# Sex and Fear: Mathematical models of mate choice, parental care, and maladaptive anxiety

Frazer Meacham

A dissertation
submitted in partial fulfillment of the
requirements for the degree of

Doctor of Philosophy

University of Washington

2018

Reading Committee:

Carl T. Bergstrom, Chair

Joseph Felsenstein

Michael Beecher

Program Authorized to Offer Degree:
Biology

University of Washington

## Abstract

Sex and Fear: Mathematical models of mate choice, parental care, and maladaptive anxiety

Frazer Meacham

Chair of the Supervisory Committee:
Carl T. Bergstrom
Department of Biology

In many contexts, animals must infer salient information about another individual indirectly by observing some other characteristic of that individual. In Chapter 1 of this thesis, a model of costly signaling is developed to investigate how stochastic signal costs influence the overall cost of communication. Chapter 2 presents a model of mate choice where females must infer from his appearance whether a potential mate will choose to be a good parent to the future offspring.

Chapters 3 and 4 deal with mathematical models of anxiety disorders. These disorders affect a huge number of people and can be tremendously disabling. But it is clear that the capacity for anxiety is an evolutionary adaptation. This presents a puzzle: why has natural selection not protected us from such a common malfunctioning of an adaptation? Chapter 3 develops a model that shows how the basic information constraints inherent in the problem of learning about an environment can unavoidably cause a subset of the population to be overly sensitive to signs of danger. Chapter 4 addresses the perplexing observation that as the society of developed countries has continually become safer, anxiety has increased rather than decreased. A model is presented that shoes how the mismatch between a modern environment and the environment to which we adapted can cause this seemingly paradoxical increase in levels of anxiety. This result is in some ways analogous to the well-known "hygiene hypothesis" of inflammatory and autoimmune diseases.

# TABLE OF CONTENTS

# LIST OF FIGURES

# LIST OF TABLES

Table Number Page

# ACKNOWLEDGMENTS

During my time as a Ph.D. student in the Biology Department at the University of Washington, I received help, support, and friendship that far exceeded my expectations. Above all was my wonderful adviser, Carl T. Bergstrom, whose excitement for science was contagious and inspiring, and who guided me personally as well as academically, offering me friendship in addition to mentorship. Collaborating with him on the first, third, and fourth chapters of this thesis was a true delight. I also received great support from my committee members, Michael Beecher, Ben Kerr, Joe Felsenstein, and Dan Eisenberg. Among other things, Michael Beecher listened to many different practice talks, offering me tremendously useful advice, Ben Kerr offered personal support as well as technical expertise, Joe Felsenstein helped keep my math rigorous and was a terrific teacher of population genetics, and Dan Eisenberg offered me his expertise in studying humans as I moved into a new field. I am also extremely grateful to Thomas Getty, my collaborator on the second chapter, whose wisdom and advice, especially in crafting the manuscript, were exceptionally helpful. In addition I am grateful to Randy Nesse, who is the only person who could give Carl and me the advice and feedback that we needed as we began studying the evolutionary causes of anxiety and depression.

As the graduate program manager, Marissa Heringer helped me throughout my graduate studies, and was particularly supportive in difficult times. Toby Bradshaw was an excellent teacher, and deserves thanks for supporting the submission process for Chapter 2. Jennifer Nemhauser, John Herron, Adam Leache and many others of the faculty and staff of the Biology Department also offered me valuable instruction, mentorship, and support.

I would also like to thank my fellow graduate students in Biology for all their friendship and encouragement. In particular, Daril Vilhena, as the student senior to me in the Bergstrom lab, gave generously of his time, advice, and friendship. I am also eternally grateful to Elisha Vilhena, Peter Conlin, and Marie Clifford, as well as so many others.

Finally, I thank my sister Zuriah for her companionship and invaluable optimism during times when it was hard for me to see things positively, and my parents Chris and Edith for their heartfelt support of all kinds, especially during hard times. I could not have done it without you.

# DEDICATION

to my parents, Chris and Edith

# INTRODUCTION

Natural selection's astonishing creative power is nowhere more evident than in the sophisticated behaviors of myriad animal species. These behaviors are produced by nervous systems that are among the most complex structures known to us, each neuron of which is built by equally complex interactions of thousands of biomolecules. Because of this nearly impenetrable complexity, we are lucky that there are approaches to understanding the evolution of behavior that are simpler.

From a theoretical perspective, some of the most efficient methods of reducing evolutionary complexity to a point where it can be understood are optimization and game-theoretic models. These models outline precisely what the limits of natural selection are, and what behaviors we can expect to evolve, while relying on a minimum of structural machinery. Once the models are analyzed, this relative simplicity gives us the greatest chance of understanding what causes led to our theoretical result. Thus, these kinds of models can tell us not only what to expect, but *why* we should expect it. The work in this thesis applies the simplicity of these models to the complexity of behavior, with the hope of capturing some essential insights that would be lost in the wash of more complicated models.

Halfway through my dissertation work, I became interested in the evolutionary cause of depression and anxiety disorders, and decided to switch the focus of my research to this problem. Because of this, the first two chapters of this dissertation are game-theoretic models of the evolution of animal behaviors, and the last two chapters are optimization models of the evolution of human behavior. The animal behavior models fall within a long-established research tradition, and reflect a fascination with natural selection's effect on animals that must both cooperate and compete, while relying on information that is never complete. The anxiety models fall within the field now known as evolutionary medicine, which emerged in the early 1990s [95, 64], and were in part my attempt to adapt the modeling paradigm I had become familiar with to a problem with importance for human wellbeing. These two subjects have more in common than is at first apparent. First, the animal studies can apply to humans, and the human studies can apply to animals. Second, all four chapters describe how individuals may have been shaped by natural selection to balance conflicting priorities under conditions of uncertainty. Uncertainty is very tough to deal with, and as we will see, even natural selection's extraordinary power will not always be able to overcome it.

It is my hope that the models of anxiety can ultimately contribute to advancements in treatment methods. Because of their abstract nature, however, they of course cannot be used to discover anything about the physiological or neurological mechanisms that underly these conditions. Instead, my hope is that they can be useful in proposing new hypotheses, in changing how psychologists understand the role of uncertainty in decision making, and perhaps even as a useful tool for how clinicians conceptualize and explain dysfunctioning anxiety.

# Chapter 1

# HONEST SIGNALLING WITH COSTLY GAMBLES

By Frazer Meacham, Aaron K. Perlmutter, and Carl T. Bergstrom

Originally published in *Journal of the Royal Society Interface*, 2013

## *1.1 Abstract*

Costly signalling theory is commonly invoked as an explanation for how honest communication can be stable when interests conflict. However, the signal costs predicted by costly signalling models often turn out to be unrealistically high. These models generally assume that signal cost is determinate. Here we consider the case where signal cost is instead stochastic. We examine both discrete and continuous signalling games and show that, under reasonable assumptions, stochasticity in signal costs can decrease the average cost at equilibrium for all individuals. This effect of stochasticity for decreasing signal costs is a fundamental mechanism that likely acts in a wide variety of circumstances.

## *1.2 Introduction*

Signalling and communication abound in nature and human society [54]. Often, communication takes place between entities that do not share entirely coincident interests. Yet honest communication frequently persists in spite of incentives to deceive. Evolutionary biologists and economists alike have developed a suite of game-theoretic models that aim to explain how communication can originate and be maintained among individuals with partially conflicting interests [13, 80]. Biologists have paid particular attention to the role of signal cost in stabilizing communication [89]. Costly signalling models propose that appropriate signal costs can facilitate honest communication by making deceptive signals so expensive that they become counter-productive. While this class of signalling models allows communication at equilibrium, honesty often comes at considerable cost. Signal costs can be so high that all participants in a costly signalling interaction end up worse off at the signalling equilibrium than in an alternative equilibrium in which no communication takes place [8]. For this reason, there has been considerable interest in understanding how honest signalling can occur without high cost. Researchers have noted that honest signals need not be costly so long as dishonest signals are expensive [38, 26, 88, 48], and proposed that mechanisms such as punishment or spatial structure can further reduce signal costs while allowing honesty to persist [81, 53, 47, 93, 9, 84, 97]. These analyses have generally assumed determinate signal costs. In this paper we study signalling models with stochastic costs and show that this simple difference can have substantial consequences for individuals in terms of their average costs at equilibrium.

We examine a type of action-response game where a *signaller* with private information may send a signal to a *receiver* who must then select a response. Sending a signal carries a cost, which depends on the condition of the signaller. We study the case when this cost is a random variable, and characterize how the average costs at equilibrium depend on the risk preferences of signallers. We show that, when signallers have *decreasing absolute risk aversion* (defined in the following section), stochasticity facilitates honest communication at lower expected cost. We present two models: a discrete action-response game with two signaller qualities, two signals, and two responses, and a continuous signalling game with a continuum of qualities, signals, and responses.

### *1.3 Measures of risk preferences*

To study the relative costliness of signals that involve risk, we must know how an individual's welfare depends on the risk taken. In a biological context, this means we must know how the resource being risked translates into reproductive success or fitness. Many types of resources exhibit diminishing returns. That is, a needy individual's fitness will increase more than a well-off individual's fitness if they both obtain the same amount of additional resources. In an economic context, this is the same as saying that individuals have concave utility functions—or equivalently that they are risk averse. An example of such a function is illustrated in figure 1.1. In the economic context, utility is the analogue of fitness and wealth is the resource of interest. We present our models within an economic framework because economics provides a well-developed theory of risk and precise terminology. However, the models we present are general, and we interpret the implications of our results for biological contexts as well as economic ones.

With the above points in mind, we describe some economic terminology for risk preferences. Any statement about the risk preferences of an individual can be translated into a statement about the shape of her utility function, $u$. If an individual prefers a sure thing of getting \$10 to a bet that has an expected payoff of \$10, i.e, the individual is risk-averse, this is equivalent to saying that the second derivative of $u$ is negative. If an individual is less willing to risk \$10 when poor than when rich, then her utility function exhibits decreasing absolute risk aversion (DARA). The geometric equivalent is that $-u''/u'$ is decreasing. The assumption of DARA is standard in economics, and is supported by empirical studies in humans [19]. In the biological context, assuming DARA means that the fitness consequences of risking resources are more grave when resources are rare. The extent to which this is the norm in nature is an empirical question, but it would be surprising if having more resources did not often put an individual in a position to be more willing to risk some of them.

## *1.4 Discrete model*

In order to understand how risk influences costly signalling, we will compare two signalling games, one in which the signals involve risk and one in which they do not.

### *1.4.1 Deterministic signalling*

We first establish the baseline for comparison: a standard costly signalling game in which individuals signal their wealth by deterministically burning some portion of that wealth. We assume that our signallers have a utility function $u(\cdot)$ that is increasing but concave in

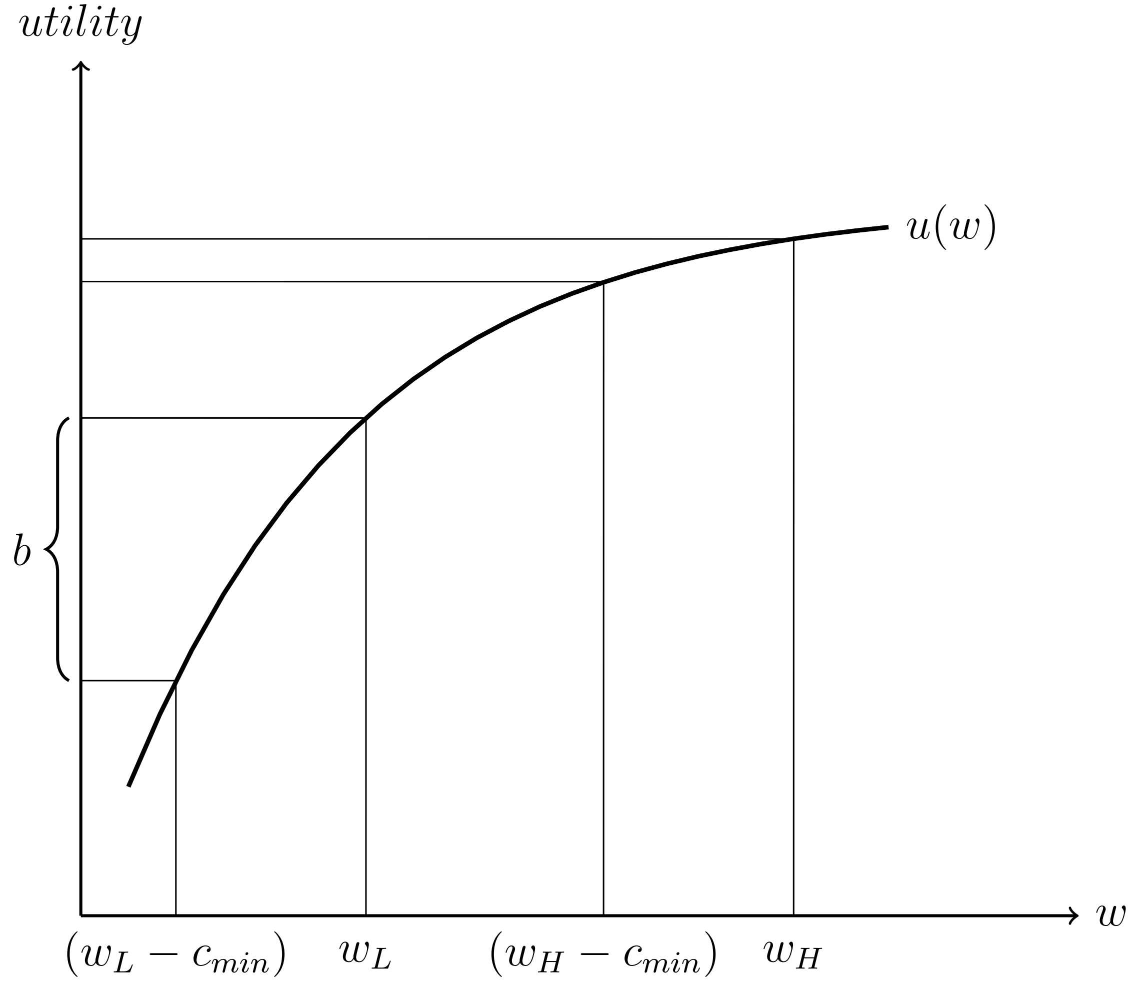


Figure 1.1: The signaller has a concave utility function $u(w)$. The benefit of being accepted by the signal receiver is a utility increment of magnitude $b$. The minimal cost for stable honest signalling is $c_{min}$.

wealth, as in figure 1.1. (Within the economic framework, this utility is conceptualized as von Neumann-Morgenstern utility.)

Our base discrete action-response game is illustrated in figure 1.2. The signaller may be in one of two conditions: High, with a high wealth $w_H$, or Low, with a low wealth $w_L$. The signaller chooses whether or not to send a costly signal by squandering a pre-set amount of money $c$ on a costly signal with no value beyond its communicative role (imagine burning money or buying cut flowers). The receiver then decides whether to accept or to reject the signaller. Receivers do best to accept High signallers and reject Low signallers. Signallers of both types do best to be accepted. Specifically, if accepted, either type of signaller receives a benefit of $b$ (in units of utility, not of wealth). We can thus obtain the signaller's payoffs for each outcome directly from the signaller's utility function. A signalling equilibrium exists when the signal cost $c$ satisfies the following condition:

$$w_L - u^{-1}\left[u(w_L) - b\right] < c < w_H - u^{-1}\left[u(w_H) - b\right].$$

At the signalling equilibrium, High signallers will send a signal and Low signallers will not. Receivers will accept those who signal and reject those who do not. The minimum signal cost that allows honest signalling in this game is thus

$$c_{min} = w_L - u^{-1}\left[u(w_L) - b\right].$$

This baseline model demonstrates that when signallers are risk averse (or equivalently, when the utility of money is concave), it is possible to signal wealth by an "ideal handicap" [32], directly burning some fraction of one's endowment.

### *1.4.2 Stochastic signalling*

To model stochastic signal cost, we alter the game described above by letting the cost for a particular signal be drawn from some probability distribution rather than being a fixed cost. Thus, instead of burning an amount of money $c$, a signaller now takes a risk where the amount of money lost, $Z$, is a random variable. In this case we can obtain the signaller's

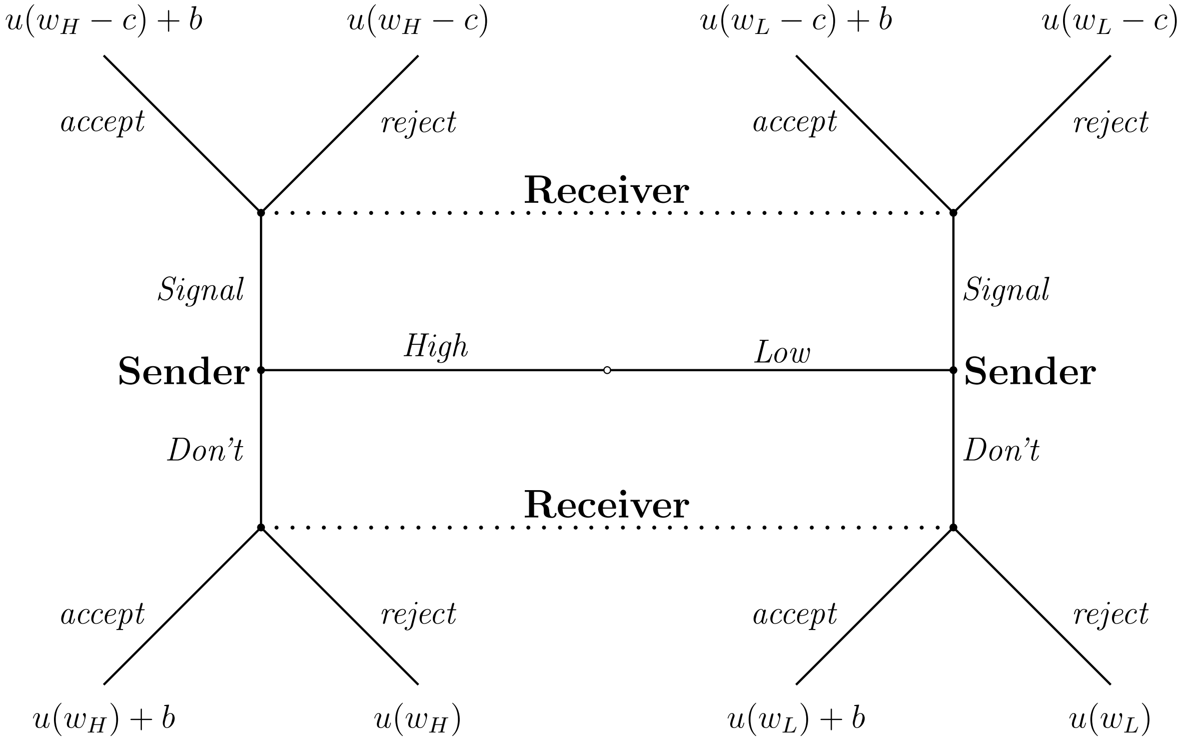


Figure 1.2: An action-response game with cost-free signals and partial conflict of interest. The game begins at the central node (open circle). The first move is a move by "nature" to determine the type of the signaller; this type is revealed to the signaller but not the receiver. In the second move, the signaller conditions its behavior on its type and chooses whether or not to send a signal. As the third move, the receiver must choose between two actions. The receiver can condition on the signal, but not the type; this uncertainty is represented by the dotted lines. Only the payoffs to the signaller are shown at the terminal nodes. Payoffs to the receiver are 1 if accepting a high individual or rejecting a low individual, and 0 otherwise.

*expected* payoffs from the utility function. Now a signalling equilibrium will exist when the lottery $Z$ that describes the stochastic signal cost satisfies

$$E(u(w_L + Z)) + b < u(w_L) \text{ and}$$

$$E(u(w_H + Z)) + b > u(w_H).$$

A lottery $Z$ that minimally allows honest signalling in this game thus satisfies

$$E(u(w_L + Z)) + b = u(w_L).$$

**Example.**

Suppose the signaller has a logrithmic utility function of the form $u(x) = \log_2(x + 1)$. Also suppose that the lottery $Z$ that describes the cost in the stochastic game takes value $-c_2$ with probability $p$ and value 0 with probability $1 - p$, where $0 < p < 1$ and $c_2 > 0$. For this example we will let $p = 1/4$. Suppose that the low-quality signallers have wealth level $w_L = 1$ and the high-quality signallers have wealth level $w_H = 2$. Finally, suppose that the benefit to a signaller of being accepted is $b = 1$ util.

In the deterministic game, the minimal cost $c_1$ needed to make the low-quality signallers have no incentive to signal is given by $u(w_L - c_1) + b = u(w_L)$ and thus,

$$\begin{aligned} c_1 &= w_L - u^{-1}(u(w_L) - b) \\ &= w_L - [2^{\log_2(w_L+1)-b} - 1] \\ &= 1 - [2^{\log_2(1+1)-1} - 1] \\ &= 1. \end{aligned}$$

This signal cost gives high-quality signallers a payoff of

$$\begin{aligned} u(w_H - c_1) + b &= \log_2(w_H - c_1 + 1) + b \\ &= \log_2(2 - 1 + 1) + 1 \\ &= 2. \end{aligned}$$

In the stochastic game, the minimal value $c_2$ needed to ensure that the low-quality signallers have no expected gain from signalling is given by

$$\begin{aligned}
E(u(w_L + Z)) + b &= u(w_L)\\
p\,u(w_L - c_2) + (1-p)u(w_L) + b &= u(w_L)\\
u(w_L - c_2) &= u(w_L) - \tfrac{1}{p}b\\
c_2 &= w_L - u^{-1}(u(w_L) - \tfrac{1}{p}b)\\
c_2 &= w_L - [2^{\log_2(w_L+1)-\frac{1}{p}b} - 1]\\
c_2 &= 1 - [2^{\log_2(1+1)-4} - 1]\\
c_2 &= \frac{15}{8}.
\end{aligned}$$

Thus the expected loss of wealth is (1/4)(15/8)=15/32, which is substantially less than the loss of 1 unit of wealth due to signalling in the deterministic game. This cost gives high-quality signallers an expected payoff of

$$\begin{aligned}
E(u(w_H + Z)) + b &= p\,u(w_H - c_2) + (1-p)u(w_H) + b\\
&= p[\log_2(w_H - c_2) + 1] + (1-p)[\log_2(w_H) + 1] + b\\
&= \tfrac{1}{4}\log_2(2 - \tfrac{15}{8}) + 1 + (1 - \tfrac{1}{4})[\log_2(2) + 1] + 1\\
&= \frac{11}{4},
\end{aligned}$$

which is greater than the payoff of 2 to a high-quality signaller in the deterministic game. So in this example, stochasticity decreases the average cost of signalling both in wealth and in utility. ■

We want to understand the differences between the stochastic and deterministic signalling games in general, and discover whether the outcome of the example above is typical. First, we can say that signallers will lose less money on average in the stochastic game than in the deterministic game. This is because signallers will not have to spend as much wealth on average in the stochastic game. Because signallers are risk averse, their expected utility from a fixed wealth is higher than their expected utility from a lottery with the same expected value. Therefore, in order to maintain the same average utility level (the level at which it is worthwhile to signal) in the stochastic game as in the deterministic game, the expected wealth loss must be less. Next, we want to know if signallers will be better off playing the stochastic game or the deterministic game. This amounts to asking in which game will there be less loss in expected utility due to the costs of signalling.

### *1.4.3 Stochasticity decreases average signal cost*

Before stating our results for the discrete case, we describe the basic economic concepts of *certainty equivalents* and the *coefficient of absolute risk aversion.* For any utility function $u$, the certainty equivalent of some lottery $X$ is the certain wealth level that has the same utility as the expected utility of the lottery $X$. We will write this as $C(X)$. An example is shown in figure 1.3. Because $u$ is concave, i.e., the second derivative is negative, the certainty equivalent $C(X)$ is less than $E(X)$, the expected value of $X$. It turns out that the certainty equivalent depends on the coefficient of absolute risk aversion, which is given by $A(x) = \dfrac{-u''(x)}{u'(x)}$.

**Proposition 1** *A successful signaller in the stochastic game will have higher expected utility than a successful signaller in the deterministic game if and only if the players have decreasing absolute risk aversion.*

**Proof.** First consider the deterministic game. Let $c$ be the maximum amount of money that a low-quality signaller can spend to obtain the reward without receiving a net loss in utility. Thus, $c$ is defined by the equation

$$u(w_L - c) = u(w_L) - b. \tag{1.1}$$

Therefore, in order to be successful, a high-quality signaller must pay a cost of $c + \epsilon_1$, for some arbitrarily small $\epsilon_1 > 0$, and will receive a utility of

$$u(w_H - c - \epsilon_1) + b.$$

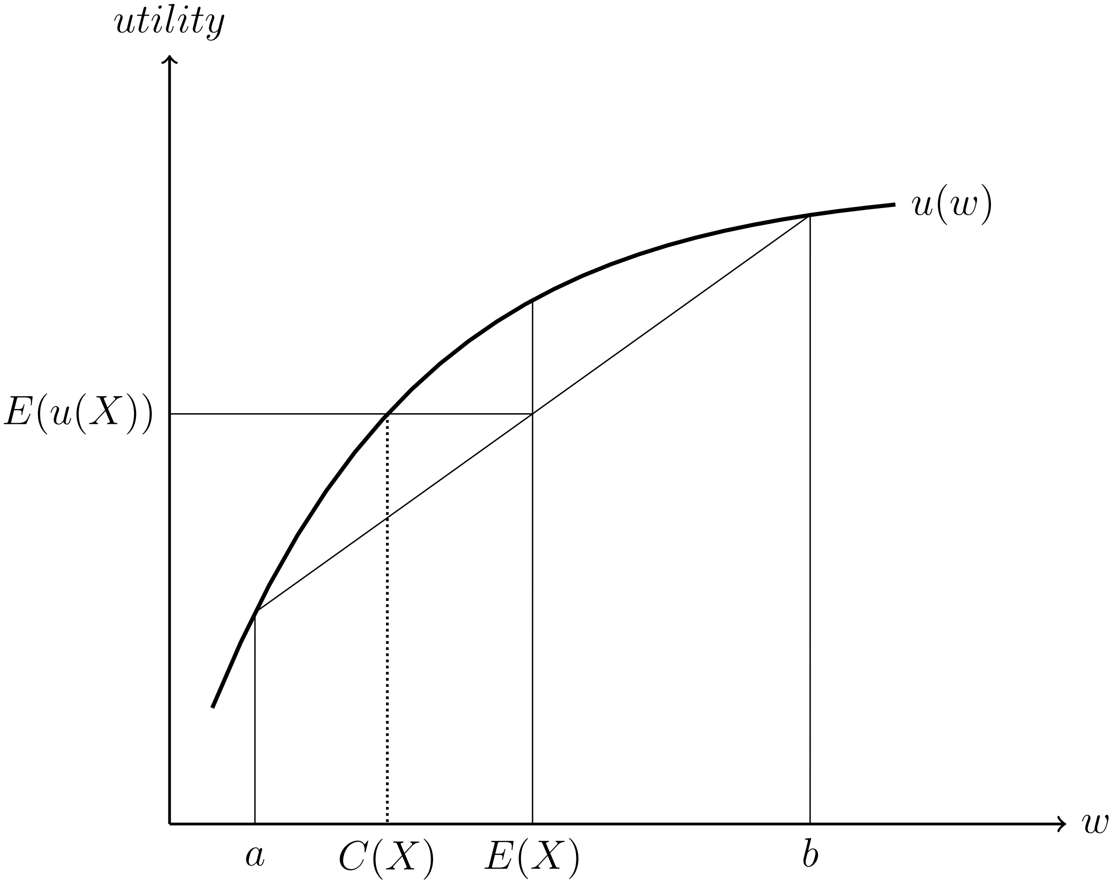


Figure 1.3: The certainty equivalent $C(X)$ of a lottery $X$ is the certain wealth amount such that its utility is equal to the expected utility of the lottery $X$. Illustrated here is the certainty equivalent of the lottery $X$ that pays $a$ with probability 1/2 and $b$ with probability 1/2.

Now consider the stochastic game. Let $Z$ be any random variable with a distribution described by some non-degenerate lottery (i.e., $Z$ takes more than one possible value) such that

$$E(u(w_L + Z)) = u(w_L) - b. \tag{1.2}$$

So if a low-quality signaller risks money in the lottery $Z$ in order to gain the reward, his expected utility will not increase. Therefore, a high-quality signaller can be successful by risking money in the lottery with outcome $Z - \epsilon_2$, for some arbitrarily small $\epsilon_2 > 0$, and will receive an expected utility of

$$E(u(w_H + Z - \epsilon_2)) + b.$$

Thus, a successful signaller in the stochastic game will have higher expected utility than a successful signaller in the deterministic game when

$$E(u(w_H + Z - \epsilon_2)) + b > u(w_H - c - \epsilon_1) + b.$$

Since the epsilons are arbitrarily small, we may move them outside the utility functions and cancel them out along with the $b$ on both sides to get

$$E(u(w_H + Z)) > u(w_H - c). \tag{1.3}$$

We now show that this condition holds when the players have decreasing absolute risk aversion. Define utility function $u_+$ by

$$u_+(x) = u(x + w_H - w_L).$$

Rewriting inequality 3 using $u_+$ we have

$$E(u_+(w_L + Z)) > u_+(w_L - c).$$

Since $u_+$ is increasing, so is $u_+^{-1}$ and we can write

$$\begin{aligned} u_+^{-1}(E(u_+(w_L + Z))) &> w_L - c \\ u_+^{-1}(E(u_+(w_L + Z))) &> u^{-1}(u(w_L - c)). \end{aligned}$$

Equation 1 allows us to rewrite the right-hand side,

$$u_+^{-1}(E(u_+(w_L + Z))) > u^{-1}(u(w_L) - b)$$

and from equation 2 this gives us

$$u_+^{-1}(E(u_+(w_L + Z))) > u^{-1}(E(u(w_L + Z))).$$

This last line says that the certainty equivalent of the lottery $w_L + Z$ is greater for utility function $u_+$ than for $u$. Since the choice of $w_L$ is arbitrary, this is equivalent to the statement that $u$ exhibits greater absolute risk aversion than $u_+$. (See, for example, *Microeconomic Theory* by Mas-Colell, Whinston and Green, 1995 [52].) Since $u_+(x) = u(x + a)$ where $a = w_H - w_L > 0$, this means that $u$ exhibits decreasing absolute risk aversion. So a successful signaller in the stochastic game will have higher expected utility than a successful signaller in the deterministic game precisely when the players have decreasing absolute risk aversion. ■

The next proposition states what probability distribution on the cost of signalling will maximize the utility and wealth level of successful signallers in the stochastic game. We suppose that the lottery $Z$ that describes this cost has a range that is restricted to some interval $[\alpha, \beta]$.

**Proposition 2** *If the signallers have DARA, the expected utility of a successful signaller is maximized when the distribution for $Z$ assigns positive probability only to the endpoints $\alpha$ and $\beta$. This also maximizes the expected wealth level of signallers with concave utility (DARA or otherwise).*

**Proof.** As a preliminary note, if $h$ is a convex function and $X$ is some random variable with $E(X)$ fixed that takes values within $[\alpha, \beta]$, then the distribution for $X$ that maximizes $E(h(X))$ assigns positive probability only to the endpoints $\alpha$ and $\beta$. For suppose that to the contrary there is a distribution for $X$ with some probability mass not at the extreme points. Suppose $c$ is a point between $\alpha$ and $\beta$ that has some positive probability $p > 0$. Let $\epsilon > 0$ be a positive number with magnitude less than the distance between $\alpha$ and $c$ and the distance between $c$ and $\beta$. Then consider the distribution where $c$ has zero probability but $c - \epsilon$ and $c + \epsilon$ each have probability increased by $\frac{1}{2}p$. Then $E(X)$ is not changed, but since $h$ is convex, $\frac{1}{2}h(c-\epsilon) + \frac{1}{2}h(c+\epsilon) > h(c)$. Thus, $E(h(X))$ is increased and so our supposition that there exists a distribution for $X$ with some probability mass not at the extreme points that maximizes $E(h(X))$ is contradicted.

Since a high-quality signaller can be successful by risking money in the lottery $Z$, a low-quality signaller must be just barely unwilling to risk money in this lottery. This gives us the constraint on $Z$

$$E(u(w_L + Z)) = u(w_L) - b - \epsilon \tag{1.4}$$

for some arbitrarily small $\epsilon > 0$. In other words, the distribution for $Z$ is constrained by the fact that $E(u(w_L + Z))$ is constant.

The expected utility of a successful signaller is then

$$E(u(w_H + Z)) + b = E(u_+(w_L + Z)) + b,$$

where we define $u_+(x) = u(x + w_H - w_L)$ as in the proof of Proposition 1.

Since $u$ is increasing, so is $u_+$, and this implies that there exists an increasing function $g$ such that $u(x) = g(u_+(x))$ for all $x$. If $u$ exhibits DARA, than $u_+$ has lower absolute risk aversion than $u$ and this means that $g$ is concave. (See again reference [52].) Therefore, there exists a *convex* function $h = g^{-1}$ such that $u_+(x) = h(u(x))$ for all $x$. This gives us

$$E(u_+(w_L + Z)) = E(h(u(w_L + Z)))$$

and so we can find the distribution of $u(w_L + Z)$ that maximizes $E(h(u(w_L + Z)))$. Now $E(u(w_L + Z))$ is constant, so because $h$ is convex, this distribution is the one that assigns positive probability only to the extreme points, which are $u(w_L + \alpha)$ and $u(w_L + \beta)$ (see note above). Therefore, the distribution for $Z$ that maximizes the expected utility of a successful signaller is the one that assigns positive probability only to the endpoints $\alpha$ and $\beta$.

Since $u^{-1}$ is also convex regardless of whether $u$ exhibits DARA, as long as $u$ is concave, an analogous argument shows that the expected wealth level of a successful signaller is also maximized when the distribution for $Z$ assigns positive probability only to the endpoints $\alpha$ and $\beta$. $\blacksquare$

### *1.5 Continuous signalling*

In our discrete model, there are only two types of signaller, two options for signalling, and two types of response. Alternatively, we imagine a situation where there are signallers with many different wealth levels, many possible signal intensities, and receivers may choose many different responses. The extreme case is when wealth levels, signal intensities, and responses may come from any point along a continuum. This produces a continuous signalling game—a class of model which has been instrumental in the development of the theory of costly signalling (for example, Grafen 1990 [27]). In Grafen's biological interpretation, each signaller has a "quality" instead of a wealth level. Receivers are typically thought of as potential mates. Receivers must gauge a signaller's quality based on the signal intensity, and do best to respond more enthusiastically the higher the signaller's quality.

Following the notation of Bergstrom et al. (2002), signallers have a payoff function $\pi(q, s, r)$ that depends on their own quality ($q$), the intensity of the signal they send ($s$), and the level of response they receive from the receiver ($r$). This payoff function is conceived of as the difference of a benefit function $H(q, r)$ and a cost function $C(q, s)$. The benefit depends on the quality of the signaller and the response it receives, and the cost depends on the quality of the signaller and the intensity of the signal that it chooses to send. Each receiver has a payoff function $G(q, r)$ that depends on how appropriate the response ($r$) is given the signaller's true quality ($q$). Of course the receiver knows only what signal intensity ($s$) the signaller chose. A strategy for signallers is a function $s = s(q)$ that specifies a choice of signal intensity for all signaller qualities $q$. A strategy for a receiver is a function $r = r(s)$ that specifies a choice of response for all signal intensities that the signaller might send. If the functions $s(q)$ and $r(s)$ make up a signalling equilibrium, this means that the function

$s$ is one-to-one and that neither player can benefit by unilaterally modifying its strategy function.

Bergstrom et al. (2002) give a method of finding the functions $s(q)$ and $r(s)$ that make up a signalling equilibrium for any particular game of the above form [10]. Building upon this method, we prove that when signallers have decreasing absolute risk aversion, stochasticity decreases average signal cost in continuous signalling games as well as in discrete games.

As we did for the discrete case, we will describe two signalling games, one deterministic and one stochastic, and compare the average payoffs at equilibrium. For both games, we assume that benefit is proportional to response level, and that signal intensity is proportional to signal cost. The receiver's payoff $G(q, r)$, for how appropriate the response is given the signaller's true quality, is also the same for both games. We will call the signal intensity functions for the deterministic and stochastic games $S_B$ and $S_G$ respectively (for Burning money or Gambling money). So for the deterministic game, a signal of intensity $S_B$ will cost $S_B$ units of wealth. For the stochastic game, a signal of intensity $S_G$ will cost $S_G$ units of wealth with probability $p$ and 0 units of wealth with probability $1 - p$, where $0 < p < 1$.

For the deterministic regime, the payoff to a successful signaller with wealth level $w_H$ in the discrete game was

$$b + u(w_H - c),$$

where $c$ is the signal cost and $b$ is the benefit of being accepted by the receiver. In the continuous case, cost is proportional to signal intensity and the benefit is proportional to response level. This gives us

$$\pi_B = r + u(q - S_B) \tag{1.5}$$

as the payoff function for signallers in the continuous game with deterministic costs.

In the stochastic regime, the payoff to a successful signaller with wealth level $w_H$ in the discrete game was $b + u(w_H + Z)$ giving an expected payoff of

$$b + E(u(w_H + Z)).$$

Exchanging $r$ for $b$ and $q$ for $w_H$ gives us the expected payoff for a signaller in the continuous case.

$$\pi_G = r + E(u(q + Z))$$

The lottery $Z$ takes value $-S_G$ with probability $p$ and value 0 with probability $1-p$. Therefore, we can write

$$\pi_G = r + p\,u(q - S_G) + (1 - p)u(q). \tag{1.6}$$

Having defined the strategy space and payoff functions for both games, the problem now is to find a general solution for the equilibrium response functions of the signal intensities and signal intensity functions of quality. We give the proof of the following proposition in the appendix.

**Proposition 3** *If the players have decreasing absolute risk aversion, then at equilibrium they will have higher expected utility in the stochastic signalling game than in the deterministic signalling game.*

### *1.6 Discussion*

Signalling models in both biology and economics have typically assumed determinate costs. In the real world, signal costs will often if not always be stochastic. This difference matters. Here we show that when signallers have realistic risk preferences, stochastic signal costs result in signals that are cheaper, on average, than when signals have determinate costs. This comparative result holds in discrete and continuous models alike.

In biology, signal costs may be stochastic for a variety of reasons. Begging calls are likely costly because of stochastic predation risk instead of determinate energy expenditures [55, 50, 14, 31]. Physical ornamentation such as long tails or colorful plumage in birds may similarly be costly due to predation risk [62]. Extravagant territorial and courtship displays can be risky as well: instead of storing resources for lean times, an individual invests time and energy in prolonged displays [96].

Many if not most biological instances of stochastic signal costs will be more complicated in form than the simple lotteries modeled here. The important point is that our analysis shows that to simply treat stochastic costs as equivalent to their expectation will often lead to a distorted picture of the true costs. And our results suggest that variable signal costs, rather than undermining honesty in costly signalling, in fact bolster it.

Empirical studies could provide evidence for the action of gambles to decrease average signal costs. The greater the variance in the stochastic cost of a signal, the more likely it is that the signal cost is being reduced by the stochasticity (see our Proposition 2). This suggests the need for empirical studies to take into account risk structure when measuring signal costs. If the risk structure has high variance, then high average cost is not as important for honest signalling.

In the last couple decades, researchers have described a number of systems in which honest communication is less costly than in traditional handicap theory. Such efforts are essential if we are to explain the large number of different contexts in which communication is found to be stable. The effects of stochasticity for decreasing signal costs is another fundamental mechanism that deserves attention because of the wide variety of circumstances in which it likely acts.

### 1.7 Acknowledgements

This work was supported by NSF grant EF-1038590 to CTB.

### 1.8 Appendix: Proof of proposition 3

We first use the method from Bergstrom et al. [10] to obtain differential equations for the signalling strategy under burning money, $S_B(q)$, and under gambling money, $S_G(q)$. For burning money, the signaller's payoff function is

$$\pi_B = r + u(q - S_B).$$

We can break this function into the difference of a benefit function $H(q,r)$, that depends on the signaller's quality $q$ and the receiver's response level $r$, and a cost function $C_B(q,S_B)$ that depends on $q$ and the signal intensity, $S_B$. Indeed, if

$$\begin{aligned} H &= r + u(q), \\ C_B &= u(q) - u(q - S_B), \end{aligned} \tag{1.7}$$

then $\pi_B = H - C_B$.

Similarly, for gambling money we have

$$\pi_G = r + p\,u(q - S_G) + (1-p)u(q),$$

which is broken down into $\pi_G = H - C_G$ as follows:

$$\begin{aligned} H &= r + u(q), \text{ same as before, and} \\ C_G &= p\big(u(q) - u(q - S_G)\big). \end{aligned} \tag{1.8}$$

Following Bergstrom et al (2002), we obtain the differential equation

$$\frac{dS}{dq} = \frac{\partial H}{\partial r}\frac{dR^*}{dq} \Big/ \frac{\partial C}{\partial S},$$

which provides the slope of the signalling strategy $S(q)$ in terms of the benefit function $H$, the cost function $C$ (which depends on $S(q)$ itself), and the equilibrium response level $R^*(q)$.

From expressions (1.7) and (1.8) we see that $\partial H/\partial r = 1$. And if we denote by $r'(q)$ the derivative $dR^*/dq$, then for burning money,

$$\begin{aligned} \frac{dS_B}{dq} &= \frac{\partial H}{\partial r}\frac{dR^*}{dq} \Big/ \frac{\partial C_B}{\partial S_B} \\ &= \frac{dR^*}{dq} \Big/ \frac{\partial C_B}{\partial S_B} \\ &= \frac{r'(q)}{u'(q - S_B)}. \end{aligned} \tag{1.9}$$

And for gambling money,

$$\begin{aligned} \frac{dS_G}{dq} &= \frac{dR^*}{dq} \Big/ \frac{\partial C_G}{\partial S_G} \\ &= \frac{r'(q)}{p\,u'(q - S_G)}. \end{aligned} \tag{1.10}$$

Since the benefit function $H$ is the same for both games, and at the separating equilibrium the receiver's response $r$ will be the same for both games, signallers in the stochastic game will do better than signallers in the deterministic game when $C_G < C_B$. Therefore, we want to show that when $u$ exhibits DARA, $C_G < C_B$. To do so, we first show that $C_G < C_B$ is equivalent to inequality (1.13) below, and then show that inequality (1.13) follows when $u$ exhibits DARA.

From equations (1.7) and (1.8), $C_G < C_B$ gives us

$$p\big(u(q) - u(q - S_G)\big) < u(q) - u(q - S_B),$$

i.e.,

$$u(q - S_B) < p\,u(q - S_G) + (1 - p)u(q). \tag{1.11}$$

Rewriting to isolate $S_G$, we have

$$\tfrac{1}{p}u(q - S_B) + (1 - \tfrac{1}{p})u(q) < u(q - S_G).$$

Since $u$ is increasing, so is $u^{-1}$. Thus, $C_G < C_B$ when

$$S_G < q - u^{-1}\big(\tfrac{1}{p}u(q - S_B) + (1 - \tfrac{1}{p})u(q)\big).$$

Define

$$S_G^* = q - u^{-1}\big(\tfrac{1}{p}u(q - S_B) + (1 - \tfrac{1}{p})u(q)\big) \tag{1.12}$$

so that $C_G < C_B$ if $S_G < S_G^*$.

Note that $S_G^*$ is a function of $q$ and consider the value of the differential equation (1.10), i.e.,

$$\frac{dS_G}{dq} = \frac{r'(q)}{p\,u'(q - S_G)}$$

along the curve $S_G^*(q)$. Note that if $\frac{dS_G}{dq}\Big|_{S_G = S_G^*} < \frac{dS_G^*}{dq}$ for $q > 0$ then $S_G < S_G^*$ for $q > 0$ since $S_G(0) = 0 = S_G^*(0)$. So $C_G < C_B$ when $\frac{dS_G}{dq}\Big|_{S_G = S_G^*} < \frac{dS_G^*}{dq}$, i.e., when $\frac{r'(q)}{p\,u'(q - S_G^*)} < \frac{dS_G^*}{dq}$.

Substituting for $S_G^*$:

$$\frac{r'(q)}{p\,u'\Big(q-\big[q-u^{-1}\big(\tfrac{1}{p}u(q-S_B)+(1-\tfrac{1}{p})u(q)\big)\big]\Big)} < \frac{d}{dq}[q-u^{-1}(\tfrac{1}{p}u(q-S_B)+(1-\tfrac{1}{p})u(q))]$$

Simplifying the left hand side and evaluating the derivative on the right hand side,

$$\frac{r'(q)}{p\,u'\Big(u^{-1}\big(\tfrac{1}{p}u(q-S_B)+(1-\tfrac{1}{p})u(q)\big)\Big)} < 1-(u^{-1})'\Big(\tfrac{1}{p}u(q-S_B)+(1-\tfrac{1}{p})u(q)\Big)\left(\tfrac{1}{p}u'(q-S_B)\left(1-\frac{dS_B}{dq}\right)+(1-\tfrac{1}{p})u'(q)\right)$$

Applying the inverse rule for derivatives, we get

$$\frac{r'(q)}{p\,u'\Big(u^{-1}\big(\tfrac{1}{p}u(q-S_B)+(1-\tfrac{1}{p})u(q)\big)\Big)} < 1-\frac{1}{u'(u^{-1}(\tfrac{1}{p}u(q-S_B)+(1-\tfrac{1}{p})u(q)))}\left(\tfrac{1}{p}u'(q-S_B)\left(1-\frac{dS_B}{dq}\right)+(1-\tfrac{1}{p})u'(q)\right)$$

Multiplying by the (positive) denominator of the left hand side yields

$$r'(q) < p\,u'(u^{-1}(\tfrac{1}{p}u(q-S_B)+(1-\tfrac{1}{p})u(q))) - \left(u'(q-S_B)\left(1-\frac{dS_B}{dq}\right)-(1-p)u'(q)\right)$$

Rearranging and substituting expression (1.9) gives

$$r'(q) < p\,u'(u^{-1}(\tfrac{1}{p}u(q-S_B)+(1-\tfrac{1}{p})u(q))) + (1-p)u'(q) - u'(q-S_B)\left(1-\frac{r'(q)}{u'(q-S_B)}\right)$$

And after a bit of algebra we obtain

$$u'(q-S_B) < p\,u'(u^{-1}(u(q)-\tfrac{1}{p}(u(q)-u(q-S_B)))) + (1-p)u'(q)$$

If we let $d = u(q) - u(q-S_B)$ be the distance between $u(q)$ and $u(q-S_B)$, the above inequality becomes

$$u'(q-S_B) < p\,u'(u^{-1}(u(q)-\tfrac{1}{p}d)) + (1-p)u'(q) \tag{1.13}$$

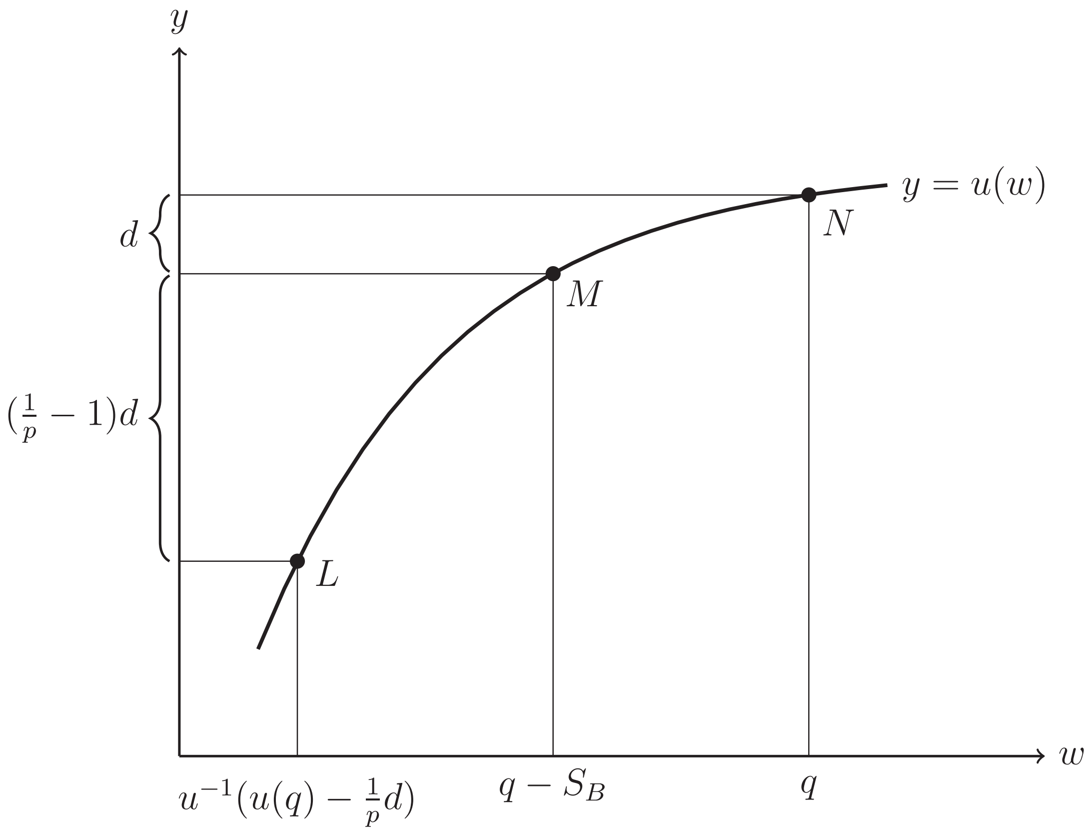


Figure 1.4: Inequality 1.13 says that the derivative of $u$ at point $M$ is less than the weighted average of the derivatives at $L$ and $N$. Notice that the weighted average of $y = u(q)$ and $y = u(q) - \frac{1}{p}d$ is $p\,(u(q) - \frac{1}{p}d) + (1-p)u(q) = u(q - S_B)$.

This inequality has a nice geometric interpretation, illustrated in figure 1.4. It says that the derivative of $u$ at point $M$ is less than the weighted average of the derivative at points $L$ and $N$.

We now show that inequality (1.13) follows when $u$ exhibits DARA. We first point out that the statement that $u$ exhibits DARA is equivalent to the statement that the rate of decrease of $u'(w)$, with respect to $y$, is decreasing. Thus, in figure 1.4, $u'$ decreases proportionately more from $L$ to $M$ than from $M$ to $N$. Thus, $u'$ at $N$ is not small enough to balance out the value of $u'$ at $L$, and so the weighted average is greater than the single value $u'(q - S_B)$.

Indeed, if $u$ exhibits DARA, then by definition $-u''(w)/u'(w)$ is decreasing in $w$. Since $u^{-1}$ is increasing, this implies that $-u''(u^{-1}(y))/u'(u^{-1}(y))$ is decreasing in $y$. Thus,

$$\begin{aligned}\frac{-u''(u^{-1}(y))}{u'(u^{-1}(y))} &= -u''(u^{-1}(y))(u^{-1})'(y)\\ &= \frac{d}{dy}[-u'(u^{-1}(y))]\end{aligned}$$

is decreasing as well (the above equalities follow from the inverse rule of derivatives and the chain rule respectively). So the rate of decrease of $u'(w)$, with respect to $y$, is decreasing.

This means that, in figure 1.4, the difference between $u'$ at $L$ and $u'$ at $M$ is more than $(\frac{1}{p} - 1)$ times the difference between $u'$ at $M$ and $u'$ at $N$. Let's call these differences $D_1$ and $D_2$ respectively, so we have $D_1 > (\frac{1}{p} - 1)D_2$. Therefore, by the definitions of $D_1$ and $D_2$,

$$u'(u^{-1}(u(q) - \tfrac{1}{p}d)) - u'(q - S_B) > (\tfrac{1}{p} - 1)(u'(q - S_B) - u'(q)).$$

Rearranging, we get

$$p\,u'(u^{-1}(u(q) - \tfrac{1}{p}d)) + (1 - p)u'(q) > u'(q - S_B),$$

which is inequality (1.13). So DARA gives us inequality (1.13), which is equivalent to $C_G < C_B$, and we have proved the result.

### *1.8.1 Equilibrium stability*

We next apply the second part of Bergstrom et al's result to show that the equilibrium strategies $S_B(q)$ and $S_G(q)$ we found above are stable (i.e., the extrema are maxima rather than minima). Their result states that the equilibrium is stable when the following second-order condition holds everywhere along the solution curve.

$$\frac{d}{dq}\frac{d}{dp}H(q, R^*(p)) > \frac{\frac{\partial^2}{\partial s \partial q}C(q, s)\,\frac{d}{dq}H(q, R^*(p))}{\frac{\partial}{\partial s}C(q, s)} \tag{1.14}$$

We have $H(q, r) = r + u(q)$. So

$$H(q, R^*(p)) = R^*(p) + u(q).$$

Thus,

$$\frac{d}{dp}H(q, R^*(p)) = \frac{d}{dp}R^*(p)$$

and so

$$\frac{d}{dq}\frac{d}{dp}H(q, R^*(p)) = 0.$$

Also,

$$\frac{d}{dq}H(q, R^*(p)) = u'(q).$$

For the cost function, we have $C(q, s) = u(q) - u(q - s)$ for burning money and $C(q, s) = p(u(q) - u(q - s))$ for gambling money. For burning money, this gives us

$$\frac{\partial}{\partial s}C(q, s) = u'(q - s)$$

and

$$\frac{\partial^2}{\partial s \partial q}C(q, s) = u''(q - s).$$

For gambling money, we have

$$\frac{\partial^2}{\partial s \partial q}C(q, s) = p\,u''(q - s).$$

For both cases, inequality 1.14 then reduces to

$$0 > \frac{\delta\,u''(q - s)\,u'(q)}{u'(q - s)}$$

where $\delta$ is either 1 or $p$. Since $u$ is increasing, $u'$ is positive, so the above inequality holds when $u''(q - s)$ is negative, i.e., when utility is concave. Thus, the second-order condition holds since we are only considering individuals with concave utility.

## Chapter 2

# DEADBEATS OR LOSERS: DISCRETIONARY MALE PARENTAL INVESTMENT CAN MAKE FEMALES LESS CHOOSY

By Frazer Meacham and Thomas Getty

Originally published in *Journal of Theoretical Biology*, 2017

### *2.1 Abstract*

Two of the most important reproductive decisions that animals face are how to choose mates and how to invest in offspring. In species where both males and females provide offspring care, these selection pressures will often be reciprocally intertwined: mate preferences may depend on parental investment patterns while parental investment patterns may depend on mate preferences. We describe and analyze a mathematical model of this interaction, in which females can choose amongst males who have high attractiveness or low attractiveness, while males can decide whether to provide offspring care. We compare the case where males decide whether to provide care to the cases where males always provide care and where they never provide care. For a wide range of parameter settings, we find that when males decide whether to provide care, females are selected to be less choosy. This reduction in female choosiness occurs even though discretionary male care leads to greater variation among males in their offspring output. This finding contrasts with previous theoretical studies, and is driven by our assumption that males can decide whether to help provide care after mating occurs. Our results show how the interdependencies between mate choice and parental care can generate outcomes that can only be understood by considering both processes simultaneously.

## *2.2 Introduction*

Two of the most important reproductive decisions that animals face are how to choose mates and how to invest in offspring. Parental investment (PI) theory describes how individuals balance investment in current offspring against future reproductive efforts [86], while sexual selection theory has extensively investigated the strategies animals use to select good mates [39, 77] and the outcomes of mate competition [46]. But studying mate choice and PI separately ignores how these processes may sometimes be inextricably intertwined [1]. Especially when both sexes care for offspring, mate preferences will depend in part on the amount of PI that different potential mates will contribute. The amount of PI that individuals contribute will depend on whether their time would be better spent pursuing new potential mates [21, 94]. But whether an individual should pursue new potential mates depends on the mate preferences of those other individuals. Thus, mate preferences depend on parental investment patterns while parental investment patterns depend on mate preferences, creating a circle that cannot be disentangled. To understand either process, we must consider both simultaneously. This means that even if we are only interested in mate preferences for physical traits, we must incorporate how these preferences interact with offspring care decisions.

Many studies at least partially address the interdependence of mate choice and parental care. For example, "good father" theories of sexual selection are specifically interested in how female choice influences male parental investment [33, 35]. However, few studies address how this interdependence tempers mate preferences for other traits, even though there is good evidence that tradeoffs exist between mate choice for parental care and for other traits [23, 75]. These tradeoffs may result in females varying their preferences [4, 17, 79], or even preferring "lower-quality" males [37, 78], instead of tending to prefer males of the best physical quality.

[20] presented a mathematical model that showed that males of lower genetic quality can be preferred by females if such males provide better parental care. However, their model

did not include the evolutionary dynamics of female strategies, and so did not model the interaction between male and female behavior. Other theoretical studies of mate choice and parental investment tend to ignore physical quality of potential mates. Instead, studies have addressed how female preference may select for observable male parental care [2], or how honest male signaling might arise if females cannot observe the male's future PI [33, 74, 45].

In this paper, we construct a model of interaction of female choice and male parental investment. The model explicitly includes selection for both mate choice decisions and parental investment decisions. It also includes effects on reproductive success of both indirect and direct benefits. This allows us to examine the interdependence of mate choice and parental care as well as how this interdependence affects mate-choice tradeoffs between different kinds of benefits. We therefore are able to show how female preferences for male physical quality affect male PI decisions, and how these male decisions reciprocally influence female preferences. Our model captures the entanglement of these two selection pressures, and gives insight into what can result.

## ***2.3 Model***

We develop a game theory model with infinite populations. Because we are interested in the interaction of several different factors, our model is designed to be as simple as possible while capturing these interactions. As well as making the analyses more tractable, this also makes the model more transparent so that we have the best chance of understanding the causal mechanisms underlying the results.

### *2.3.1 Assumptions*

Females are all of the same quality when it comes to their intrinsic ability to provide direct and indirect maternal contributions to offspring fitness. We assume that females must always stay with their young to provide parental care. In contrast, males can choose after mating whether or not to help provide offspring care. We assume that development progresses according to a time schedule that does not depend on whether the male helps provide care.

Adult males come in two types: high-quality and low-quality. High quality males have better indirect contributions to offspring fitness, as well as having more valuable parental investment potential (if they should choose to provide it). Before adulthood, each male randomly becomes either high-quality (with probability $p$) or low-quality (with probability $1-p$). We assume infinite population sizes, so the fraction of all males that are high-quality is $p$ and the fraction that are low-quality is $1-p$.

Male quality is perfectly discernible to females. A male's inclination towards providing future parental care is not.

*Strategies*

Each female has three possible strategies available to her: only mate with high-quality males, only mate with low-quality males, or be indiscriminate and mate with both. We denote the proportion of each of these strategies by $x_H$, $x_L$, and $x_I$ respectively. These proportions sum to 1.

After mating, each male, whether high- or low-quality, chooses between staying to help his mate care for their offspring or abandoning his mate to search for other females immediately. Males take into account their own quality when making this decision. (Recall that females do not vary in quality.) Thus, we have four male strategies: the male strategy that provides care whether or not it is a high quality male (High:Care; Low:Care), the strategy that abandons the female if it is high quality but provides care if it is low quality (High:Abandon; Low:Care), the strategy that abandons the female whether or not it is high quality (High:Abandon; Low:Abandon), and the strategy that provides care if it is high quality but abandons the female if it is low quality (High:Care; Low:Abandon). We denote the proportions of males that follow these strategies by $y_{CC}$, $y_{AC}$, $y_{AA}$, $y_{CA}$, respectively (see Table 2.1). The male proportions also sum to 1.

### *Offspring*

The average number of offspring that a pair raises to adulthood depends on both the quality of the father and on whether he stays to help the female care for the offspring. (We assume that individuals reach adulthood as soon as they become independent.) The model does not include any extra pair paternity, so by definition the mother of the brood and the father of the brood get the same number of offspring from it.

We let $B$ be the average number of offspring produced by a pair where the male is low quality and abandons. Then we let $m_H > 1$ be the multiplier for having the male be high quality, and $m_C > 1$ be the multiplier for having the male provide offspring care. Thus, the average number of offspring produced by a pair where the male is both high quality and helps care is $m_C m_H B$, the average number for a pair where the male is high quality and does not help care is $m_H B$, and the average number for a pair where the male is low quality but helps care is $m_C B$.

We express the benefits of the male being high quality or helping care for the offspring as multipliers that are independent of each other because we do not want there to be a difference in the intrinsic incentive to care for offspring for low quality and high quality males. Otherwise, females might prefer low quality males for the straightforward reason that low quality males intrinsically provide greater reproductive benefit. Allowing this kind of situation would, it seems to us, start to blur the notion of what “low quality” would mean.

In general we assume that $m_C > m_H$. In other words, a female will do better to mate with a low quality male who provides care than with a high quality male who does not. In section 2.4.6 however, we briefly look at what happens if we relax this assumption.

### *Reproductive success and mate search*

We calculate an individual’s reproductive success as its long-term average rate of offspring production. In other words, we have an infinite time horizon, and are only interested in an individual’s long-term rate of offspring production. So, not only does brood size matter, but so does the average time it takes an individual to find mates between raising broods. To model this, we assume that at any time, all individuals are either caring for their current offspring or searching for their next mate. We let $T$ be the amount of time it takes for offspring to reach adulthood and independence. (Recall that development progresses according to a time schedule that does not depend on whether the male helps provide care.) Once individuals are done caring for their offspring, they return to the search pool.

In the search pool, males and females encounter each other at a rate proportional to their densities. We will let $\lambda$ be the parameter that scales the encounter rate. (The critical relationship is the amount of time it takes to search for mates compared to the amount of time it takes to raise offspring.) We let $r$ be the sex ratio of males per female. Let $s_H$, $s_L$, and $s_I$ be the densities of the three female strategies in the search pool. These are the proportion of females that are searching and prefer high-quality males, the proportion that are searching and prefer low-quality males, and the proportion that are searching and indiscriminate, respectively. Let $s_{H:C}$, $s_{H:A}$, $s_{L:C}$, $s_{L:A}$ be the relative density (proportion multiplied by the sex ratio) of the four possible kinds of male. These are the relative searching density of males that are high-quality and will provide care, high-quality and will abandon, low-quality and will provide care, and low-quality and will abandon, respectively. Table 2.1 lists all the notation introduced above.

### *2.3.2 Derivations*

For simplicity, we separate the time scales of the search dynamics and the evolutionary dynamics. So for the purposes of calculating the fitness of each strategy, we assume that the search dynamics are at equilibrium.

#### *Search pool densities*

Because they spend all their time in the search pool, the density of high-quality males who abandon equals the sex ratio times the fraction of males that are high quality times the

| Parameters | |
|---|---|
| Fraction of males that are high quality | $p$ |
| Sex ratio (males per female) | $r$ |
| Encounter rate | $\lambda$ |
| Time to raise offspring | $T$ |
| RS when male is low quality and abandons | $B$ |
| High quality multiplier | $m_H$ |
| Caring multiplier | $m_C$ |

| Strategy frequencies | |
|---|---|
| Female strategy frequencies | |
| Prefer high quality | $x_H$ |
| Prefer low quality | $x_L$ |
| Indiscriminate | $x_I$ |
| Male strategy frequencies | |
| High:Care; Low:Care | $y_{CC}$ |
| High:Abandon; Low:Care | $y_{AC}$ |
| High:Abandon; Low:Abandon | $y_{AA}$ |
| High:Care; Low:Abandon | $y_{CA}$ |

| Searching densities | |
|---|---|
| Searching density of female strategies | |
| Prefer high quality | $s_H$ |
| Prefer low quality | $s_L$ |
| Indiscriminate | $s_I$ |
| Searching density of male types | |
| High quality, provide care | $s_{H:C}$ |
| High quality, abandon | $s_{H:A}$ |
| Low quality, provide care | $s_{L:C}$ |
| Low quality, abandon | $s_{L:A}$ |

fraction of high-quality males who abandon:

$$s_{H:A} = r\,p\,(y_{AC} + y_{AA}) \tag{2.1}$$

Similarly for the density of low-quality males who abandon, we have

$$s_{L:A} = r\,p\,(y_{AA} + y_{CA}) \tag{2.2}$$

The density in the search pool of males who are high-quality and help care equals the sex ratio ($r$) multiplied by the proportion of males of that type ($p\,(y_{CC}+y_{CA})$) multiplied by the proportion of time that a male of that type spends searching for mates rather than caring for offspring. This proportion is given by the average time it takes for a male of that type to find a willing mate divided by the average total time spent per mate. We have

$$s_{H:C} = r\,p\,(y_{CC} + y_{CA})\,\frac{\frac{1}{\lambda\,(s_H+s_I)}}{\frac{1}{\lambda\,(s_H+s_I)} + T}$$

Rearranging,

$$s_{H:C} = \frac{r\,p\,(y_{CC} + y_{CA})}{1 + \lambda\,(s_H + s_I)\,T} \tag{2.3}$$

Similarly for low-quality males who help care,

$$s_{L:C} = \frac{r\,(1-p)\,(y_{CC} + y_{AC})}{1 + \lambda\,(s_L + s_I)\,T} \tag{2.4}$$

The same reasoning gives us the searching densities for the female strategies:

$$s_H = \frac{x_H}{1 + \lambda\,(s_{H:C} + s_{H:A})\,T} \tag{2.5}$$

$$s_L = \frac{x_L}{1 + \lambda\,(s_{L:C} + s_{L:A})\,T} \tag{2.6}$$

$$s_I = \frac{x_I}{1 + \lambda\,(s_{H:C} + s_{H:A} + s_{L:C} + s_{L:A})\,T} \tag{2.7}$$

Given any strategy frequencies, equations (2.1)–(2.7) determine the searching densities of each of the 4 male and 3 female strategies.

*Strategy payoffs*

The fitness payoff for a strategy is the average reproductive benefit obtained per mate divided by the average time spent per mate. We denote by $\pi_\star$ the fitness of the strategy of type $\star$. For the females who prefer high-quality males, we have

$$\pi_{x_H} = \frac{\frac{s_{H:C} B\, m_C\, m_H + s_{H:A} B\, m_H}{s_{H:C} + s_{H:A}}}{\frac{1}{\lambda\,(s_{H:C} + s_{H:A})} + T}$$

Rearranging,

$$\pi_{x_H} = \frac{s_{H:C} B\, m_C\, m_H + s_{H:A} B\, m_H}{\frac{1}{\lambda} + (s_{H:C} + s_{H:A})\, T} \tag{2.8}$$

For females who prefer low-quality males and females who will mate with both, we have:

$$\pi_{x_L} = \frac{s_{L:C} B\, m_C + s_{L:A} B}{\frac{1}{\lambda} + (s_{L:C} + s_{L:A})\, T} \tag{2.9}$$

$$\pi_{x_I} = \frac{s_{H:C} B\, m_C\, m_H + s_{H:A} B\, m_H + s_{L:C} B\, m_C + s_{L:A} B}{\frac{1}{\lambda} + (s_{H:C} + s_{H:A} + s_{L:C} + s_{L:A})\, T} \tag{2.10}$$

Each male strategy has two terms, each scaled according the the probability of being high quality. We have

$$\pi_{y_{CC}} = p\, \frac{(s_H + s_I) B\, m_C\, m_H}{\frac{1}{\lambda} + (s_H + s_I)\, T} + (1-p) \frac{(s_L + s_I) B\, m_C}{\frac{1}{\lambda} + (s_L + s_I)\, T} \tag{2.11}$$

$$\pi_{y_{AC}} = p\, \lambda\, (s_H + s_I) B\, m_H + (1-p) \frac{(s_L + s_I) B\, m_C}{\frac{1}{\lambda} + (s_L + s_I)\, T} \tag{2.12}$$

$$\pi_{y_{AA}} = p\, \lambda\, (s_H + s_I) B\, m_H + (1-p) \lambda\, (s_L + s_I) B \tag{2.13}$$

$$\pi_{y_{CA}} = p\, \frac{(s_H + s_I) B\, m_C\, m_H}{\frac{1}{\lambda} + (s_H + s_I)\, T} + (1-p) \lambda\, (s_L + s_I) B \tag{2.14}$$

### *2.3.3 Analysis*

We modeled evolutionary dynamics with the replicator-mutator equation [36]. Mutations were included to eliminate unrealistic neutral equilibria. Thus, the strategy frequencies are thought of as functions of time, $t$, and we have a mutation rate parameter, $\mu_x$, for the female strategies, and a mutation rate parameter, $\mu_y$, for the male strategies. The dynamics for the female strategy of preferring high quality males is described by

$$x'_H(t) = (1-\mu_x)\, x_H(t)\, \pi_{x_H} + \tfrac{1}{2}\mu_x\, x_L(t)\, \pi_{x_L} + \tfrac{1}{2}\mu_x\, x_L(t)\, \pi_{x_L} - x_H(t)\, \phi_x, \tag{2.15}$$

where

$$\phi_x = x_H(t)\, \pi_{x_H} + x_L(t)\, \pi_{x_L} + x_L(t)\, \pi_{x_L}$$

is the average fitness. Similar expressions describe the dynamics for the other two female strategies. For the male strategy of providing care whether or not it is high quality, the dynamics is described by

$$y'_{CC}(t) = (1-\mu_y)\, y_{CC}(t)\, \pi_{y_{CC}} + \tfrac{1}{3}\mu_y\, y_{AC}(t)\, \pi_{y_{AC}} + \tfrac{1}{3}\mu_y\, x_{AA}(t)\, \pi_{y_{AA}} + \tfrac{1}{3}\mu_y\, x_{CA}(t)\, \pi_{y_{CA}} - y_{CC}(t)\, \phi_y, \tag{2.16}$$

where again,

$$\phi_y = y_{CC}(t)\, \pi_{y_{CC}} + y_{AC}(t)\, \pi_{y_{AC}} + x_{AA}(t)\, \pi_{y_{AA}} + x_{CA}(t)\, \pi_{y_{CA}}$$

is the average fitness. Again, similar expressions describe the dynamics for the other three male strategies.

Because we separated the time scales of the mate search and evolutionary dynamics, contained within these equations is the algebraic system of equations that give the search densities of the strategies at each time point. This means that our dynamical system is a system of differential-algebraic equations, rather than simply a system of differential equations. This makes it difficult to perform standard stability analyses of equilibria, so we instead simply numerically solved the differential-algebraic system, (see Figure 2.1 for an example), and observed end points of the dynamics. It turns out that under the wide range of parameter settings we explored, the dynamics settle towards a single point equilibrium.

We wrote scripts in Mathematica to automate the process of finding the endpoint of the dynamics given any parameter settings and initial conditions. For each numerical solution, the initial conditions were drawn uniformly at random from the unit simplex over female strategy frequencies and the unit simplex over male strategy frequencies. We can then look at how the end point of the evolutionary dynamics changes as we vary different parameters of the model. (Figure 2.2, below, will be an example of this.)

### *2.4 Results*

To discover how discretionary male parental investment influences female preferences for male physical quality, our approach is to compare female preferences at equilibrium across three different models. The first model, which is described above, involves males having the choice between providing care or not. The two models we compare this with are one in which male parental care is obligate, and one in which males never provide offspring care. We get these two models simply by making simplifications to our first model, by removing the male option to abandon, or removing the male option to provide care, respectively.

We make this comparison in subsection 2.4.4 below, exploring how the evolutionary equilibrium changes with respect to various parameters. But first we make some preliminary comments. In the version of the model where males never provide care, as would be expected, females often only accept high quality males. This is because, in that model, there is no advantage to mating with a low quality male, and high quality males are readily available.

However, in the version of the model where all males provide care, even though again there is no advantage to mating with a low quality male, often some females are indiscriminate. The reason is that in this case high quality males become scarce in the search pool as they are removed from it to care for offspring. Because some females only mate with high quality males, more high quality males are removed from the search pool than low quality males. But if every female prefers high quality males, indiscriminate females can invade the population because it is so much quicker for them to find a mate. Thus, the female population generally reaches a mixed equilibrium, with some females being indiscriminate, even though all males provide care. This same observation was made by [58], but in the context of a different question. This characteristic of the version of the model where males always provide care is important to keep in mind when we compare the three versions of the model in subsection 2.4.4 below.

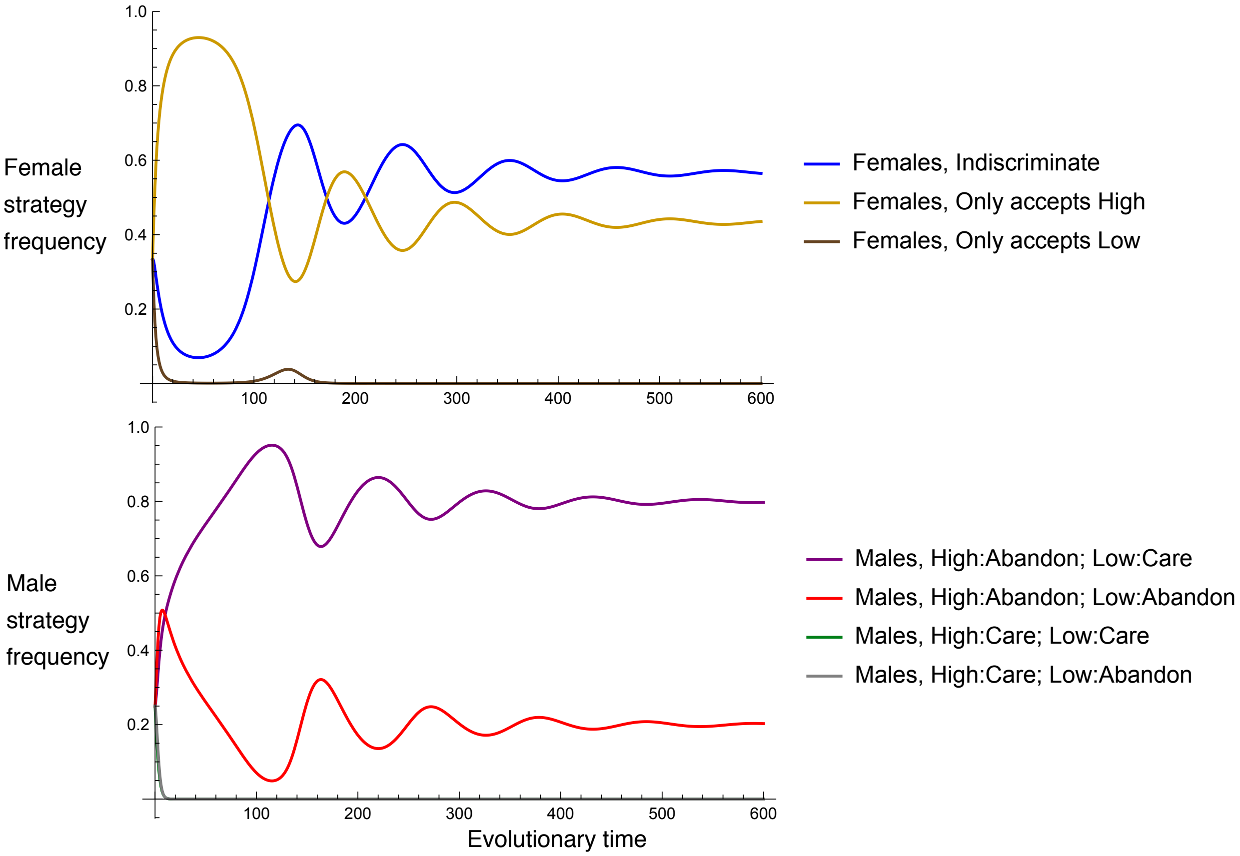


Figure 2.1: An example trajectory of the evolutionary dynamics. Female strategies are shown in the upper panel, males in the lower panel. For both males and females, the population starts at the point where all strategies are equally common. The dynamics eventually stabilize around the point where males never care for offspring when they are high quality, 8/10ths of males care for offspring when they are low quality, 13/30ths of females mate with only high-quality males, 17/30ths of females are indiscriminate, and no females mate with only low-quality males. The parameter settings in this example are: sex ratio $r = 1$, fraction of males that are high quality $p = 0.5$, relative encounter rate 100 (that is, the product $\lambda T = 100$), value of male care and male quality $m_C = 5/3$, $m_H = 3/2$.

### *2.4.1 Evolutionary dynamics*

Before exploring the effect of different parameter settings and comparing the case when male care is discretionary against the other two possibilities, we provide an example of one particular evolutionary trajectory (Figure 2.1). For this example, the sex ratio is even ($r = 1$), half of the males are high quality and half low quality ($p = 0.5$), it takes 100 days to raise offspring ($T = 100$), and the encounter rate is $\lambda = 1$. (Remember that a male's quality is perfectly discriminable to females but his paternal care strategy is not.) We let the multiplier for male care be $m_C = 5/3$ and the multiplier for the male being high quality be $m_H = 3/2$. These settings make male care more valuable than male quality, so it is possible for a female to get greater benefit out of a low quality than a high quality male if the low quality male provides care and the high quality male does not. We set the mutation rate parameters, $\mu_x$ and $\mu_y$, each at $10^{-6}$.

Figure 2.1 shows that with the above parameter settings, the dynamics stabilize towards a point where low-quality males are more likely to care for offspring than high-quality males. A slight majority of females will mate with either high- or low-quality males, a minority only mates with high-quality males, and the female strategy of only mating with low-quality males has died out.

### *2.4.2 No females refuse high-quality males at equilibrium*

In the example dynamics of Figure 2.1, females who prefer low-quality males over high-quality males are not present at equilibrium. This turns out to be the case generally. We can show why by the following argument: Suppose that we have an equilibrium with a nonzero frequency of females who only mate with low-quality males. Then either (A) there are also females who prefer high-quality males in this equilibrium or (B) there are no females who prefer high-quality males. If (B) there are no females that prefer high-quality males, then low-quality males are in greater demand than high-quality males. In this case, selection will act so that high-quality males will be at least as likely to care for offspring as low-quality males are. But this means that both indiscriminate females and females who prefer high-quality males will have higher fitness than females who prefer low-quality males, and so this cannot be an equilibrium. On the other hand, if (A) there are females who prefer high-quality males as well as females who prefer low-quality males at this equilibrium, then they must have equal payoffs. But then indiscriminate females will have a higher payoff, because they find mates more quickly while receiving the same average payoff per mate. Thus, it cannot be an equilibrium. Therefore, no equilibria include a nonzero frequency of females who only mate with low-quality males.

In what follows we will compare the equilibrium frequencies between just the two female strategies of preferring high-quality males and being indiscriminate. We nevertheless continue to include in the model the female strategy of preferring low-quality males for the following analyses because it can influence the dynamics. We just omit this strategy from our description of the equilibria because it is always absent at equilibrium.

### *2.4.3 Varying the encounter rate*

Figure 2.2 shows how the equilibrium strategy frequencies depend on the search rate parameter, $\lambda$. At the left hand side of the plots, $\lambda$ is very small, and it takes searching individuals a long time to find potential mates. Because it takes so long to find a mate, females cannot afford to be choosy, and males always provide care, even if they are high quality, because they are unlikely to find new mates any time soon.

As we move right along the plot, at $\lambda$ just above 0.01, mate search begins to be fast enough that males start to sometimes abandon females. At $\lambda = 0.02$, males now always abandon whether or not they are high quality. At $\lambda = 0.04$, it suddenly (discontinuously) becomes worthwhile for some females to be choosy. At this point, mate search is quick enough that if a female rejects a low quality male, she still has a chance of encountering a high quality male in a reasonable amount of time.

The fact that some females are now choosy causes low quality males to start having different incentives from high quality males. Low quality males are now less likely to find

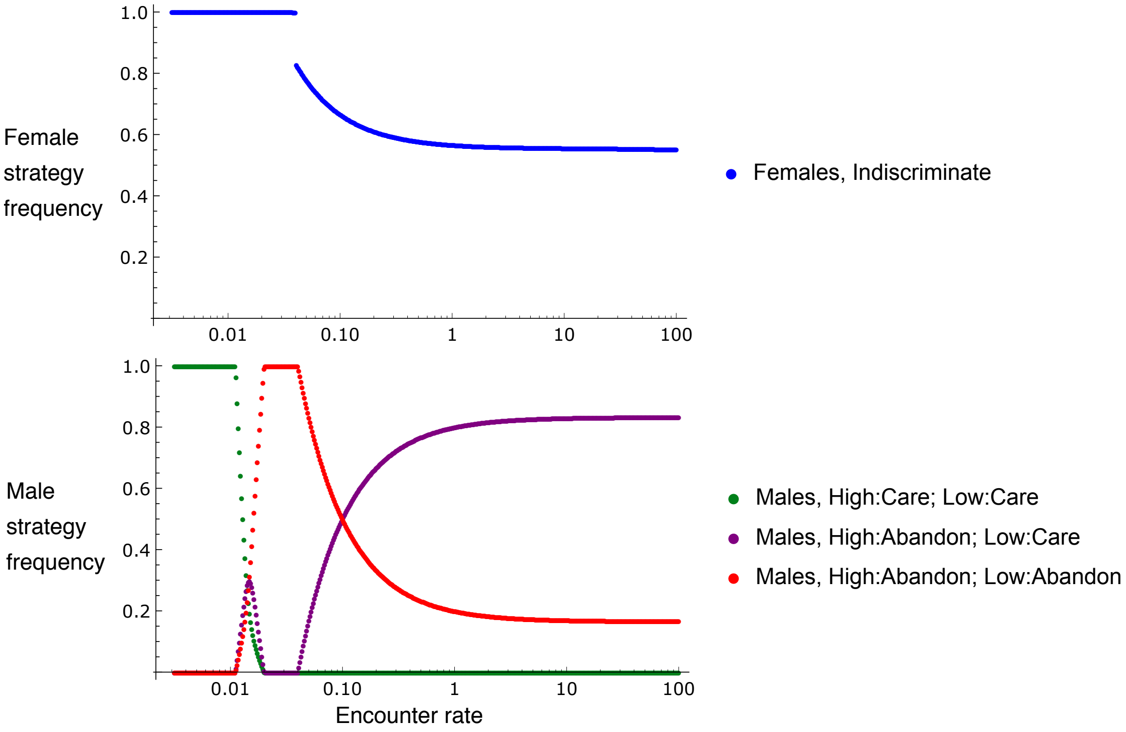


Figure 2.2: Equilibrium strategy frequencies as a function of the encounter rate, $\lambda$. The other parameter settings are the same as for Figure 2.1. The female equilibrium frequencies can be read from the upper plot; the male equilibrium frequencies can be read from the lower plot. To reduce clutter, only 1 of the 3 female strategies and 3 of the 4 male strategies are shown. For the males, the frequency of the 4th strategy at any point equals one minus the frequencies of the others. As explained in subsection 2.4.2, at equilibrium, all females either prefer high quality males or are indiscriminate. Thus, we only need the frequency of indiscriminate females in the upper plot, because the frequency of females that prefer high quality males is one minus the frequency of those that are indiscriminate. Subsection 2.4.3 interprets the pattern in these plots.

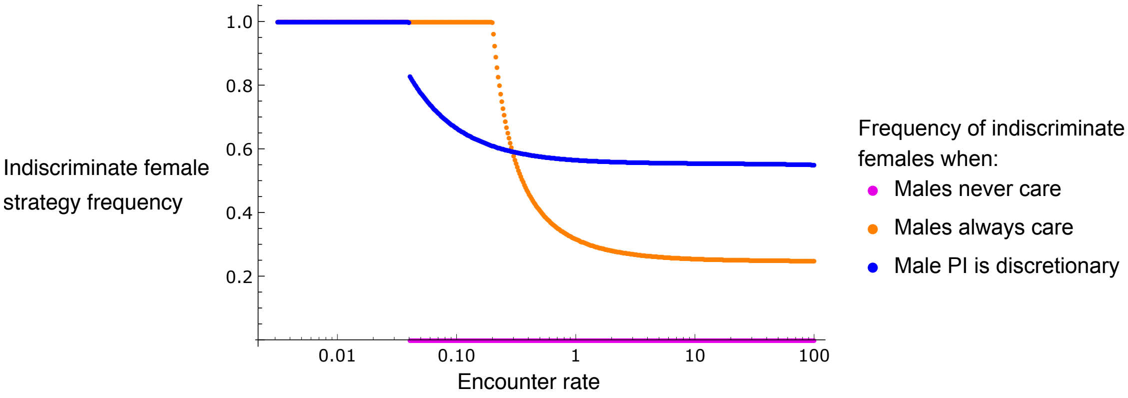


Figure 2.3: Equilibrium female strategy frequencies for the three versions of the model. When males never provide care for offspring (magenta), all females prefer high-quality males except at very low encounter rates, in which case all females mate with both male types. When males always care (orange), a minority of females accept low-quality males at higher encounter rates, while at lower encounter rates all females mate with both male types. When males are free to choose whether to provide care (blue), no matter the encounter rate, a majority of females will mate with both male types.

new willing mates if they abandon their partner, and so males start to sometimes provide care when they are low quality. As $\lambda$ continues to increase, the search friction for finding mates diminishes to nothing. Low quality males provide care more than 80 percent of the time, high quality males never provide care, and females are choosy less than half of the time. The next subsection compares what happens here to the two alternative models where males cannot choose whether to provide care.

### *2.4.4 The effect of discretionary male PI on female preferences*

Figure 2.3 shows the equilibrium frequency of female preferences for each of the 3 models, plotted against $\lambda$. When males never provide care, females always prefer high quality males (females are always choosy) unless the encounter rate is so small that they cannot afford to refuse low quality males. At high encounter rates, fewer females accept low quality males (females are more choosy), when all males care than when male care is discretionary. This is because, when male care is discretionary, low quality males provide care more often than

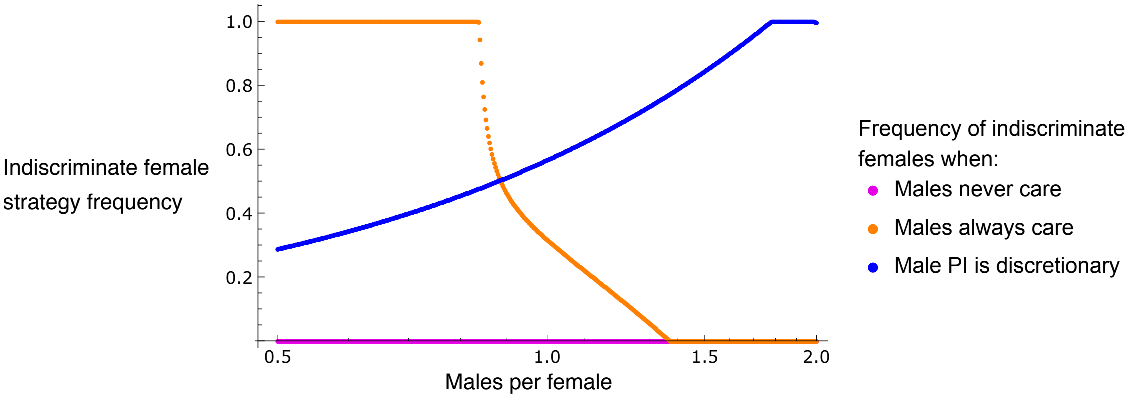


Figure 2.4: Equilibrium female strategy frequencies as a function of the sex ratio, $r$. The other parameter settings are the same as for Figure 2.1.

high quality males, and so the difference to a female between getting a high quality male and getting a low quality male is not as important.

At encounter rates lower than about $\lambda = 0.4$, the pattern reverses, and more females accept low quality males when all males provide care than when male care is discretionary. This is because high quality males are scarce in the search pool when all males provide care, but they are not scarce in the search pool when male care is discretionary. When male care is discretionary, high quality males do not provide care and are therefore always searching for mates.

### *2.4.5 Varying the sex ratio*

We next provide the same comparison as above but while varying the sex ratio instead of the encounter rate. (The operational sex ratio of individuals in the search pool varies depending on the strategies of both sexes. When we speak of the "sex ratio" in this section though, we are referring to the fundamental sex ratio of total males per total females. This is the parameter $r$ in the model.)

Figure 2.4 shows equilibrium frequencies of female preferences for the three versions of the model, where now the $x$-axis is the sex ratio instead of the encounter rate. The $x$-axis ranges from a sex ratio of 2:1 females per male ($r = 0.5$) to 2:1 males per female ($r = 2$). When males never provide care, females are always choosy (prefer high quality males) throughout this range. When males always provide care, female preferences are strongly sensitive to the sex ratio. When the sex ratio is much above even, there are enough extra males that females always prefer high quality males. On the other hand, when the sex ratio is much below even, females quickly shift to being entirely indiscriminate.

This contrasts strongly with the case where male parental care is discretionary. In this case, female preferences are not as sensitive to changes in the sex ratio, and throughout most of the range there is a mixed equilibrium with some females preferring high quality males and some females being indiscriminate. And, in contrast to the case where all males provide offspring care, females actually become less choosy as the number of males per female increases. To understand what is going on here, Figure 2.5 displays the male strategy frequencies alongside the female, for the case where male care is discretionary. The range is chosen to encompass all the different types of equilibria that are found along this cross section of parameter space.

Figure 2.5 exhibits 6 regions with different equilibrium characteristics. We will go through them as the sex ratio increases, from left to right on the plot. In the first region, at about 1 male per 25 females and fewer ($r \leq 0.04$), the equilibrium consists of all females being indiscriminate and males always abandoning. As $r$ increases, we enter a region where females have discontinuously begun to be mostly choosy, and males start to sometimes provide care when they are low quality. This region continues from an $r$ value of about 0.04 all the way to an $r$ value of a bit over 1.7. As we saw in Figure 2.4, in this region the proportion of females who are indiscriminate increases as males become more common, counter to what may be expected. The reason that this happens is that the proportion of males that provide care when they are low quality is also increasing.

Between a sex ratio of about $r = 1.7$ and a sex ratio of $r = 2$, all females have become indiscriminate again. At this point, because females are the sex in short supply, males begin to sometimes provide care even when they are high quality. This behavior rises in frequency

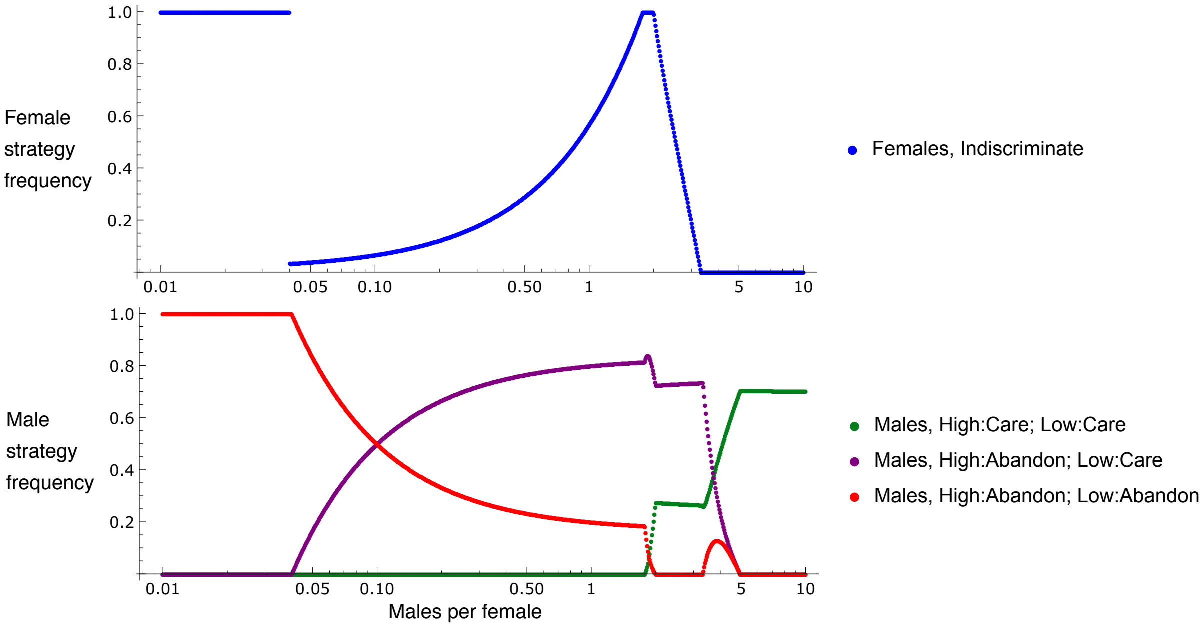


Figure 2.5: Equilibrium strategy frequencies as a function of the sex ratio, $r$. The other parameter settings are the same as for Figure 2.1. As in Figure 2.2, only one of the three female strategies and three of the four male strategies are shown.

until $r = 2$, at which point females can again afford to be choosy, because now both high and low quality males sometimes provide care. (Because all females are indiscriminate in this range, one might expect to see males behave the same whether they are high or low quality. Yet males are more often providing care when they are low quality than when they are high quality. The reason is that any infinitesimal increase in the proportion of high quality males providing care upsets the equilibrium, and would lead to some females preferring high quality males again. This would lead to males who abandon when they are high quality getting a higher payoff, leading to fewer high quality males providing care, leading to females going back to being indiscriminate. Thus, the equilibrium remains with females being indiscriminate and males providing care more often when they are low quality.)

Note that in this range males still more often provide care when they are low quality then when they are high quality, even though all females are indiscriminate. At first this may appear surprising because

Between $r = 2$ and an $r$ value of about 3.2, the proportion of females who are indiscriminate falls steadily, until, for the first time, all females are choosy. At this point, the male strategy of providing care when low quality but abandoning when high quality (High:Abandon, Low:Care), begins to fall in frequency, as a greater proportion of males begins to provide care when high quality. Because all females are now choosy, there is no selection on what males do when low quality, because they never get the opportunity to reproduce. Finally, at a sex ratio of 5 males per female and above ($r \geq 5$), females only mate with high quality males, and all high quality males provide care.

### *2.4.6 Varying the importance of offspring care*

Finally, in Figure 2.6 we show what happens when we vary the value of male care for offspring ($m_C$), including values where care is less important than quality ($m_C < m_H$). For both of the alternative models, where males always provide care or never provide care, the proportion of females that are indiscriminate does not vary with the value of male care. This is because when all males behave the same way, changing the value of male care does not cause any relative change between high and low quality males.

For the case when male care is discretionary, the proportion of females that are indiscriminate is 0 when $m_C$ is less than $m_H$ (our parameter settings have $m_H = 1.5$). When $m_C < m_H$, low quality males are less valuable as mates than high quality males, even when comparing a low quality male who provides care to a high quality male who does not. Thus, as long as the encounter rate ($\lambda$) is high enough that females can afford to discriminate, (here $\lambda = 1$), females will never have reason to select a low quality male. Once male care becomes more valuable than male quality ($m_C > m_H$), we immediately see a substantial fraction of females becoming indiscriminate. As the value of male care continues to increase, the payoff to mating with a low quality male who cares becomes greater, so the proportion of females that are indiscriminate increases. Finally, at very high values of $m_C$, all females are indiscriminate.

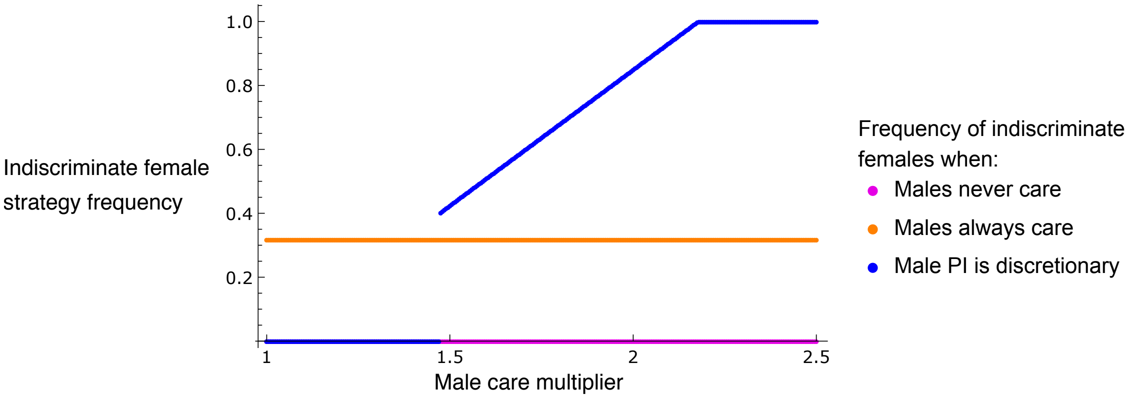


Figure 2.6: Equilibrium female strategy frequencies as a function of the value of male care, $m_C$. The other parameter settings are the same as for Figure 2.1. In particular, the value of the male being high quality, $m_H$, equals 1.5, which can be seen as the location of the discontinuity for the female strategy frequency in the discretionary male care model.

## *2.5 Discussion*

Our main finding is that when males are free to choose whether to provide parental care, females often become more willing to mate with low-quality males. This is as compared to the case when males are not free to make this choice. This increase in female acceptance of low quality males is not so large that it removes the overall preference for high-quality males, it only reduces the magnitude of that preference.

In many species, females are selected to be choosy because males vary with respect to physical traits that affect the female's fitness [3]. When males vary not only physically but also with respect to providing PI, it would seem that a female's choice of a mate would be all the more important. But we show that in fact it is often the reverse: for species in which males vary with respect to PI as well as with respect to physical quality, females will often evolve to be less choosy.

It should be emphasized that in this model, lower quality males are worse not only in their fertility but also in their ability to care for the offspring. They are lower quality in every way. Nonetheless, because low-quality males can choose to help the female, they can increase their value as mates even though their PI is less valuable than the PI of a high quality male. But they do this even though they are free to abandon a female after mating with her. In our model, females cannot directly select for males who care, so male care does not result from females preferring good fathers. Instead, reduced female choosiness is a result of how the mating market induces greater opportunity costs to providing offspring care on high-quality males than on low-quality males. Because males of lower-quality have less to gain from searching for additional mates, they gain relatively more by helping the female care for offspring. This happens despite the fact that the value of low quality males' offspring care is less than what it would be for a high quality male.

One area that addresses both offspring care and mate attractiveness is studies of the differential allocation (DA) hypothesis, which predicts that less attractive mates may provide more offspring care [15]. But here the emphasis is usually on how mates should allocate care and not on how these care decisions influence mate choice. Also, while much empirical work has addressed the DA hypothesis [16, 76], theoretical studies justifying the DA interpretation are relatively scarce, while other models predict the opposite effect, in which high-quality mates provide more PI instead of less [76].

By modeling the tradeoffs faced by males of different genetic qualities and with different PI costs, [20] showed that males of lower genetic quality can be preferred by females if such males provide better parental care. However, their model did not explicitly consider the evolutionary dynamics acting on female strategies, and so they could not compare shifts in relative female preference when low quality males are not intrinsically more incentivized to provide offspring care. In contrast, by including both female and male strategies, our model is able to compare relative strengths of female preferences. Thus, we can show how males of genuinely lower quality become more valuable as mates even though females do not switch completely to preferring low quality males. This allows us to discover how females sometimes become less choosy when males control their own PI decisions.

One straightforward way in which our model agrees with empirical observations is that in species where males provide significant parental care, males should invest less in physical displays for courtship [72]. This is likely often due to reduced female choice because males are also exercising choice. But our model shows that this can result even if there is no mutual mate choice and females monopolize mating decisions. Even in such cases, and even when males vary greatly in their mate value, female choice may be weak.

Some of the predictions of the model could be tested in different species by determining how much males vary with respect to fertility and if males of lower fertility provide more offspring care. If males vary greatly in fertility even in species with weaker female choice, and if males of lower fertility provide more offspring care, the model predictions would be supported. In some species it may also be possible to directly test whether increased variation in male PI leads to decreased female choice, but such experiments would be difficult.

Other factors can also affect the strength of mate-choice selection, including variation in female quality, promiscuity, extra pair paternity etc. Future modeling work is needed to provide predictions for how these phenomena should interact with the differential PI and decreased female preferences described here.

### *2.6 Acknowledgements*

The authors thank Carl T. Bergstrom for helpful discussions and comments, and Christopher Wells, Jeff Smith, Sonia Singhal, Natalie Miller, and an anonymous reviewer for comments on earlier versions of the manuscript. This work was funded in part by the BEACON consortium, an NSF Science and Technology Center. FM was also supported in part through NSF grant no. EF-1038590 to Carl T. Bergstrom. This material is based in part upon work supported by the National Science Foundation under Cooperative Agreement No. DBI-0939454. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.

# Chapter 3

# ADAPTIVE BEHAVIOR CAN PRODUCE MALADAPTIVE ANXIETY DUE TO INDIVIDUAL DIFFERENCES IN EXPERIENCE

By Frazer Meacham and Carl T. Bergstrom

Originally published in *Evolution, Medicine, and Public Health*, 2016

## *3.1 Abstract*

Normal anxiety is considered an adaptive response to the possible presence of danger, but is susceptible to dysregulation. Anxiety disorders are prevalent at high frequency in contemporary human societies, yet impose substantial disability upon their sufferers. This raises a puzzle: why has evolution left us vulnerable to anxiety disorders? We develop a signal detection model in which individuals must learn how to calibrate their anxiety responses: they need to learn which cues indicate danger in the environment. We derive the optimal strategy for doing so, and find that individuals face an inevitable exploration-exploitation tradeoff between obtaining a better estimate of the level of risk on one hand, and maximizing current payoffs on the other. Because of this tradeoff, a subset of the population can become trapped in a state of self-perpetuating over-sensitivity to threatening stimuli, even when individuals learn optimally. This phenomenon arises because when individuals become too cautious, they stop sampling the environment and fail to correct their misperceptions, whereas when individuals become too careless they continue to sample the environment and soon discover their mistakes. Thus, over-sensitivity to threats becomes common whereas under-sensitivity becomes rare. We suggest that this process may be involved in the development of excessive anxiety in humans.

## *3.2 Introduction*

Motile animals have evolved elaborate mechanisms for detecting and avoiding danger. Many of these mechanisms are deeply conserved evolutionarily [61]. When an individual senses possible danger, this triggers a cascade of physiological responses that prepare it to deal with the threat. Behavioral ecological models treat the capacity for anxiety as a mechanism of regulating how easily these defensive responses are induced [51, 65, 67, 34, 5, 69]. Greater anxiety causes an individual to be alert to more subtle signs of potential danger, while lowered anxiety causes the individual to react only to more obvious signs [18]. As unpleasant as the experience of anxiety may be, the capacity for anxiety is helpful in tuning behavior to environmental circumstance. This viewpoint is bolstered by epidemiological evidence suggesting that long-term survival is worse for people with low anxiety-proneness than for those in the middle of the distribution, due in part to increased rates of accidents and accidental death in early adulthood [49, 63].

While the capacity for anxiety is adaptive, dysregulated anxiety is also common, at least in humans. Of all classes of mental disorders, anxiety disorders affect the largest number of patients [42]. The global prevalence of individuals who suffer from an anxiety disorder at some point in their life is commonly estimated at around 15 percent [42, 85], with 5 to 10 percent of the population experiencing pathological anxiety in any given year [42, 85, 6]. The consequences can be drastic: in a 12 month period in the US, 4 percent of individuals had an anxiety disorder that was severe enough to cause work disability, substantial limitation, or more than 30 days of inability to maintain their role [44]. The prevalence and magnitude of anxiety disorders is also reflected in the aggregate losses they cause to economic productivity: in the 1990s the annual cost was estimated at $42 billion in the US alone [28].

Episodes of clinically-significant anxiety are distributed broadly across the lifespan, and anxiety disorders typically manifest before or during the child-rearing years [43]. Because of the severity of impairment that often results from anxiety disorders, and the fact that onset occurs before or during reproduction, these disorders will often have a substantial effect on

Darwinian fitness. Thus, the prevalence of anxiety disorders poses an apparent problem for the evolutionary viewpoint. If the capacity for anxiety is an adaptation shaped by natural selection, why is it so prone to malfunction?

One possible explanation invokes the so-called smoke detector principle [65, 67]. The basic idea is to think about how anxiety serves to help an organism detect danger, and to note the asymmetry between the low cost of a false alarm and the high cost of failing to detect a true threat. This allows us to frame anxiety in the context of signal detection theory. Because of asymmetry in costs of false alarms versus false complacency, the theory predicts that optimized warning systems will commonly generate far more false positives than false negatives. This provides an explanation for why even optimal behavior can produce seemingly excessive sensitivity in the form of frequent false alarms [64, 67]. More recently, the signal detection framework has been expanded to describe how the sensitivity of a warning system should track a changing environment and become more easily triggered in dangerous situations [69]. This approach, together with error management theory [40], begins to provide an account of how anxiety and mood regulate behavior over time, and why high levels of anxiety may be adaptive even when true threats are scarce. Better to be skittish and alive than calm but dead.

The smoke detector principle cannot be the whole story, however. There are a number of aspects of anxiety that it does not readily explain. First, the smoke detector principle deals with evolutionarily adaptive anxiety — but not with the issue of why evolution has left us vulnerable to anxiety *disorders*. A fully satisfactory model of anxiety and anxiety disorders should explain within-population variation: Why does a small subset of the population suffer from an excess of anxiety, while the majority regulate anxiety levels appropriately? Second, a critical component of anxiety disorders is the way they emerge from self-reinforcing negative behavior patterns. Individuals with anxiety disorders often avoid situations or activities that are in fact harmless or even beneficial. Effectively, these individuals are behaving too pessimistically, treating harmless situations as if they were dangerous. We would like to explain how adaptive behavior might lead to self-reinforcing pessimism. Third, if the evolutionary function of anxiety is to modulate the threat response according to environmental circumstances [68], evolutionary models of anxiety will need to explicitly treat that modulation process—that is, such models should incorporate the role of learning explicitly.

In this paper, we show that optimal learning can generate behavioral over-sensitivity to threat that is truly harmful to the individual's fitness, but expressed in only a subset of the population. Our aim is not to account for the specific details of particular anxiety disorders—phobias, generalized anxiety disorder, post-traumatic stress disorder, and so forth—but rather to capture some of the general features of how anxiety is regulated and how this process can go awry.

In section 2, we illustrate the basic mechanism behind our result using a very simple model borrowed from foraging theory [56] in which an actor must learn by iterative trial and error whether taking some action is unacceptably dangerous or sufficiently safe. (Trimmer et al. [91] independently developed a related model to study clinical depression. Also see Frankenhuis and Panchanathan [25] as well as [73] for closely related models of developmental plasticity in general.) In section 3, we extend the model into the domain of signal detection theory and consider how an actor learns to set the right threshold for responding to an indication of danger. In most signal detection models, the agent making the decision is assumed to know the distribution of cues generated by safe and by dangerous situations. But where does this knowledge come from? Unless the environment is homogeneous in time and space over evolutionary timescales, the distributions of cues must be learned. In our model, therefore, the agent must actively learn how the cues it observes relate to the presence of danger. We show that under these circumstances, some members of a population of optimal learners will become overly pessimistic in their interpretations of cues, but fewer will become overly optimistic.

### *3.3 Learning about an uncertain world*

If we want to explain excess anxiety from an evolutionary perspective, we must account for why only a subset of the population is affected. Although genetic differences may be partly responsible, random variation in individual experience can also lead to behavioral differences among individuals. In particular, if an individual has been unfortunate during its early experience, it may become trapped in a cycle of self-reinforcing pessimism. To demonstrate this, we begin with a simple model that shows how responses to uncertain conditions are shaped by individual learning. The model of this section does not include the possibility of the individual observing cues of the potential danger. Thus, it does not capture anxiety's essential characteristic of threat detection. But this model does serve to illustrate the underlying mechanism that can lead a subset of the population to be overly pessimistic.

### *3.3.1 Model*

Because our aim is to reveal general principles around learned pessimism, rather than to model specific human pathologies, we frame our model as a simple fable. Our protagonist is a fox. In the course of its foraging, it occasionally comes across a burrow in the ground. Sometimes the burrow will contain a rabbit that the fox can catch and eat, but sometimes the burrow will contain a fierce badger that may injure the fox. Perhaps our fox lives in an environment where badgers are common, or perhaps it lives in an environment where badgers are rare, but the fox has no way of knowing beforehand which is the case. Where badgers are rare, it is worth taking the minor risk involved in digging up a burrow to hunt rabbits. Where badgers are common, it is not worth the risk and the fox should eschew burrows in favor of safer foraging options: mice, birds, fruits, berries etc. The fox encounters burrows one at a time, and faces the decision of whether to dig at the burrow or whether to slink away. The only information available to the fox at each decision point is the prior probability that badgers are common, and its own experiences with previous burrows.

To formalize this decision problem, we imagine that the fox encounters a sequence of burrows, one after the other. The fox makes a single decision of whether to explore each burrow before encountering the next burrow, and each burrow contains either a rabbit or a badger. We let $R$ be the payoff to the fox for digging up a burrow that contains a rabbit and $C$ be the cost of digging up a burrow that contains a badger. If the fox decides to leave a burrow undisturbed, its payoff is zero. When the fox decides to dig up a burrow, the probability of finding a badger is $p_g$ if badgers are rare, and $p_b$ if badgers are common, where $p_g < p_b$. If badgers are rare it is worthwhile for the fox to dig up burrows, in the sense that the expected payoff for digging is greater than zero. That is, we assume that

$$(1 - p_g)R - p_g C > 0.$$

If badgers are common, burrows are best avoided, because the expected payoff for digging is less than zero:

$$(1 - p_b)R - p_b C < 0.$$

We let $q_0$ be the prior probability that badgers are common and we assume that the correct prior probability is known to the fox. We assume a constant extrinsic death rate $d$ for the fox (and we assume that badger encounters are costly but not lethal), so that the present value of future rewards is discounted by $\lambda = 1 - d$ per time step.

If the fox always encountered only a single burrow in its lifetime, calculating the optimal behavior would be straightforward. If the expected value of digging exceeds the expected value of not doing so, the fox should dig. That is, the fox should dig when

$$(1 - q_0)\big((1 - p_g)R - p_g C)\big) + q_0\big((1 - p_b)R - p_b C\big) > 0.$$

But the fox will very likely encounter a series of burrows, and so as we evaluate the fox's decision at each stage we must also consider the value of the information that the fox gets from digging. Each time the fox digs up a burrow, it gets new information: did the burrow contain a rabbit or a badger? Based on this information, the fox can update its estimate of the probability that the environment is favorable. If the fox chooses not to dig, it learns nothing and its beliefs remain unchanged. Thus even if the immediate expected value of digging at the first burrow is less than 0, the fox may still benefit from digging because it may learn that the environment is good and thereby benefit substantially from digging at subsequent burrows. In other words, the fox faces an exploration-exploitation tradeoff [41] in its decision about whether to dig or not. Because of this tradeoff, the model has the form of a one-armed bandit problem [12], where the bandit arm returns a payoff of either $R$ or $-C$, and the other arm always returns a payoff of zero.

### *3.3.2 Optimal behavior*

As an example, suppose good and bad environments are equally likely *a priori* ($q_0 = 0.5$) and foxes die at a rate of $d = 0.05$ per time step. For simplicity we set the costs and rewards to be symmetric: $C = 1$, $R = 1$, $p_g = 1/4$, $p_b = 3/4$. In a good environment where badgers are less common, the expected value of digging up a burrow is positive ($-0.25 + (1 - 0.25) = 0.5$) whereas in a bad environment where badgers are common, the expected value of digging up a burrow is negative ($-0.75 + (1 - 0.75) = -0.5$). (Recall that the fox also has other foraging options available, and therefore will not necessarily starve if it avoids the burrows.)

Applying dynamic programming to this scenario (see Appendix B), we find that the fox's optimal behavior is characterized by a threshold value of belief that the environment is bad, above which the fox does not dig at the burrows. (This threshold is the same at all time steps.) Figure 3.1 illustrates two different outcomes that a fox might experience when using this optimal strategy. Along the upper path, shown in gray, a fox initially encounters a badger. This is almost enough to cause the fox to conclude he is in a bad environment and stop sampling. But not quite—the fox samples again, and this time finds a rabbit. In his third and fourth attempts, however, the fox encounters a pair of badgers, and that's enough for him—at this point he does give up. Since he does not sample again, he gains no further information and his probability estimate remains unchanged going forward. Along the lower path, shown in black, the fox initially encounters a series of rabbits, and his probability estimate that he is in a bad environment becomes quite low. Even the occasional encounter with a badger does not alter this probability estimate enough that the fox ought to stop sampling, so he continues to dig at every hole he encounters and each time adjusts his probability estimate accordingly.

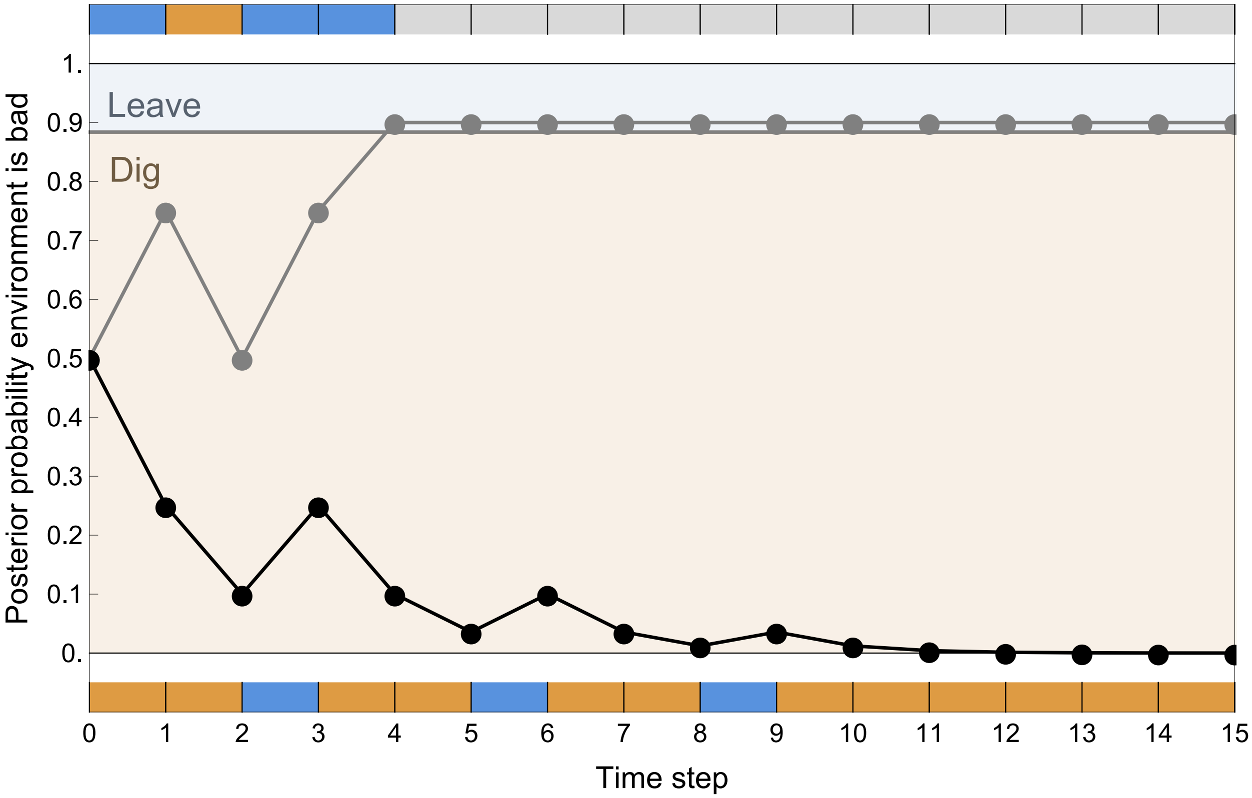


Figure 3.1: Two examples of optimal behavior by the fox. The vertical axis indicates the fox's posterior subjective probability that it is in a bad environment. In the tan region, the fox should dig. In the blue region, the fox should avoid the burrow. The grey path and black path trace two possible outcomes of a fox's foraging experience. The colored bars above and below the graph indicate the fox's experience along the upper and lower paths respectively: brown indicates that the fox found a rabbit and blue indicates that the fox found a badger. Along the grey path, the fox has a few bad experiences early. This shifts the fox's subjective probability that the environment is bad upward, into the blue region. The fox stops sampling, its probability estimate stays fixed, and learning halts. Along the black path, the fox finds two or more rabbits between each encounter with a badger. Its subjective probability remains in the tan zone throughout, and the fox continues to sample—and learn—throughout the experiment.

### *3.3.3 Population outcomes*

After solving for the optimal decision rule, we can examine statistically what happens to an entire population of optimally-foraging foxes. To see what the foxes have learned, we can calculate the population-wide distribution of individual subjective posterior probabilities that the environment is bad. We find that almost all of the foxes who are in unfavorable environments correctly infer that things are bad, but a substantial minority of foxes in favorable circumstances fail to realize that things are good. In Appendix A we show that the general pattern illustrated here is generally robust to variation in model parameters.

Figure 3.2 shows the distribution of posterior subjective probabilities that the environment is good among a population of optimally learning foxes for the above parameter choices. We can see that a non-negligible number of individuals in the favorable environment come to the false belief that the environment is probably bad. This occurs because even in a favorable environment, some individuals will uncover enough badgers early on that it seems to them probable that the environment is unfavorable. When this happens those individuals will stop digging up burrows. They will therefore fail to gain any more information, and so their pessimism is self-perpetuating.

### *3.3.4 Comments*

This self-perpetuating pessimism is not a consequence of a poor heuristic for learning about the environment; we have shown that this phenomenon occurs when individuals are using the *optimal* learning strategy. Because of the asymmetry of information gain between being cautious and being exploratory, there results an asymmetry in the numbers of individuals who are overly pessimistic versus overly optimistic. Even when individuals follow the optimal learning rule, a substantial subset of the population becomes too pessimistic but very few individuals become too optimistic.

One might think, knowing that the current learning rule leads to excessive pessimism on average, that we could do better on average by altering the learning rule to be a bit more

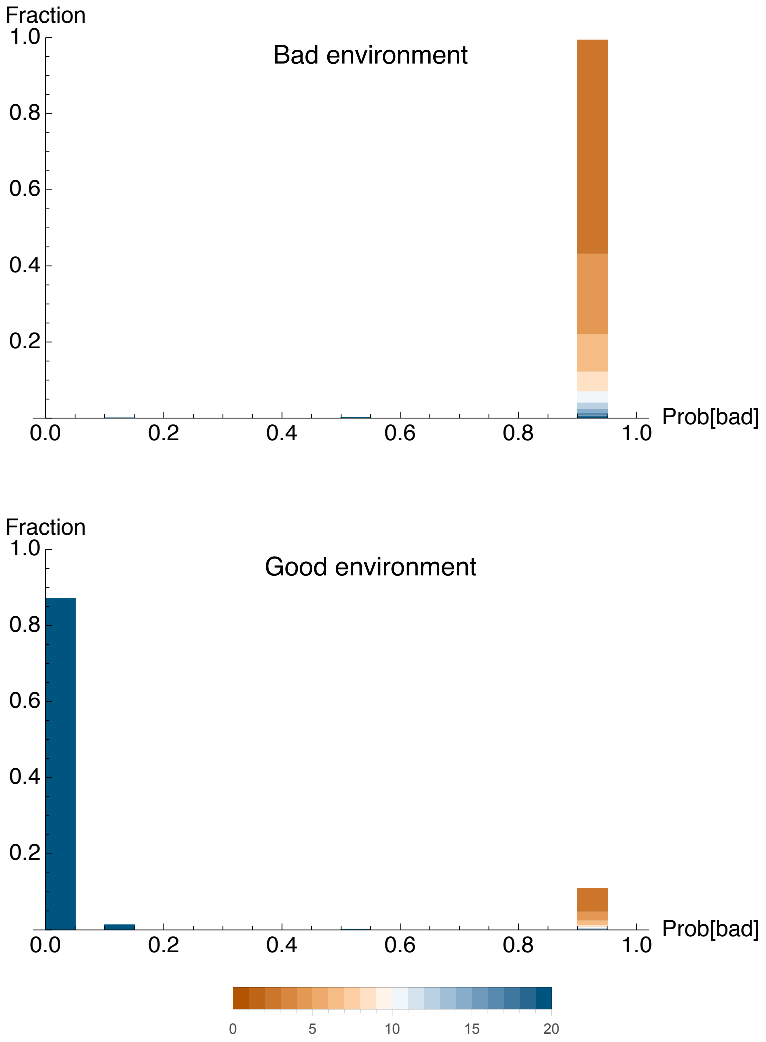


Figure 3.2: Population distribution of individual posterior probabilities that the environment is bad when the environment is indeed bad (upper panel), and when the environment is actually good (lower panel). The horizontal axis is the individual's posterior probability estimate that environment is bad after 20 opportunities to dig at a burrow. (This is among foxes who have lived that long. Conditioning in this way introduces no sampling bias because survival is independent of environment and behavior in the model.) Frequency is plotted on the vertical axis. Color indicates the number of times an individual has sampled the environment. All individuals began with a prior probability of 0.5 that the environment is bad. When the environment is indeed bad, only 0.2% of the population erroneously believe the environment is likely to be good. When the environment is good, 11.1% of the population erroneously believes that it is likely to be bad. The majority of these individuals have sampled only a few times and then given up after a bit of bad luck.

optimistic. This is not the case. Any learning rule that is more optimistic will result in lower expected payoffs to the learners, and thus would be replaced under natural selection by our optimal learning rule.

This scenario may reflect an important component of pathological human pessimism or anxiety. For example, many people think that they “can’t sing” or “are no good at math” because early failures, perhaps during childhood, led to beliefs that have never been challenged. When someone believes he can’t sing, he may avoid singing and will therefore never have the chance to learn that his voice is perfectly good. Thus, attitudes that stem from earlier negative experiences become self-perpetuating.

### *3.4 Modeling anxiety by including cues*

In the model we have just explored, the fox knows nothing about a new burrow beyond the posterior probability it has inferred from its past experience. In many situations, however, an individual will be able to use additional cues to determine the appropriate course of action. For example, a cue of possible danger, such as a sudden noise or looming object, can trigger a panic or flight response, and anxiety can be seen as conferring a heightened sensitivity to such signs of threat. In this view, the anxiety level of an individual determines its sensitivity to indications of potential danger. The higher the level of anxiety, the smaller the cue needed to trigger a flight response [65, 67, 5, 69]. To model anxiety in this sense, we extend our model of fox and burrow to explore how individuals respond to signs of potential threat. We will find that even with the presence of cues, a substantial fraction of individuals will fall into a self-perpetuating pattern where their anxiety levels are set too high.

The key consideration in our model is that individuals must *learn* how cues correspond to potential threats. In other words, individuals need to calibrate their responses to environmental cues, setting anxiety levels optimally to avoid predators without wasting too much effort on unnecessary flight. Admittedly, if the environment is homogeneous in space and extremely stable over many generations, then natural selection may be able to encode the correspondence between cues and danger into the genome. But when the environment is less predictable, the individual faces the problem of learning to properly tune its responses to cues of possible threat.

### *3.4.1 Model*

We return to our story of the fox, who we now suppose can listen at the entrance to the burrow before deciding whether to dig it up. Rabbits typically make less noise than badgers, so listening can give the fox a clue as to the contents of the burrow. When the burrow is relatively silent it is more likely to contain a rabbit, and when the fox hears distinct snuffling and shuffling noises it is likely that the burrow contains a badger. But the sounds aren't fully reliable. Sometimes rabbits can be noisy, and sometimes badgers are quiet. So although the amount of noise coming from the burrow gives the fox some information about how likely the burrow is to contain a badger, the information is probabilistic and the fox can never be certain.

In contrast to the model of the previous section, the difference between environments is now a matter of how easy it is for the fox to distinguish between dangerous and safe situations, rather than how common danger is. If the environment is good, the fox only needs to be cautious if a burrow is quite noisy. But if the environment is bad, then the fox should be cautious even if faint noises emanate from a burrow. This is because when the environment is bad, it is too risky to dig up a burrow unless the burrow is nearly silent. The fox does not know beforehand whether the environment is good or bad, and therefore it does not know how the probability of finding a badger in the burrow depends on the amount of noise it hears. The only way for it to gain information is to learn by experience.

To formalize the problem, we extend the model in section 2 by supposing that the fox observes a cue before each decision. The cue is a continuous random variable drawn from Gaussian distributions that depend on the environment and what is in the burrow. We first consider the good environment. As before, we let $p_g$ be the probability that any given burrow contains a badger. When the burrow contains a badger, the cue strength is drawn from a Gaussian distribution with mean $\mu_{g,c}$ and standard deviation $\sigma_{g,c}$. When the burrow contains a rabbit, the cue strength is drawn from a Gaussian distribution with mean $\mu_{g,r}$ and standard deviation $\sigma_{g,r}$. Similarly for the bad environment, we let $p_b$ be the probability that any given burrow contains a badger, with a cue strength drawn from a Gaussian distribution with mean $\mu_{b,c}$ and standard deviation $\sigma_{b,c}$ when the burrow contains a badger, and mean $\mu_{b,r}$ and standard deviation $\sigma_{b,r}$ when the burrow contains a rabbit.

After observing the cue, the fox decides whether to dig or leave. If the fox decides to leave, its payoff is zero. As before, the cost of encountering a badger is $C$ and the reward for finding a rabbit is $R$. The prior probability that the environment is bad is $q_0$ and future decisions are discounted at a rate of $\lambda$ per time step. Although not as simple as before, we can again use dynamic programming to calculate the optimal behavior (see Appendix B).

### *3.4.2 Optimal behavior*

In this extended model, the good and bad environments can differ not only in the frequency of badgers, but also in how readily badgers can be distinguished from rabbits by sound alone. Here we will investigate what happens when in good environments, badgers are much louder than rabbits, but in bad environments they are only a little bit louder. We are particularly interested in this case because we want to know what happens when the fox must learn how cues correspond to potential threats.

To model this situation, we set the mean loudness of rabbits to 0 in both good and bad environments ($\mu_{g,r} = \mu_{b,r} = 0$). (The scale is arbitrary; we have chosen the value 0 for convenience.) In the good environment, badgers are much louder than rabbits ($\mu_{g,c} = 2$), and are therefore usually easy to detect. In the bad environment, they are only a bit louder than rabbits ($\mu_{b,c} = 1$) which can make them more difficult to detect. Everything else about the signal detection problem in the two environments is the same: $\sigma_{g,r} = \sigma_{g,c} = \sigma_{b,r} = \sigma_{b,c} = 0.5$, and $p_g = p_b = 0.2$. Figure 3.3A shows the distributions of cue intensities for the two environments. The punishment for encountering a badger is greater than the reward for finding a rabbit ($R = 1$, $C = 19$) and as in the previous model, future rewards are discounted at a rate of $\lambda = 0.95$ per time step and good and bad environments are equally common

($q_0 = 0.5$).

The optimal decision rule for the fox, as found by dynamic programming, is illustrated in Figure 3.3B. The fox now takes into account both its subjective probability that the environment is bad and the intensity of the cue it observes. A curve separates the (cue, probability) pairs at which the fox should dig from the (cue, probability) pairs at which the fox should not. For cues below 0.11, the fox should dig irrespective of the state of the environment; for cues above 0.81, the fox should not dig under any circumstance. In between, the fox must balance the strength of the cue against its subjective probability that the environment is bad. Here we can see the exploration-exploitation tradeoff in action. Given the large payoff to be gained from exploiting a good environment over many time steps, the possibility of discovering that the environment is good may compensate for the risk of punishment—even when it is more likely than not that the environment is bad.

### *3.4.3 Population outcomes*

In this signal detection model the fox has two ways to learn about its environment. As before, the fox gains information from exploring a burrow and discovering either a rabbit or badger. But even when the fox chooses not to dig, the fox still gains a small amount of information from observing the cue itself, because the probability of observing a given cue is generally different between the two environments. As a result, individuals will not become stuck forever with an incorrect belief that the environment is bad the way they could in the model of section 2. However, an asymmetry remains between the two kinds of mistakes: it is easier for a fox to learn that it has mistakenly inferred that the environment is good than it is for the fox to learn that it has mistakenly inferred that the environment is bad.

In this model, we observe a qualitatively similar pattern to what we found in the simpler model without cues. Figure 3.4 shows the outcome for the whole population when individuals follow the optimal strategy depicted in Figure 3.3B. When the environment is bad, the majority of foxes correctly learn this. The population distribution of beliefs forms a curve that increases roughly monotonically from left to right, with very few individuals believing

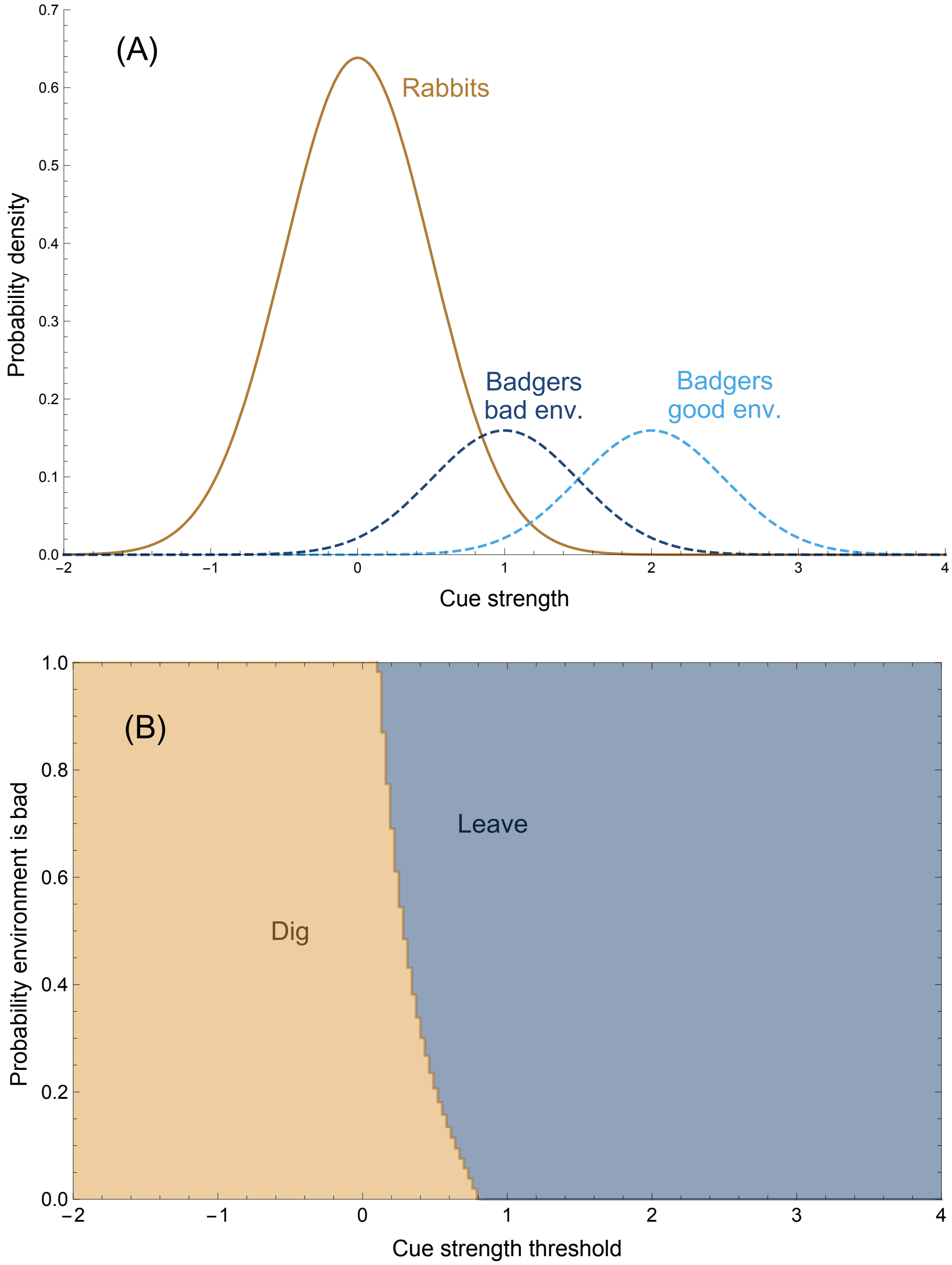


Figure 3.3: The two environments differ in how loud badgers are (A). In the good environment, badgers are easier to detect than they are in the bad environment. The optimal decision rule is computed using dynamic programming and illustrated in the lower panel (B). The decision about whether to dig depends on the value $x$ of the cue and the subjective probability that the environment is bad. A curve separates the region in which one should dig (tan) from the region in which one should not (blue).

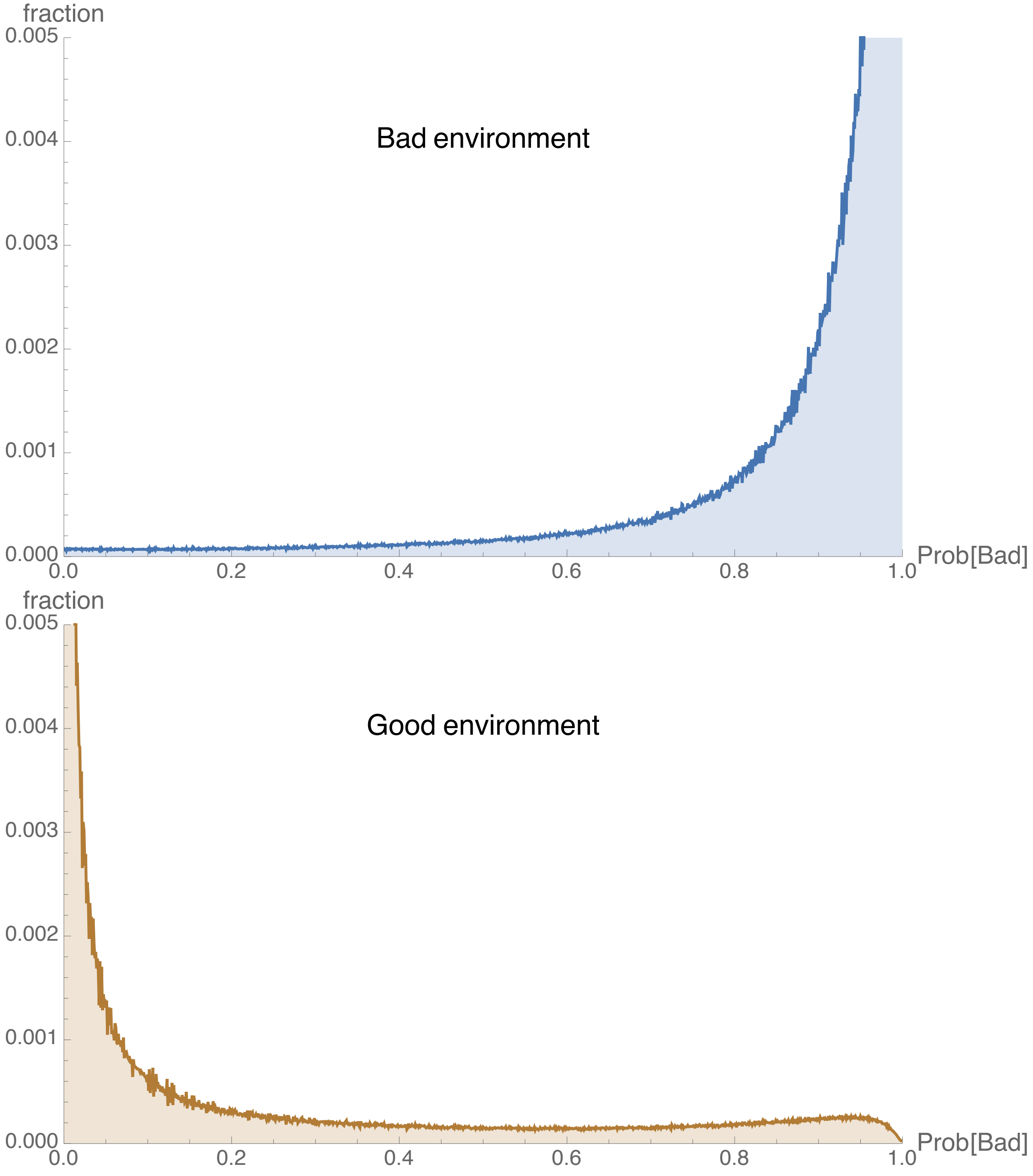


Figure 3.4: Population distribution of subjective probabilities that the environment is bad after 20 time steps, among foxes who have lived that long. When the environment is actually bad (upper panel), all but 4.5 percent of the population accurately come to believe that the environment is more likely to be bad than good. But when the environment is actually good (lower panel), 8.8 percent of the population erroneously come to believe that it is more likely that the environment is bad. All individuals began with a prior probability of 0.5 on the environment being bad.

that the environment is good, and the great majority correctly believing that the environment is bad. When the environment is good, the majority of foxes learn this as well. But a substantially minority reach the incorrect conclusion that the environment is bad. We see this in the fatter tail of the population distribution of beliefs, and in the existence of a small peak corresponding to the false conclusion that the environment is bad. In this example, roughly twice as many individuals become overly sensitive to loud sounds because they think the environment is bad as become insufficiently sensitive to loud sounds because they think the environment is good (8.8 percent versus 4.5 percent).

One might have thought that having informative cues would always enable the individual to learn to respond appropriately. The reason that it doesn't is that if a fox is in a good environment but is initially unlucky, and receives punishments after observing intermediate cues, then the individual will no longer dig when faced with cues of similar or greater strength. It thus becomes difficult for the fox to correct its mistake and learn that these cues indicate a lower risk of danger than it believes. So this particular fox becomes stuck with an over-sensitivity to the cues of potential danger. Its anxiety level is set too high. The same thing does not happen when a fox in a bad environment is initially lucky. In that situation, the fox continues to dig at burrows and is soon dealt a harsh punishment by the law of large numbers.

### 3.5 Discussion

Researchers are discovering many ways in which adaptive behavior can result in seemingly perverse consequences, such as apparent biases or "irrational" behavior [40, 22]. Examples include contrast effects [57], state-dependent cognitive biases [30, 69], optimism and pessimism [59], and superstition [24].

The results of these studies generally explain that the apparently irrational behavior is actually adaptive when understood in its appropriate evolutionary context. In this paper we take a different approach by separating the question of optimal learning rules from the question of whether each individual following such rules ends up behaving optimally. (See

Trimmer et al. [91] for a similar approach applied to clinical depression.) We show how behavior that is truly dysfunctional (in the sense that it reduces fitness) can arise in a subset of a population whose members follow the optimal behavioral rule, i.e., the rule that generates the highest expected payoff and would thus be favored by natural selection. This approach is well suited to providing insight into behavioral disorders, since they afflict only a subset of the population and are likely detrimental to fitness. We find that because an exploration-exploitation tradeoff deters further exploration under unfavorable circumstances, optimal learning strategies are vulnerable to erroneously concluding that an environment is bad. A major strength of the model is that it predicts excessive anxiety in a subset of the population, rather than in the entire population as we would expect from "adaptive defense mechanism" or "environmental mismatch" arguments [66].

An interesting aspect of our model is that it predicts the effectiveness of exposure therapy for anxiety disorders [70]. In the model, the individuals that are overly anxious become stuck because they no longer observe what happens if they are undeterred by intermediate-valued cues. If these individuals were forced to take risks in response to the cues that they believe are dangerous but are actually safe, then they would learn that their beliefs were mistaken and would correct their over-sensitivity. This exactly corresponds with the approach employed in exposure therapy.

Of course such a simple model cannot explain the myriad specific characteristics of real anxiety disorders. One example is that our model fails to capture the self-fulfilling prophecy, or vicious circle aspect, common to excessive anxiety. Being afraid of badgers does not make a fox more likely to encounter badgers in the future. But if a person is nervous because of past failures, that nervousness may be a causal component of future failure. Test anxiety is an example: a student performs poorly on one or more tests, becomes anxious about subsequent tests, and that anxiety contributes to poor performance in the future. Though it is challenging to see how such self-fulfilling anxiety fits into a framework of evolutionary adaptation, modeling the runaway positive feedback aspect of anxiety is an intriguing area for future work.

Another interesting direction for future work would be to investigate the case when the environment varies over time. Our current model is well suited to address a situation in which offspring disperse to different patches in the environment that remain constant over time (that is, when there is spacial variation but not temporal variation). But some environments will also vary over the duration of an individual's lifetime.

Before concluding, we want to point out a consequence in the second model of foxes being able to learn about their environment even when they only observe the cue itself. This means that there are actually two ways that a fox can end up being overly afraid despite living in a good environment. The first parallels our example in the first model: the fox could have had an unlucky early experience with a badger despite detecting only a modest signal, and from this could have mistakenly concluded that it lives in a bad environment. But there is another way that has no analog in the first model: It could be that or fox has never actually encountered a badger firsthand, but rather has received a series of cues more consistent with a bad environment then with a good one, and from these cues alone concluded that he lives in a bad environment even though he's never actually met a badger.

We speculate that these two different scenarios may correspond at least somewhat to different types of anxiety disorders. In the former scenario, present anxiety is the result of past trauma. Post-traumatic stress disorder would appear to be a very straightforward example of such a situation. In the latter, present anxiety would be the result of the mistaken belief that one lives in an unpredictable world, specifically one in which future trauma is difficult to detect and avoid. In both cases, the excessive anxiety on the part of the fox is a consequence of bad luck. But the bad luck can take different forms. In the former case the bad luck comes in the form of a badger observed despite a low signal. In the latter, the bad luck comes in the form of the unsampled signals taking a lower distribution than would be expected given the state of the world.

In general, signal detection models of threat such as these can have a number of moving parts. The degree to which the distribution of cues resulting from good events in bad worlds, and bad events in good worlds happens to overlap is one important factor, and the one we focused on here. Another factor that we've mentioned is when the frequency of good and bad events vary. A further possibility is that the benefits and costs of good and bad events could vary as well. One might even consider mismatch models in which foxes have evolved to distinguish between good and bad worlds but in fact badgers are entirely extinct. Here, the fox might conclude that he lives in a bad world with low discriminability because he has't seen any of the high magnitude signals that he would see in a good world with high discriminability. Considering this range of model possibilities one might be able to demarcate a number of different types of anxiety with different etiology and different predicted forms of treatment. We are currently developing models to explore these possibilities.

In this paper we have illustrated a fundamental design compromise: If an anxiety system is able to learn from experience, even the most carefully optimized system is vulnerable to becoming stuck in a state of self-perpetuating over-sensitivity. This effect is driven by the tradeoff an individual faces between gaining information by experience and avoiding the risk of failure when circumstances are likely unfavorable. Our results provide a new context for thinking about anxiety disorders: rather than necessarily viewing excessive anxiety as a result of dysregulated or imperfectly adapted neurological systems, we show that many of the features of anxiety disorders can arise from individual differences in experience, even when individuals are perfectly adapted to their environments. We suggest that this phenomenon may be an important causal component of anxiety disorders.

## 3.6 Acknowledgments

The authors thank Corina Logan, Randy Nesse and two anonymous referees for helpful suggestions and discussions. This work was supported in part by NSF grant EF-1038590 to CTB and by a WRF-Hall Fellowship to FM.

## 3.7 Appendix A: Sensitivity analysis for Model 1

A central point of this paper is that there is an asymmetry between the fraction of individuals who are wrong about the environment when it is in fact good, and the fraction who are wrong about it when it is bad. In the example we chose in section 2, only 0.2 percent of the population were optimistic in a bad environment, but 11.1 percent of the population were pessimistic in a good environment. In this appendix we investigate the extent to which changes in the model parameters affect this result.

There are 4 important independent values that parametrize the model. They are: the probability $p_g$ of encountering a badger when the environment is good, the probability $p_b$ of encountering a badger when the environment is bad, the discount factor $\lambda$, and the magnitude of the cost of encountering a badger relative to the reward for finding a rabbit, $C/R$.

We first investigate the effect of varying $p_g$ and $p_b$. In order for the state of the environment to matter—for there to be any use of gaining information—we must have the expected payoff be positive when the environment is good, $(1-p_g)R-p_gC > 0$, and be negative when the environment is bad, $(1-p_b)R-p_bC < 0$. Rearranging these inequalities gives us the constraints

$$p_b > \frac{R}{R+C} > p_g. \tag{3.1}$$

When $R = C$, as in section 3, these constraints, along with the constraint that $p_g$ and $p_b$ are probabilities that must lie between 0 and 1, restrict us to the square $0.5 < p_b \leq 1$, $0 \leq p_g < 0.5$. Figure 3.5 displays the results of analyzing the model over a grid of values for $p_g$ and $p_b$ within this square. Plotted is the fraction of the population that is wrong about the environment, as measured after 20 time steps among foxes who have survived that long.

In the upper left panel of Figure 3.5 (bad environment) the fraction of the population that is wrong is negligible everywhere except for the lower left corner of the plot, where the probabilities of encountering a badger in the good environment and in the bad environment are so similar that 20 trials simply does not provide enough information for accurate discrimination. But when the environment is actually good (upper right of Figure 3.5), it is almost the entire parameter space in which a substantial fraction of the population is wrong about the environment.

Instead of being smooth, the plots are textured by many discontinuities. Optimal behavioral rules cease to explore after small numbers of failures. But these small numbers depend on the parameter values and so discontinuities result around curves in parameter space that are thresholds for different optimal behavioral rules. However, in spite of the rugged shape of the plot, the basic trend in the upper right-hand panel of Figure 3.5 is that the fraction of the population that believes the environment is bad when it is actually good increases with $p_g$. In the lower panel of Figure 3.5 we see that for over 93 percent of the points in the parameter grid more of the population is wrong in the good environment than in the bad environment. And the small fraction of parameter combinations where this is not the case all occur towards the edge of the parameter space (on the left side in the plot).

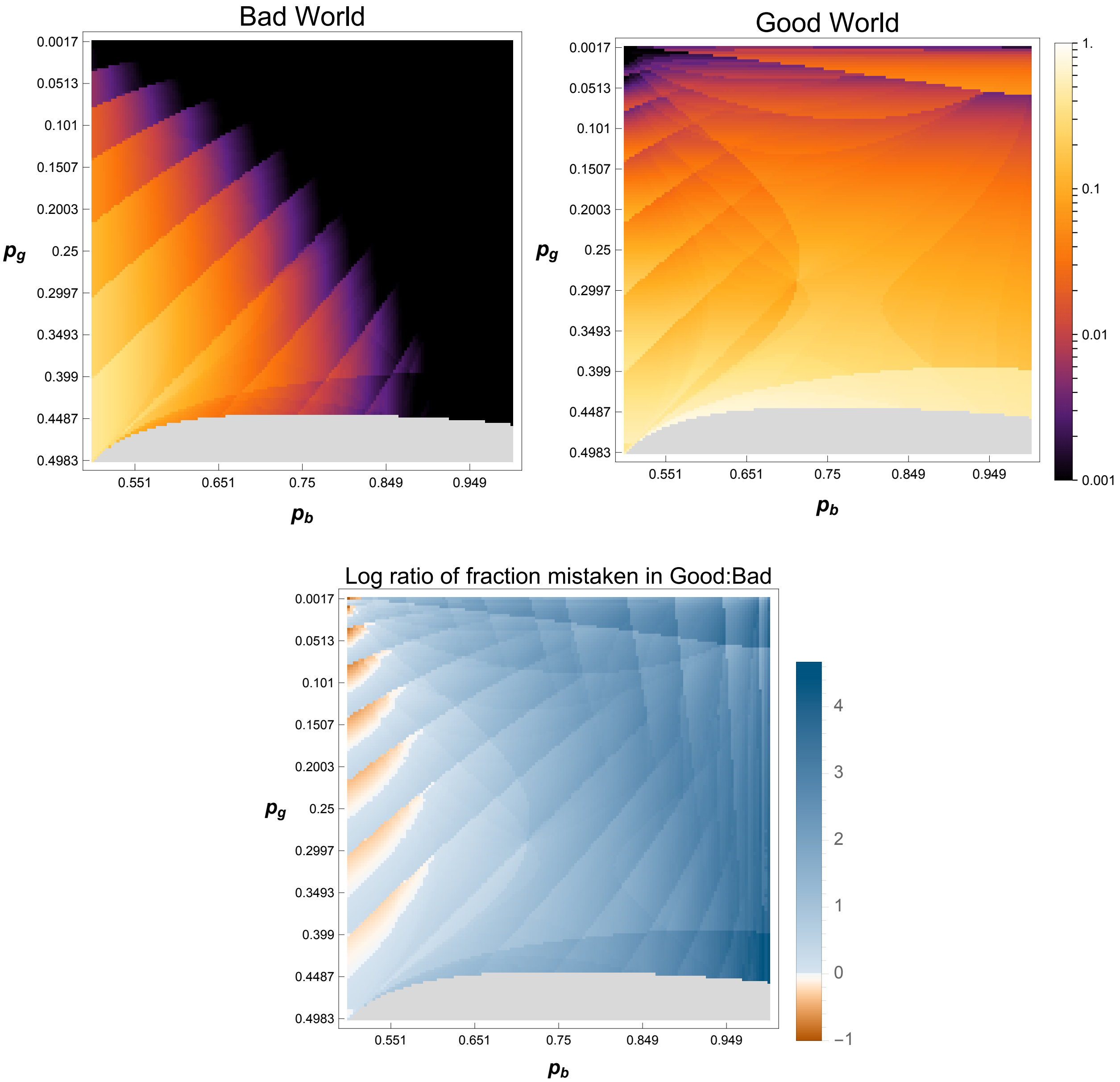
We next investigate the effect of varying the discount factor $\lambda$ and the cost to reward ratio $C/R$, while keeping $p_g$ and $p_b$ constant. Again, for it to matter whether the environment is good or bad, our parameters must satisfy inequalities (3.1). Rearranging these gives us the following constraint on the cost/reward ratio:

$$\frac{1-p_b}{p_b} < \frac{C}{R} < \frac{1-p_g}{p_g}. \tag{3.2}$$

When $p_g = 0.25$ and $p_b = 0.75$ this gives us $\frac{1}{3} < \frac{C}{R} < 3$. Figure 3.6 shows results for the model with values of $\frac{C}{R}$ sampled within this interval and values of $\lambda$ ranging from 0.75 to 0.99. The beliefs are measured at the time step that is closest to $\frac{1}{1-\lambda}$, the average lifespan given a discount factor of $\lambda$.

Figure 3.6 shows that, similar to the pattern in Figure 3.5, the fraction of the population that is wrong when the environment is bad is negligible except when there are not enough time steps in which to make accurate discriminations (in the lower part of Figure 3.6). By contrast, the fraction of the population that is wrong when the environment is good is non-negligible throughout most of the parameter space.

Discontinuities due to optimal behavior being characterized by small integer values are especially striking here, especially in the upper right panel of Figure 3.6. What is happening is that the number of failures it takes before it is optimal to cease to explore is the main outcome distinguishing different parameter choices. With $p_g$ and $p_b$ fixed, that number also determines the fraction of the population that will hit that number of failures. And so


Bad World

Good World

Log ratio of fraction mistaken in Good:Bad

Figure 3.5: Varying the probability of encountering badgers in each environment. With $\lambda = 0.95$, $C = 1$, and $R = 1$, the upper panels show how the fraction of the population that is wrong about the environment varies as a function of the parameters $p_g$ and $p_b$. The upper left shows the fraction that thinks the environment is good when it is actually bad. The upper right panel shows the fraction that thinks the environment is bad when it is actually good. This fraction is measured conditional on survival to the 20th time step, which is the average lifespan when $\lambda = 0.95$. The lower panel illustrates the log (base 10) of the ratio of incorrect inference rates in good and bad environments. For a small set of parameter values (shown in orange), incorrect inferences are more common in the bad environment. The gray area in each plot is a region in which it is not worthwhile to start exploring at all.

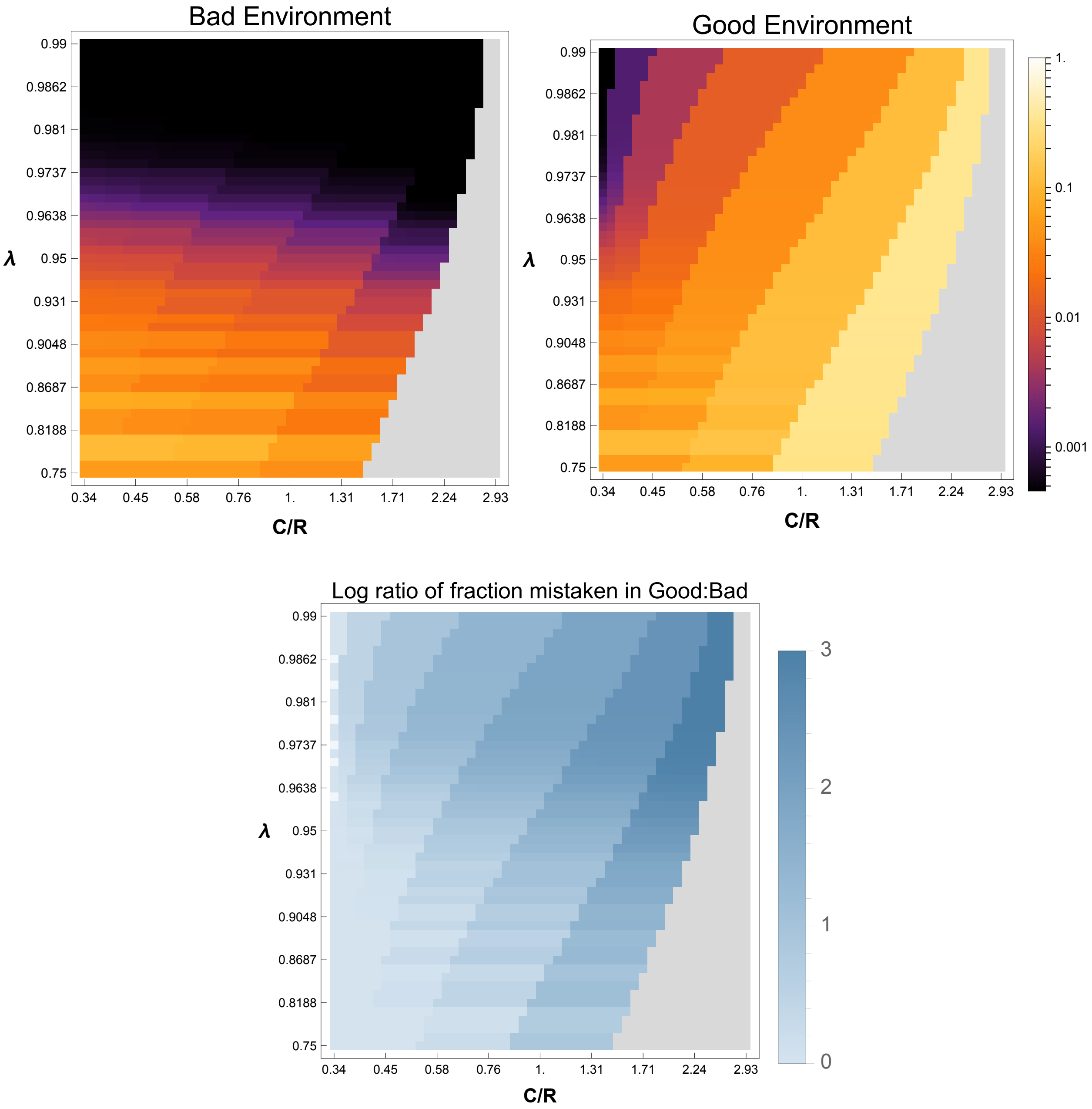


Figure 3.6: Varying the discount factor and cost/reward ratio. With $p_g = 0.25$ and $p_b = 0.75$, the upper panels show how the fraction of the population that is wrong about the environment varies as a function of $\lambda$ and $C/R$. The upper left plot displays the fraction that thinks the environment is good when it is actually bad; the upper right plot displays the fraction that thinks the environment is bad when it is actually good. This fraction is measured at the time step that is closest to $\frac{1}{1-\lambda}$, the average lifespan given $\lambda$. (The faint horizontal bands towards the lower part of the plots are due to the fact that $\frac{1}{1-\lambda}$ must be rounded to the nearest integer-valued time step.) The lower plot illustrates the log (base 10) of the ratio of incorrect inference rates in good and bad environments. Here, incorrect inferences are more common in good environments for all parameter values. The gray area in each plot is a region in which it is not worthwhile to start exploring at all.

the plot is characterized by a small number of curved bands in which the fraction of the population that is wrong about the environment is nearly constant. Although the value is nearly constant within each band, we can still describe the trend across these bands. We see that the fraction of the population that believes that the environment is bad when it is actually good increases with increasing relative cost or decreasing discount factor.

### *3.8 Appendix B: Finding optimal behavior*

The signal detection model of section 3, in which the fox uses environmental cues, is defined by the discount factor $\lambda$, the cost of encountering a badger $C$, the reward from catching a rabbit $R$, the initial subjective probability of being in a bad environment $q_0$, the probabilities of badgers in the good environment ($p_g$) and in the bad environment ($p_b$), and the Gaussian distribution parameters $\mu_{g,c}$, $\sigma_{g,c}$, $\mu_{g,r}$, $\sigma_{g,r}$, $\mu_{b,c}$, $\sigma_{b,c}$, $\mu_{b,r}$, and $\sigma_{b,r}$. The simpler model of section 2 can be seen as a special case of the more complex model in which the cues carry no information (because the means are all the same). Thus, analyzing the model with cues will also provide an analysis of the simpler model. The problem can be framed as a Markov decision process, and can be analyzed with a dynamic programming approach [11].

The fox knows the initial prior probability that the environment is bad, and at time step $t$ will also know the outcome of any attempts made before $t$. For each time step $t$ and all possible previous experience, a behavioral rule specifies the threshold cue level $u_t$ such that the fox will not dig at the burrow if the observed cue intensity, $x_t$, is greater than $u_t$. The only relevant aspect of previous experience is how this experience changes the current conditional probability $q_t$ that the environment is bad. So an optimally behaving agent will calculate $q_t$ using Bayes' rule, and use this value to determine the threshold level $u_t$. Thus, we can express a behavioral rule as the set of functions $u_t(q_t)$.

Let

$$f_{g,c}(x) = \frac{1}{\sigma_{g,c}\sqrt{2\pi}}\, e^{-\frac{(x-\mu_{g,c})^2}{2\sigma_{g,c}^2}}$$

be the Gaussian distribution function with mean $\mu_{g,c}$ and standard deviation $\sigma_{g,r}$ and similarly let $f_{g,r}(x)$, $f_{b,c}(x)$, and $f_{b,r}(x)$ be the other three corresponding Gaussian distributions.

### *3.8.1 Expected immediate payoff*

Let $\iota_t$ be the indicator random variable that equals 1 if the burrow contains a badger and equals 0 if the burrow contains a rabbit at time $t$. (Note that the random variables $\iota_t$ and $x_t$ covary.) We now define $y(q_t, u_t, x_t, \iota_t)$ to be the payoff the fox receives at time $t$ as a function of its threshold ($u_t$), the probability $q_t$ that the environment is bad, and the random variables $x_t$ and $\iota_t$. So

$$y(q_t, u_t, x_t, \iota_t) = \begin{cases} 0 & \text{if } x_t > u_t \\ R & \text{if } x_t \le u_t \text{ and } \iota_t = 0 \\ -C & \text{if } x_t \le u_t \text{ and } \iota_t = 1 \end{cases}$$

In the bad environment, the probability density of badgers and a cue strength of $x_t$ is $f_{b,c}(x_t)\, p_b$. Likewise, $f_{g,c}(x_t)\, p_g$ gives the probability density of badgers and a cue strength of $x_t$ in the good environment. Similarly $f_{b,r}(x_t)\,(1-p_b)$, and $f_{g,r}(x_t)\,(1-p_g)$ give the same probability densities for rabbits. This allows us to calculate the expected immediate payoff for the strategy of threshold $u_t$ as

$$\begin{aligned} \mathrm{E}\{y(q_t, u_t, x_t, \iota_t)\} = \int_{-\infty}^{u_t} \big[ &\big(f_{b,c}(x_t)\, p_b\, q_t + f_{g,c}(x_t)\, p_g\, (1-q_t)\big)(-C) \\ &+ \big(f_{b,r}(x_t)\,(1-p_b)\, q_t + f_{g,r}(x_t)\,(1-p_g)\,(1-q_t)\big) R\big]\, \mathrm{d}x_t. \end{aligned}$$

### *3.8.2 Bayesian updating*

We now describe how the Bayesian probability that the environment is bad, $q_t$, changes with time $t$. That is, we show how $q_{t+1}$ stochastically depends on $q_t$ and the threshold $u_t$.

The probability that the cue intensity is less than or equal to the threshold ($x_t \le u_t$) is given by

$$\int_{-\infty}^{u_t} \big[\big(f_{b,c}(x_t)\, p_b + f_{b,r}(x_t)\,(1-p_b)\big)\, q_t + \big(f_{g,c}(x_t)\, p_g + f_{g,r}(x_t)\,(1-p_g)\big)(1-q_t)\big]\, \mathrm{d}x_t.$$

The probability that the cue intensity is greater than the threshold ($x_t > u_t$) is the complement (the above quantity subtracted from 1).

If $x_t > u_t$, then the fox does not explore the burrow, but gains information about the environment from the cue itself, $x_t$. In the bad environment, the probability density on cues $x$ is given by

$$p_b\, f_{b,c}(x) + (1 - p_b) f_{b,r}(x).$$

Similarly, in the good environment it is given by

$$p_g\, f_{g,c}(x) + (1 - p_g) f_{g,r}(x).$$

So according to Bayes' rule, the posterior distribution on $q_{t+1}$ given an $x_t > u_t$ is

$$q_{t+1} = \frac{q_t \big(p_b\, f_{b,c}(x_t) + (1 - p_b) f_{b,r}(x_t)\big)}{q_t \big(p_b\, f_{b,c}(x_t) + (1 - p_b) f_{b,r}(x_t)\big) + (1 - q_t)\, \big(p_g\, f_{g,c}(x_t) + (1 - p_g) f_{g,r}(x_t)\big)}.$$

On the other hand if $x_t \leq u_t$, the fox decides to dig. The fox will then observe both the cue and whether the burrow contains a badger or a rabbit. The conditional probability that the burrow contains a badger given the cue $x_t$ and a bad environment is

$$\begin{aligned}
\mathrm{P}(\text{badger} \,|\, x_t, \text{bad env.}) &= \frac{\mathrm{P}(x_t \,|\, \text{badger}, \text{bad env.})\, \mathrm{P}(\text{badger} \,|\, \text{bad env.})}{\mathrm{P}(x_t \,|\, \text{bad env.})} \\
&= \frac{f_{b,c}(x_t)\, p_b}{f_{b,c}(x_t)\, p_b + f_{b,r}(x_t)\, (1 - p_b)}.
\end{aligned}$$

(Note that technically some of these quantities are probability densities rather than probabilities.) Similarly, if the environment is good, then

$$\begin{aligned}
\mathrm{P}(\text{badger} \,|\, x_t, \text{good env.}) &= \frac{\mathrm{P}(x_t \,|\, \text{badger}, \text{good env.})\, \mathrm{P}(\text{badger} \,|\, \text{good env.})}{\mathrm{P}(x_t \,|\, \text{good env.})} \\
&= \frac{f_{g,c}(x_t)\, p_g}{f_{g,c}(x_t)\, p_g + f_{g,r}(x_t)\, (1 - p_g)}.
\end{aligned}$$

Thus, the total probability of encountering a badger ($\iota_t = 1$) given an $x_t \leq u_t$ is

$$q_t\, \frac{f_{b,c}(x_t)\, p_b}{f_{b,c}(x_t)\, p_b + f_{b,r}(x_t)\, (1 - p_b)} + (1 - q_t) \frac{f_{g,c}(x_t)\, p_g}{f_{g,c}(x_t)\, p_g + f_{g,r}(x_t)\, (1 - p_g)}.$$

If the burrow does contain a badger, then by Bayes' rule we can express the posterior probability that the environment is bad as follows.

$$
\begin{aligned}
\mathrm{P}(\text{bad env.} \,|\, x_t, \text{badger}) &= \frac{\mathrm{P}(x_t, \text{badger} \,|\, \text{bad env.})\, \mathrm{P}(\text{bad env.})}{\mathrm{P}(x_t, \text{badger})} \\
\mathrm{P}(\text{bad env.} \,|\, x_t, \text{badger}) &= \frac{\mathrm{P}(x_t \,|\, \text{badger}, \text{bad env.})\, \mathrm{P}(\text{badger} \,|\, \text{bad env.})\, \mathrm{P}(\text{bad env.})}{\mathrm{P}(x_t, \text{badger})} \\
q_{t+1} &= \frac{f_{b,c}(x_t)\, p_b\, q_t}{f_{b,c}(x_t)\, p_b\, q_t + f_{g,c}(x_t)\, p_g\, (1 - q_t)}.
\end{aligned}
$$

Perfectly analogous calculations hold for the case when the burrow contains a rabbit ($\iota_t = 0$).

Below, we will express $q_{t+1}$ as a function,

$$q_{t+1} = w(q_t, u_t, x_t, \iota_t),$$

that depends on the threshold $u_t$, the probability $q_t$, and the random variables $x_t$ and $\iota_t$, as described above.

### *3.8.3 Dynamic programming*

The dynamic programming algorithm now consists of recursively calculating the maximum payoff attainable over all time steps subsequent to $t$, as a function of the current probability that the environment is bad. This maximum payoff is denoted $V_t(q_t)$, and the recursive formula is

$$V_t(q_t) = \max_{u_t} \mathrm{E}\{y(q_t, u_t, x_t, \iota_t) + \lambda\, V_{t+1}(w(q_t, u_t, x_t, \iota_t))\}.$$

And the optimal decision rule functions, $u_t^*$, are given by

$$u_t^*(q_t) = \arg\max_{u_t} \mathrm{E}\{y(q_t, u_t, x_t, \iota_t) + \lambda\, V_{t+1}(w(q_t, u_t, x_t, \iota_t))\}.$$

Because $q_t$ is a continuous variable, a discrete approximation must be used for the actual computation. Then the table of values for $V_{t+1}$ is used to compute the values for $V_t$, indexed by $q_t$. For $q_t$ we used 1001 discrete values $(0, 0.001, 0.002, \ldots, 1)$. As mentioned in section 3, we set $\mu_{g,c} = 2, \mu_{b,c} = 1$, $\mu_{g,r} = \mu_{b,r} = 0$, and $\sigma_{g,r} = \sigma_{g,c} = \sigma_{b,r} = \sigma_{b,c} = 0.5$. To discretize $x_t$, we must pick minimum and maximum values, which we set at $-2$ and 4, respectively.

Within this interval we discretized $x_t$ to 200 values. Because there is a tiny area lost at the ends of the distributions, we renormalized the total probabilities to 1.

The algorithm then gives us two tables: one containing the expected values and the other containing the optimal decision rule functions, or thresholds, $u_t^*(q_t)$, which are indexed by our grid of values for $q_t$. To find the optimal behavior in the limit as the possible lifetime extends towards infinity, the recursion is repeated until the optimal decision rules converge [11]. The algorithm was implemented in python.

Once we have found the optimal decision rule, for each time step we can calculate the expected proportion of the population that has each value of $q_t$ as its estimate. Since the behavioral rule specifies the threshold for each value of $q_t$, we can use the distribution derived above for $q_{t+1} = w(q_t, u_t, x_t, \iota_t)$ to calculate the proportions for time $t$+1 given the proportions for time $t$. Because we discretized the $q_t$ values, we round the calculation of $q_{t+1}$ to the nearest one thousandth.

# Chapter 4

# A HYGIENE HYPOTHESIS FOR ANXIETY?

By Frazer Meacham and Carl T. Bergstrom

## *4.1 Abstract*

Anxiety disorders are a leading cause of disability and morbidity, especially in developed countries, with rising incidence and limited effective treatments. The ultimate cause for the frequency of anxiety disorders remains largely a mystery. Why has natural selection not done a better job of preventing us from being vulnerable to such a common, chronic condition? The evolutionary medicine approach has provided great insight into other kinds of common chronic disorders. For example, the "hygiene hypothesis" for autoimmune and inflammatory diseases (whether in its original form or an updated "old friends" version) has provided an environmental mismatch explanation for the high prevalence of such diseases. This hypothesis now has extensive empirical support showing that the decreased prevalence of microorganisms and parasites has led, paradoxically, to increased immune activity that causes these diseases of the immune system. In this paper, we explore an analogous explanation for anxiety disorders, whereby the decreased prevalence of threats in our modern environment is the cause of increased levels of anxiety. As an illustration, we develop and analyze a mathematical model, based on previous canonical models of anxiety, to capture this evolutionary mismatch phenomenon. We show that the agents of this simple model show paradoxical excessive anxiety when exposed to an environment with fewer threats than the environment in which they evolved. Although this work serves only to propose and explore this hypothesis, if future empirical work supports it, this will have important implications for the prevention and possibly the treatment of anxiety disorders.

## *4.2 Introduction*

By a myriad of measures, the modern environment of developed societies is safer and contains vastly fewer threats to our survival than the environment of our ancestors during most of our evolutionary history. Not only are we safer from accidents, diseases, and predators, but evidence from contemporary primitive societies strongly suggests we are also safer from threats from other humans such as homicides and warfare [29]. Our ancestors likely coped with their dangerous environment by being vigilant to all sorts of potential threats. This necessity for vigilance is the evolutionary explanation for our capacity to experience anxiety: anxiety is an adaptive response to situations that involve potential threats, and it serves to make us more alert and primed to deal with threats.

When an environment becomes safer, such as the modern environment of developed societies, threatening situations become rarer, and the protective role of anxiety will not need to be triggered as often. So given the evolutionary explanation for anxiety, one would expect that a safer modern environment would lead to anxiety being less common. But this is not what we see. Anxiety is extremely common in modern society, with a large subset of the population suffering from anxiety disorders [85, 42, 6]. How do we make sense of this?

The evolutionary medicine approach to anxiety and anxiety disorders begins with Nesse's framing the problem in the context of signal detection theory [65, 67]. The central insight was that even an adaptive anxiety response should be expected to exhibit false alarms the majority of the time. Further modeling work has described how individuals may become stuck in overly pessimistic states because of asymmetries of the learning process [91, 60]. However, none of these models explain why anxiety would increase as the environment becomes safer. In the absence of the dangers that anxiety was designed to protect us from, anxiety itself is now causing us harm, and it is not clear why.

We notice a parallel in the biology of inflammatory and autoimmune diseases. These diseases pose another example where a system that is built to protect us—in this case the immune system—is now the cause of a disease. Recent history has shown a striking increase in the prevalence of these types of disease. One well supported explanation for this increase, based on principles of evolutionary medicine, is the so-called “hygiene hypothesis” [87, 71]. This hypothesis proposes that our immune system is adapted to an environment where parasites and infectious agents are relatively common, and that it is not adapted to the extremely low level of parasites and pathogens in modern developed societies. This mismatch results in over-activation of the immune system. The hypothesis has been well supported by many studies, such as studies showing how increasing rates of parasitism and microbial infection correspond to decreasing rates of inflammatory diseases [92, 7]. The upshot is that these diseases may actually be caused in part by our attempts to make the world more free of disease.

We wonder if the same thing might be happening with anxiety. Could the increased safety of modern society actually be a causal factor for anxiety disorders? This paper develops a mathematical model to explore this question. In particular, our approach is to discover whether this seemingly paradoxical outcome can occur within the standard evolutionary medicine modeling framework used to model anxiety.

## *4.3 Modeling approach*

The signal detection model of anxiety was proposed by Nesse and has since been expanded upon by other researchers [65, 5, 69]. In its basic form, the model consists of an individual who is trying to distinguish safe situations from dangerous situations. With each of these types of situations is associated a normal distribution of cue intensities. These cue intensities might represent how loud a noise is that might indicate an approaching predator, or how threatening a rival is behaving that might indicate an imminent fight, etc. The cue distribution for dangerous situations has a higher mean than the cue distribution for safe situations, so an individual trying to avoid danger will flee from cues above some threshold level. If the threshold is set relatively low, the individual is quite sensitive to signs of potential danger, and this is conceptualized as corresponding to higher anxiety. If the individual sets its threshold relatively high, it is not sensitive to signs of threat, which corresponds to low anxiety.

Given specific costs for false negatives and false positives, along with the means and variances of both normal distributions, the simple version of the model described above can be used to calculate optimal behavior. But individuals will not have all of these pieces of information, and instead must learn some of them from observation. For example, an individual might not know the cost of getting hurt, or the frequency of danger, or how easy it is to distinguish dangerous situations from safe situations. Especially relevant to the problem of how changing the amount of threatening stimuli affects anxiety is the process whereby individuals learn from experience how much threat stimulus to expect. Thus, in this paper we investigate a version of the model in which individuals must learn from observation whether they are in an environment in which danger is easy to detect (more and stronger threat stimuli) or in an environment in which danger is hard to detect (fewer and weaker threat stimuli).

In the signal detection paradigm, danger is hard to detect when the means of the cue distributions for safe and dangerous situations are close together. Danger is easy to detect when the means of the distributions are far apart (see Figure 4.1). If an individual does not know which kind of environment it has been born into, it must learn from experience how easy danger is to detect. Natural selection will act towards optimizing the learning behavior that the individual follows in this trial and error process. Therefore, after setting up our model, we will solve for the optimal learning behavior, and then investigate how this optimal behavioral strategy responds to a safer environment from the one to which it is adapted. In this way, our model should give insight into how we should expect an anxiety system to respond to a safer modern environment.

## *4.4 Model*

We suppose that natural selection has acted to optimize individuals' responses to the problem of distinguishing between two environments: one in which danger is difficult to discriminate from safety (the "hard" environment), and one in which danger is relatively easy to discrim-

Figure 4.1: Cue distributions for safe and dangerous events in the easy environment and hard environment. Cues generated from safe events have the same distribution in both environments. Cues generated by dangerous events in the easy environment have a distribution that is far outside the distribution for safe events. Cues generated by dangerous events in the hard environment, however, have a distribution with substantial overlap with the distribution for safe events.

inate from safety (the “easy” environment). In these ancestral conditions, upon birth an individual finds itself in a hard environment with probability $q_0$, and in an easy environment with probability $1 - q_0$. In both environments, the probability that an event is dangerous is $p$, the probability that an event is safe is $1 - p$, and the mean of the cue distribution for safe events is $\mu$. In the hard environment, the mean of the cue distribution for danger is $\theta_H$, whereas in the easy environment it is $\theta_E$ (where $\theta_E > \theta_H$). We assume that all distributions have the same variance, $\sigma^2$.

The individual will encounter a sequence of cues at successive time points during its lifetime, and must decide at each one whether to flee from the cue or not. We let $R > 0$ be the reward for staying when the event was safe, and $C < 0$ be the cost of staying when the event was actually dangerous. If the individual decides to flee, the payoff is zero (regardless of whether the event that the individual flees was actually dangerous or safe). Future rewards are discounted at a rate of $\lambda$ per time step, and the individual's success is measured by its total lifetime payoff. We discount the rewards to account for the death of individuals. A discount rate of $\lambda$ corresponds to individuals dying at a constant probability of $1 - \lambda$ per time step. We assume infinite population size.

At each time step, the individual must make a decision as to what level of cue it will explore. The individual gains information about the kind of environment it is in at each time step by observing the cue, and also from the outcome of exploring the event if it does decide to explore. So a strategy consists of a threshold cue setting for any sequence of previous experiences. Luckily, the analysis is made simpler by the fact that all the relevant information from the sequence of previous experiences is contained in the current conditional probability that the individual's environment is easy versus hard. Thus, a strategy consists of a threshold setting for any given conditional probability that the environment is easy versus hard. This optimal behavior is found with a dynamic programming algorithm (see appendix).

Once we have found the optimal learning rule, we can look at a population of optimal learners within any particular environment, and describe their threshold settings after experiencing a certain number of events. Because of the stochasticity involved, the cue thresholds within a population of individuals will follow some distribution. This distribution is based on the distribution of individual conditional probabilities that the environment is easy versus hard, which result from the different specific experiences each individual has.

In the next section, we calculate these distributions for both the hard and easy environment, as well as for two cleansed environments where instead of the probability of danger being $p$, it is $p_m < p$ (see Figure 4.2). These environments represent modern environments that are mismatched to our ancestral conditions in that the great majority of dangerous threats have been removed. One is a less dangerous version of the easy environment (described below and shown in Figure 4.2), the other is a less dangerous version of the hard environment (results described in the text).

### *4.5 Results*

Figure 4.2 (upper panel) shows the distribution of thresholds for individuals who have experienced 20 events in a hard environment, and Figure 4.2 (middle panel) shows the distribution of thresholds for individuals who have experienced 20 events in an easy environment. (This value is chosen because 20 time steps is the life expectancy given our value for $\lambda$.) The parameter settings in this example are: safe cue distribution mean $\mu = 4$, mean of the cue distribution in the hard environment $\theta_H = 6$, mean of the cue distribution in the easy environment $\theta_E = 9$, distribution variances $\sigma^2 = 1$, reward for staying when the event is safe $R = 1$, cost of staying when the event is dangerous $C = -20$. The rate at which future rewards are discounted is $\lambda = 0.95$. The probability that an event is dangerous in the ancestral environments is $p = 0.25$, in the modern environments it is $p_m = 0.025$.

In the third panel of Figure 4.2 we show what happens when individuals are placed in an environment mismatched to the ones in which they evolved. This environment is like the one where danger is easy to detect, except that the frequency of danger has been reduced by 90%. This is analogous to a modern developed society where imminent threats to basic survival have been mostly removed. What we see is that, after individuals are exposed to

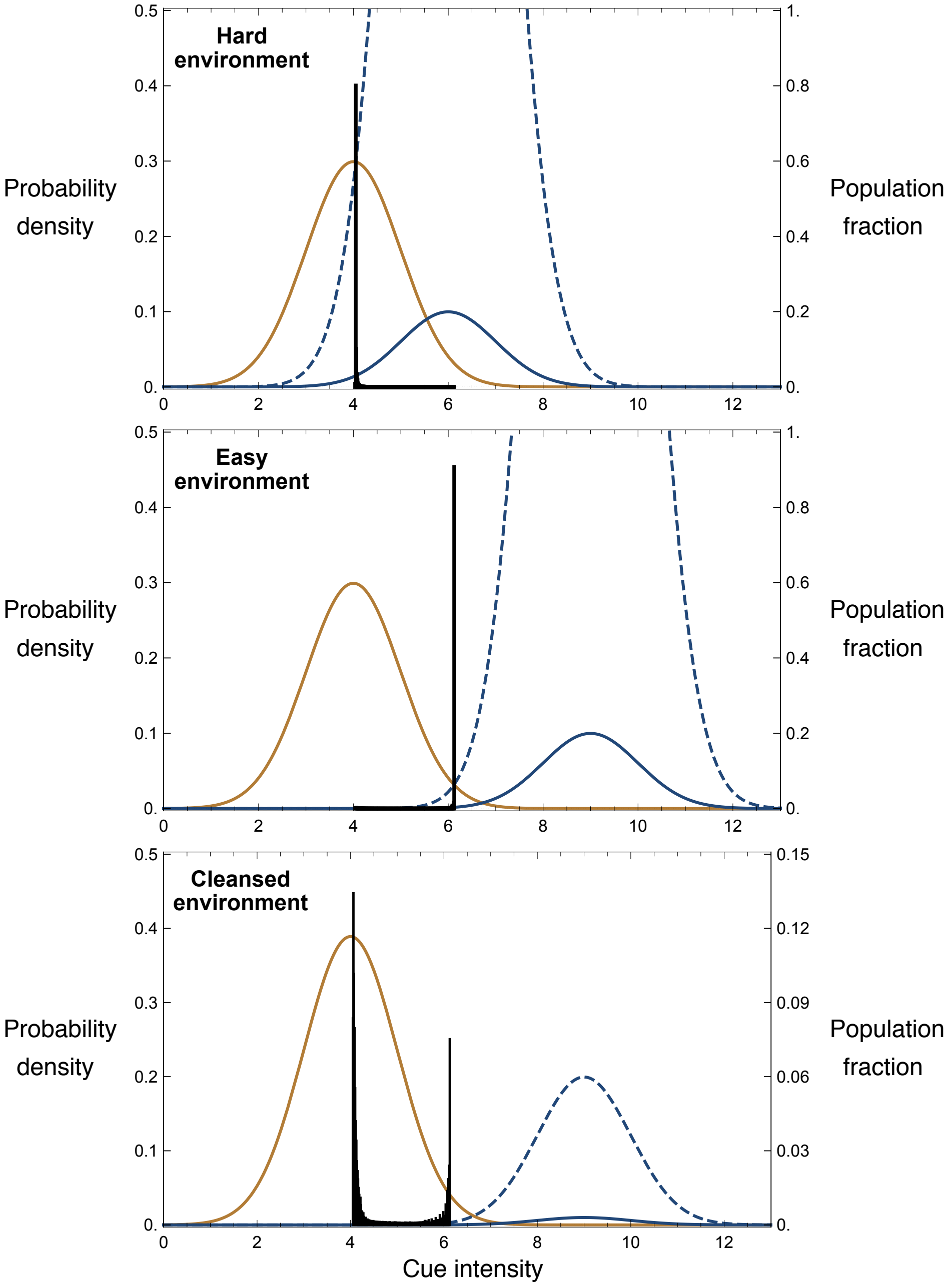


Figure 4.2: Cue distributions and population thresholds after individuals observe 20 events. The upper and middle panels show the hard environment and the easy environment, respectively. The yellow curves give the cue distribution for safe events, the blue curves give the cue distribution for dangerous events, and the dashed blue curves give the cue distribution for dangerous events multiplied by the relative cost of a false negative. The black area plot spikes give the population thresholds. For both the hard and easy environment, the great majority of individuals learn how to set their thresholds close to optimally, placed where the yellow curve intersects the blue dashed curve. In the cleansed environment (lower panel), a great many of the individuals incorrectly believe that the environment is hard, and consequently end up much too sensitive to cues of potential danger.

20 events, a great many (over 50%) believe that the environment is one in which danger is hard to detect, and thus are far too sensitive to cues of potential danger.

Another possible environment is one like the ancestral hard environment but with less frequent danger. When we calculate expected behavior in such an environment where danger has been reduced by 90%, the results are very similar to the ancestral hard environment, with over 95% of individuals believing that the environment is hard (data not plotted). This still causes individuals to set their thresholds too low because this cleansed environment is of course safer than the ancestral hard environment.

The results we see here are driven by the way individuals in the model learn from the observations they make. Because individuals must avoid danger, they set their thresholds low enough so that it is rare that they stay in a situation that turns out to be dangerous. Because of this, instead of learning about the environment by directly experiencing harmful events, the majority of an individual's learning comes from observing the intensities of the cues that they do not explore. In the ancestral conditions, what they learn from observing the cues is how often there are cues that strongly indicate danger. If cues that indicate danger strongly are relatively common, this means that the individual is more likely to be in an environment in which danger is easy to distinguish from safety. On the other hand, if cues that indicate danger strongly are relatively rare, this means they are more likely to be in an environment in which danger is difficult to detect. In the ancestral environments, this strategy works. But when these individuals are placed in the environment in which danger has been greatly reduced, their strategy malfunctions. We are left with the observation that removing danger has caused anxiety to increase.

## *4.6 Discussion*

We have demonstrated the surprising result that *decreasing* the amount of threatening stimuli has the potential to *increase* an individual's anxiety. Another mechanism that could produce a similar effect was recently demonstrated by [90]. As in previous work [91, 60], both our model and their's expand the basic signal detection framework by including multiple successive trials. But rather than including environmental mismatch to the evolved anxiety response, as we do, they add an assumption that the cost of fleeing is a loss of a foraging opportunity. This loss takes away from the individual's energy reserves. With enough such losses, energy reserves drop below a critical level and the individual dies. With these assumptions, they find that an individual might become more sensitive to cues of danger when such cues are rarer, not because they misinterpret the frequency of danger, as in our model, but because the individual can afford to be more skittish because future safer foraging opportunities will be more common. In this way, their model depends on varying energy reserves, whereas our model depends on environmental mismatch. Because of this, the two hypotheses could be distinguished by testing whether, for example, lower energy reserves lead the individual to be less sensitive to cues of danger.

An important point in the background of our hypothesis, as well as the traditional hygiene hypothesis for inflammatory and autoimmune diseases, is that the environment of modern society has not changed from our ancestral environment in an arbitrary way. Instead, we humans have purposely modified our environment in ways that are consistent with our evolved desires. We attempt to increase access to things that we are innately attracted to, and eliminate things we are innately averse to. In the case of parasites, pathogens, and threats of danger, our society has generally tried to remove them, and has had much success in doing so.

The above point has guided the development of our model. Our approach of modeling an environment mismatched to the one that individuals are adapted to contrasts with the traditional optimality approach used in much previous work [56, 91, 90, 60]. If our choice of mismatched environments were arbitrary, then the objection could be raised that our results would be equally arbitrary. But this is not the case, as our choice was constrained by the fact that we humans have modified our environment to remove threatening stimuli. The mismatch in our model reflects this exact case.

In recent years, alternative formulations have been proposed to revise the original hygiene hypothesis. In particular, the "old friends" hypothesis points out that many autoimmune and inflammatory diseases may be caused not by the absence of pathogenic organisms but by the absence of benign or beneficial microorganisms [83, 82]. But whichever version of the hypothesis one accepts, the basic cause of the increase in prevalence of these diseases is our own modification of our environment. And these modifications are usually a result of our attempts to satisfy basic human wants. In this view, inflammatory and autoimmune diseases are diseases of society's over-indulgence of our evolved desires. Our suggestion is that anxiety disorders may be another example of such a disease.

## 4.7 Acknowledgments

The authors thank Randy Nesse, Joe Felsenstein, and Michael Beecher for helpful comments and suggestions.

## 4.8 Appendix: Model analysis

The model and analysis of this paper is similar to that of [60]. Similar to that paper, our signal detection model is defined by the discount factor $\lambda$, the cost of failing to flee true danger $C$, the reward for staying during safety $R$, the initial subjective probability of being in the hard environment $q_0$, the probability of danger in both ancestral environments $p$, the Gaussian distribution parameters (cue variances $\sigma^2$, mean cue for safety in both environments $\mu$, mean cue for danger in the hard environment $\theta_H$, and in the easy environment $\theta_E$). The problem of finding optimal behavior can be framed as a Markov decision process, and can be analyzed with a dynamic programming approach [11]. After finding optimal behavior, we can calculate the expected population distribution of anxiety levels for both these ancestral environments and for a mismatched modern environment where the probability of danger is $p_m < p$.

An individual is adapted to believe that the initial prior probability that the environment is hard is $q_0$, and at time step $t$ will also know the outcome of any attempts made before $t$. For each time step $t$ and all possible previous experience, a behavioral rule specifies the threshold cue level $u_t$ such that the individual will flee for any cue $x_t$ where $x_t > u_t$. The only relevant aspect of previous experience is how this experience changes the current conditional probability $q_t$ that the environment is hard. So an optimally behaving agent will calculate $q_t$ using Bayes' rule, and use this value to determine the threshold level $u_t$. Thus, we can express a behavioral rule as the set of functions $u_t(q_t)$.

### *4.8.1 Expected immediate payoff*

For what follows, let

$$g_\theta(x) = \frac{1}{\sigma^2\sqrt{2\pi}}\, e^{-\frac{(x-\theta)^2}{2\sigma^2}}$$

be the Gaussian distribution function with mean $\theta$ and standard deviation $\sigma^2$.

We let $\iota_t$ be the indicator random variable that equals 1 if the event was dangerous and equals 0 if the event was safe at time $t$. (Note that the random variables $\iota_t$ and $x_t$ covary.) We now define $y(q_t, u_t, x_t, \iota_t)$ to be the payoff the individual receives at time $t$ as a function of its threshold ($u_t$), the probability $q_t$ that the environment is hard, and the random variables $x_t$ and $\iota_t$. So

$$y(q_t, u_t, x_t, \iota_t) = \begin{cases} 0 & \text{if } x_t > u_t \\ R & \text{if } x_t \leq u_t \text{ and } \iota_t = 0 \\ -C & \text{if } x_t \leq u_t \text{ and } \iota_t = 1 \end{cases}$$

In the hard environment, the probability density of danger and a cue strength of $x_t$ is $g_{\theta_H}(x_t)\, p$. Likewise, $g_{\theta_E}(x_t)\, p$ gives the probability density of danger and a cue strength of $x_t$ in the easy environment. In both environments, the cue distribution for safety is $g_\mu(x_t)\,(1-p)$. This allows us to calculate the expected immediate payoff for the strategy of threshold $u_t$ as

$$\mathrm{E}\{y(q_t, u_t, x_t, \iota_t)\} = \int_{-\infty}^{u_t} \left[\left(g_{\theta_H}(x_t)\, q_t + g_{\theta_E}(x_t)\,(1-q_t)\right) p\,(-C) + g_\mu(x_t)\,(1-p)\, R\right] \mathrm{d}x_t. \tag{4.1}$$

### *4.8.2 Bayesian updating*

We now describe how the individual's subjective Bayesian probability that the environment is hard, $q_t$, changes with time $t$. That is, we show how $q_{t+1}$ stochastically depends on $q_t$ and the threshold $u_t$. Keep in mind that $q_t$ is the probability that the environment is hard, which we could also express as P(hard env.).

The probability that the cue intensity is greater than the threshold ($x_t > u_t$) is given by

$$\mathrm{P}(x_t > u_t) = \int_{u_t}^{\infty} \left[ \big( g_{\theta_H}(x_t)\, q_t + g_{\theta_E}(x_t)\, (1 - q_t) \big)\, p + g_{\mu}(x_t)\, (1-p) \right] \mathrm{d}x_t. \tag{4.2}$$

The probability that the cue intensity is less than or equal to the threshold, $\mathrm{P}(x_t \le u_t)$, is the above quantity subtracted from 1.

If $x_t > u_t$, the individual flees, but gains information about the environment from the cue itself, $x_t$. In the hard environment, the probability density on cues $x_t$ (regardless of $t$) is given by

$$f(x_t \,|\, \text{hard env.}) = g_{\theta_H}(x_t)\, p + g_{\mu}(x_t)\, (1-p).$$

Similarly, in the easy environment it is given by

$$f(x_t \,|\, \text{easy env.}) = g_{\theta_E}(x_t)\, p + g_{\mu}(x_t)\, (1-p).$$

So according to Bayes' rule, the posterior value for $q_{t+1}$ given an $x_t > u_t$ is

$$\begin{aligned} \mathrm{P}(\text{hard env.} \,|\, x_t) &= \frac{f(x_t \,|\, \text{hard env.})\, \mathrm{P}(\text{hard env.})}{f(x_t \,|\, \text{hard env.})\, \mathrm{P}(\text{hard env.}) + f(x_t \,|\, \text{easy env.})\, \mathrm{P}(\text{easy env.})} \\ q_{t+1} &= \frac{\big( g_{\theta_H}(x_t)\, p + g_{\mu}(x_t)\, (1-p) \big)\, q_t}{\big( g_{\theta_H}(x_t)\, p + g_{\mu}(x_t)\, (1-p) \big)\, q_t + \big( g_{\theta_E}(x_t)\, p + g_{\mu}(x_t)\, (1-p) \big)\, (1-q_t)}. \end{aligned} \tag{4.3}$$

On the other hand if $x_t \le u_t$, the individual decides to stay. The individual will then observe both the cue and whether the event was actually dangerous or safe. The conditional probability that the event was dangerous given the cue $x_t$ and a hard environment is

$$\begin{aligned} \mathrm{P}(\text{danger} \,|\, x_t, \text{hard env.}) &= \frac{f(x_t \,|\, \text{danger}, \text{hard env.})\, \mathrm{P}(\text{danger} \,|\, \text{hard env.})}{f(x_t \,|\, \text{hard env.})} \\ &= \frac{g_{\theta_H}(x_t)\, p}{g_{\theta_H}(x_t)\, p + g_{\mu}(x_t)\, (1-p)}. \end{aligned}$$

Similarly, if the environment is easy, then

$$\begin{aligned}\mathrm{P}(\text{danger} \mid x_t, \text{easy env.}) &= \frac{\mathrm{P}(x_t \mid \text{danger}, \text{easy env.})\, \mathrm{P}(\text{danger} \mid \text{easy env.})}{\mathrm{P}(x_t \mid \text{easy env.})} \\ &= \frac{g_{\theta_E}(x_t)\, p}{g_{\theta_E}(x_t)\, p + g_\mu(x_t)\, (1-p)}.\end{aligned}$$

Thus, the total probability of the event being dangerous ($\iota_t = 1$) given an $x_t \leq u_t$ is

$$\mathrm{P}(\iota_t = 1 \mid x_t) \;=\; q_t \,\frac{g_{\theta_H}(x_t)\, p}{g_{\theta_H}(x_t)\, p + g_\mu(x_t)\, (1-p)} + (1-q_t)\frac{g_{\theta_E}(x_t)\, p}{g_{\theta_E}(x_t)\, p + g_\mu(x_t)\, (1-p)}. \tag{4.4}$$

If the event indeed was dangerous, then by Bayes' rule we can express the posterior probability that the environment is hard as follows.

$$\begin{aligned}\mathrm{P}(\text{hard env.} \mid x_t, \text{danger}) &= \frac{f(x_t, \text{danger} \mid \text{hard env.})\, \mathrm{P}(\text{hard env.})}{f(x_t, \text{danger})} \\ \mathrm{P}(\text{hard env.} \mid x_t, \text{danger}) &= \frac{f(x_t \mid \text{danger}, \text{hard env.})\, \mathrm{P}(\text{danger} \mid \text{hard env.})\, \mathrm{P}(\text{hard env.})}{f(x_t, \text{danger})} \\ q_{t+1} &= \frac{g_{\theta_H}(x_t)\, p\, q_t}{g_{\theta_H}(x_t)\, p\, q_t + g_{\theta_E}(x_t)\, p\, (1-q_t)}. \end{aligned} \tag{4.5}$$

On the other hand if the event was safe, then the posterior probability that the environment is hard is

$$\begin{aligned} q_{t+1} &= \frac{g_\mu(x_t)\, p\, q_t}{g_\mu(x_t)\, p\, q_t + g_\mu(x_t)\, p\, (1-q_t)}. \\ q_{t+1} &= q_t. \end{aligned} \tag{4.6}$$

In other words, when the agent stays and the event was safe, this information does not change the individual's subjective probability that the environment is hard. This is because the distribution on cues for safe events is the same in both the hard and easy environments.

In total, the above calculations allow us to express $q_{t+1}$ as a function,

$$q_{t+1} = w(q_t, u_t, x_t, \iota_t),$$

where

$$w(q_t, u_t, x_t, \iota_t) = \begin{cases} \text{[the quantity in eq. 4.3]} & \text{with probability } \mathrm{P}(x_t > u_t) \\ \frac{g_{\theta_H}(x_t)\, p\, q_t}{g_{\theta_H}(x_t)\, p\, q_t + g_{\theta_E}(x_t)\, p\, (1-q_t)} & \text{with prob. } \mathrm{P}(\iota_t = 1 \mid x_t)\, (1 - \mathrm{P}(x_t > u_t)) \\ q_t & \text{with prob. } (1 - \mathrm{P}(\iota_t = 1 \mid x_t)\, (1 - \mathrm{P}(x_t > u_t)) \end{cases}$$

from equations 4.3, 4.5, and 4.6 and where the values for $\mathrm{P}(x_t > u_t)$ and $\mathrm{P}(\iota_t = 1 \,|\, x_t)$ are given in equations 4.2 and 4.4, respectively.

### *4.8.3 Dynamic programming*

The dynamic programming algorithm now consists of recursively calculating the maximum payoff attainable over all time steps, $t$, as a function of $q_t$, the probability that the environment is hard. The calculation is carried out over a discretization spanning all values of $q_t$, moving backwards in time. The maximum payoff is denoted $V_t(q_t)$, and the recursive formula is

$$V_t(q_t) = \max_{u_t} \mathrm{E}\{y(q_t, u_t, x_t, \iota_t) + \lambda\, V_{t+1}(w(q_t, u_t, x_t, \iota_t))\}.$$

And the optimal decision rule functions, $u_t^*$, are given by

$$u_t^*(q_t) = \arg\max_{u_t} \mathrm{E}\{y(q_t, u_t, x_t, \iota_t) + \lambda\, V_{t+1}(w(q_t, u_t, x_t, \iota_t))\}.$$

Where the functions $y$ and $w$ are described in the previous subsections.

Because $q_t$ is a continuous variable, a discrete approximation must be used for the actual computation. Then the table of values for $V_{t+1}$ is used to compute the values for $V_t$, indexed by $q_t$. For $q_t$ we used 1001 discrete values $(0, 0.001, 0.002, \ldots, 1)$.

As for the cues, $x_t$, as mentioned in section 3, we set the parameters of the distributions to $\mu = 4, \theta_H = 6$, $\theta_E = 9$, and $\sigma^2 = 1$. To discretize $x_t$, we must pick minimum and maximum values, which we set at 0 and 13, respectively. Because there is a tiny area lost at the ends of the distributions beyond $[0, 13]$, we renormalized the total probabilities to 1. Because the optimal threshold given $q_t$ will never be lower than the optimum for the hard environment (when $q_t = 1$), nor higher than the optimum for the easy environment (when $q_t = 0$), precision matters more for the values of $x_t$ within this interval than beyond it. So, to increase computational efficiency, we discretized $x_t$ to intervals of size 0.0125 within $[4.0, 6.2]$ but only to size 0.05 outside of $[4.0, 6.2]$.

The algorithm then gives us two tables: one containing the expected payoff values and the other containing the optimal decision rule functions, or thresholds, $u_t^*(q_t)$, which are indexed by our grid of values for $q_t$. To find the optimal behavior in the limit as the possible lifetime extends towards infinity, the recursion is repeated until the optimal decision rules converge [11]. The algorithm was implemented in python.

Once we have found the optimal decision rule, for each time step we can calculate the expected proportion of the population that has each value of $q_t$ as its estimate. Since the behavioral rule specifies the threshold for each value of $q_t$, we can use the distribution derived above for $q_{t+1} = w(q_t, u_t, x_t, \iota_t)$ to calculate the proportions for time $t+1$ given the proportions for time $t$. This can be done for different probabilities or distributions on $\iota_t$ and $x_t$ than were used to find the optimal behavior, so for the mismatched environments we used $p_m$ instead of $p$ as the marginal probability that $\iota_t = 1$. Because we discretized the $q_t$ values, we rounded the calculations of $q_{t+1}$ to the nearest one thousandth.

# BIBLIOGRAPHY

[1] Suzanne H. Alonzo. Social and coevolutionary feedbacks between mating and parental investment. *Trends in Ecology & Evolution*, 25(2):99–108, February 2010.

[2] Suzanne H. Alonzo. Sexual selection favours male parental care, when females can choose. *Proceedings of the Royal Society of London B: Biological Sciences*, 279(1734):1784–1790, May 2012.

[3] M. B. Andersson. *Sexual Selection.* Princeton University Press, 1994.

[4] Alexandra L. Basolo. Variation between and within the sexes in body size preferences. *Animal Behaviour*, 68(1):75–82, 2004.

[5] Melissa Bateson, Ben Brilot, and Daniel Nettle. Anxiety: an evolutionary approach. *Canadian journal of psychiatry. Revue canadienne de psychiatrie*, 56(12):707–715, 2011.

[6] A. J. Baxter, K. M. Scott, T. Vos, and H. A. Whiteford. Global prevalence of anxiety disorders: a systematic review and meta-regression. *Psychological Medicine*, 43(05):897–910, May 2013.

[7] Meike Bendiks and Matthias Volkmar Kopp. The relationship between advances in understanding the microbiome and the maturing hygiene hypothesis. *Current allergy and asthma reports*, 13(5):487–494, 2013.

[8] C. T. Bergstrom and M. Lachmann. Signalling among relatives. I. Is costly signalling too costly? *Philosophical Transactions of the Royal Society of London, Series B*, 352:609–617, 1997.

[9] C. T. Bergstrom and M. Lachmann. Signalling among relatives. III. Talk is cheap. *Proceedings of the National Academy of Sciences USA*, 95:5100–5105, 1998.

[10] C.T. Bergstrom, S. Számadó, and M. Lachmann. Separating equilibria in continuous signalling games. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, 357(1427):1595–1606, 2002.

[11] Dimitri P. Bertsekas. *Dynamic Programming and Optimal Control.* Athena Scientific, 2005.

[12] Dimitri P. Bertsekas. *Dynamic Programming and Optimal Control: Approximate dynamic programming*. Athena Scientific, 2012.

[13] J. W. Bradbury and S. L. Vehrencamp. *Principles of Animal Communication, Second Edition*. Sinauer Associates, 2011.

[14] J. V. Briskie, P. R. Martin, and T. E. Martin. Nest predation and the evolution of nestling begging calls. *Proceedings of the Royal Society of London B*, 266:2153–2159, 1999.

[15] Nancy Burley. Sexual Selection for Aesthetic Traits in Species with Biparental Care. *The American Naturalist*, 127(4):415–445, April 1986.

[16] Nancy Burley. The Differential-Allocation Hypothesis: An Experimental Test. *The American Naturalist*, 132(5):611–628, November 1988.

[17] Nancy Tyler Burley and Valerie S. Foster. Variation in female choice of mates: condition influences selectivity. *Animal Behaviour*, 72(3):713–719, September 2006.

[18] Oliver H. P. Burman, Richard M. A. Parker, Elizabeth S. Paul, and Michael T. Mendl. Anxiety-induced cognitive bias in non-human animals. *Physiology & Behavior*, 98(3):345–350, September 2009.

[19] P.A. Chiappori and M. Paiella. Relative risk aversion is constant. *Economic Review*, 93(2):383–91, 2008.

[20] Codina Cotar, John M. McNamara, E. J. Collins, and Alasdair I. Houston. Should females prefer to mate with low-quality males? *Journal of Theoretical Biology*, 254(3):561–567, October 2008.

[21] S. T. Emlen and L. W. Oring. Ecology, sexual selection, and the evolution of mating systems. *Science (New York, N.Y.)*, 197(4300):215–223, July 1977.

[22] Tim W Fawcett, Benja Fallenstein, Andrew D Higginson, Alasdair I Houston, Dave EW Mallpress, Pete C Trimmer, and John M McNamara. The evolution of decision rules in complex environments. *Trends in cognitive sciences*, 18(3):153–161, 2014.

[23] Elisabet Forsgren. Female sand gobies prefer good fathers over dominant males. *Proceedings of the Royal Society of London B: Biological Sciences*, 264(1386):1283–1286, September 1997.

[24] Kevin R Foster and Hanna Kokko. The evolution of superstitious and superstition-like behaviour. *Proceedings of the Royal Society B: Biological Sciences*, 276(1654):31–37, 2009.

[25] Willem E Frankenhuis and Karthik Panchanathan. Balancing sampling and specialization: an adaptationist model of incremental development. *Proceedings of the Royal Society of London B: Biological Sciences*, page rspb20110055, 2011.

[26] T. Getty. Reliable signalling need not be a handicap. *Animal Behaviour*, 56:253–255, 1998.

[27] A. Grafen. Biological signals as handicaps. *Journal of theoretical biology*, 144(4):517–546, 1990.

[28] P. E. Greenberg, T. Sisitsky, R. C. Kessler, S. N. Finkelstein, E. R. Berndt, J. R. Davidson, J. C. Ballenger, and A. J. Fyer. The economic burden of anxiety disorders in the 1990s. *The Journal of Clinical Psychiatry*, 60(7):427–435, July 1999.

[29] Michael Gurven and Hillard Kaplan. Longevity among hunter-gatherers: a cross-cultural examination. *Population and Development review*, 33(2):321–365, 2007.

[30] Emma J Harding, Elizabeth S Paul, and Michael Mendl. Animal behaviour: cognitive bias and affective state. *Nature*, 427(6972):312–312, 2004.

[31] D. G. Haskell. The effect of predation on begging-call evolution in nestling wood warblers. *Animal Behaviour*, 57:893–901, 1999.

[32] O. Hasson. Towards a general theory of biological signaling. *Journal of Theoretical Biology*, 185(2):139–156, 1997.

[33] John S. Heywood. Sexual Selection by the Handicap Mechanism. *Evolution*, 43(7):1387–1397, 1989.

[34] Andrea L. Hinds, Erik Z. Woody, Ana Drandic, Louis A. Schmidt, Michael Van Ameringen, Marie Coroneos, and Henry Szechtman. The psychology of potential threat: properties of the security motivation system. *Biological Psychology*, 85(2):331–337, October 2010.

[35] Guy A. Hoelzer. The good parent process of sexual selection. *Animal Behaviour*, 38(6):1067–1078, December 1989.

[36] Josef Hofbauer. The selection mutation equation. *Journal of mathematical biology*, 23(1):41–53, 1985.

[37] Marie-Jeanne Holveck and Katharina Riebel. Low-quality females prefer low-quality males when choosing a mate. *Proceedings of the Royal Society B: Biological Sciences*, 277(1678):153–160, January 2010.

[38] P. L. Hurd. Communication in discrete action-response games. *Journal of Theoretical Biology*, 174(2):217–222, 1995.

[39] Anthony C. Janetos. Strategies of female mate choice: A theoretical analysis. *Behavioral Ecology and Sociobiology*, 7(2):107–112, July 1980.

[40] Dominic DP Johnson, Daniel T Blumstein, James H Fowler, and Martie G Haselton. The evolution of error: error management, cognitive constraints, and adaptive decision-making biases. *Trends in ecology & evolution*, 28(8):474–481, 2013.

[41] Leslie Pack Kaelbling, Michael L Littman, and Andrew W Moore. Reinforcement learning: A survey. *Journal of Artificial Intelligence Research*, 4:237–285, 1996.

[42] Ronald C Kessler, Sergio Aguilar-Gaxiola, Jordi Alonso, Somnath Chatterji, Sing Lee, Johan Ormel, T Bedirhan Üstün, and Philip S Wang. The global burden of mental disorders: an update from the who world mental health (wmh) surveys. *Epidemiologia e psichiatria sociale*, 18(01):23–33, 2009.

[43] Ronald C Kessler, Patricia Berglund, Olga Demler, Robert Jin, Kathleen R Merikangas, and Ellen E Walters. Lifetime prevalence and age-of-onset distributions of dsm-iv disorders in the national comorbidity survey replication. *Archives of general psychiatry*, 62(6):593–602, 2005.

[44] Kessler RC, Chiu W, Demler O, and Walters EE. Prevalence, severity, and comorbidity of 12-month dsm-iv disorders in the national comorbidity survey replication. *Archives of General Psychiatry*, 62(6):617–627, June 2005.

[45] Hanna Kokko. Should advertising parental care be honest? *Proceedings of the Royal Society of London. Series B: Biological Sciences*, 265(1408):1871–1878, October 1998.

[46] Hanna Kokko, Michael D. Jennions, and Robert Brooks. Unifying and Testing Models of Sexual Selection. *Annual Review of Ecology, Evolution, and Systematics*, 37:43–66, December 2006.

[47] D. C. Krakauer and M. Pagel. Spatial structure and the evolution of honest cost-free signalling. *Proceedings of the Royal Society of London B*, 260:365–372, 1995.

[48] M. Lachmann, Sz. Számadó, and C. T. Bergstrom. Cost and conflict in animal signals and human language. *Proceedings of the National Academy of Sciences USA*, 98:13189–13194, 2001.

[49] W. E. Lee, M. E. J. Wadsworth, and M. Hotopf. The protective role of trait anxiety: a longitudinal cohort study. *Psychological Medicine*, null(03):345–351, March 2006.

[50] S. M. Leech and M. L. Leonard. Begging and the risk of predation in nestling birds. *Behavioral Ecology*, 8:644–646, 1997.

[51] Isaac M. Marks and Randolph M. Nesse. Fear and fitness: An evolutionary analysis of anxiety disorders. *Ethology and Sociobiology*, 15(56):247–261, September 1994.

[52] A. Mas-Colell, M.D. Whinston, J.R. Green, et al. *Microeconomic theory*. Oxford university press New York, 1995.

[53] J. Maynard Smith. Must reliable signals always be costly? *Animal Behaviour*, 47:1115–1120, 1994.

[54] J. Maynard Smith and D. Harper. *Animal signaling*. Oxford University Press, 2003.

[55] J. P. McCarty. The energetic cost of begging in nestling passerines. *The Auk*, 113(1):1–275, 1996.

[56] John McNamara and Alasdair Houston. The application of statistical decision theory to animal behaviour. *Journal of Theoretical Biology*, 85(4):673–690, August 1980.

[57] John M McNamara, Tim W Fawcett, and Alasdair I Houston. An adaptive response to uncertainty generates positive and negative contrast effects. *Science*, 340(6136):1084–1086, 2013.

[58] John M McNamara, Lutz Fromhage, Zoltan Barta, and Alasdair I Houston. The optimal coyness game. *Proceedings of the Royal Society of London B: Biological Sciences*, 276(1658):953–960, 2009.

[59] John M McNamara, Pete C Trimmer, Anders Eriksson, James AR Marshall, and Alasdair I Houston. Environmental variability can select for optimism or pessimism. *Ecology letters*, 14(1):58–62, 2011.

[60] Frazer Meacham and Carl T Bergstrom. Adaptive behavior can produce maladaptive anxiety due to individual differences in experience. *Evolution, Medicine, and Public Health*, 2016(1):270–285, 2016.

[61] Michael Mendl, Oliver HP Burman, and Elizabeth S Paul. An integrative and functional framework for the study of animal emotion and mood. *Proceedings of the Royal Society B: Biological Sciences*, 277(1696):2895–2904, 2010.

[62] A. P. Moller and J. T. Nielsen. Differential predation cost of a secondary sexual character: sparrowhawk predation on barn swallows. *Animal Behaviour*, 54:1545–1551, 1997.

[63] Arnstein Mykletun, Ottar Bjerkeset, Simon verland, Martin Prince, Michael Dewey, and Robert Stewart. Levels of anxiety and depression as predictors of mortality: the HUNT study. *The British Journal of Psychiatry*, 195(2):118–125, August 2009.

[64] Randolph M. Nesse. *Why We Get Sick: The New Science of Darwinian Medicine*. Times Books, 1994.

[65] Randolph M Nesse. The smoke detector principle. *Annals of the New York Academy of Sciences*, 935(1):75–85, 2001.

[66] Randolph M Nesse. Maladaptation and natural selection. *The Quarterly review of biology*, 80(1):62–70, 2005.

[67] Randolph M. Nesse. Natural selection and the regulation of defenses: A signal detection analysis of the smoke detector principle. *Evolution and Human Behavior*, 26(1):88–105, January 2005.

[68] Randolph M Nesse and Phoebe C Ellsworth. Evolution, emotions, and emotional disorders. *American Psychologist*, 64(2):129, 2009.

[69] Daniel Nettle and Melissa Bateson. The evolutionary origins of mood and its disorders. *Current Biology*, 22(17):R712–R721, 2012.

[70] Peter J Norton and Esther C Price. A meta-analytic review of adult cognitive-behavioral treatment outcome across the anxiety disorders. *The Journal of nervous and mental disease*, 195(6):521–531, 2007.

[71] H Okada, C Kuhn, H Feillet, and J-F Bach. The hygiene hypothesis for autoimmune and allergic diseases: an update. *Clinical & Experimental Immunology*, 160(1):1–9, 2010.

[72] I. P. F. Owens and I. R. Hartley. Sexual dimorphism in birds: why are there so many different forms of dimorphism? *Proceedings of the Royal Society of London B: Biological Sciences*, 265(1394):397–407, March 1998.

[73] Karthik Panchanathan and Willem E Frankenhuis. The evolution of sensitive periods in a model of incremental development. *Proceedings of the Royal Society of London B: Biological Sciences*, 283(1823):20152439, 2016.

[74] Nicolas Perrini. Signalling, Mating Success and Paternal Investment in Sticklebacks (Gasterosteus Aculeatus): a Theoretical Model. *Behaviour*, 132(13):1037–1057, January 1995.

[75] Anna Qvarnstrm and Elisabet Forsgren. Should females prefer dominant males? *Trends in Ecology & Evolution*, 13(12):498–501, December 1998.

[76] Irja Ida Ratikainen and Hanna Kokko. Differential allocation and compensation: who deserves the silver spoon? *Behavioral Ecology*, 21(1):195–200, January 2010.

[77] L. Real. Search theory and mate choice. I. Models of single-sex discrimination. *American Naturalist*, 136(3):376–405, 1990.

[78] Katharina Riebel, Marie-Jeanne Holveck, Simon Verhulst, and Tim Fawcett. Are high-quality mates always attractive? State-dependent mate preferences in birds and humans. *Communicative & Integrative Biology*, 3(3):271–273, 2010.

[79] Katharina Riebel, Marc Naguib, and Diego Gil. Experimental manipulation of the rearing environment influences adult female zebra finch song preferences. *Animal Behaviour*, 78(6):1397–1404, 2009.

[80] John G. Riley. Silver signals: Twenty-five years of screening and signaling. *Journal of Economic Literature*, 39:432–478, 2001.

[81] A. J. Robson. Efficiency in evolutionary games: Darwin, Nash, and the secret handshake. *Journal of Theoretical Biology*, 144:37–396, 1990.

[82] GAW Rook. 99th dahlem conference on infection, inflammation and chronic inflammatory disorders: Darwinian medicine and the hygieneor old friends hypothesis. *Clinical & Experimental Immunology*, 160(1):70–79, 2010.

[83] Graham AW Rook, Roberta Martinelli, and Laura Rosa Brunet. Innate immune responses to mycobacteria and the downregulation of atopic responses. *Current opinion in allergy and clinical immunology*, 3(5):337–342, 2003.

[84] J. B. Silk, E. Kaldor, and R. Boyd. Cheap talk when interests conflict. *Animal Behaviour*, 59:423–432, 2000.

[85] Julian M Somers, Elliot M Goldner, Paul Waraich, and Lorena Hsu. Prevalence and incidence studies of anxiety disorders: a systematic review of the literature. *Canadian Journal of Psychiatry*, 51(2):100, 2006.

[86] Stephen C. Stearns. *The Evolution of Life Histories.* OUP Oxford, May 1992.

[87] David P Strachan. Hay fever, hygiene, and household size. *BMJ: British Medical Journal*, 299(6710):1259, 1989.

[88] Sz. Számadó. The validity of the handicap principle in discrete action-response games. *Journal of Theoretical Biology*, 198:593–602, 1999.

[89] Sz. Számadó. The cost of honesty and the fallacy of the handicap principle. *Animal Behaviour*, 81:3–10, 2011.

[90] Pete C Trimmer, Sean M Ehlman, John M McNamara, and Andrew Sih. The erroneous signals of detection theory. In *Proc. R. Soc. B*, volume 284, page 20171852. The Royal Society, 2017.

[91] Pete C Trimmer, Andrew D Higginson, Tim W Fawcett, John M McNamara, and Alasdair I Houston. Adaptive learning can result in a failure to profit from good conditions: implications for understanding depression. *Evolution, medicine, and public health*, 2015(1):123–135, 2015.

[92] Mathilde Versini, Pierre-Yves Jeandel, Tomer Bashi, Giorgia Bizzaro, Miri Blank, and Yehuda Shoenfeld. Unraveling the hygiene hypothesis of helminthes and autoimmunity: origins, pathophysiology, and clinical applications. *BMC medicine*, 13(1):81, 2015.

[93] H. Viljugrein. The cost of dishonesty. *Proceedings of the Royal Society of London Series B*, 264(1383):815–821, 1997.

[94] Michael S. Webster. Male Parental Care and Polygyny in Birds. *The American Naturalist*, 137(2):274–280, February 1991.

[95] George C Williams and Randolph M Nesse. The dawn of darwinian medicine. *The Quarterly review of biology*, 66(1):1–22, 1991.

[96] A. Zahavi and A. Zahavi. *The Handicap Principle.* Oxford University Press, Oxford, 1997.

[97] K. J. S. Zollman, C. T. Bergstrom, and S. M. Huttegger. Cheaper-than-costly signaling. *Proceedings of the Royal Society of London B*, 2012 In press.