliography

Agrawal, A. A., Laforsch, C., and Tollrian, R. (1999). Transgenerational induction of defences in animals and plants. *Nature*, 401(September):60–63.

Alicke, M. D. and Govorun, O. (2005). The better-than-average effect. In Alicke, M. D., Dunning, D. A., and Krueger, J., editors, *The Self in Social Judgment. Studies in self and identity*, chapter 5, pages 85–106. Psychology Press, New York, New York, USA.

Backwell, P. R., Christy, J. H., Telford, S. R., Jennions, M. D., and Passmore, N. I. (2000). Dishonest signalling in a fiddler crab. *Proceedings. Biological sciences / The Royal Society*, 267(1444):719–24.

Baker, J. E. (1987). Reducing bias and inefficiency in the selection algorithm. In Grefenstette, J. J., editor, *Proceedings of the Second International Conference on Genetic Algorithms on Genetic algorithms and their application*, pages 14–21. L. Erlbaum Associates, Hillsdale, NJ, USA.

Baldwin, J. M. (1896). A new factor in evolution. *American Naturalist*, 30(355):441–451.

Bandura, A. (2011). Open peer commentary on the evolution and psychology of self-deception. *Behavioral and Brain Sciences*, 34(1):16–41.

Belsky, J., Steinberg, L., and Draper, P. (1991). Childhood experience, interpersonal development, and reproductive strategy: An evolutionary theory of socialization. *Child Development*, 62(4):647–670.

Black, A. R. and Dodson, S. I. (1990). Demographic costs of Chaoborus-induced phenotypic plasticity in Daphnia pulex. *Oecologia*, 83(1):117–122.

Boeing, W. J., Wissel, B., and Ramcharan, C. W. (2005). Costs and benefits of Daphnia defense against Chaoborus in nature. *Canadian Journal of Fisheries and Aquatic Sciences*, 62:1286–1294.

Bradshaw, A. D. (1965). Evolutionary significance of phenotypic plasticity in plants. *Advances in Genetics*, 13:115–155.

Brooks, M. L. and Swann, W. B. (2011). Is social interaction based on guile or honesty? *Behavioral and Brain Sciences*, 34(1):17–18.

Chambers, J. R. and Windschitl, P. D. (2004). Biases in social comparative judgments: the role of nonmotivated factors in above-average and comparative-optimism effects. *Psychological Bulletin*, 130(5):813–838.

Charnov, E. L. (1976). Optimal foraging, the marginal value theorem. *Theoretical Population Biology*, 9(2):129–136.

Dall, S. R. X., Giraldeau, L.-A., Olsson, O., McNamara, J. M., and Stephens, D. W. (2005). Information and its use by animals in evolutionary ecology. *Trends in Ecology & Evolution*, 20(4):187–93.

Darwin, C. R. (1859). *On the Origin of Species by Means of Natural Selection, or the Preservation of Favoured Races in the Struggle for Life.* Murray, London.

Davis, M. and Hadiks, D. (1995). Demeanor and credibility. *Semiotica*, 106(1-2):5–54.

De Jong, G. (2005). Evolution of phenotypic plasticity: patterns of plasticity and the emergence of ecotypes. *New Phytologist*, 166(1):101–118.

Dennett, D. (2003). The Baldwin effect: A crane, not a skyhook. In Weber, B. H. and Depew, D. J., editors, *Evolution and Learning: The Baldwin Effect Reconsidered*, chapter 4, pages 69–71. Cambridge, Mass.: MIT Press.

DePaulo, B. M. (1992). Nonverbal behavior and self-presentation. *Psychological Bulletin*, 111(2):203–243.

DePaulo, B. M., Stone, J. I., and Lassiter, G. D. (1985). Deceiving and detecting deceit. *The Self and Social Life*, 323.

DeWitt, T. J. and Scheiner, S. M. (2004). *Phenotypic Plasticity: Functional and Conceptual Approaches.* Oxford University Press Oxford.

DonaldsonMatasci, M. C., Bergstrom, C. T., and Lachmann, M. (2010). The fitness value of information. *Oikos*, 119(2):219—-230.

Dougherty, M. R., Gettys, C. F., and Ogden, E. E. (1999). MINERVA-DM: A memory processes model for judgments of likelihood. *Psychological Review*, 106(1):180–209.

Ekman, P. (1989). Why lies fall and what behaviors betray a lie. In Yuille, J. C., editor, *Credibility Assessment*, pages 71–82. Kluwer, Dordrecht, The Netherlands.

Ekman, P., O'Sullivan, M., Friesen, W. V., and Scherer, K. R. (1991). Face, voice, and body in detecting deceit. *Journal of Nonverbal Behavior*, 15(2):125–135.

Ellis, B. J. (2004). Timing of pubertal maturation in girls: an integrated life history approach. *Psychological Bulletin*, 130(6):920.

Fawcett, T. W., Marshall, J. A. R., and Higginson, A. D. (2015). The evolution of mechanisms underlying behaviour. *Current Zoology*, 61(2):221–225.

Fischer, B., van Doorn, G. S., Dieckmann, U., and Taborsky, B. (2014). The evolution of age-dependent plasticity. *The American Naturalist*, 183(1):108–125.

Forsman, A. (2014). Rethinking phenotypic plasticity and its consequences for individuals, populations and species. *Heredity*, (August):1–9.

Frankenhuis, W. E. and Del Giudice, M. (2012). When do adaptive developmental mechanisms yield maladaptive outcomes? *Developmental Psychology*, 48(3):628–642.

Frankenhuis, W. E. and Panchanathan, K. (2011). Balancing sampling and specialization: an adaptationist model of incremental development. *Proceedings. Biological sciences / The Royal Society*, 278(1724):3558–3565.

Frey, U. and Voland, E. (2011). The evolutionary route to self-deception: Why offensive versus defensive strategy might be a false alternative. *Behavioral and Brain Sciences*, 34(1):21–22.

Funder, D. C. (2011). Directions and beliefs of self-presentational bias. *Behavioral and Brain Sciences*, 34(1):23.

Gabriel, W. (2006). Selective advantage of irreversible and reversible phenotypic plasticity. *Archiv für Hydrobiologie*, 167(1-4):1–20.

Gittins, J. C., Glazebrook, K., and Weber, R. (2011). *Multi-armed Bandit Allocation Indices*. Wiley, 2nd edition.

Gluckman, P. D., Hanson, M. A., Spencer, H. G., and Bateson, P. (2005). Environmental influences during development and their later consequences for health and disease: implications for the interpretation of empirical studies. *Proceedings. Biological sciences / The Royal Society*, 272(1564):671–677.

Hammill, E., Rogers, A., and Beckerman, A. P. (2008). Costs, benefits and the evolution of inducible defences: a case study with Daphnia pulex. *Journal of Evolutionary Biology*, 21(3):705–15.

Havel, J. E. and Dodson, S. I. (1984). Chaoborus predation on typical and spined morphs of Daphnia pulex: Behavioral observations. *Limnology and Oceanography*, 29(3):487–494.

Havel, J. E. and Dodson, S. I. (1987). Reproductive costs of Chaoborus-induced polymorphism in Daphnia pulex. *Hydrobiologia*, 150(3):273–281.

Hayden, B. Y. and Platt, M. L. (2009). Gambling for Gatorade: risk-sensitive decision making for fluid rewards in humans. *Animal Cognition*, 12(1):201–207.

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

Jablonka, E., Oborny, B., Molnár, I., Kisdi, E., Hofbauer, J., and Czárán, T. (1995). The adaptive advantage of phenotypic memory in changing environments. *Philosophical Transactions of The Royal Society of London. Series B, Biological sciences*, 350(1332):133–141.

Johnson, D. D. P. and Fowler, J. H. (2011). The evolution of overconfidence. *Nature*, 477(7364):317–20.

Johnson, D. D. P. and Fowler, J. H. (2013). Complexity and simplicity in the evolution of decision-making biases. *Trends in Ecology & Evolution*, 28(8):446–447.

Kacelnik, A. and Bateson, M. (1996). Risky theoriesthe effects of variance on foraging decisions. *American Zoologist*, 36(4):402–434.

Kikuchi, D. W. and Pfennig, D. W. (2010). Predator cognition permits imperfect coral snake mimicry. *The American Naturalist*, 176(6):830–834.

Klar, Y. and Giladi, E. E. (1997). No one in my group can be below the group's average: a robust positivity bias in favor of anonymous peers. *Journal of Personality and Social Psychology*, 73(5):885–901.

Krebs, J. R., Stephens, D. W., and Sutherland, W. J. (1983). Perspectives in optimal foraging. In *Perspectives in Ornithology: Essays Presented for the Centennial of the American Ornithologists' Union*, page 165.

Krueger, D. A. and Dodson, S. I. (1981). Embryological induction and predation ecology in Daphnia pulex. *Limnology and Oceanography*, 26(2):219–223.

Kruger, J. and Dunning, D. (1999). Unskilled and unaware of it: how difficulties in recognizing one's own incompetence lead to inflated self-assessments. *Journal of Personality and Social Psychology*, 77(6):1121–1134.

Kuzawa, C. W. (2005). Fetal origins of developmental plasticity: are fetal cues reliable predictors of future nutritional environments? *American Journal of Human Biology : The Official Journal of the Human Biology Council*, 17(1):5–21.

Levins, R. (1968). *Evolution in Changing Environments: Some Theoretical Explorations.* Number 2. Princeton University Press.

Marshall, J. A. R., Trimmer, P. C., and Houston, A. I. (2013a). Unbiased individuals use valuable information when making decisions: a reply to Johnson and Fowler. *Trends in Ecology & Evolution*, 28(8):444–445.

Marshall, J. A. R., Trimmer, P. C., Houston, A. I., and McNamara, J. M. (2013b). On evolutionary explanations of cognitive biases. *Trends in Ecology & Evolution*, pages 1–5.

Martins, M. (1989). Deimatic behavior in Pleurodema brachyops. *Journal of Herpetology*, 23(3):305–307.

Maynard Smith, J. (1982). *Evolution and the Theory of Games.* Cambridge University Press.

McCormick, I. A., Walkey, F. H., and Green, D. E. (1986). Comparative perceptions of driver ability a confirmation and expansion. *Accident Analysis & Prevention*, 18(3):205–208.

McKay, R. and Efferson, C. (2010). The subtleties of error management. *Evolution and Human Behavior*, 31(5):309–319.

McNamara, J. M. (1982). Optimal patch use in a stochastic environment. *Theoretical Population Biology*, 21(2):269–288.

McNamara, J. M. (1985). An optimal sequential policy for controlling a Markov renewal process. *Journal of Applied Probability*, 22(2):324–335.

McNamara, J. M., Green, R. F., and Olsson, O. (2006). Bayes' theorem and its applications in animal behaviour. *Oikos*, 112(2):243–251.

McNamara, J. M. and Houston, A. I. (1980). The application of statistical decision theory to animal behaviour. *Journal of Theoretical Biology*, 85(4):673–690.

McNamara, J. M. and Houston, A. I. (2009). Integrating function and mechanism. *Trends in Ecology and Evolution*, 24(12):670–675.

McNamara, J. M., Trimmer, P. C., Eriksson, A., Marshall, J. A. R., and Houston, A. I. (2011). Environmental variability can select for optimism or pessimism. *Ecology Letters*, 14(1):58–62.

Meyers, L. A. and Bull, J. J. (2002). Fighting change with change: adaptive variation in an uncertain world. *Trends in Ecology & Evolution*, 17(12):551–557.

Moller, A. P. (1987). Social control of deception among status signalling house sparrows Passer domesticus. *Behavioral Ecology and Sociobiology*, 20(5):307–311.

Moran, N. A. (1992). The evolutionary maintenance of alternative phenotypes. *American Naturalist*, 139(5):971–989.

Nepomnaschy, P. and Flinn, M. (2009). Early life influences on the ontogeny of the neuroendocrine stress response in the human child. In Gray, P. and Ellison, P., editors, *The Endocrinology of Social Relationships*, chapter 16, pages 364–384. Harvard University Press Cambridge, MA.

Nettle, D. (2011). Flexibility in reproductive timing in human females: integrating ultimate and proximate explanations. *Philosophical Transactions of the Royal Society of London. Series B, Biological sciences*, 366(1563):357–365.

Ord, T. J., Stamps, J. A., and Losos, J. B. (2010). Adaptation and plasticity of animal communication in fluctuating environments. *Evolution*, 64(11):3134–3148.

Padilla, D. K. and Adolph, S. C. (1996). Plastic inducible morphologies are not always adaptive: the importance of time delays in a stochastic environment. *Evolutionary Ecology*, 10(1):105–117.

Pallier, G., Wilkinson, R., Danthiir, V., Kleitman, S., Knezevic, G., Stankov, L., and Roberts, R. D. (2002). The role of individual differences in the accuracy of confidence judgments. *The Journal of General Psychology*, 129(3):257–299.

Parejko, K. (1991). Predation by chaoborids on typical and spined Daphnia pulex. *Freshwater Biology*, 25(2):211–217.

Parejko, K. and Dodson, S. I. (1991). The evolutionary ecology of an antipredator reaction norm: Daphnia pulex and Chaoborus americanus. *Evolution*, pages 1665–1674.

Piersma, T. and Drent, J. (2003). Phenotypic flexibility and the evolution of organismal design. *Trends in Ecology & Evolution*, 18(5):228–233.

Price, T. D., Qvarnström, A., and Irwin, D. E. (2003). The role of phenotypic plasticity in driving genetic evolution. *Proceedings. Biological sciences / The Royal Society*, 270(February):1433–1440.

Pyke, G. H. (1984). Optimal foraging theory: a critical review. *Annual Review of Ecology and Systematics*, 15(1):523–575.

Pyke, G. H., Pulliam, H. R., and Charnov, E. L. (1977). Optimal foraging: a selective review of theory and tests. *Quarterly Review of Biology*, pages 137—-154.

Ramachandran, V. S. (1996). The evolutionary biology of self-deception, laughter, dreaming and depression: some clues from anosognosia. *Medical Hypotheses*, 47(5):347–62.

Riessen, H. P. (1992). Cost-benefit model for the induction of an antipredator defense. *The American Naturalist*, 140(2):349–362.

Riessen, H. P. (1999). Predator-induced life history shifts in Daphnia: a synthesis of studies using meta-analysis. *Canadian Journal of Fisheries and Aquatic Sciences*, 56:2487–2494.

Riessen, H. P. and Sprules, W. G. (1990). Demographic costs of antipredator defenses in Daphnia pulex. *Ecology*, 71(4):1536–1546.

Roff, D. A. (2002). *Life History Evolution*, volume 7. Sinauer Associates Sunderland.

Rohwer, S. (1977). Status signaling in Harris sparrows: some experiments in deception. *Behaviour*, 61(1):107–129.

Rohwer, S. and Rohwer, F. C. (1978). Status signalling in Harris sparrows: experimental deceptions achieved. *Animal Behaviour*, 26:1012–1022.

Scheiner, S. M. (1993). Genetics and evolution of phenotypic plasticity. *Annual Review of Ecology and Systematics*, pages 35–68.

Scheiner, S. M. (2014). The Baldwin effect: neglected and misunderstood. *The American Naturalist*, 184(4).

Schlichting, C. D. (1986). The evolution of phenotypic plasticity in plants. *Annual Review of Ecology and Systematics*, (143):667–693.

Schlichting, C. D. and Pigliucci, M. (1998). *Phenotypic Evolution: A Reaction Norm Perspective.* Sinauer Associates Incorporated.

Semple, S. and McComb, K. (1996). Behavioural deception. *Trends in Ecology & Evolution*, 11(10):434–437.

Sharot, T. (2011a). The optimism bias. *Current Biology*, 21(23):R941–R945.

Sharot, T. (2011b). *The Optimism Bias: A Tour of the Irrationally Positive Brain.* Pantheon Books, New York, New York, USA.

Spitze, K. (1992). Predator-mediated plasticity of prey life history and morphology: Chaoborus americanus predation on Daphnia pulex. *The American Naturalist*, 139(2):229–247.

Stamps, J. A. and Krishnan, V. V. (2014a). Combining information from ancestors and personal experiences to predict individual differences. *The American Naturalist*, 184(5):647—-657.

Stamps, J. A. and Krishnan, V. V. (2014b). Individual differences in the potential and realized developmental plasticity of personality traits. *Frontiers in Ecology and Evolution*, 2(October):1–15.

Stephens, D. W. (1991). Change, regularity, and value in the evolution of animal learning. *Behavioral Ecology*, 2(1):77–89.

Sultan, S. E. (2000). Phenotypic plasticity for plant development, function and life history. *Trends in Plant Science*, 5(12):537–542.

Svenson, O. (1981). Are we all less risky and more skillful than our fellow drivers? *Acta Psychologica*, 47(2):143–148.

Tenenbaum, J. B., Griffiths, T. L., and Kemp, C. (2006). Theory-based Bayesian models of inductive learning and reasoning. *Trends in Cognitive Sciences*, 10(7):309–318.

Tollrian, R. (1993). Neckteeth formation in Daphnia pulex as an example of continuous phenotypic plasticity: morphological effects of Chaoborus kairomone concentration and their quantification. *Journal of Plankton Research*, 15(11):1309–1318.

Tollrian, R. (1995). Chaoborus crystallinus predation on Daphnia pulex: can induced morphological changes balance effects of body size on vulnerability? *Oecologia*, 101(2):151–155.

Trimmer, P. C., Houston, A. I., Marshall, J. A. R., Mendl, M. T., Paul, E. S., and McNamara, J. M. (2011). Decision-making under uncertainty: biases and Bayesians. *Animal Cognition*, 14:1–12.

Trivers, R. (1991). Deceit and self-deception: The relationship between communication and consciousness. In Robinson, M. and Tiger, L., editors, *Man and Beast Revisited*, pages 175–191. Smithsonian, Washington.

Trivers, R. (2000). The elements of a scientific theory of self-deception. *Annals of the New York Academy of Sciences*, 907(1):114–131.

Trivers, R. (2011). *The Folly of Fools: The Logic of Deceit and Self-Deception in Human Life.* Basic Books, New York, New York, USA.

von Hippel, W. and Trivers, R. (2011a). Reflections on self-deception. *Behavioral and Brain Sciences*, 34(1):41–56.

von Hippel, W. and Trivers, R. (2011b). The evolution and psychology of self-deception. *Behavioral and Brain Sciences*, 34(1):1–56.

Vrij, A. (1995). Behavioral correlates of deception in a simulated police interview. *The Journal of Psychology*, 129(1):15–28.

Vrij, A. and Semin, G. R. (1996). Lie experts' beliefs about nonverbal indicators of deception. *Journal of Nonverbal Behavior*, 20(1):65–80.

Vrij, A., Semin, G. R., and Bull, R. (1996). Insight into behavior displayed during deception. *Human Communication Research*, 22(4):544–562.

Walls, M., Caswell, H., and Ketola, M. (1991). Demographic costs of Chaoborus-induced defences in Daphnia pulex: a sensitivity analysis. *Oecologia*, 87(1):43–50.

Walls, M. and Ketola, M. (1989). Effects of predator-induced spines on individual fitness in Daphnia pulex. *Limnology and Oceanography*, 34(2):390–396.

West-Eberhard, M. J. (2003). *Developmental Plasticity and Evolution.* Oxford University Press.

Whitman, D. W. and Agrawal, A. A. (2009). What is phenotypic plasticity and why is it important. *Phenotypic Plasticity of Insects*, 10:1–63.

Williams, G. C. (1966). Adaptation and natural selection: a critique of some current evolutionary thoughts. *Princeton, New Jersey.*

Zahavi, A. (1975). Mate selection - A selection for a handicap. *Journal of Theoretical Biology*, 53(1):205–214.

Zuckerman, M., DePaulo, B. M., and Rosenthal, R. (1981). Verbal and nonverbal communication of deception. *Advances in Experimental Social Psychology*, 14(1):1–59.

Zuckerman, M., Driver, R. E., Siegman, A. W., and Feldstein, S. (1985). Telling lies: verbal and nonverbal correlates of deception. *Multichannel Integrations of Nonverbal Behavior*, pages 129–147.

Appendix A

Supplementary information for Chapter 4

A.1 Expected payoff of an owner in the simplified model presented in Section 4.2

The expected payoff of any owner with internal bias k is calculated as a function of k, as shown in Equation A.1.

$$F(k) = \sum_{A \in S_A} P(A) I[\widehat{p_W}(A,k) > \frac{c}{r}][I(A > 0)r - c] \tag{A.1}$$

This function approximates the owner's payoff as a summation of the weighted partial payoffs the owner receives in simulated encounters with challengers in a set S_A of uniformly sampled values of A. The weighting factor is the probability of each A, denoted by $P(A)$. Each partial payoff in the summation is expressed in terms of an auxiliary identity function of the form $I(C_x)$, which returns unity if the condition represented by C_x holds true and zero otherwise. The expression $\widehat{p_W}(A,k)$ denotes the individual's estimated probability of winning, given A and k.[1] Partial payoffs are added in the calculation of expected payoff if and only if $\widehat{p_W}(A,k) > \frac{r}{c}$, *i.e.*, if the owner decides to fight back. The partial payoff is $r - c$ if the owner is stronger than its opponent in the simulated owner-challenger encounter and $-c$ otherwise.

---

[1] The estimate $\widehat{p_W}(A,k)$ is given by $P(\widehat{\theta_j} < \theta_i + k) = P(\widehat{\theta_j} - \theta_i - k < 0)$. Since $\widehat{\theta_j}$ is normally distributed with mean θ_j and standard deviation σ_ϵ then the sum $\widehat{\theta_j} - \theta_i - k$ is normally distributed with mean $\theta_j - \theta_i - k$ and standard deviation σ_ϵ. Therefore $P(\widehat{\theta_j} - \theta_i - k < 0)$ can be restated as $P(\Theta < 0)$, where $\Theta \sim \mathcal{N}(-A - k, \sigma_\epsilon)$.

A.2 Optimal owner behaviour in the simplified model presented in Section 4.2

In the model introduced in Section 4.2 it holds that if an owner i accepts an opponent j's challenge then the *marginal probability* that i defeats j is given by $p_W = P(\theta_i > \theta_j) = P(A > 0) = \frac{1}{2}$ and therefore an owner's expected payoff from a single fight is $\frac{1}{2}r - c$. An unsophisticated owner that ignores relevant information and fights randomly each time with *fighting probability* p_F is then expected to receive a mean payoff $F_{\min} = p_F(\frac{1}{2}r - c)$ after repeated encounters with random challengers. Owners in this model, however, choose to fight only when their estimated expected payoff is positive, therefore their payoffs in the long term should be higher than $F_{\min}$. Theorem A.1 shows that an owner may fight only if $\frac{r}{c}$ is in the interval $(1, \infty)$. In addition to this, Theorem A.2 shows that for each $\frac{r}{c}$ there is a *capability superiority threshold for conflict*, denoted by z, and that every owner i decides to repel a challenger if and only if $A > z - k_i$. Thus the fighting probability of an owner with bias k_i is given by $p_F = P(A > z - k_i)$ and it increases as $\frac{r}{c}$ and k_i increase and as z and $\frac{c}{r}$ decrease.

Theorem A.1

When $\frac{r}{c} \in (0, 1]$ an owner never fights back.

Proof: If $\frac{r}{c} \in (0, 1]$ then $\frac{c}{r} \in [1, \infty)$ and $\widehat{p_W}$ can never be above $\frac{c}{r}$. Therefore the owner's decision rule (*i.e.*, "fight back if and only if $\widehat{p_W} > \frac{c}{r}$") is never satisfied.

Theorem A.2

Every owner i fights j if and only if $\theta_i + k_i > \theta_j + z$ where z, the *capability superiority threshold for conflict*, is the solution to the equation $\int_{-\infty}^{z} \Phi'(x)dx = \frac{c}{r}$ and $\Phi'(x)$ is the density function of a normal distribution with mean zero and standard deviation σ_ϵ.

Proof: From the fact that $\widehat{p_W} = P(\widehat{\theta_j} < \theta_i + k_i)$ and $\widehat{\theta_j} \sim \mathcal{N}(\theta_j, \sigma_\epsilon)$ it follows that $\widehat{p_W} = \int_{-\infty}^{\theta_i + k_i} \Phi(x)dx$, where $\Phi(x)$ is the density function of the normally-distributed $\widehat{\theta_j}$. Let v be the solution to equation $\int_{-\infty}^{v} \Phi(x)dx = \frac{c}{r}$ and let $z = v - \theta_j$. Therefore $\widehat{p_W} > \frac{c}{r}$ (*i.e.*, the owner fights back) if and only if $\theta_i + k_i > v = \theta_j + z$, where z satisfies the equation $\frac{c}{r} = \int_{-\infty}^{v} \Phi(x)dx = \int_{-\infty}^{\theta_j + z} \Phi(x)dx = \int_{-\infty}^{z} \Phi'(x)dx$ and $\Phi'(x)$ is the density function of a normal distribution with mean zero and standard deviation σ_ϵ.

Marshall et al. (2013b) show that an optimally-behaving individual in the model proposed by Johnson and Fowler (2011) whose opponent is known to fight necessarily retaliates if its marginal probability of winning satisfies the inequality $p_W > \frac{c}{r}$. An optimal owner in the owner-challenger model presented in Section 4.2 should exhibit exactly the same behaviour because its decision is made only in the knowledge that its opponent is determined to fight. If every owner had access to perfect information it would be able to compute accurately its probability of winning and follow the decision rule "fight if and only if $p_W > \frac{c}{r}$" (Marshall et al., 2013b). However, each owner only has access to its own estimated probability of winning ($\widehat{p_W}$), which is likely to deviate from the actual value (p_W) due to the individual's perception error (σ_ϵ) and internal bias (k). The individual uses this information to make a rational decision but due to the uncertainty present it is possible that at some point the individual makes the wrong choice.

Given any randomly selected owner i and challenger j, then by definition i's marginal probability of winning is $p_W = P(A > 0) = \frac{1}{2}$ and this means i is objectively expected to be stronger than half the opponents it encounters. Therefore an optimal owner i that hypothetically takes part in repeated encounters with random challengers should decide to fight in approximately half of those encounters, otherwise i would be missing opportunities to defeat and increase its fitness; or it would risk itself in fights that are bound to be lost, which would in turn have a detrimental effect on its long term fitness. Then it can be predicted that the owners with the highest long-term payoffs must have internal biases that make p_F equal to $\frac{1}{2}$. That is to say, optimal behaviour is a sufficient (but not necessary) condition for $p_F = \frac{1}{2}$. Similarly, a fighting probability $p_F = \frac{1}{2}$ is a necessary (but not sufficient) condition of optimality.

A.3 Details of the evolutionary model presented in Section 4.3

The evolutionary model used in Section 4.3 is as follows. A population of 500 individuals with role-dependent biases is initialized with random standard normal biases. Every generation the population is assorted in such a way that every individual x is paired at random with exactly one adversary y in the population. The capability difference between x and y (denoted by A) is a randomly-chosen standard normal value. Two encounters between x and y are simulated. In the first encounter x plays owner and y plays challenger. In the second encounter the roles are inverted. In each encounter each individual expresses only the biases corresponding to the role played by the individual. That is to say, when x plays the role of an owner it estimates its own capability as $\theta_x + k_{o,x}$ and attempts to deceive y into believing that x's capability is $\theta_x + s_{o,x}$. On the other hand, when x plays the role of a challenger it estimates its own capability as $\theta_x + k_{c,x}$ and attempts to project onto y a capability equal to $\theta_x + s_{c,x}$. The biases of y are expressed analogously. The fitness of x is calculated as the average of the payoff received by this individual in the two encounters. The fitness of y is calculated in the same manner.

One half of the population are selected through stochastic universal sampling (Baker, 1987). A new population is formed consisting of the selected individuals. Those that fail to be selected are replaced by randomly-chosen copies of the selected ones to complete the new population. Approximately 1% of the new population members have their genetically-encoded biases mutated with Gaussian noise. Evolution runs until the $5,000$-th generation, when no considerable changes are observed and the population is assumed to be in equilibrium.

Appendix B

Technical details of the model in Chapter 2

B.1 Source files

The source code that implements the model presented in Chapter 2 can be found in the compressed file `Chapter2.zip` downloadable from `http://goo.gl/fdRBy5`. These sources were written in `C++` (compiled with `g++` version `4.9.1`) and `Matlab` (version `8.1.0.604`). The operating system used was `Scientific Linux release 6.5 (Carbon)`.

The `C++` sources must be compiled using the `makefile` provided.

B.2 Figure 2.2

The subplots in Figure 2.2 are generated by compiling all the `C++` sources, as described in Section B.1, and then running script `Run1.sh`.

B.3 Figure 2.3

The subplots in Figure 2.3 are generated by compiling all the `C++` sources, as described in Section B.1, and then running script `Run2.sh`.

B.4 Figure 2.4

The subplots in Figure 2.4 are generated by compiling all the `C++` sources, as described in Section B.1, and then running script `Run3.sh`.

B.5 Figure 2.5

The subplots in Figure 2.5 are generated by compiling all the `C++` sources, as described in Section B.1, and then running script `Run4.sh`.

B.6 Figure 2.6

The subplots in Figure 2.6 are generated by compiling all the `C++` sources, as described in Section B.1, and then running script `Run5.sh`.

Appendix C

Technical details of the model in Chapter 3

C.1 Source files

The source code that implements the model presented in Chapter 3 can be found in folder `Chapter3/Sources` inside the compressed file `Chapter3.zip` downloadable from `http://goo.gl/usFGkO`. These sources were written in `C++` (compiled with `g++` version `4.9.1`), `R` (version `3.0.2`) and `Matlab` (version `8.1.0.604`). The operating system used was `Scientific Linux release 6.5 (Carbon)`.

Folder `Chapter3/Data` inside the compressed file `Chapter3.zip` contains the data produced with the source code and can be reused to generate the plots in Chapter 3 without having to rerun the code, since some of the programs may take a long time to finish. Details on how to reuse these data appear further below.

C.2 C++ sources

The `C++` sources must be compiled using the `makefile` provided. The source code uses libraries `Boost (Version 1.53.0)` and `GAlib (Version 2.4.7)`. In order to compile the `C++` sources these libraries must be downloaded and installed by following their respective sets of instructions. Once installed, the variables `BOOST_DIR` and `GALIB_DIR` in the `makefile` must be updated to point to the installation folders of the `Boost` and `GAlib` libraries. After doing this, compilation of all the `C++` sources is done by calling the `make` command in `Linux`.

C.3 Figure 3.3

Each subplot in Figure 3.3 is generated by compiling all the `C++` sources, as described in Section C.2, and then running script `Run1.sh`. Doing this will run four new evolutionary simulations with the same parameters as those illustrated in Figure 3.3. Different runs are likely to result in different evolutionary histories thus the newly generated subplots will probably be different from the ones depicted in the current figure.

C.4 Figure 3.4

Each subplot in Figure 3.4 is generated by compiling all the `C++` sources, as described in Section C.2, and then running script `Run2.sh`. Doing this will run four new evolutionary simulations with the same parameters as those illustrated in Figure 3.4. Different runs are likely to result in different evolutionary histories thus the newly generated subplots will probably be different from the ones depicted in the current figure.

C.5 Figure 3.5

Each subplot in Figure 3.5 is generated by compiling all the `C++` sources, as described in Section C.2, and then running script `Run3.sh`. Doing this will run four new evolutionary simulations with the same parameters as those illustrated in Figure 3.5. Different runs are likely to result in different evolutionary histories thus the newly generated subplots will probably be different from the ones depicted in the current figure.

C.6 Figure 3.6

Each subplot in Figure 3.6 is generated by compiling all the `C++` sources, as described in Section C.2, and then running script `Run4.sh`. Doing this will run four new evolutionary simulations with the same parameters as those illustrated in Figure 3.6. Different runs are likely to result in different evolutionary histories thus the newly generated subplots will probably be different from the ones depicted in the current figure.

C.7 Figure 3.7

Each subplot in Figure 3.7, and its corresponding subfigure in Figure 3.8, is generated by compiling all the `C++` sources, as described in Section C.2, and then running script `Run5.sh`.

Doing this will run four new evolutionary simulations with the same parameters as those illustrated in Figure 3.7. Different runs are likely to result in different evolutionary histories thus the newly generated subplots will probably be different from the ones depicted in the current figure.

C.8 Figure 3.9

Each subplot in Figure 3.9 is generated by compiling all the `C++` sources, as described in Section C.2, and then running script `Run7.sh`. Doing this will run four new evolutionary simulations with the same parameters as those illustrated in Figure 3.9. Different runs are likely to result in different evolutionary histories thus the newly generated subplots will probably be different from the ones depicted in the current figure.

C.9 Figure 3.10

Each subplot in Figure 3.10 is generated by compiling all the `C++` sources, as described in Section C.2, and then running script `Run8.sh`. Doing this will run four new evolutionary simulations with the same parameters as those illustrated in Figure 3.10. Different runs are likely to result in different evolutionary histories thus the newly generated subplots will probably be different from the ones depicted in the current figure.

C.10 Figure 3.11

Each subplot in Figure 3.11 is generated by compiling all the `C++` sources, as described in Section C.2, and then running script `Run9.sh`. Doing this will run four new evolutionary simulations with the same parameters as those illustrated in Figure 3.11. Different runs are likely to result in different evolutionary histories thus the newly generated subplots will probably be different from the ones depicted in the current figure.

C.11 Figure 3.12

Figure 3.12 is generated by compiling the `C++` code, as described in Section C.2, and then running `Run11.sh`. This script calls a `C++` program to simulate the evolutionary computation between Bayesians and frequentists with each combination of environment hyperparameters. The average number of Bayesians is output to a `.CSV` file. Then the script calls a *Matlab* program to read the `.CSV` file and plot the data.

The evolutionary simulations can take long. The `.CSV` file generated by `Run11.sh` is stored in compressed file `Run11.zip`. This file can be used to generate the plot more easily by calling only the *Matlab* program.

C.12 Figure 3.13

Figure 3.13 is generated by compiling the `C++` code, as described in Section C.2, and then running `Run12.sh`. This script calls a `C++` program to simulate the evolutionary computation between Bayesians and frequentists with each combination of environment hyperparameters. The average number of Bayesians is output to a `.CSV` file. Then the script calls a *Matlab* program to read the `.CSV` file and plot the data.

The evolutionary simulations can take long. The `.CSV` file generated by `Run12.sh` is stored in compressed file `Run12.zip`. This file can be used to generate the plot more easily by calling only the *Matlab* program.

C.13 Figure 3.14

Figure 3.14 is generated by compiling the `C++` code, as described in Section C.2, and then running `Run21.sh`. This script calls a `C++` program to simulate the evolutionary computation between Bayesians and frequentists with each combination of environment hyperparameters. The average number of Bayesians is output to a `.CSV` file. Then the script calls a *Matlab* program to read the `.CSV` file and plot the data.

The evolutionary simulations can take long. The `.CSV` file generated by `Run21.sh` is stored in compressed file `Run21.zip`. This file can be used to generate the plot more easily by calling only the *Matlab* program.

C.14 Figure 3.15

Figure 3.15 is generated by compiling the `C++` code, as described in Section C.2, and then running `Run22.sh`. This script calls a `C++` program to simulate the evolutionary computation between Bayesians and frequentists with each combination of environment hyperparameters. The average number of Bayesians is output to a `.CSV` file. Then the script calls a *Matlab* program to read the `.CSV` file and plot the data.

The evolutionary simulations can take long. The `.CSV` file generated by `Run22.sh` is stored in compressed file `Run22.zip`. This file can be used to generate the plot more easily by calling only the *Matlab* program.

C.15 Figure 3.16

Figure 3.16 is generated by compiling the `C++` code, as described in Section C.2, and then running `Run23.sh`. This script calls a `C++` program to simulate the evolutionary computation between Bayesians and frequentists with each combination of environment hyperparameters. The average number of Bayesians is output to a `.CSV` file. Then the script calls a *Matlab* program to read the `.CSV` file and plot the data.

The evolutionary simulations can take long. The `.CSV` file generated by `Run23.sh` is stored in compressed file `Run23.zip`. This file can be used to generate the plot more easily by calling only the *Matlab* program.

C.16 Figure 3.17

Figure 3.17 is generated by compiling the `C++` code, as described in Section C.2, and then running `Run24.sh`. This script calls a `C++` program to simulate the evolutionary computation between Bayesians and frequentists with each combination of environment hyperparameters. The average number of Bayesians is output to a `.CSV` file. Then the script calls a *Matlab* program to read the `.CSV` file and plot the data.

The evolutionary simulations can take long. The `.CSV` file generated by `Run24.sh` is stored in compressed file `Run24.zip`. This file can be used to generate the plot more easily by calling only the *Matlab* program.

C.17 Figure 3.18

In order to generate the four subplots in Figure 3.18 the `C++` code must be compiled first, as described in Section C.2.

Figure 3.18a is generated by running `Run13.sh`. This script calls a `C++` program to simulate the evolutionary computation between Bayesians and frequentists with each combination of environment hyperparameters and generate two `.CSV` files with the Bayesian population numbers per environment. Then the script calls a *Matlab* program to read the `.CSV` files and plot the data.

The evolutionary simulations can take long. The `.CSV` files generated by `Run13.sh` are stored in compressed file `Run13.zip`. These files can be used to generate each subplot more easily by calling only the *Matlab* program using the instruction in `Run13.sh`.

Figure 3.18b is generated by running `Run14.sh`. This script calls a `C++` program to simulate the evolutionary computation between Bayesians and frequentists with each combination

bination of environment hyperparameters and generate two `.CSV` files with the Bayesian population numbers per environment. Then the script calls a *Matlab* program to read the `.CSV` files and plot the data.

The evolutionary simulations can take long. The `.CSV` files generated by `Run14.sh` are stored in compressed file `Run14.zip`. These files can be used to generate each subplot more easily by calling only the *Matlab* program using the instruction in `Run14.sh`.

Figure 3.18c is generated by running `Run15.sh`. This script calls a `C++` program to simulate the evolutionary computation between Bayesians and frequentists with each combination of environment hyperparameters and generate two `.CSV` files with the Bayesian population numbers per environment. Then the script calls a *Matlab* program to read the `.CSV` files and plot the data.

The evolutionary simulations can take long. The `.CSV` files generated by `Run15.sh` are stored in compressed file `Run15.zip`. These files can be used to generate each subplot more easily by calling only the *Matlab* program using the instruction in `Run15.sh`.

Figure 3.18d is generated by running `Run16.sh`. This script calls a `C++` program to simulate the evolutionary computation between Bayesians and frequentists with each combination of environment hyperparameters and generate two `.CSV` files with the Bayesian population numbers per environment. Then the script calls a *Matlab* program to read the `.CSV` files and plot the data.

The evolutionary simulations can take long. The `.CSV` files generated by `Run16.sh` are stored in compressed file `Run16.zip`. These files can be used to generate each subplot more easily by calling only the *Matlab* program using the instruction in `Run16.sh`.

C.18 Figure 3.19

In order to generate the four subplots in Figure 3.19 the `C++` code must be compiled first, as described in Section C.2.

Figure 3.19a is generated by running `Run17.sh`. This script calls a `C++` program to simulate the evolutionary computation between Bayesians and frequentists with each combination of environment hyperparameters and generate two `.CSV` files with the Bayesian population numbers per environment. Then the script calls a *Matlab* program to read the `.CSV` files and plot the data.

The evolutionary simulations can take long. The `.CSV` files generated by `Run17.sh` are stored in compressed file `Run17.zip`. These files can be used to generate each subplot more easily by calling only the *Matlab* program using the instruction in `Run17.sh`.

Figure 3.19b is generated by running `Run18.sh`. This script calls a `C++` program to simulate the evolutionary computation between Bayesians and frequentists with each combination of environment hyperparameters and generate two `.CSV` files with the Bayesian population numbers per environment. Then the script calls a *Matlab* program to read the `.CSV` files and plot the data.

The evolutionary simulations can take long. The `.CSV` files generated by `Run18.sh` are stored in compressed file `Run18.zip`. These files can be used to generate each subplot more easily by calling only the *Matlab* program using the instruction in `Run18.sh`.

Figure 3.19c is generated by running `Run19.sh`. This script calls a `C++` program to simulate the evolutionary computation between Bayesians and frequentists with each combination of environment hyperparameters and generate two `.CSV` files with the Bayesian population numbers per environment. Then the script calls a *Matlab* program to read the `.CSV` files and plot the data.

The evolutionary simulations can take long. The `.CSV` files generated by `Run19.sh` are stored in compressed file `Run19.zip`. These files can be used to generate each subplot more easily by calling only the *Matlab* program using the instruction in `Run19.sh`.

Figure 3.19d is generated by running `Run20.sh`. This script calls a `C++` program to simulate the evolutionary computation between Bayesians and frequentists with each combination of environment hyperparameters and generate two `.CSV` files with the Bayesian population numbers per environment. Then the script calls a *Matlab* program to read the `.CSV` files and plot the data.

The evolutionary simulations can take long. The `.CSV` files generated by `Run20.sh` are stored in compressed file `Run20.zip`. These files can be used to generate each subplot more easily by calling only the *Matlab* program using the instruction in `Run20.sh`.

C.19 Figure 3.20

In order to generate the four subplots in Figure 3.20 the `C++` code must be compiled first, as described in Section C.2. Then the figures are produced by running `Run25.sh`.

C.20 Figure 3.21

In order to generate the four subplots in Figure 3.21 the `C++` code must be compiled first, as described in Section C.2. Then the figures are produced by running `Run26.sh`.

C.21 Tests of evolutionary stability

The tests of evolutionary stability used in Section 3.10 are run by compiling the `C++` code first, as described in Section C.2, and then by running `Run27.sh`.

Appendix D

Technical details of the model in Chapter 4

D.1 Source files

The source code that implements the model presented in Section 4.2 and the generalized model presented in Section 4.3 can be found in the compressed file `Chapter4.zip` downloadable from `http://goo.gl/vkyy3U`. These sources were written in the `R` programming language (version `3.0.2`) and `Matlab` (version `8.1.0.604`). The operating system used was `Scientific Linux release 6.5 (Carbon)`.

The simplified model with only internal biases and always-aggressive challengers described in Section 4.2 and the numerical analysis described in Section 4.4 are implemented in source file `Model_1.R` whereas the generalized model with role-dependent internal and external biases described in Section 4.3 and the simulations and numerical analyses described in Section 4.5 are implemented in source file `Model_2.R`. These sources are adaptations from the `R` code by Johnson and Fowler (2011).

The `Matlab` source files `SurfacePlot.m`, `SurfacePlot2.m` and `SurfacePlot3.m` are used for producing Figures 4.2, 4.3, 4.4 and 4.5 from data generated by the `R` source code. The instructions for doing this appear in the following sections.

D.2 Figure 4.2

Figure 4.2a is produced following the steps below.

1. Executing function `Model_1_1()` in the `R` source file `Model_1.R`, which outputs the data file `Expected_payoffs.csv`.

2. Executing function `SurfacePlot()` in `Matlab` source file `SurfacePlot.m` in the same directory as `Expected_payoffs.csv`. The plot is output in the same directory as a file named `Expected_payoffs.pdf`.

Figure 4.2b is produced by running function `Model_1_2()` in the `R` source file `Model_1.R`. The plot is output in the same directory as a file named `optimal_owner_biases.pdf`.

D.3 Figure 4.3

Figure 4.3a and Figure 4.3b are produced following the steps below.

1. Running `Model_2_1()` in the `R` source file `Model_2.R`, which outputs a folder named `Model_2_1` containing two `.csv` files named `1_owners.csv` and `2_challengers.csv`. In order to generate these two files, the program needs to output several auxiliary files into this folder first. While the program is running the size of the folder can reach over `300 MB`, but these auxiliary files are deleted by the program upon completion.

2. Running the `Matlab` source file `SurfacePlot2.m` in the same directory as `1_owners.csv` and `2_challengers.csv` (these two files can copied from folder `Model_2_1`). This program reads the two `.csv` files and produces the plots. The plots are output in the same directory as two files named `owners_sf_plot.pdf` and `challengers_sf_plot.pdf`.

Figure 4.3c and Figure 4.3d are produced following the steps below.

1. Running `Model_2_2()` in the `R` source file `Model_2.R`, which outputs a folder named `Model_2_2` containing two `.csv` files named `1_owners.csv` and `2_challengers.csv`. In order to generate these two files, the program needs to output several auxiliary files into this folder first. While the program is running the size of the folder can reach over `100 MB`, but these auxiliary files are deleted by the program upon completion.

2. Running the `Matlab` source file `SurfacePlot2.m` in the same directory as `1_owners.csv` and `2_challengers.csv` (these two files can copied from folder `Model_2_2`). This program reads the two `.csv` files and produces the plots. The plots are output in the same directory as two files named `owners_sf_plot.pdf` and `challengers_sf_plot.pdf`.

Figure 4.3e and Figure 4.3f are produced following the steps below.

1. Running `Model_2_3()` in the `R` source file `Model_2.R`, which outputs a folder named `Model_2_3` containing two `.csv` files named `1_owners.csv` and `2_challengers.csv`. In order to generate these two files, the program needs to output several auxiliary

files into this folder first. While the program is running the size of the folder can reach over `100 MB`, but these auxiliary files are deleted by the program upon completion.

2. Running the `Matlab` source file `SurfacePlot2.m` in the same directory as `1_owners.csv` and `2_challengers.csv` (these two files can copied from folder `Model_2_3`). This program reads the two `.csv` files and produces the plots. The plots are output in the same directory as two files named `owners_sf_plot.pdf` and `challengers_sf_plot.pdf`.

D.4 Figure 4.4

Figure 4.4a and Figure 4.4b are produced following the steps below.

1. Running `Model_2_4()` in the `R` source file `Model_2.R`, which outputs a folder named `Model_2_4` containing two `.csv` files named `1_owners.csv` and `2_challengers.csv`.

2. Running the `Matlab` source file `SurfacePlot3.m` in the same directory as `1_owners.csv` and `2_challengers.csv`. This program reads the two `.csv` files and produces the plots. The plots are output in the same directory as two files named `owners_sf_plot.pdf` and `challengers_sf_plot.pdf`.

D.5 Figure 4.5

Figure 4.5a and Figure 4.5b are produced following the steps below.

1. Running `Model_2_5()` in the `R` source file `Model_2.R`, which outputs a folder named `Model_2_5` containing two `.csv` files named `1_owners.csv` and `2_challengers.csv`.

2. Running the `Matlab` source file `SurfacePlot3.m` in the same directory as `1_owners.csv` and `2_challengers.csv`. This program reads the two `.csv` files and produces the plots. The plots are output in the same directory as two files named `owners_sf_plot.pdf` and `challengers_sf_plot.pdf`.