

**Change and Reliability in the
Evolution of Learning and Memory**

A DISSERTATION
SUBMITTED TO THE FACULTY OF THE GRADUATE SCHOOL
OF THE UNIVERSITY OF MINNESOTA
BY

Aimee Sue Dunlap-Lehtilä

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
OF THE DEGREE OF
DOCTOR OF PHILOSOPHY

David W. Stephens

May 2009

Acknowledgements

I want to thank committee members and others who have contributed to the intellectual development of the work in this dissertation. My advisor Dave Stephens, and committee members Marla Spivak, Celia Gershenson, and Craig Packer all had great comments and thankfully passed me in my exam. I hatched the fly ideas during a semester of talking with Bruce Overmier and Ben Kerr. Mark Bee and Ann Pusey have been very supportive with reviewing drafts and asking questions. And the EEB Behavior Group every Friday has been educational, helpful, and sometimes extremely hilarious. I especially want to thank Holly MacCormick and the 2003 cohort of Emily Wroblewski, Mike Simone, and Meggan Craft.

It is hard to properly thank my advisor, Dave Stephens, his Holiness. We have both experienced a great deal of character building during the past 6 years. Dave mentored me through my research, attempted to thwack me with various projectiles aimed at my head, taught me how to do a roll cast with a fly rod, and literally picked me up off the floor twice. He also tried to drown me in Utah, got me drenched in Wisconsin during a tornado (“No Aimee, it’s not going to rain.”), and repeatedly tried to get me to jump from his moving car and land in a rolling tuck onto the pavement. Despite these memories, I am glad to have Dave as my friend.

Money is important. I was lucky to have plenty of it for my research. For that I must acknowledge a pre-doctoral traineeship from the Center for Cognitive Sciences, multiple mini-grants for research expenses and travel from the Center for Cognitive Sciences, the

Florence Rothman Fellowship, the Frank McKinney Fellowship, the Doctoral Dissertation Fellowship from the Graduate School, a Grant-in-Aid to Dave Stephens from the Graduate School, and multiple summer grants for research expenses and travel from the Department of Ecology, Evolution, and Behavior.

One of the best parts of working on my research was being able to work with a hugely diverse group of undergraduate students. Some of the people helping me out with flies were Tyler Blazey, Will Blanco, Mark Peterson, Nissma Eldeeb, Karrie Kasperbauer, Jenny Miller, Kim Nguyen, Justin Nelson, Sarah Ranallo, Kelly Ryberg, and Nick Stephens. People helping with jays included Heather Oehler, Sam Henly, Rebekah Wilson, Jackie Wentworth, Brooke Gilbert, and Alison Ostdiek. I also thank our two jay lab managers, Carmen Silvers and Matt Scott. I mention that Tyler Blazey gets special thanks for starting the Lab Tape Ball from the masking tape we used to seal the fly cages. The tape ball retired at a diameter of nearly 2 feet.

Support on the home front is important. I had the best support ever. My husband, Juho, helped me with everything from laundry to experimental tasks at night. He retired at the level of Fly Assistant IV, refusing to ever kill a single fly. Juho carried me home when I couldn't walk anymore and learned to make meatloaf and roast a turkey. Juho onhan maailman vahvin nalle. I also want to thank Milja, who is no longer with us, for helping me get to bed on time every night, and for keeping me company during all those horrible months of bedridden illness, after the fruit flies tried to kill me off. And finally, I thank my family for being patient while we missed holidays and trips to the Keys, the Bahamas, Disney, and all kinds of other fun places, and for being so very supportive throughout the entire process.

Abstract

Why do animals learn to perform some behaviors while others are innate? Why do animals learn some things more easily than others? And, why do animals remember some things better than others? Theoreticians argue that patterns of environmental change explain these patterns, but we have little data to support these claims. I used statistical decision theory to model behaviors and fitness consequences, and experimental evolution studies with fruit flies where I manipulated patterns of environmental change across evolutionary time, to address the first two of these fundamental questions about the evolution of learning. The first experiment tested the effects of the reliability of experience and the fixity of the best action upon the evolution of learning and non-learning across 30 generations. I found that indeed, the interaction of these two variables determined when learning, and when non-learning evolved. The second study was a full factorial experiment manipulating the reliabilities of two modes of stimuli: olfactory and visual. After 40 generations, I found that as predicted, flies in environments where olfactory stimuli are reliable learned better about olfactory than color stimuli, with the same being true for color stimuli. Finally, I addressed the question of why animals remember some things better than others using a dynamic programming technique and experiment with blue jays, finding interactions between rates of change and time. These novel studies show the importance of reliability and change in evolution of learning and memory.

Table of Contents

Acknowledgements	i
Abstract	iii
List of Tables	viii
List of Figures	ix
Introduction	1
Chapter 1 – Components of change in the evolution of learning and non-learning.....	5
Introduction.....	5
The Components of Change Model.....	8
The parameters p and q	8
Model predictions: learning vs. non-learning	9
Experimental Approach and Methods.....	11
Initial populations and Treatment assignments	12
Implementing the experimental selection regimes	13
End of Selection Assays	14
Results.....	14
Evolution of non-learning (preference).....	14
Evolution of learning	17
General Selection Results	19

Discussion.....	19
Summary.....	22
Chapter 2 – Stimulus reliability in the evolution of “prepared” learning	23
Introduction.....	23
Model & Predictions	25
Methods	28
Flies and Husbandry.....	28
Aversion Learning & Selection on Populations	28
Treatments and Lines	29
Results.....	30
Selection Data	30
End of Selection Assays	33
Discussion.....	37
Summary.....	39
Chapter 3 – Why some memories do not last a lifetime: dynamic long-term retrieval in changing environments	40
Introduction.....	40
The Model	43
Results.....	49
Proxy for Optimal Memory.....	51

Graphical Approach	55
Convergence.....	58
Discussion.....	61
Chapter 4 – Choice in a variable environment: an experimental test of tracking, sampling, and memory length	67
Introduction.....	67
Model	69
Materials & Methods.....	74
Subjects & Experiment Apparatus.....	74
Basic Experimental Design.....	76
Testing a Single Retention Interval	77
Dependent Measures	79
Results.....	80
Sampling the Environment.....	80
Tracking Change in the Environment With Learning	81
Memory and Choice of the Last Experienced Stimulus.....	81
Balancing Recent with Past Information.....	85
Discussion.....	88
Literature Cited.....	92
Appendix 1 – Modeling the Evolution of Learning versus Non-Learning.....	99

Appendix 2 – Modeling Prepared Learning to Color Versus Odor	102
Appendix 3 – Derivation of the Dynamic Programming Recursion.....	107
Appendix 4 – The Proxy for Optimal Memory	113
Appendix 5 – Extensions and Explanations of the Dynamic Programming Model for Optimal Memory.....	116
5A: The Effect of Stimulus Change on Optimal Memory	116
5B: Exploration of the Cost of Living.....	117
5C: Concavity and Jensen’s Inequality	121
5D: Iteroparity	123
5E : Explicit Costs to Memory	124

List of Tables

Table 3-1. Table of parameters, variables, and functions used in the model.	48
Table 4-1. Possible outcomes concerning the state variable.....	70
Table 4-2. Events observable by the subject.....	71
Table A1-1. Contingencies and fitnesses for learners and non-learners.....	100
Table A2-1. Table of all possible outcomes in the consequence phase.....	103
Table A2-2. Possible pairings and outcomes in the two stimulus modality system.....	106

List of Figures

Figure 1-1. Best-action fixity and reliability of experience influence the fitness value of learning. Natural selection favors learning most strongly when the reliability of experience is high, but best-action fixity is low (Point D); selection favors non-learning when best-action fixity is high, but reliability is low (Point A). We test points A. and D. 11

Figure 1-2. Interaction of treatment by initial best assigned environment during following the 30th generation of selection. Error bars are 95% confidence intervals.16

Figure 1-3. Significant effect of treatment during the final experiment assays following the 30th generation of selection. Error bars are 95% confidence intervals. The difference between reliable/unpredictable and the other treatments is statistically significant.18

Figure 2-1. Assume the best environment is indicated by the box with the dashed lines. In the first possibility, both A and α reliably predict the best environment when paired with quinine. In the second possibility, only the pairing of α with quinine predicts the best environment; an individual attending to/ learning the A pairing would be making an incorrect choice. There are two additional possibilities: that A, but not α predicts, and that neither predict. 26

Figure 2-2. Predictions from the model. Whenever the probability that odor predicts is greater than the probability that color predicts ($O > C$), then learning about odor is favored. We tested each the points at the four corners of the graph. Two of these points fall on the line between color learning favored and odor learning favored, and

thus either or both could be favored. Intuitively, we predict that when both modalities of stimuli predict equally well, learning about both should be favored, while learning about neither should be favored when neither modality of stimuli predict the best environment. 27

Figure 2-3. Selection data of following the quinine pairing with odor. The x-axis represents means of 2-generation blocks (consistent with the randomization scheme of the experiment), taken at 10 generation increments. Error bars are standard errors. .31

Figure 2-4. Selection data of following the quinine plus color cues. Error bars are standard errors. The x-axis represents means of 2-generation blocks (consistent with the randomization scheme of the experiment), taken at 10 generation increments..... 32

Figure 2-5. This figure presents the data in the factorial form of the experiment's design. Learning about color was enhanced when color was a reliable predictor of the best environment across evolutionary time. The effect of color reliability was statistically significant ($F_{1,36}=4.189$, $P=.048$). The effect of odor reliability was NS, and did not differ across color reliability levels. The interaction between color and odor reliability nears significance ($F_{1,36}=3.435$, $P=0.072$)..... 34

Figure 2-6. Learning to odor alone is enhanced in both treatments where odor is reliable. The effect of odor reliable is statistically significant ($F_{1,36}=7.8829$; $P=0.008$), neither the effect of color reliable nor the interaction between color and odor reliability are statistically significant. 35

Figure 2-7. The same data as in Figures 2-5 and 2-6 (end of experiment assay data) presented together in a single graph to demonstrate the relationship between learning the two sensory modes of stimuli. As expected, in each treatment where

odor reliably predicted the best environment, the flies learned better to odor than to color. Only the lines of flies in the color-reliable treatment showed better learning to color than to odor. Error bars are standard errors..... 36

Figure 3-1. Optimal memory $m^*(x,t)$ with $X=50$, $N=10$, $T=40$, $\delta=0.1$, $\kappa=12$, $\pi_i=5$, $\pi_u=10$, and $\pi_c=15$. Each surface corresponds to a different value of the rate of stimulus change (ρ) and either a convex final fitness function $\Phi(x)=100(x/X)^4$ or a concave final fitness function $\Phi(x)=100(x/X)^{0.25}$ 50

Figure 3-2. The probability of a correct response to a familiar stimulus (blue/medium gray bars), the probability of a response to an unfamiliar stimulus (green/light gray bars), and the probability of an incorrect response to a familiar stimulus (red/dark gray bars) as a function of memory size (with $N=10$). Each chart corresponds to a different value of the rate of stimulus change, ρ . These probabilities are $\omega(1-\beta^m)$, α^m , $1-\alpha^m-\omega(1-\beta^m)$, respectively. 54

Figure 3-3. Schematics to show the costs and benefits of memory. (a) We start by assuming $F(x, t+1, T)$ is a linear function (although generally it is not). We assume that the state of our organism at age t is x' . A solid red arrow points to $i(x')$, the state acquired through an incorrect response to a familiar stimulus. A dotted green line and arrow points to $u(x')$, the state acquired through a response to an unfamiliar stimulus. A dashed blue line and arrow points to $c(x')$, the state acquired through a correct response to a familiar stimulus. The potential benefit $B(x',t)$ of being familiar with a stimulus is $F(c(x'),t+1,T)-F(u(x'),t+1,T)$ and potential cost $C(x',t)$ of being familiar with a stimulus is $F(u(x'),t+1,T)-F(i(x'),t+1,T)$. (b) If π_u decreases, then $u(x')$ can decrease and B can increase while C decreases. (c) If the state of the organism is very low and the cost of living is very high ($\kappa > \pi_u$), then $i(x')=u(x')=0$

while $c(x') > 0$. Consequently $C = 0$ and $B > 0$; that is, there are no costs to memory, whereas there are benefits—thus, the memory window should be maximally long.

(d) If the cost of living is low ($\kappa < \pi_u$, see Appendix D), then $C > 0$ and $B > 0$ and maximal memory may not be favored. (e) If $F(x, t+1, T)$ is convex, the ratio of B to C increases (relative to the linear F function in (a)) and thus an increased memory window is favored. (f) If $F(x, t+1, T)$ is concave, the ratio of B to C decreases (relative to the linear function in (a)) and thus a decreased memory window is favored..... 56

Figure 3-4. (a) The final fitness function $\Phi(x) = 100(x/X)^4$ is given in grey and $\Phi(x) = 100(x/X)^{0.25}$ is given in black. These functions are normalized to range from 0 to 1. (b) The corresponding renormalized $F(x, t, T)$ functions at time $t = 20$, generated through backwards iteration of equation [1] (with $X = 50$, $N = 10$, $T = 40$, $\delta = 0.1$, $\pi_i = 5$, $\pi_u = 10$, $\pi_c = 15$, $\kappa = 12$, and $\rho = 0.1$) starting with $F(x, T, T) = 100(x/X)^4$ (in grey) and $F(x, T, T) = 100(x/X)^{0.25}$ (in black). (c) The corresponding renormalized $F(x, t, T)$ functions at $t = 5$. Both functions are approaching an equilibrium function $H(x)$ given in black dots in (a)-(c). (d) Using $H(x)$ in place of $F(x, t+1, T)$ we calculate $m_H(x, t)$ by equation [7] for all x and t values. Here we show the difference between m_H and m^* to show the appropriateness of the m_H approximation. We see two “types” of deviations between m_H and m^* . The constraint deviations occur where the m_H approximation is outside of the allowed range for memory (e.g. $m_H(x, t) > t$) and the signature deviations occur when m_H is near attainable memory values, but other memory values are optimal (these deviations are due to the fact that $H(x)$ is an inappropriate approximation for $F(x, t, T)$). In the case shown, the signature deviations ($m_H - m^*$) are positive, suggesting that there is optimal decline in memory

for large values of t (note that we are dealing with the concave final fitness function shown in black in (a)).	60
Figure 4-1. The effect of increasing change (lower values of q) is to depress the value of memory.	73
Figure 4-2. Results for $q=0.98186$ and $\epsilon=0.01$. We chose the q value based upon the rate of change in the experiment in this paper. In 2b, the sharpest V-shape represents the beginning time step, with each of the next time steps following in order with more and more shallow V-shapes.	74
Figure 4-3. The experimental apparatus.	75
Figure 4-4. Design of the experimental trials. Birds experience variability in the form of stimulus reward reversals. In the low change treatment, birds experience stimulus A correct for five hours, then stimulus B is correct for the next five elapsed hours of trials. In the high change treatment, stimulus A and B reverse reward on average every 30 min. Stimuli rewards reverse for a final time in both treatments during the Day 2, and this final learning set continues until the bird reaches a criterion of accuracy. Birds then began a simple task following white lights and we inserted a single probe trial to test one of 5 retention intervals. At the end of the experimental day of the probe trial, the bird began the process again with a new stimuli pair.	77
Figure 4-5. The mean performance of birds when making choices relative to the last stimulus they experienced as correct, over five retention intervals. Each set of birds was exposed to different rates of environmental change (low or high). Significant interactions were observed between the low and high change treatments at 24	

hours, and within the high change treatment between 4 and 24 hours. The bars reflect the standard deviations. 84

Figure 4-6. The deviation from expected for birds choosing according to the stimulus that was correct after the first day of learning, when that stimulus conflicted with the most recently correct stimulus. Values above zero on the Y-axis are tend towards choosing the overnight stimulus (past information) more than expected by chance, while negative values are moving towards choosing the most recently-rewarded stimulus (recent information) more than expected by chance. 87

Figure A5-1. This figure corresponds to Figure 1 except $\kappa=8$ here (and $X=50$, $N=10$, $T=40$, $\delta=0.1$, $\pi_i=5$, $\pi_u=10$, and $\pi_c=15$). Each surface gives optimal memory $m^*(x,t)$ and corresponds to a different value of the rate of stimulus change (ρ) and either a convex final fitness function $\Phi(x)=100(x/X)^4$ or a concave final fitness function $\Phi(x)=100(x/X)^{0.25}$119

Figure A5-2. The arrows and terminology in this figure correspond to Figure 3. (a) If the state of the organism is very high and the cost of living is low ($\kappa < \pi_u$), then $u(x')=c(x')=X$ while $i(x') < X$. Consequently $C > 0$ and $B = 0$; that is, there are no benefits to memory, whereas there are costs—thus, the organism should have no memory. (b) If the cost of living is high ($\kappa > \pi_u$), then $C > 0$ and $B > 0$ and some memory may be favored.....121

Figure A5-3. This surface gives optimal memory $m^*(x,t)$ as a function of age (t) and state (x) under the iteroparous extension of Appendix 5D. The other parameters are $X=50$, $N=10$, $T=40$, $\delta=0.1$, $\pi_i=5$, $\pi_u=10$, $\pi_c=15$, $\kappa=12$, $\rho=0.1$, and $t_{RM}=20$. The fitness function is concave: $\Phi(x)=100(x/X)^{0.25}$125

Figure A5-4. This surface gives optimal memory $m^*(x,t)$ as a function of age (t) and state (x) when memory is explicitly costly as laid out above. The other parameters are $X=50$, $N=10$, $T=40$, $\delta=0.1$, $\pi_i=10$, $\pi_u=5$, $\pi_c=15$, $\rho=0.1$, $\kappa_0=4$ and $\kappa_1=0.15$. The final fitness function is concave: $\Phi(x)=100(x/X)^{0.25}$126

Introduction

Animals are not omniscient. It is a rare situation when an individual has full knowledge of its environment and of the outcomes of its potential decisions, and thus always being able to choose the optimal behavior. However, animals can reduce uncertainty about their world by acquiring information. Animals can use two types of information: information gained genetically and inherited from successful ancestors, and information gained from the environment through experience. When should an individual rely on experience and when should it come into the world prepared with the appropriate behavior? An animal may gain experience to decrease its uncertainty about the world, but the experience itself becomes unreliable with additional change in the world, requiring the acquisition of more information to update its estimates. But, reducing uncertainty by gaining new information is not always the best option. If the costs of gaining new information outweigh the costs of acting without it, then acting without further information is best. When and how should animals acquire new information, and how should this information guide their future behavior? In this dissertation, I address these basic questions from an evolutionary ecological approach.

In chapter 1, I address the problem of when animals should use a genetically fixed response or a learned response to guide their behavior. When should learning evolve, and when should non-learning evolve? Learning is thought of primarily as an adaptation to a changing environment and evolutionary treatments of learning have stressed the importance of environmental variability. But theorists disagree on the exact role of that

variability. Too much variability reduces the selective value of learning because the state of the environment may have no relation to the state of the environment tomorrow (e.g. Staddon, 1983). However, in an environment that never changes, we might not expect learning to evolve because a genetically fixed behavior can meet the challenges of a static environment without the opportunity costs or risks associated with learning (e.g. Johnston, 1982; Papaj and Prokopy, 1989). A third view resolves this problem by arguing that intermediate levels of change promote the evolution of learning (e.g. Slobodkin and Rapoport, 1974), something Kerr & Feldman have coined the “Goldilocks principle” (2003). Stephens (1991; 1993) provided a more nuanced view, showing that the pattern of change is what matters. Moving from theory to experiment poses challenges. First, “change” as a principle is rather vague. By breaking change up into well-defined components, we can use statistical decision theory to create very specific predictions of when animals should use genetically or experientially acquired information. I present a model and an experimental evolution study with fruit flies (*Drosophila melanogaster*), testing how two components of change, the fixity of the best action to take and the reliability of experience, interact to favor the evolution of enhanced learning in some situations, and the evolution of non-learning in other situations.

In chapter 2, I address the interaction between inherited and experienced information to explain why animals are more prepared to learn some things over others. In a large number of studies, animals appear to learn about some relationships better than others (e.g. LoLordo, 1979; Seligman, 1970; Shettleworth, 1972). There have been a number of attempts to explain these phenomena, but many questions about the evolution of prepared learning remain. We can generally agree that learning should enable animals to predict when biologically significant events will occur. Throughout the evolutionary

history of a given organism, some aspects of the environment will have more reliably indicated fitness consequences than others. The most likely hypothesis for the evolution of prepared learning is that some stimuli pairings have remained reliable throughout the lineage of a given organism, and natural selection has, therefore, favored learning of these types of associations. In chapter 2, I describe a simple model and an experimental evolution study with *D. melanogaster*, in which I test predictions about the evolution of prepared learning based upon patterns of stimulus reliability.

With chapters 3 and 4, I consider patterns of change in the environment on a smaller scale: how they affect how animals acquire experience and manage the information they gain in multiple iterations over the course of their lifetimes. In chapter 3, I focus on how long an animal should retain information in a theoretical treatment of adaptive memory length. In this chapter I describe a dynamic programming model addressing the role of change in optimal memory length across the lifespan. There is a general consensus that remembering everything forever is not ideal: there are physiological costs as well as costs with potential interference with the recall of memories. With limited resources, what information should animals retain, and for how long? Here the question of what makes information valuable becomes important: more reliable information is more valuable because it has greater predictive power for future events (Stephens, 1989). The reliability of information relates to the pattern of change: when cues in the environment vary, they become unreliable. How long learning lasts—memory length—should track the patterns of change in the environment.

While chapter 3 looks solely at memory length, in chapter 4 I address the entire learning process, from acquisition to application. How should animals track changes in the

environment? A family of tracking models provides guidance in what to expect (Stephens, 1987; Stephens and Krebs, 1986). A key aspect of this work is that reducing uncertainty is not necessarily always the best action. Animals can sample the environment to acquire new information, but this sampling is not the optimal behavior in every case. Thus the task is more a matter of managing uncertainty than reducing it at all costs. As rates of change in the environment increase or decrease, when should animals sample for new information, and when should they change their behavior to track changes indicated by that information? While the theory is well defined, only a few empirical studies have manipulated change itself. I present a model and an experiment testing how blue jays (*Cyanocitta cristata*) should sample, track and remember in environments that differ in their rates of change. This experiment addresses the role of change throughout the acquisition, retention, and application of experienced information.

Chapter 1 — Components of change in the evolution of learning and non-learning¹

Introduction

In a well-known study, John Garcia and his colleagues showed that rats learned an association between a sweet tasting water and gastric illness much more readily than they learned an analogous association between bright-noisy water and gastric illness (Garcia and Koelling, 1966). In the 20 year period following Garcia et al's influential work, investigators identified many similar learning phenomena, variously called 'preparedness,' 'selective association' or 'biological constraints on learning' (e.g. LoLordo, 1979; Seligman, 1970; Shettleworth, 1972). Taken together these observations suggested that natural selection has shaped the properties of animal learning. In recent years, however, interest in this topic has waned, in part because the key explanatory variables seemed experimentally inaccessible.

To understand the variables that have seemed so problematic, consider why learning may make more sense in some situations than others. Although many models have addressed this question, they all focus on the nature of statistical relationships between stimuli and consequences in the animal's environment (Kerr and Feldman, 2003; Papaj, 1994; Stephens, 1991; Stephens, 1993). The simplest and oldest of these ideas focuses on change. Learning, the argument goes, exists because environments change and it follows

¹ This chapter is currently under review as Dunlap, A.S. & Stephens, D.W. Components of change in the evolution of learning and non-learning.

that animals must use experience to adjust to this change (e.g. Johnston, 1982; Papaj and Prokopy, 1989; Plotkin and Odling-Smee, 1979; Thorpe, 1963). This logic leads to a simple claim that we call the ‘learning folk theorem:’ changing environments favor learning, but stable-fixed environments favor non-learning (innate or fixed behavior) (e.g. Anderson, 1995; Mery and Kawecki, 2004a; Plotkin and Odling-Smee, 1979).

While the folk theorem continues to influence the thinking of casual students of learning evolution, recent models argue that it oversimplifies the problem (Borenstein et al., 2008; Kerr and Feldman, 2003; Stephens, 1991). According to these models the folk theorem mistakenly lumps all forms of environmental change together, when in reality—these models argue—some components of environmental change select for learning (as the folk theorem suggests) but others select against learning. Regardless of which model one favors, controlling or observing the statistical relationships in an animal’s environment presents a significant empirical challenge. This paper develops a simple, experimentally tractable ‘components of change’ model, and presents results from a study testing this model experimentally. This experiment varies relevant components of ‘environmental change’ across many generations and it offers the first experimental confirmation of the claim that some types of change promote learning while others promote non-learning.

To develop our model and test its predictions we focus on an experimental preparation developed by Mery and Kawecki (2002; 2004a). In this preparation, the experimenter presents two types of egg-laying media to a small group of female *Drosophila melanogaster*, one option is flavored with orange juice and the other is flavored with pineapple. The experimenter offers this choice twice. In the first presentation the

investigator pairs one of the media types with the aversive chemical quinine, so the flies experience either (i) orange plus quinine vs. pineapple without quinine, or (ii) orange without quinine vs. pineapple plus quinine. We call this first presentation the ‘experience’ phase, because flies experience the pairing with quinine at this stage. In the second presentation the experimenter again offers the orange and pineapple media, but now without quinine in either media type. We call this second presentation the consequence phase, because—as we will explain—this is when the investigator imposes a fitness consequence. Notice that this arrangement creates a relatively simple aversion learning problem in which a ‘learning’ fly can use the experience of quinine-pairing in the first phase (hence the name ‘experience’ phase) to adjust its egg-laying behavior in the second phase.

Using this preparation an experimenter can control two variables that, according to our components of change model, influence the fitness value of aversion learning: the reliability of experience and the fixity of the best action. First, the experimenter creates the next generation of flies by rearing eggs from one of the media types in the consequence phase (eggs laid in the experience phase are discarded). The investigator can, for example, create an ‘orange best’ situation by rearing only those eggs laid on the orange media. Second, the experimenter can control the extent to which the quinine pairing in the experience phase reliably indicates the best action in the consequence phase.

Mery and Kawecki used this preparation to test the learning folk theorem. They created a changing environment that should favor learning by alternating orange-best (rear eggs only from orange) and pineapple best from one generation to the next. In addition,

pairing with quinine in the experience phase reliably indicated the media type that flies should avoid in the consequence phase. In agreement with the folk theorem, Mery and Kawecki found enhanced learning (that is increased sensitivity to the experience of quinine pairing) in this changing environment. In a second study, Mery and Kawecki created a fixed environment in which they always reared eggs (for example) from the orange media type in the consequence phase. Contrary to expectations of the folk theorem, Mery and Kawecki also found increased responsiveness to experience in this condition. Here, again, pairing with quinine in the experience phase reliably indicated the best action in the consequence phase. As we will explain below, according to the ‘components of change’ view of learning Mery and Kawecki’s ‘fixed environment’ situation did not, in fact, select against learning. This is because while fixity of the best action does select against learning as the folk theorem claims, fixity of the relationship between experience and consequence favors learning.

The Components of Change Model

The parameters p and q

Here, we develop a model based on the Mery-Kawecki preparation (the appendix presents the algebraic details). Let p represent the overall probability that the experimenter rears eggs from the orange flavored media (so that laying eggs on orange is the best action). We focus on ‘orange’ to simplify the model development; focusing on pineapple yields identical results. The parameter p ($0.5 \leq p \leq 1.0$) specifies the fixity of the best action, and we call it the *best-action fixity*. This is our first component of change. For example, $p=1.0$ gives the highest possible best-action fixity because it means that the experimenter always rears eggs from orange, and ‘lay eggs on orange’ is always

the best policy. In contrast, $p=0.5$ gives the lowest meaningful value of best action fixity because it means that laying eggs on orange is the best half the time and laying on pineapple is the best half the time. Similarly, we use the parameter q to represent the fixity of the relationship between experience and best action. Mathematically, q is the conditional probability that the experimenter rears eggs from the substrate type that was NOT paired with quinine in the first or 'experience' phase of the experiment. The parameter q ($0.5 \leq q \leq 1.0$) therefore measures the fixity of the relationship between experience and the best action. As before we simplify the terminology by calling this variable the *reliability of experience*, which is our second component of change. If $q=1$, the flies can reliably select the best action by avoiding the substrate that was paired with quinine in the experience phase; however if $q=0.5$ pairing with quinine carries no information about the fitness consequences of egg-laying choices in the second (or consequence) phase of the experiment.

Model predictions: learning vs. non-learning

To evaluate the effects of these parameters we compare the fitness of a non-learning genotype that always lays eggs on orange (since we have arbitrarily assumed that 'lay on orange' is the most common best action, i.e. $1/2 < p < 1.0$) to the fitness of a learning genotype that uses the pairing with quinine in the experience phase to guide its behavior in the 'consequence' phase. Figure 1-1 shows the results of these calculations. The figure shows how the two fixity parameters influence the relative fitness obtained by our learning and non-learning genotypes. As the figure shows, a diagonal line [running from $(p,q)=(1/2,1/2)$ to $(1, 1)$] separates the learning and non-learning regions; the learning genotype does better above the line while the non-learning genotype should prevail below the line. For example, the point where $p=0.5$ and $q=1.0$ strongly promotes

learning because the best action changes randomly (i.e., there is low best-action fixity, $p=0.5$), yet the experience of quinine-pairing reliably signals the best action (i.e. there is high reliability of experience, $q=1.0$). This crudely corresponds to Mery and Kawecki's (2002) 'changing environment' study, in which they found that learning abilities improved within 20 generations of selection. Notice however that the condition that most strongly selects against learning is the point where there is high best action fixity ($p=1.0$) and low reliability of experience ($q=0.5$). Mery and Kawecki did not test this situation. Instead they tested the 'completely fixed' case (i.e., $p=q=1.0$), which following the 'folk theorem' they argued should select against learning. However, as the figure shows this situation is actually selectively neutral (see Appendix 1 for mathematical rationale). In the absence of a learning cost, complete fixity neither favors nor disfavors our learning genotype.

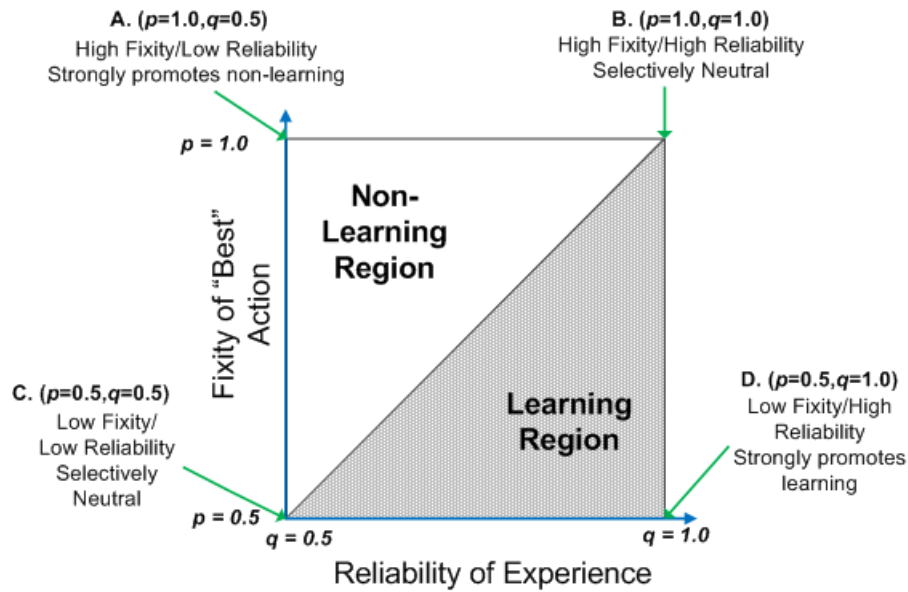


Figure 1-1. Best-action fixity and reliability of experience influence the fitness value of learning. Natural selection favors learning most strongly when the reliability of experience is high, but best-action fixity is low (Point D); selection favors non-learning when best-action fixity is high, but reliability is low (Point A). We test points A. and D.

Experimental Approach and Methods

The experiment presented here compares selection in the two regimes that most strongly favor and disfavor learning. We assigned small populations of *Drosophila melanogaster* to three conditions: i) High best-action fixity ($p=1.0$), low reliability of experience ($q=0.5$). This strongly disfavors learning because experience is unreliable and the same action is always best; ii) Low best-action fixity ($p=0.5$) and high reliability of experience ($q=1.0$). This strongly favors learning because sensitivity to the experience of quinine

pairing allows flies to consistently track the best action; iii) In addition, we established control populations with the same initial population size and rearing procedures as the two experimental groups, but these flies never experienced the fruit-flavored media or quinine. Notice that one can, in principle, fix the best action in two ways: orange always best or pineapple always best. We included both possibilities in our design by randomly assigning half of the populations in each treatment to ‘orange initially’ best and half to ‘pineapple initially’ best conditions. Although we made this assignment for all three treatments, it has different implications for each of the three treatments. For the ‘high best action fixity’ treatment it fixes the best action (lay on orange or lay on pineapple) across all thirty generations; for the ‘low best action fixity’ it determines the initial state but this changes randomly on subsequent generations; for the control lines—which never experienced orange or pineapple—it is simply an arbitrary designation.

Initial populations and Treatment assignments

To create our initial stock of flies, we mixed 400 males and 400 females from each of four lab-adapted, wild-caught populations from Minnesota and Wisconsin (USA). We maintained them in overlapping generations in a large population cage for five months prior to the start of the experiment. We housed all flies at 24C. We reared all eggs at a density of 80 eggs per vial, with six vials per line per generation. We established 36 lines of 400 flies and randomly assigned twelve lines to each of three treatments. For each treatment we randomly assigned six of the twelve lines to ‘orange best’ and six to ‘pineapple best.’ As outlined above, our three selection treatments are: (1) best action fixed ($p=1.0$)/experience unreliable ($q=0.5$). (2) ‘best action changing ($p=0.5$)/experience reliable ($q=1.0$)’ treatment. (3) Control.

Implementing the experimental selection regimes

Every generation, we transferred 200 female flies (along with a similar number of males) from each line to a test cage. The cages were approximately shoe box size (33.3 cm L x 21 cm W x 12 cm H), and we equipped each of them with a sliding drawer that could hold two petri dishes. We presented petri dishes with standard cornmeal and molasses media until we were ready to begin the selection (3 days). As the introduction explains, the experimental selection regime consists of two phases: an experience phase (in which we paired quinine with one type of media) and a consequence phase (where quinine was never present). The experience phase exposed flies to two fruit flavors of agar-based media in a single three-hour session (reconstituted frozen orange or pineapple juice, 12 g agar/1 L juice, with 20 mL of juice agar placed in the bottom of each 100 mm x 15 mm petri dish). Following our experimentally-determined schedule we paired quinine with one of the two flavors (4 g quinine/1 L agar). In the consequence phase, we presented fresh petri dishes of the two flavors of media (using the sliding drawer to change the media). We randomized the locations of orange and pineapple plates within each cage, but kept the location the same in the experience and consequence phases for a given line in a given generation. An interval of 30 min separated the removal of the experience phase plates and the introduction of the consequence phase plates.

Following an experimentally-determined schedule we reared eggs laid on only one of the media flavors in the consequence phase, and we discarded all other eggs. We removed eggs selected for propagation from the substrate using a needle and placed them in vials on standard cornmeal-based fly food for incubation.

End of Selection Assays

Following 30 generations of selections, we tested each line in a series of assays. We reared the flies used in these assays from eggs collected on standard (unflavored) media. We conducted two types of assays: learning assays and preference (or non-learning) assays. Our learning assays consisted of two tests. First, we tested a group of 200 naïve females (with a similar number of males) from each line with a 3 hr experience phase of quinine paired with orange. Second, we tested a different group of naïve flies with quinine paired with pineapple. We followed both with a 1.5 hr consequence phase in which neither flavor was paired with quinine. (Note that we use the terminology “experience phase” and “consequence phase” for simplicity here even though these assays differ in some details from the experimental selection procedures). Our preference assay tested flies with no quinine present during a 3 hr experience phase, and no quinine present during the 1.5 hr consequence phase.

Results

Evolution of non-learning (preference)

We tested the effect of our three selection regimes on unlearned preferences by simultaneously presenting orange and pineapple flavored media to naïve groups of flies from each of our treatments, and observing the number of eggs laid on each type of media. Figure 1-2 shows these data expressed as the proportional preference for the orange media. As our model predicts, the figure shows changes in unlearned preferences for the ‘best action fixed ($p=1.0$)/experience unreliable ($q=0.5$)’ lines, but not for the ‘best action changing ($p=0.5$)/experience reliable ($q=1.0$)’ treatment. Focusing on the ‘best action fixed/experience unreliable’ lines, we see a striking difference between the

lines assigned to the orange-best and pineapple-best conditions. As we would expect, lines assigned to the pineapple-best treatments showed a decreased preference for orange. An analysis of variance of all three selection regimes supports this interpretation by showing a significant interaction between treatment and best assignment ($F_{2,30}=3.381$, $P=0.0474$). In addition, post hoc analyses show a difference between the orange-best and pineapple best lines for the 'best action fixed/experience unreliable' treatments but not for the other treatments.

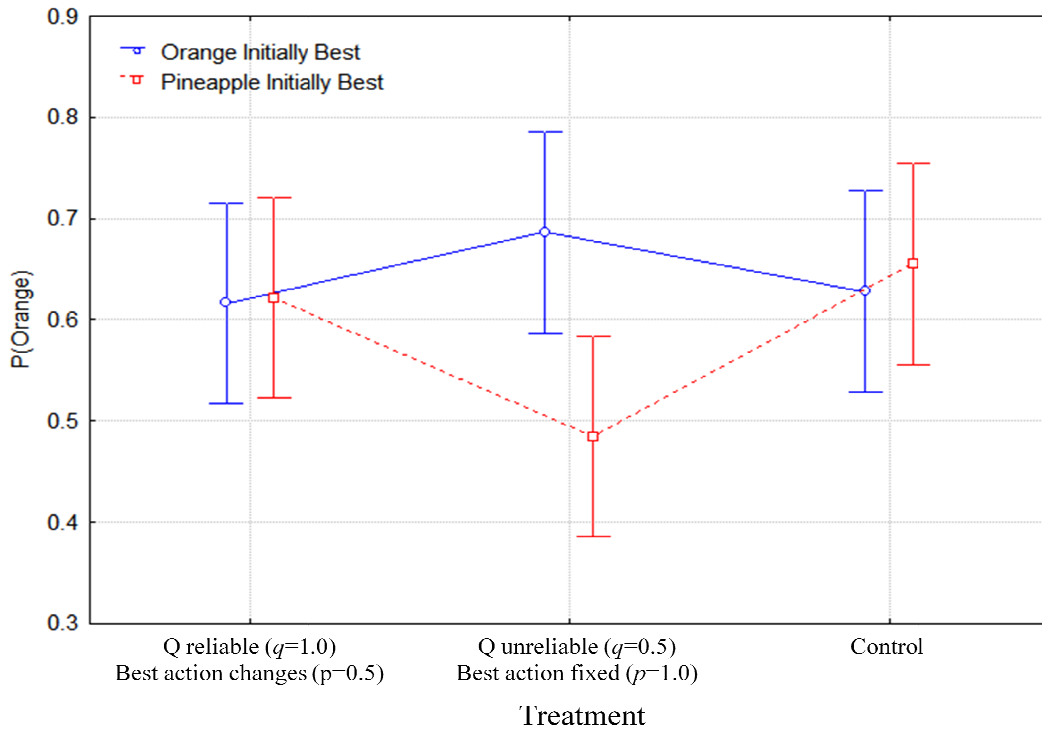


Figure 1-2. Interaction of treatment by initial best assigned environment during following the 30th generation of selection. Error bars are 95% confidence intervals.

Evolution of learning

To assess differences in sensitivity to experience, we exposed groups of flies from each of our treatments to an assay that closely paralleled our selection procedures. In this assay, we paired quinine with either orange or pineapple, and then scored oviposition preferences in a second presentation of the two types of media without quinine. By testing a separate group of flies from each line in both an orange paired with quinine and a pineapple paired with quinine condition, we can derive a contingency score for each line using Cramer's Phi. This score measures the extent to which oviposition preferences in the second stage of the assay depend on the experience of quinine pairing in the first stage. Figure 1-3 shows these data. As our model predicts, we see enhanced sensitivity to experience in the 'best action changing ($p=0.5$)/experience reliable ($q=1.0$)' treatment (compared to the control group) and no difference in sensitivity to experience between the 'best action fixed/experience unreliable' treatment and the control. A one-way analysis of variance confirms a significant effect of treatment ($F_{2,33}=4.17, P=0.02$). In addition, post hoc analyses (using Tukey's LSD) confirm the pattern shown in the figure. Specifically, the 'best action changing ($p=0.5$)/experience reliable' ($q=1.0$) treatment shows a higher sensitivity to experience (as measured by phi) than either the control or 'best action fixed ($p=1.0$)/experience unreliable ($q=0.5$)' treatments.

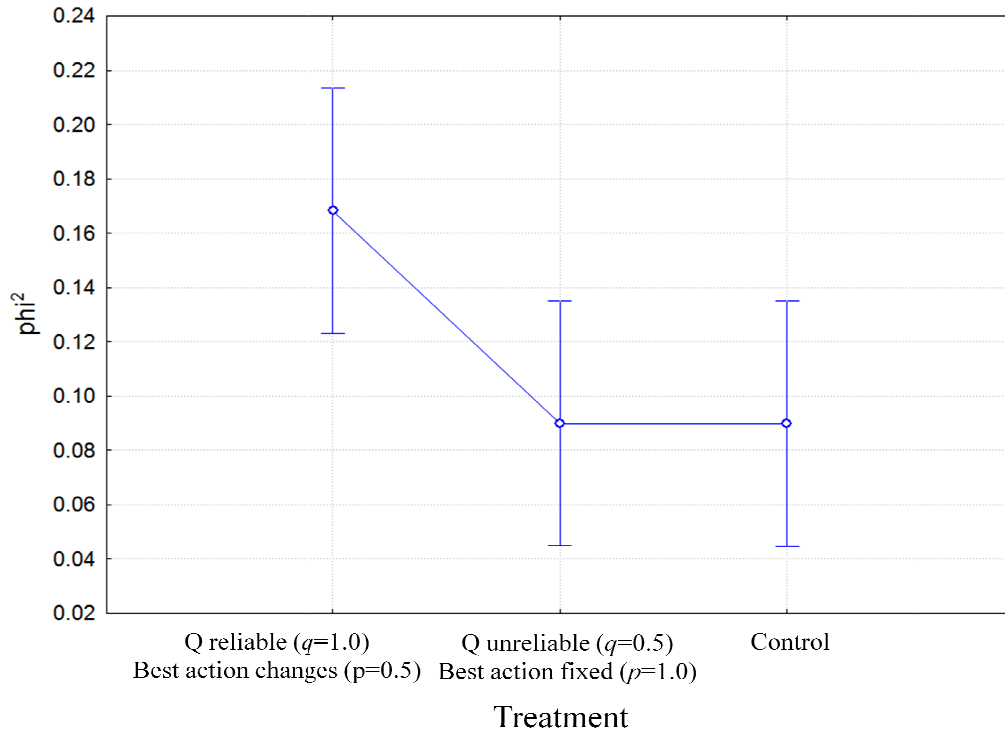


Figure 1-3. Significant effect of treatment during the final experiment assays following the 30th generation of selection. Error bars are 95% confidence intervals. The difference between reliable/unpredictable and the other treatments is statistically significant.

General Selection Results

Flies never oviposited on a substrate when quinine was present. This avoidance of quinine was the same for all treatments, and it did not change during the course of selection. Although, the analysis is less tidy, an analysis of data from selection trials is in broad agreement with the analyses presented above. Specifically, we calculated the extent to which flies avoided the media type that had been paired with quinine in the experience phase of selection trials using the proportion all all eggs laid on this type of media (a variable we call $P(\text{Response to Experience})$). To account for changes in these measure across generations, we calculated $P(\text{Response to experience})$ values for each line in blocks of 2 generations each. Finally, we analyzed these scores in an ANOVA with factors of treatment and block, with repeated measures on each line. This analysis showed main effects of treatment ($F_{1,22} = 6.51, P = 0.018$), with ‘best action changing ($p=0.5$)/experience reliable ($q=1.0$)’ lines showing higher learning scores than ‘best action fixed ($p=1.0$)/experience unreliable ($q=0.5$)’ lines; and a main effect of effect of block is also statistically significant ($F_{4,308} = 2.31, P = 0.005$), but the interaction between the two is not quite significant ($F_{14,308} = 1.62, P = 0.071$).

Discussion

This study offers an experimental analysis of the selective value of learning. Specifically, it asks how two components of change (the reliability of experience, and underlying uncertainty about the appropriate action) affect the value of learning. It is, to our knowledge, the first experimental confirmation of the insight that that these two statistical relationships can select both for and against learning. Our result illustrates the weakness of the influential claim of the ‘learning folk theorem’ that “change favors

learning” while “fixity favors non-learning.” Our results suggest that randomness and not fixity is the most powerful and plausible way to select against learning. Consider, for example, the Garcia effect (Garcia and Koelling, 1966) that we discussed in the introduction, which shows that rats learn associations between tastes and gastric illness more easily than the association between bright-noisy and gastric illness. Surely this does not happen because the relationship between visual stimuli and gastric illness has been fixed throughout rat evolutionary history. It is much more plausible that visual stimuli have varied unpredictably in relation to gastric consequences.

Our study, of course, owes much to the ground-breaking work of Mery and Kawecki (2002; 2004a). Mery and Kawecki’s two studies using this experimental system tested the role of change in the evolution of learning, motivated by the learning folk theorem. Our ‘best action changing/experience reliable’ treatment replicates Mery and Kawecki’s first study in that both studies found that this condition selected for enhanced learning. Our study, however, introduced random change while Mery and Kawecki strictly alternated orange-best and pineapple-best conditions. The key difference between our approaches, however, follows from different perspectives about the condition that selects against learning. Following the learning folk theorem, Mery and Kawecki tested an absolutely fixed condition in which the best action was always the same, and where quinine reliably predicted the best action. Contrary to their expectation, they found enhanced learning in this situation. In contrast, following the ‘components of change’ view of learning we tested a condition where the best action was always the same, but where there was no predictable relationship between quinine pairing and the best action. As predicted, we find reduced sensitivity to experience and increased reliance on unlearned preferences in this selection regime.

While Mery and Kawecki's work represents the only similar empirical studies, our work has deep connections to theoretical work on the selective value of learning (Dukas, 1998; Johnston, 1982; Kerr and Feldman, 2003; Papaj, 1994; Stephens, 1991). As a group these papers emphasize the role of change and other statistical properties of the environment in learning evolution. Early work by Johnston emphasizes the learning folk theorem even though it acknowledges that animals should not learn in some changing environments (e.g. under complete unpredictability). The later papers take an increasingly nuanced view that either recognizes different components of change (e.g. Papaj, 1994; Stephens, 1991) or argues that intermediate levels of change favor learning (Kerr and Feldman, 2003). This paper, perhaps unsurprisingly, is most clearly connected to the Stephens (1991) model. The parallels between the Stephens model and the Mery-Kawecki experimental preparation (used here) are striking. Stephens modeled a hypothetical organism with a two-stage life history. In the first stage, the organism can choose to obtain experience; while in the second stage the animal can act in response to its experience in the first stage. Although the Stephens model characterized the components of environmental change in a different way, its predictions closely follow the model presented here with one key difference: the Stephens model predicts non-learning for the absolutely-fixed condition. This difference occurs because the Stephens model imposed an opportunity cost on learning. Specifically, in the Stephens model a learner can waste time acquiring experience in the 'experience phase' of its life history, when the analogous non-learner can begin to acquire fitness benefits in the experience phase. This cannot happen in the Mery-Kawecki preparation, because choices made in the experience phase do not affect fitness. Natural learning surely imposes some costs (both opportunity costs and physiological costs), however models suggest—in agreement with

our experimental results—that unpredictability is a much more powerful and robust way to select against learning than fixity, even when learning imposes costs.

Summary

The experimental analysis presented here exploits Mery and Kawecki's pioneering empirical paradigm to test a logically coherent model of learning evolution. This model recognizes two distinct types of 'fixity' that have opposing effects on the selective value of learning: 1) fixity of the best actions (e.g. it's always best to lay eggs on orange) selects against learning (as the 'folk theorem' claims), and 2) fixity of the relationship between stimuli and best action (e.g. quinine is always paired with the worst type of media) favors learning. Our results support this more complicated claim. In treatments with a fixed 'best action' and an unreliable (changing) relationship between stimuli and best action, we observed increased non-learning (i.e. simple preferences for media type). On the other hand in treatments where the best action changed and we created a reliable (fixed) relationship between stimuli and best action, we observed increased learning.

Learning is a fundamental mechanism for adjusting behavior to change in the environment. Our results emphasize a richer and more realistic view of the evolutionary advantages of this flexibility, recognizing that different components of environmental change can have different effects on the evolution of learning and phenotypic plasticity. This perspective is significant because it is immediately relevant to the explanation of variation in animal learning abilities such the Garcia effect and other examples of selective association in animal learning.

Chapter 2 — Stimulus reliability in the evolution of “prepared” learning

Introduction

For animals to learn, they must form associations among various stimuli. But in a world full of potential stimuli, why does a ‘special relationship’ form between a given stimulus and consequence in a way that actually allows the animal to predict future events?

Animals seem to solve this problem by being born better able to learn some things than others. Although the most notable example of this “special” learning is the Garcia effect (Garcia and Koelling, 1966), through the years, hundreds of additional examples have been cataloged (e.g. Dobrzecka et al., 1966; LoLordo, 1979; Rozin and Kalat, 1971; Shettleworth, 1972). Although these “exceptions” originally made up a challenge to the general process view of learning (e.g. Logue, 1979), whereby all animals learn all things in the same way, explanations eventually centered around animals being born “knowing” they should learn about certain things and not others. These explanations include ideas of “belongingness” (Thorndike, 1932), species-specific defense reactions (Bolles, 1970), biological constraints (e.g. Shettleworth, 1972; Shettleworth, 1979), adaptive specializations, and “preparedness” (Seligman, 1970). Clearly an evolutionary approach called for to explain these patterns in learning.

Despite the detailed discussion and thought from many scientists over the years, many questions about the evolution of prepared learning remain. Although there are multiple

potential explanations, the central question of why differences in associability occur has not been resolved. We can generally agree that learning should enable animals to predict when biologically significant events will occur. These predictions occur because an individual has ascertained causal relationships within the environment. Throughout the evolutionary history of a given organism, some aspects of the environment will have influenced behavior more than others aspects. The most likely hypothesis for the evolution of prepared learning is that some stimuli and consequences have remained both reliable and important throughout the lineage of a given organism, and that these stimuli pairings are selectively learned over other pairings that do not have such a history of being paired. Such a reliable linking across time is what would make some stimuli salient, and others a mere part of the background. This type of story explains, for instance, the well-cited example that young rhesus monkeys learn fear about snake-like objects more quickly than they do about flower-like objects (Cook and Mineka, 1990; Cook et al., 1985).

For such a central unanswered question in the psychology of animal learning, both theory and empirical results that test this evolutionary story are completely lacking. Even more broadly, however, the question raises a fundamental question about information use in animals: how does information an animal is born with interact with the information an animal gains through experience? We present a mathematical model describing how prepared learning of one type of stimulus over another will evolve, and then test the predictions of this model in an experimental evolution study with fruit flies (*Drosophila melanogaster*).

Model & Predictions

Our model is constructed very similarly to the model described in chapter 1. Again, we use the basic experimental preparation adapted from Mery & Kawecki (2002). This preparation provides a two stage learning process that is both experimentally tractable and can be generalized to learning in the broader sense. In this preparation, flies experience two phases. The first is an experience phase, where one substrate type is paired with quinine (an aversive stimulus), while the second substrate type is not. The second is a consequence phase, where neither substrate is paired with quinine, but where an experimenter can enact selection consequences by rearing some eggs, but not others. Female flies can use their experience from the first phase to guide their choices in the second phase. If flies are choosing in the consequence phase consistently with learning, they will lay eggs on the substrate that had not been paired with quinine in the experience phase.

In this case, we envision a system where stimuli from two different sensory modalities can predict the “best” substrate on which a female can lay her eggs. Thus, one mode, say color, would have stimulus A and stimulus B, while the other mode, odor, would have two stimuli, α and β . Thus the potential quinine pairings could be $Q+A\alpha$ or $Q+B\beta$ or $Q+A\beta$ or $Q+B\alpha$. However, these pairings need not remain the same during the consequence phase. Figure 2-1 describes possible ways these pairings can occur.

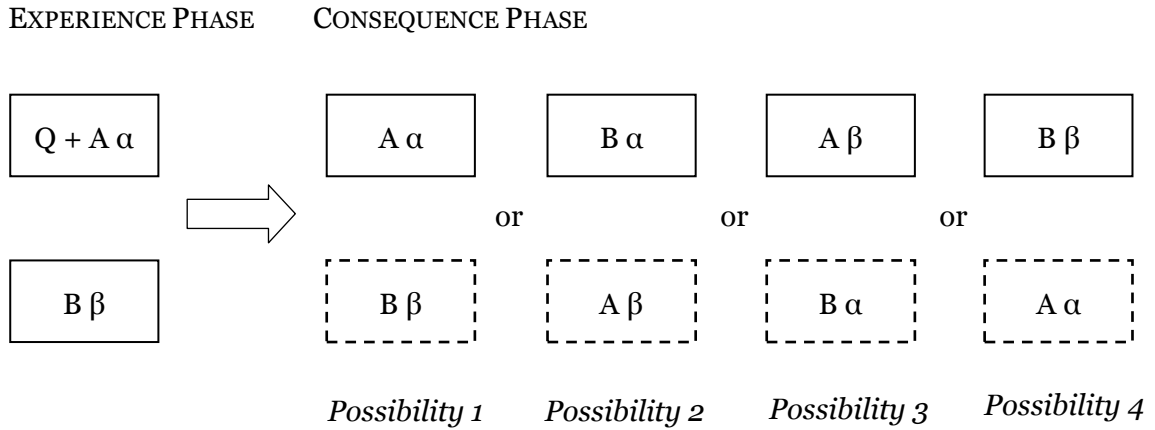


Figure 2-1. Assume the best environment is indicated by the box with the dashed lines. In the first possibility, both A and α reliably predict the best environment when paired with quinine. In the second possibility, only the pairing of α with quinine predicts the best environment; an individual attending to/ learning the A pairing would be making an incorrect choice. There are two additional possibilities: that A, but not α predicts, and that neither predict.

We can assign a probability to the reliability of each stimulus with regards to the quinine pairing predicting the best environment, where O is the probability the odor pairing predicts the best environment, and C is the probability that the color pairing predicts the best environment. The details of this model are in Appendix 2. The predictions, which take the same basic shape as those from Chapter 1, are in Figure 2-2 below.

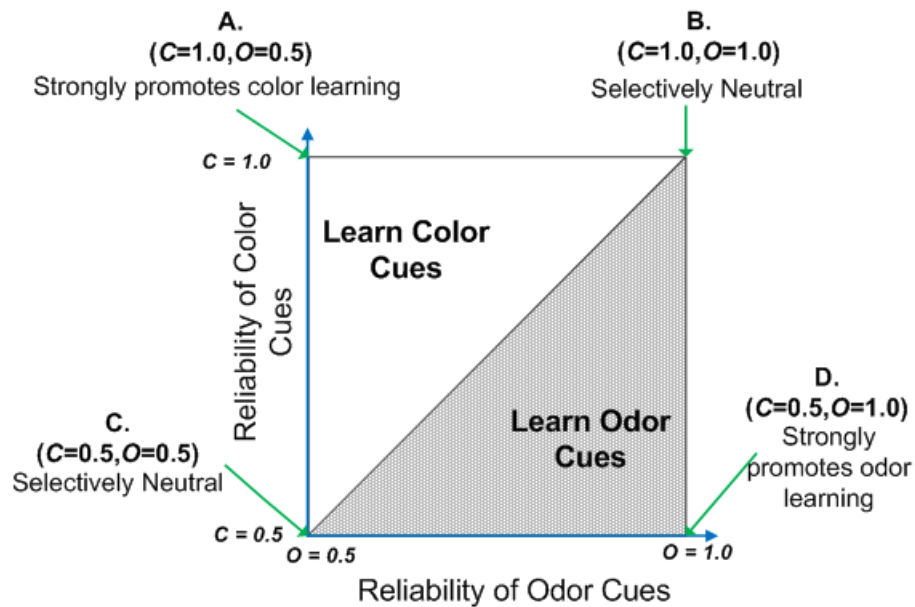


Figure 2-2. Predictions from the model. Whenever the probability that odor predicts is greater than the probability that color predicts ($O > C$), then learning about odor is favored. We tested each the points at the four corners of the graph. Two of these points fall on the line between color learning favored and odor learning favored, and thus either or both could be favored. Intuitively, we predict that when both modalities of stimuli predict equally well, learning about both should be favored, while learning about neither should be favored when neither modality of stimuli predict the best environment.

Methods

Flies and Husbandry

Our starting population was a mix of wild-caught, lab-adapted flies from four different locations in Minnesota and Wisconsin, USA. We combined 400 male and 400 female adults from each population and maintained them in overlapping generations in a large population cage for 14 months prior to the start of the experiment. We housed all flies at 24°C and tested them at 14 days old (post egg). We reared all eggs at a density of 80 eggs per vial, and six vials per line per generation. For each generation, we moved flies to population test cages (3.3 cm L x 21 cm W x 12 cm H) upon eclosion as adults, setting up populations of 200 females and a comparable number of males per line. Each cage featured a removable tray on which we placed two fresh petri dishes of standard cornmeal and molasses-based food, and after three days, we tested each cage of flies.

Aversion Learning & Selection on Populations

We tested each generation of flies once, testing as groups in the population cages. As described in our model, each test consists of two phases: an experience phase and a consequence phase. In the first phase, the experience phase, we exposed flies to two petri dishes of agar-based media in a single three-hour session (10 mL of agar placed into the bottom of each 100 mm x 15 mm petri dish). We introduced color into the substrate by placing painted disks underneath the petri dishes, using cobalt blue and aqua blue color. We introduced odor by mixing amyl acetate and benzaldehyde into the agar (we first diluted each into a mixture of 35% odorant, 65% ethyl alcohol, and added each to agar (20g sucrose, 10g agar, 1L water; 1mL for amyl acetate and 0.1mL for benzaldehyde). Prior to the experiment, we conducted pilot studies to demonstrate learning to the colors

and to the odors chosen, and tested that neither mode of stimuli completely overshadowed the other during learning trials. Finally, we added quinine at 4g / 1L agar). To start the experience phase, we positioned the petri dishes on a sliding tray at the bottom of each cage; we could replace these dishes without moving the flies. In the second phase, a consequence phase, we present new plates of agar without quinine for five hours. The pairing of color and odor could be different, depending on the assigned reliability of each (Figure 2-1). We randomized the locations of the plates, with visual stimuli always remaining in the same location in both experience and consequence phases, and the corresponding odors changing location (depending, again, on the assignment of reliability for each modality). We separated the experience and consequence phases with a 30 min period of no stimuli.

We imposed different selective regimes by rearing eggs from one substrate type (e.g. the one not paired with quinine in the experience phase) and discarding eggs from the other (e.g., the one paired with quinine in the experience phase). We removed eggs selected for propagation from the substrate using a needle and placed them in vials on standard cornmeal-based fly food for incubation.

Treatments and Lines

We set up 40 lines of 200 females and approximately the same number of males from the source population, and then randomly assigned 10 lines to each of the four experimental treatments. The treatments were 1) both color and odor reliable ($C=1.0, O=1.0$), 2) color but not odor reliable ($C=1.0, O=0.5$), 3) odor but not color reliable ($C=0.5, O=1.0$), and 4) neither color nor odor reliable ($C=0.5, O=0.5$). At the start of the experiment, we randomized the reliability of the quinine cue for each stimulus modality for each

generation within blocks of two generations each and did this separately for each line, according to its assigned treatment.

Results

Selection Data

We calculated the selection data in terms of the reliability of color cues and in terms of the reliability of odor cues. For instance, to calculate scores for flies laying eggs consistent with following the odor and quinine pairing, we took the proportion of eggs laid in the consequence phase on the substrate with the odor that had not been paired with quinine in the experience phase. Because we randomized the experiment in blocks of two generations each, we took means for each block. We completed an ANOVA for each, with factors of odor-reliable or color-reliable, and repeated measures on the lines. As predicted, there was a statistically significant effect of reliability of odor cue (Fig. 2-3; $F_{1,144} = 4.153$, $P = 0.048$). Treatments in which odor was reliable showed an increase in learning to odor across the course of the experiment, whereas treatments where odor was unreliable did not. The same was true for color, where there was a statistically significant effect of reliability of color cue (Fig. 2-4; $F_{1,144} = 15.370$, $P < 0.001$). However, the pattern across the course of selection for color was not as orderly as for odor cues. Neither analysis gave a significant interaction of block by reliability, as we would expect in a selection process where treatments are selected in different directions (Learning to odor, block*odor reliability interaction: $F_{4,144} = 1.905$, $P = 0.112$; learning to color, block*color reliability interaction: $F_{4,144} = 1.596$, $P = 0.178$).

Because these data reflect the selection regime, the scores for the Both Reliable treatment resulted in the same calculations in both graphs (because color and odor were both always reliable). Thus we have no way of knowing whether flies in this treatment were attending to

cues about color, odor, or both together. For this reason, we look to assays we ran following the end of the experiment.

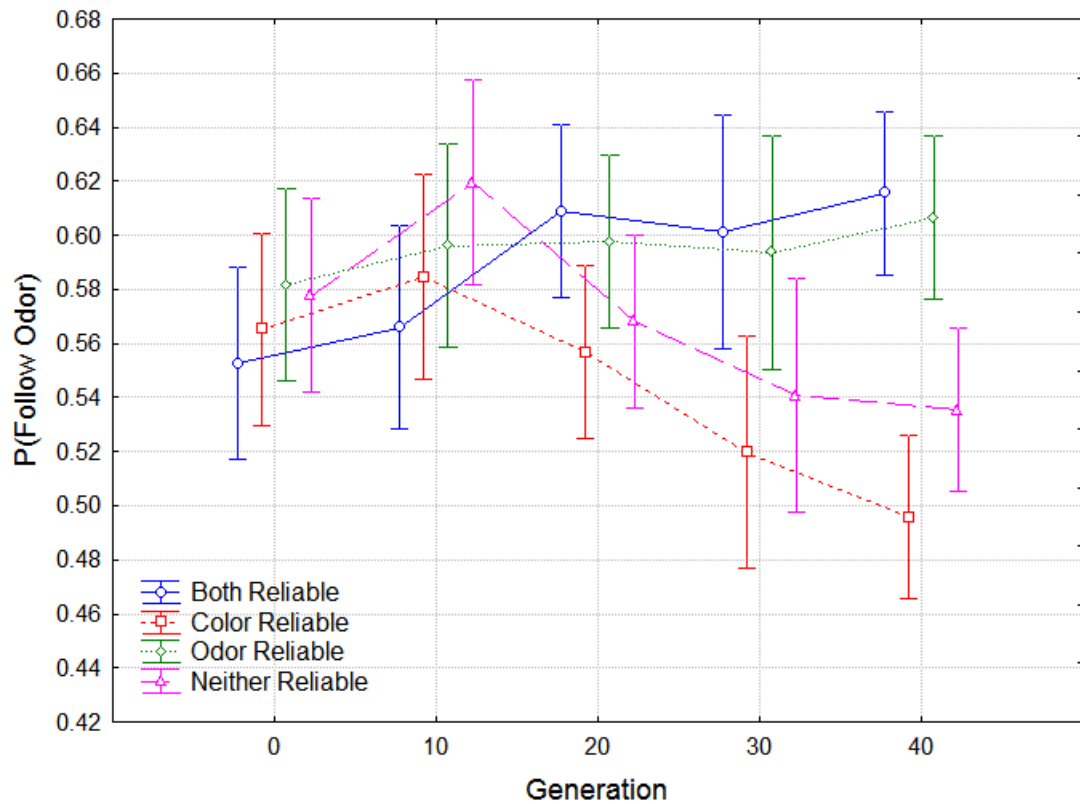


Figure 2-3. Selection data of following the quinine pairing with odor. The x-axis represents means of 2-generation blocks (consistent with the randomization scheme of the experiment), taken at 10 generation increments. Error bars are standard errors.

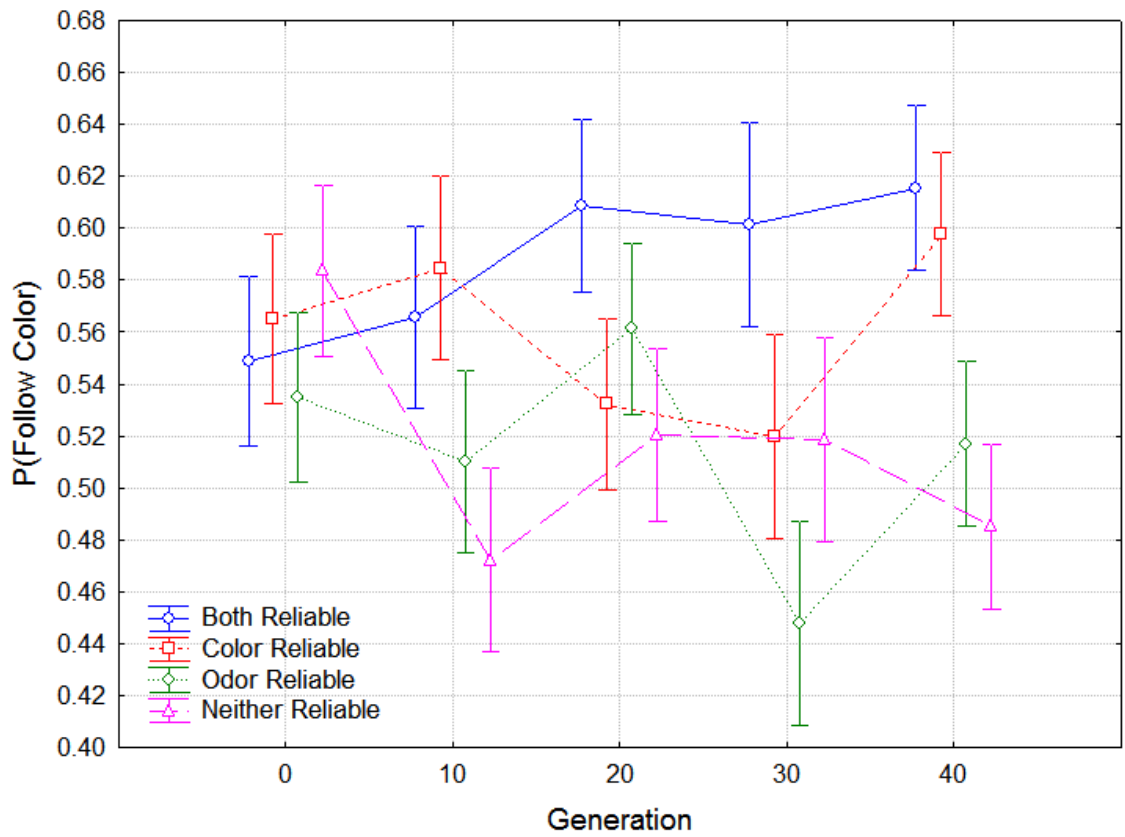


Figure 2-4. Selection data of following the quinine plus color cues. Error bars are standard errors. The x-axis represents means of 2-generation blocks (consistent with the randomization scheme of the experiment), taken at 10 generation increments.

End of Selection Assays

In the set of assays we describe here, we tested different naïve subsets of flies from each line in two types of tests. In the first type of test, we assayed learning abilities to color stimuli only, with no odors present. We tested one set of flies for quinine with blue and one for quinine with aqua. In the second type of test, we assayed learning abilities to odor only, testing a set of flies for each of quinine with amyl acetate and quinine with benzaldehyde. These assays were similar to the selection procedures, except the “consequence” phase lasted 1.5 hours, and there were no consequences since no selection occurred. (Note that we use the terminology “consequence phase” for simplicity here even though these assays differed from the experimental selection procedures). Figure 2-5 shows the results of the assays in which we tested learning to color alone and Figure 2-6 shows the results of the assays in which we tested learning to odor only.

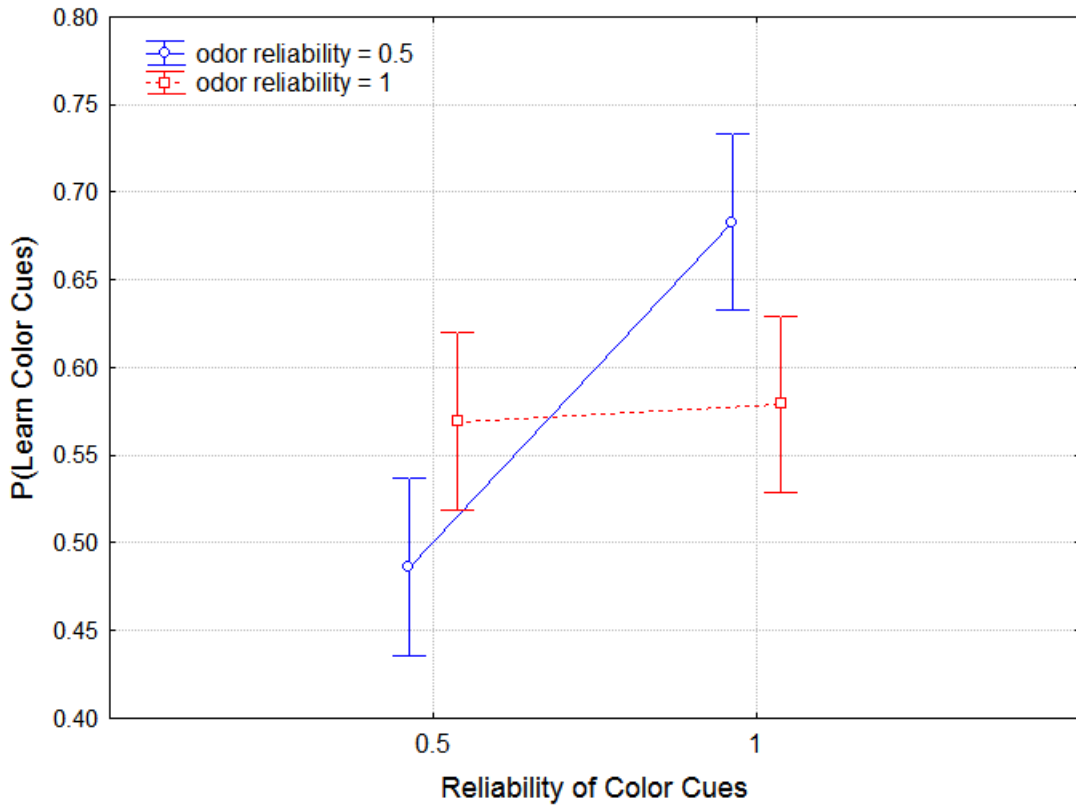


Figure 2-5. This figure presents the data in the factorial form of the experiment's design. Learning about color was enhanced when color was a reliable predictor of the best environment across evolutionary time. The effect of color reliability was statistically significant ($F_{1,36}=4.189$, $P=.048$). The effect of odor reliability was NS, and did not differ across color reliability levels. The interaction between color and odor reliability nears significance ($F_{1,36}=3.435$, $P=0.072$).

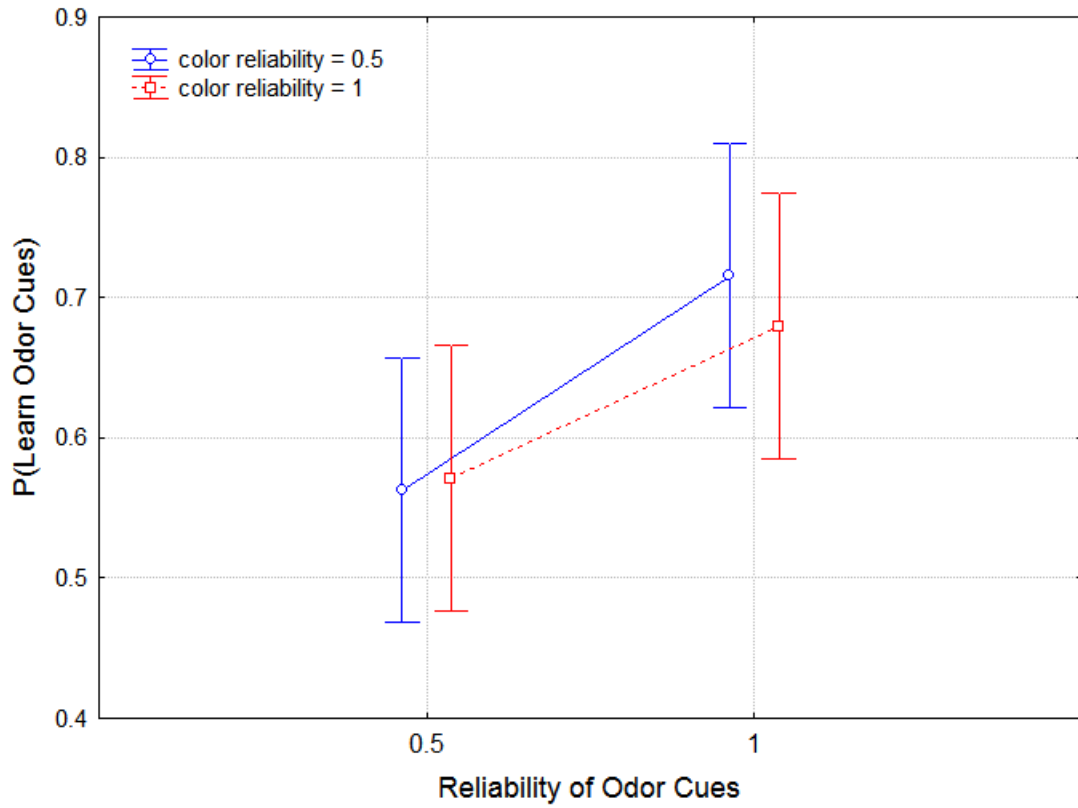


Figure 2-6. Learning to odor alone is enhanced in both treatments where odor is reliable. The effect of odor reliable is statistically significant ($F_{1,36}=7.8829$; $P=0.008$), neither the effect of color reliable nor the interaction between color and odor reliability are statistically significant.

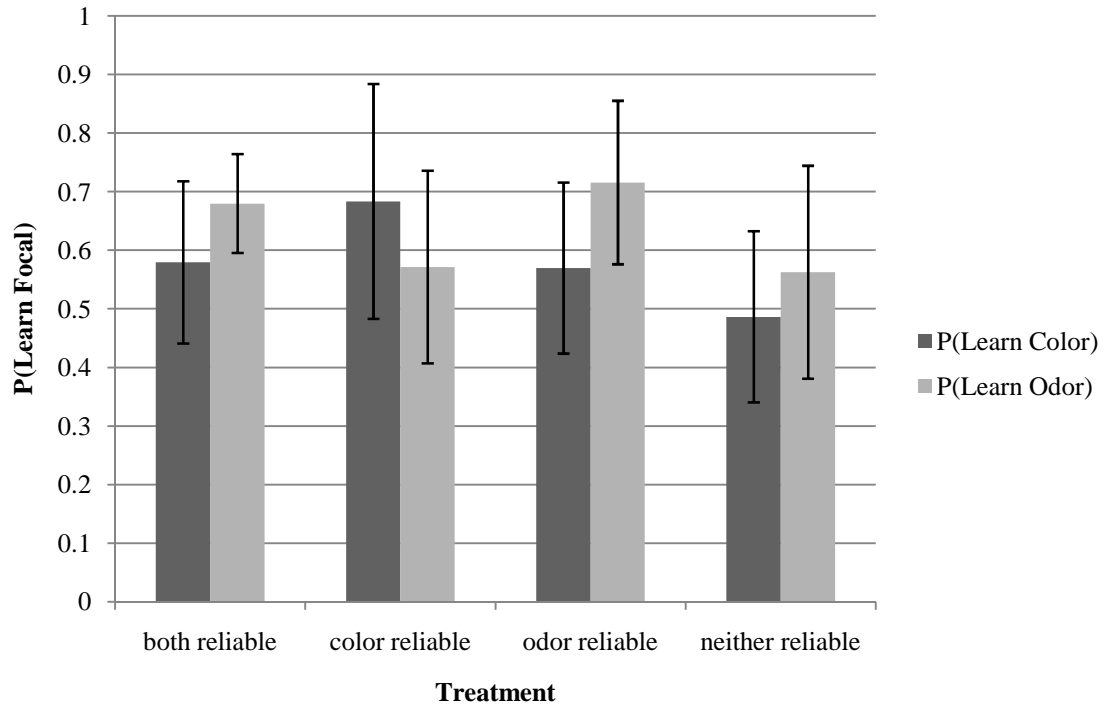


Figure 2-7. The same data as in Figures 2-5 and 2-6 (end of experiment assay data) presented together in a single graph to demonstrate the relationship between learning the two sensory modes of stimuli. As expected, in each treatment where odor reliably predicted the best environment, the flies learned better to odor than to color. Only the lines of flies in the color-reliable treatment showed better learning to color than to odor. Error bars are standard errors.

Discussion

This study asks how reliability influences which types of stimuli are most easily learned about. We have shown that the reliabilities of stimuli from different sensory modalities do influence how flies learn about those stimuli over evolutionary time. When color is a reliable predictor of the best action, this reliability is a significant factor in how flies learn about color; the same result holds for odor. This result was consistent in the selection data across the generations and in separate assays of each line after 40 generations of selection. This is, to our knowledge, the first model and the first multigenerational study demonstrating that patterns of reliability in evolutionary history do in fact influence the stimuli that are most easily learned.

Learning about color was better than learning about odor for flies when color and not odor were reliable predictors across time. For each of the other selection regimes, learning about odor was better than learning about color. This makes sense for the regime where only odor was reliable. The clearest results occurred where we expected them, at the areas where we predicted the strongest selective pressure for each type of learning (points A and D on figure 2-2). For the remaining two points on the figure, which fell on the line separating odor-learning favored and color-learning favored, we predicted that neither type of learning would be favored over the other. In the selection data, where both sensory modalities of stimuli were simultaneously present, the learning results followed the reliability assignments: when color was reliable, color learning was comparatively better than where color was not reliable; the same was true for odor. However, when we tested learning to each sensory modality alone, the treatments falling on that line both showed better learning to odor than to color. A possible explanation for

the higher scores in general for learning odor stimuli is that flies are following a rule such as, “when in doubt, follow odor.”

Our basic hypothesis that stimulus reliability over time will predict which sensory modalities are most easily learned about was supported the data, however what accounted for these differences in learning, and how do the two modes of stimuli interact to affect learning in the evolved lines? One interesting question related to the data we presented here has to do with multimodal learning and signaling. Why have so many signals in nature evolved to use more than one sensory modality (e.g. Guilford and Dawkins, 1991)? For instance, floral signals are multimodal: including shape, color, and scent, and bumblebees make more accurate decisions when presented with multimodal choices than single modality choices (Kulahci et al., 2008). Signals can be costly to produce, and arguments about the evolution of multimodal signals generally centers around the role of the receiver (e.g. Rowe, 1999). In this experiment, we evolved lines in environments where only a single sensory mode was reliable and in environments where two sensory modes of stimuli were reliable. During the selections, lines in the treatment where both modalities of stimuli were reliable consistently had the highest learning. However, when we tested learning to stimuli in one sensory modality at a time, we found that flies in these lines seem learned odor best (Figure 2-7). This suggests that in our system one modality is sufficient for learning, but the addition of a second improves performance.

This is the first study to demonstrate learning about color during oviposition in fruit flies. The vast majority of learning assays in *Drosophila* use olfactory stimuli, and those using visual stimuli are procedurally difficult, involving flight simulators (e.g. Brembs

and Heisenberg, 2001; Brembs and Wiener, 2006). The oviposition procedure used here is a promising one for studying sensory integration and learning in *Drosophila*. We have studies ongoing to test decision-making and learning in multiple sensory mode settings.

Summary

This novel study tested specific predictions about the role of the stimulus reliability across evolutionary time in the formation of relationships between particular stimuli and actions. Our central hypothesis for the evolution of prepared learning is that some stimuli pairings have remained both reliable throughout the lineage of a given organism, and natural selection has, therefore, favored learning of these types of associations. Our results throughout 40 generations of experimental evolution with fruit flies support this hypothesis. This work provides an important step forward in understanding both prepared learning and how inherited tendencies interact with information gained through experience.

Chapter 3 – Why some memories do not last a lifetime: dynamic long-term retrieval in changing environments²

Introduction

Learning is responsible for many of the flexible behaviors animals use when adjusting to changing environments. Ultimately, learning depends on an animal's ability to store and retrieve information about its world; that is, learning is tightly entwined with memory (Bouton, 1994; Shettleworth, 1998). Consequently, the value of memory often relates to the value of learning. On a simple level, learning is valuable when information stored in memory (e.g., a good response to a stimulus) remains useful over time (e.g., the same response to the same stimulus continues to be good). However, learning (and memory) loses value when the environment changes in ways that make learned responses unreliable. Indeed, possessing memory might be costly if it repeatedly leads an organism to employ inappropriate behaviors (e.g., if the environment changes extremely rapidly making remembered responses perpetually inappropriate). If memory is evolutionarily adaptive, one would expect that various properties of memory (encoding, consolidation, retrieval, etc.) would be tuned to the historical rate of environmental change. Several mathematical models have suggested that the amount of environmental variability

² This chapter is a collaborative work with Benjamin Kerr. It is currently under review as Dunlap, A. S., McLinn, C. M., MacCormick, H. A., Scott, M. E. & Kerr, B., "Why some memories do not last a lifetime: dynamic long-term retrieval in changing environments."

influences the evolutionary advantages of learning and memory (Arnold, 1978a; Bergman and Feldman, 1995; Cohen, 1991; Feldman et al., 1996; Kerr and Feldman, 2003; Mangel, 1990; McNamara and Houston, 1987; Stephens, 1987; Stephens, 1991).

Given that memory is employed throughout the lifetime of an individual, should optimal memory length change within that lifetime in response to particular patterns of environmental variability? For simplicity and tractability, most learning models have incorporated memory as a fixed quantity (Kerr and Feldman, 2003; Mangel, 1990; McNamara and Houston, 1987; Shafir and Roughgarden, 1996), often as a “sliding window” of constant length, but see (Hirvonen et al., 1999).

However, overwhelming evidence suggests that memory changes over the lifetimes of individuals. For instance, aspects of memory have been shown to decline with age in a large number of species, from humans to crustaceans (Davis et al., 2003; Punzo and Chavez, 2003; Shukitt-Hale et al., 2004; Tomsic et al., 1998). In humans, this trend is present in healthy aged adults—it is not attributable solely to the effects of Alzheimer’s or dementia (e.g. Craik and Jennings, 1992; Davis et al., 2001; Zelinski and Burnight, 1997). One explanation for memory loss suggests that aged individuals acquire such a large number of memories that retrieval errors (Cohen, 1996; Spear, 1973; Squire, 1989), perhaps caused by mechanisms such as memory interference, become more common (Bouton and Moody, 2004; Mery and Kawecki, 2004b) Another explanation attributes memory loss to the inevitable decay of synaptic connections with age (Salthouse, 1996). Poor memory or learning abilities are also found at very young ages (Guo et al., 1996; Spreng et al., 2002). Known as infantile amnesia, this is often explained as the result of the developing brain (Neissen, 2004). In addition to age, the physical state of the

individual (such as its energy reserves or stress level) may influence learning and memory (Guo et al., 1996; McEwen and Sapolsky, 1995; Xia et al., 1997; Yanai et al., 2004). For instance, while acute stress might actually improve memory in some cases, such as with post-traumatic stress disorder, (McGaugh and Roozendaal, 2002), chronic stress and increased cortisol levels can result in memory deficits (Lupien et al., 1998; McEwen and Sapolsky, 1995; Seeman et al., 1997).

In the aforementioned explanations, memory loss results from the inevitable degradation or inherent constraints in retrieval caused by advanced age or prolonged stress. Another compelling possibility is that memory change might serve some evolutionarily adaptive role. Kraemer and Golding (1997) suggest the primary mechanism for adaptive forgetting may be in the retrievability of information. Importantly, new evidence suggests that forgetting need not be a passive process (Anderson et al., 2004; Wylie et al., 2008).

In order to explore the possibility of evolutionarily adaptive forgetting, we use a dynamic programming approach—an established method for determining optimal strategies across lifetimes (Clark, 1993; Houston and McNamara, 1999; Mangel and Clark, 1988). Within this modeling framework, we focus on the long-term retrieval of information, assuming that all experiences of the organism have been successfully encoded and consolidated. For effective memory, an organism needs both successful acquisition and retrieval. A failure at any point along the way results in forgetting. We do not consider mechanism, though in our framework, forgetting may be due to an active process like suppression (Anderson et al., 2004), the decreasing availability of stored information over time, which Schacter calls transience (1999), or due to a simple failure to retrieve

the information (Loftus and Loftus, 1980). We explore the effects of age, physical state, and environmental conditions on optimal retrieval. Our model suggests that even if the environment changes at a constant rate, optimal memory length may actually decline with age.

The Model

Here we consider a learning organism that lives for a maximum of T discrete time steps, with its age at any point in time given by the variable t (the number of time steps lived so far). At each time step, there is a fixed positive probability of death, δ . This organism inhabits an environment with N stimuli to which it may respond. At every time step, the organism is exposed to one of the N stimuli (chosen at random with equal likelihood and independently of past exposures). If the stimulus is unfamiliar, the organism will need to spend some time learning how to respond appropriately to the stimulus. This learning process within a time step can be thought of as being on the time scale of short-term memory. For simplicity, we assume that the organism that learns about an unfamiliar stimulus always arrives at a behavior that yields a constant payoff, which we label π_u .

If the stimulus is familiar, then the organism must have interacted with it in a previous time step and remembers it from that time step. Here, we use a “window” model for memory retrieval—the organism can remember and retrieve information from the previous m time steps. If the organism remembers the behavior it employed previously for a stimulus, and employs this behavior again when revisiting the familiar stimulus, there are two possible outcomes: (1) the appropriate response to the stimulus is the same as it was before and the organism receives a payoff π_c (a payoff for a correct response) or

(2) the appropriate response to the stimulus has changed and the organism receives a payoff of π_i (a payoff for an incorrect response). Again, for simplicity, we assume that the payoffs for an appropriate or inappropriate remembered response each have constant value.

Thus, the expected payoff of a familiar stimulus depends on the rate of change of stimuli. Here, we assume that every stimulus changes independently with probability ρ every time step (as ρ increases, the stimuli become less reliable). We also assume $\pi_i < \pi_u < \pi_c$, as a correctly remembered behavior saves the sampling time spent on learning to respond appropriately to an unfamiliar stimulus (Lewis, 1986). However, an incorrectly remembered behavior wastes more time than learning about a stimulus from scratch (e.g., time spent employing the now inappropriate behavior in addition to time spent learning the stimulus again). We assume that the payoffs are in terms of energy reserves, such that time spent learning about stimuli or time wasted employing incorrect behaviors results in a lower energy payoff for the time step. At any time we let the energy reserves of our organism be x . We restrict x to a range of values, $0 \leq x \leq X$. We also assume that there is a cost of living, κ , imposed each time step (i.e., energy reserves removed from the organism due to the metabolic demands of living). The parameter κ is measured in the same units as x (e.g., calories) and we focus on cases in which the cost of living makes memory necessary for x to increase over a time step, that is,

$$\pi_i < \pi_u < \kappa < \pi_c.$$

In this manuscript, we explore the idea that it may be advantageous for the length of the memory window, m , to vary as a function of the energy reserves of the organism, x , as

well as its age, t (i.e., $m=m(x,t)$). Here, we use a dynamic programming approach to find the function that maximizes reproductive success of our model learner, $m^*(x,t)$, where the asterisk denotes the optimal memory. That is, we ask how the memory window might optimally change with state and age of the organism.

To this end, it helps to produce another function that measures maximal future expected reproduction. We call this function $F(x,t,T)$, the maximal future expected reproductive success of an organism at age t in state x (recall that T defines the maximum lifetime of an organism). Using F as a “common currency,” the trick to dynamic programming involves writing a backwards recursion in time. In Appendix 3 we derive the following recursion,

$$F(x,t,T) = (1 - \delta) \max_{\substack{m \\ 0 \leq m < t}} \left\{ \begin{array}{l} \alpha^m F(u(x), t+1, T) \\ + \omega [1 - \beta^m] F(c(x), t+1, T) \\ + (1 - \alpha^m - \omega [1 - \beta^m]) F(i(x), t+1, T) \end{array} \right\}, \quad [3.1]$$

with

$$\alpha = \frac{N-1}{N}, \quad \beta = \frac{(N-1)(1-\rho)}{N}, \quad \omega = \frac{1-\rho}{1+(N-1)\rho}, \quad [3.2]$$

$$u(x) = chop(x + \pi_u - \kappa; 0, X), \quad [3.3]$$

$$c(x) = chop(x + \pi_c - \kappa; 0, X), \quad [3.4]$$

$$i(x) = chop(x + \pi_i - \kappa; 0, X). \quad [3.5]$$

The function *chop* simply keeps the state of the organism in the allowed range between 0 and X . Generally, $chop(x; a, b)$ gives a if $x < a$, b if $x > b$, and x otherwise (Mangel and Clark, 1988).

Equation [3.1] can be understood roughly as follows: If the stimulus at time step t is unfamiliar, then the maximal future expected reproductive success from time $t+1$ on is given by $F(u(x), t+1, T)$. With memory length m at time t , this occurs with probability α^m (see Appendix 3). Similarly, if the stimulus at time t is familiar and if the organism employs an appropriate response, then the maximal future expected reproductive success from $t+1$ on is $F(c(x), t+1, T)$; whereas, an inappropriate response at time t would give $F(i(x), t+1, T)$. With memory length m at time t , the probabilities of appropriate and inappropriate responses to a familiar stimulus are $\omega(1 - \beta^m)$ and $1 - \alpha^m - \omega(1 - \beta^m)$, respectively. Thus, the quantity in braces in equation [3.1] corresponds to the future expected reproductive success of an organism with memory length m at time t given that the organism survives the time step (which occurs with probability $1 - \delta$) and remembers optimally from time $t+1$ onwards. To calculate the *maximal* expected future reproductive success at time t (that is, $F(x, t, T)$) we simply need to find the memory length that maximizes the weighted average in braces—see Appendix 3 for details.

To begin, we assume that our organism is semelparous—it reproduces at the end of its lifetime (at age T). The expected number of offspring will be some function of the final state of the parent—we call this function $\Phi(x)$. At age T , maximal future reproductive success is equal to the expected number of offspring such that $F(x, T, T) = \Phi(x)$. If we

specify $\Phi(x)$, then we can use recursion [3.1] to obtain $F(x, T-1, T)$ for all possible values of x . In the process of calculating $F(x, T-1, T)$ we will also reveal $m^*(x, T-1)$. This optimal memory (for a given x value) is simply the m value that maximizes the quantity in braces on the right-hand side of equation [1] with $t=T-1$. Once we have discovered $F(x, T-1, T)$ for all values of x , we can find both $F(x, T-2, T)$ and $m^*(x, T-2)$ for all values of x by applying recursion [3.1] again (with $t=T-2$). Then it is a simple matter to generate both $F(x, T-3, T)$ and $m^*(x, T-3)$ using recursion [3.1] yet again. We can do this over and over again, so that these generated F values become the stepping stones enabling us to walk backwards further and further in the organism's lifetime, revealing the optimal memory for each possible age and each possible state along the way.

Parameters	Description
X	number of states for the organism
T	maximum lifetime for the organism (total number of time steps)
N	number of stimuli in the environment
ρ	probability of stimulus change per time step
κ	cost of living for the organism per time step
δ	probability of death for the organism per time step
π_u	payoff to the organism for a response to an unfamiliar stimulus
π_c	payoff to the organism for a correct response to a familiar stimulus
π_i	payoff to the organism for an incorrect response to a familiar stimulus
Variables	
x	state of organism (current level of energy reserves)
t	age of the organism (current number of time steps)
m	memory “window” (the number of time steps back that are remembered)
Functions	
$u(x)$	state of organism after employing a response to an unfamiliar stimulus, given that its state before responding was x
$c(x)$	state of organism after employing a correct response to familiar stimulus, given that its state before responding was x
$i(x)$	state of organism after employing an incorrect response to familiar stimulus, given that its state before responding was x
$F(x,t,T)$	maximal expected future reproductive success of an organism in state x at age t
$\Phi(x)$	fitness of an organism in state x at age T
$H(x)$	The function to which $F(x,t,T)$ converges as t decreases
$B(x,t)$	potential benefit of memory for an organism in state x at age t
$C(x,t)$	potential cost of memory for an organism in state x at age t
$m^*(x,t)$	optimal memory for an organism in state x at age t (an integer)
$m^{**}(x,t)$	proxy for the optimal memory for an organism in state x at age t (a real number)
$m_H(x)$	approximate optimal memory (the m^{**} proxy using $H(x)$ in lieu of $F(x,t,T)$)

Table 3-1. Table of parameters, variables, and functions used in the model.

Results

In Figure 3-1, we give some examples of the optimal memory surface (the m^* function) for different rates of environmental change (ρ) and final fitness functions (Φ). In general, as the stimuli become more reliable (as ρ decreases), optimal memory increases (the surfaces tend to be higher). This is expected to occur because, as ρ decreases, the probability of employing a correct response increases. Thus, there is increasing value to approaching a stimulus as familiar rather than unfamiliar because the remembered response becomes more likely to be correct. This means there is increasing incentive to lengthen the memory window.

For many parameter combinations, we can partition the lifetime into two periods: a memory lift at the beginning of the lifetime and a memory plateau in the rest of the lifetime that gently drops or lifts towards the end of the lifetime (see Figure 3-1). For intermediate values of stimulus change ($0.05 < \rho < 0.25$), if Φ is convex ($d^2\Phi/dx^2 > 0$), the optimal memory window can increase at the end of the lifetime (Figures 3-1c and 3-1d). On the other hand, if Φ is concave ($d^2\Phi/dx^2 < 0$), then the optimal memory window can actually decrease at the end of the lifetime (Figures 3-1a and 3-1b).

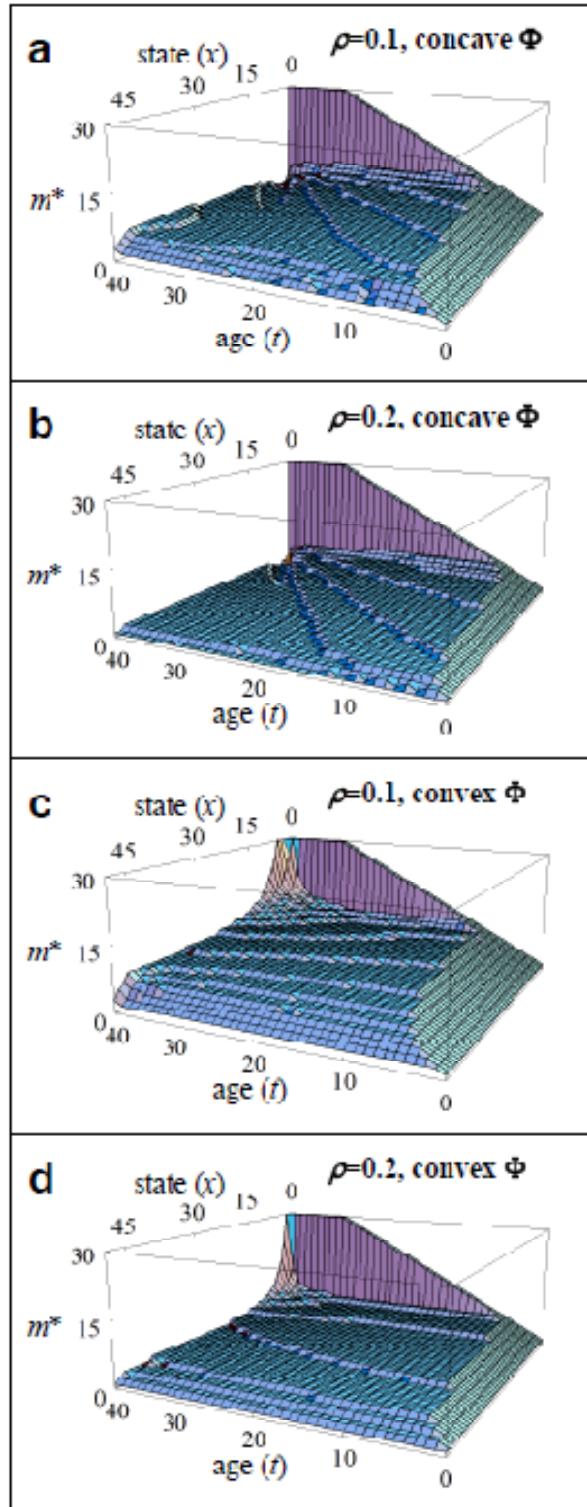


Figure 3-1. Optimal memory $m^*(x,t)$ with $X=50$, $N=10$, $T=40$, $\delta=0.1$, $\kappa=12$, $\pi_i=5$, $\pi_0=10$, and $\pi_c=15$. Each surface corresponds to a different value of the rate of stimulus change (ρ) and either a convex final fitness function $\Phi(x)=100(x/X)^4$ or a concave final fitness function $\Phi(x)=100(x/X)^{0.25}$.

Proxy for Optimal Memory

To understand the effects of stimulus change (ρ), payoff structure (the π 's), and the final fitness function (Φ) it helps to consider the following function:

$$G(m, x, t) = \left\{ \begin{array}{l} \alpha^m F(u(x), t+1, T) \\ + \omega [1 - \beta^m] F(c(x), t+1, T) \\ + (1 - \omega [1 - \beta^m] - \alpha^m) F(i(x), t+1, T) \end{array} \right\}, \quad [3.6]$$

with α , β , and ω given above in [3.2]. From equation [3.1], the integer m value (with $0 \leq m \leq t$) that gives a maximum for $G(m, x, t)$ is $m^*(x, t)$. One way to find a proxy for m^* is to treat $G(m, x, t)$ as a continuous differentiable function of m and find critical points (by setting $\partial G / \partial m = 0$). A single critical point exists at

$$m^{**}(x, t) = \frac{\ln(1 + \nu(x, t)) + \ln \frac{\omega \ln \beta}{\ln \alpha}}{\ln \frac{1}{1 - \rho}}, \quad [3.7]$$

with $\nu(x, t)$ being a benefit-to-cost ratio,

$$\nu(x, t) = \frac{B(x, t)}{C(x, t)} = \frac{F(c(x), t+1, T) - F(u(x), t+1, T)}{F(u(x), t+1, T) - F(i(x), t+1, T)}. \quad [3.8]$$

The potential benefit of using memory is given by the improvement in future expected reproductive success by employing a correct response to a familiar stimulus in lieu of a response to an unfamiliar stimulus—this is precisely $B(x, t)$ from equation [3.8]. The potential cost of using memory is given by the loss in future expected reproductive

success by employing an incorrect response to a familiar stimulus in lieu of a response to an unfamiliar stimulus—this is precisely $C(x,t)$ from equation [3.8]. Thus, without attention to the probabilities of correct and incorrect responses, ν gives a benefit to cost ratio of memory use.

The quantity m^{**} in equation [3.7] is a proxy for the actual optimal memory window (m^* , which must be an integer). If m^{**} is between 0 and t , then the integer below or above m^{**} gives m^* (m^* is the integer which yields a larger value for equation [3.6]). If m^{**} is negative, then optimal memory is zero and if m^{**} is greater than t , then optimal memory is t (see Appendix 4).

This proxy for optimal memory depends not only on the benefit to cost ratio (ν), but also on the probabilities that memory brings benefits versus costs. These probabilities depend on the number of stimuli, N , and the rate of stimulus change, ρ (which is why there are α , β , and ω terms in equation [3.7]—see equation [3.2]). In Figure 3-2, we show how the probability of incorrect familiar responses (red bars), unfamiliar responses (green bars), and correct familiar responses (blue bars) change with the size of the memory window for different values of ρ . From equation [3.1], as m grows larger, the probability of a correct response approaches ω , the probability of an incorrect response approaches $1-\omega$ and the probability of an unfamiliar response goes to zero. As ρ increases, the probabilities of incorrect responses at any memory level also increases (the fraction of red increases), which should favor lower memory values. Indeed, we show in Appendix 5A that when our optimal memory proxy, m^{**} , is greater than 0, increasing the rate of stimulus change (i.e., decreasing reliability) will promote shorter memory

windows; that is, $\partial m^{**} / \partial \rho < 0$. This result has been derived by treating m in function G as continuous; however, when m is constrained to integer values, the same general pattern emerges: stimulus changeability promotes shorter memory windows (see Figure 3-1).

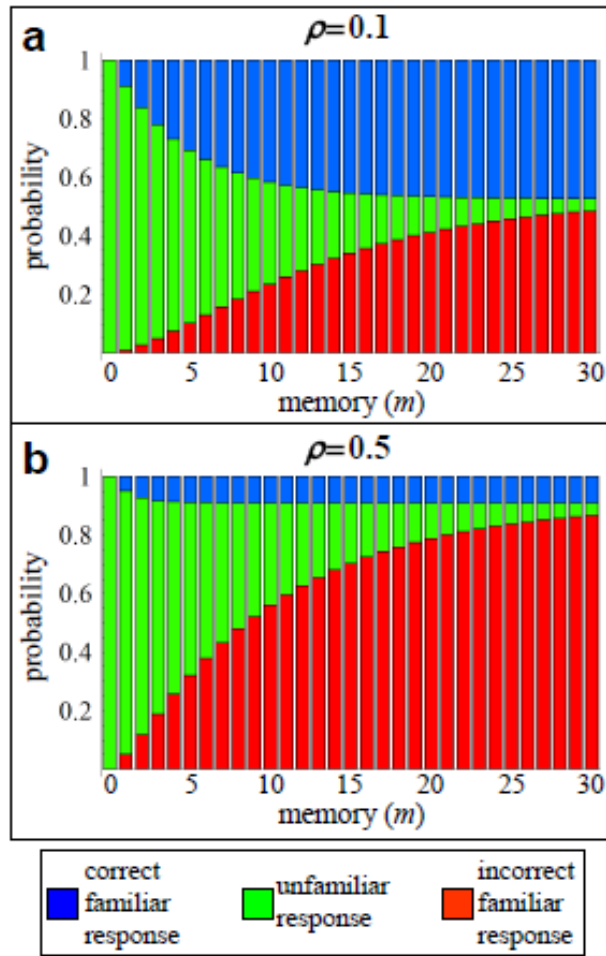


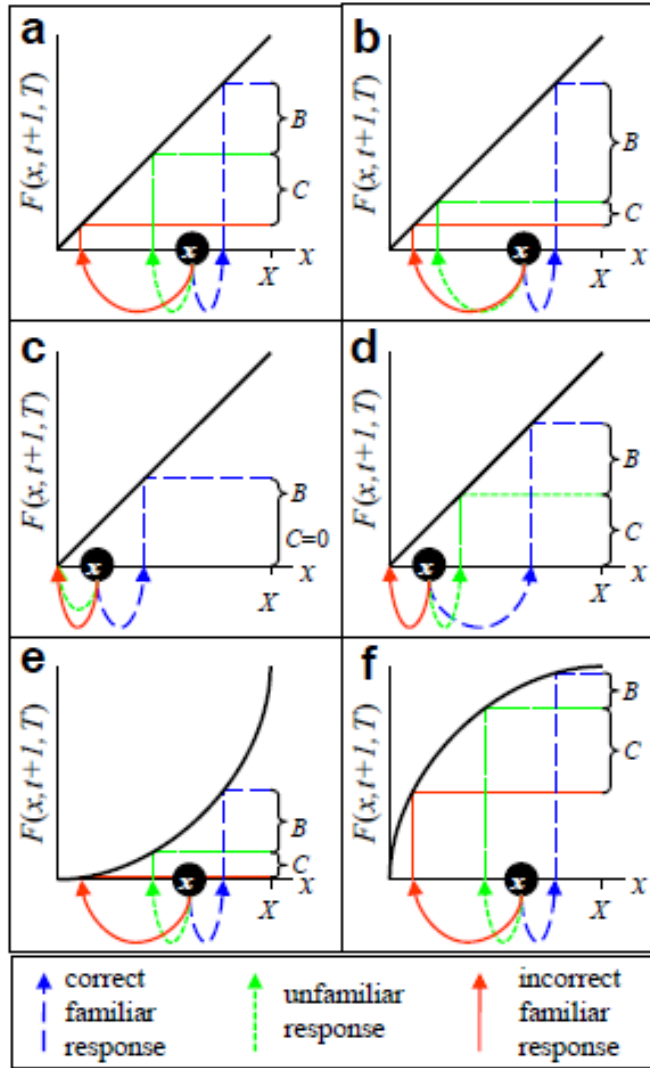
Figure 3-2. The probability of a correct response to a familiar stimulus (blue/medium gray bars), the probability of a response to an unfamiliar stimulus (green/light gray bars), and the probability of an incorrect response to a familiar stimulus (red/dark gray bars) as a function of memory size (with $N=10$). Each chart corresponds to a different value of the rate of stimulus change, ρ . These probabilities are $\omega(1-\beta^m)$, α^m , $1-\alpha^m-\omega(1-\beta^m)$, respectively.

Also, m^{**} is an increasing function of ν (i.e., $\partial m^{**} / \partial \nu > 0$). In order to explore the effects of ν , we need to know something about $F(x,t,T)$. While the precise analytical form of $F(x,t,T)$ is difficult to derive, it will generally be a non-decreasing function of x , given that we assume $\Phi(x)$ is a non-decreasing function of x . For pedagogical purposes in the next section we will pretend F assumes simple functional forms (e.g., a linear function), however this pretence will not affect most of the conclusions for more complicated non-decreasing forms of F .

Graphical Approach

Let us focus on an organism in state x' at age t (the black circles in Figure 3-3). If it employs an incorrect response to a familiar stimulus, its state changes to $i(x')$ (the red solid arrows in Figure 3-3). If it employs a response to an unfamiliar stimulus, its state changes to $u(x')$ (the green dotted lines and arrows). And if it employs a correct response to a familiar stimulus, its state changes to $c(x')$ (the blue dashed lines and arrows). We always have $i(x) \leq u(x) \leq c(x)$ because we assume $\pi_i < \pi_u < \pi_c$. In each case, the maximal future expected reproductive success is given by the value of $F(x,t+1,T)$ at $x=i(x')$, $x=u(x')$, or $x=c(x')$. From the positions of $c(x')$, $u(x')$, and $i(x')$, we can visualize the benefit, $B(x,t)$, and cost, $C(x,t)$, of memory use and thus isolate factors that will affect ν .

Figure 3-3. Schematics to show the costs and benefits of memory. (a) We start by assuming $F(x, t+1, T)$ is a linear function (although generally it is not). We assume that the state of our organism at age t is x' . A solid red arrow points to $i(x')$, the state acquired through an incorrect response to a familiar stimulus. A dotted green line and arrow points to $u(x')$, the state acquired through a response to an unfamiliar stimulus. A dashed blue line and arrow points to $c(x')$, the state acquired through a correct response to a familiar stimulus. The potential benefit $B(x',t)$ of being familiar with a stimulus is $F(c(x'),t+1,T) - F(u(x'),t+1,T)$ and potential cost $C(x',t)$ of being familiar with a stimulus is $F(u(x'),t+1,T) - F(i(x'),t+1,T)$.



(b) If π_u decreases, then $u(x')$ can decrease and B can increase while C decreases. (c) If the state of the organism is very low and the cost of living is very high ($\kappa > \pi_u$), then $i(x') = u(x') = 0$ while $c(x') > 0$. Consequently $C = 0$ and $B > 0$; that is, there are no costs to memory, whereas there are benefits—thus, the memory window should be maximally long. (d) If the cost of living is low ($\kappa < \pi_u$, see Appendix D), then $C > 0$ and $B > 0$ and maximal memory may not be favored. (e) If $F(x, t+1, T)$ is convex, the ratio of B to C increases (relative to the linear F function in (a)) and thus an increased memory window is favored. (f) If $F(x, t+1, T)$ is concave, the ratio of B to C decreases (relative to the linear F function in (a)) and thus a decreased memory window is favored.

We see that for constant values of correct and incorrect payoffs, π_c and π_i (i.e., $c(x')$ and $i(x')$ remain constant), as the value of the payoff for an unfamiliar stimulus, π_u , decreases (i.e., as $u(x')$ decreases), ν increases and longer memory windows are favored (compare Figures 3-3a and 3-3b). The quantity $\pi_u - \pi_i$ generally scales with the potential cost of memory, $C(x,t)$. The potential benefit of memory scales with $\pi_c - \pi_u$. It stands to reason that as the potential cost of memory decreases and the potential benefit of memory increases (π_u drops while π_c and π_i both stay constant), the optimal memory window should be longer.

We can also start to understand some of the effects of the state of the organism on memory. Given that we assume that the cost of living is high (i.e., $\kappa > \pi_u$), if the state of the organism is very low, then both $i(x')$ and $u(x')$ will be zero (remember these functions “chop” x into the range of 0 to X). This can be seen in Figure 3-3c, where the solid and dotted arrows point to the origin. Because $i(x')=u(x')=0$, this means that $C(x',t)=0$ and ν is infinite. As $\nu \rightarrow \infty$, we know that $m^{**} \rightarrow \infty$. Thus, when cost of living is high, organisms in a very low state should possess maximal memory. Again the idea is that employing a correct response to a familiar stimulus is the only way to survive, thus there is a premium placed on a maximal memory window. If we had assumed that the costs of living are lower (i.e., $\kappa < \pi_u$), then $u(x') > 0$ for an organism in a very low state and optimal memory is no longer necessarily maximal memory (see Figure 3-3d; in Appendix 5B, we explore the case when $\kappa < \pi_u$).

The benefit-to-cost ratio of memory, ν , also depends on the shape of the expected future reproductive value function, F . For instance, ν increases as F is made more convex

(compare Figure 3-3e with Figure 3-3a). Thus, optimal memory windows are longer for convex F functions and shorter for concave F functions (compare Figure 3-3f with Figure 3-3a). This prompts the question: what is the general shape of the F function at any arbitrary age of our organism? For one age the answer to this question is clear. At the end of the organism's lifetime $F(x,T,T)=\Phi(x)$ and we specify this final fitness function. In Figure 11, we see that a convex Φ leads to a longer optimal memory window at the end of the lifetime than a concave Φ (see Appendix 5C).

Convergence

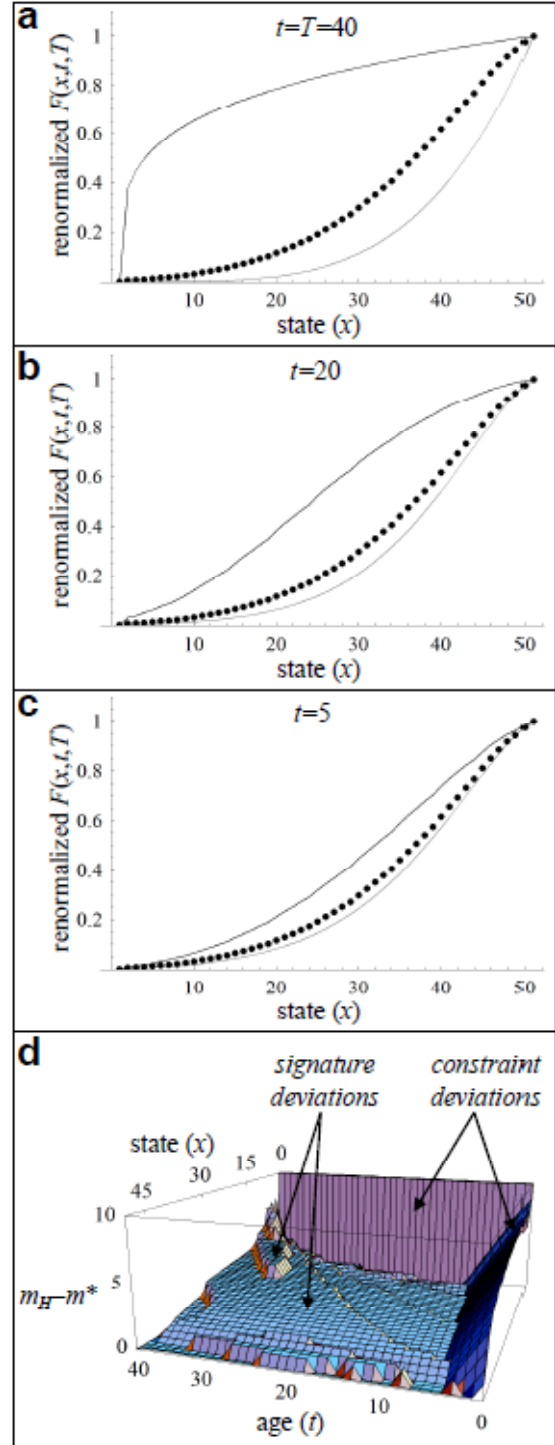
As mentioned, an analytical solution of the general form of $F(x,t,T)$ is difficult to derive. However, we do observe a convergence property in our model that is common to many dynamic programming routines (Houston and McNamara, 1999; Mangel and Clark, 1988). As we step backwards in time through repeated application of equation [3.1], $F(x,t,T)$ assumes a form (once renormalized) that is independent of the final fitness function, $\Phi(x)$. If we do not constrain m to remain below t (mathematically possible, but biologically unjustified), this function is also independent of time (this is a form of strong convergence, see (Houston and McNamara, 1999)). The convergence is shown in Figure 3-4 for the examples from Figures 3-1a and 3-1c. If we assume that F does strongly converge to some function, call it $H(x)$, then the proxy for optimal memory $m^{**}(x,t)$ is approximated by replacing $F(x,t,T)$ in v with $H(x)$ in equation [3.7]. Let us call this approximate optimal memory $m_H(x)$. We show an example of $H(x)$ in black dots in Figures 3-4a-c. For a given x value, $m^*(x,t) \approx m_H(x)$ for several different t values.

However, there are a few ages (t values) where the $m_H(x)$ approximation fails. First, for any state x , if $t < m_H(x)$, we know the memory must be smaller than the approximation

because the organism cannot remember stimuli from before it was born. Generally, for t values below $m_H(x)$, we find $m^*(x,t)=t$, that is, the optimal memory window is the maximum in the allowable range (incidentally, the constraints on m , which were ignored to compute $H(x)$, force $F(x,t,T)$ to deviate from $H(x)$ for the young ages $t < m_H(x)$). Second, for ages close to the end of the lifetime, $t \approx T$, there is a strong signature of $\Phi(x)$ on $F(x,t,T)$ (i.e., backwards convergence to $H(x)$ has not proceeded very far). If $\Phi(x)$ is more convex than $H(x)$, then the optimal memory window bends up at the end of the lifetime. However, if $\Phi(x)$ is more concave than $H(x)$, then the optimal memory window bends down at the end of the lifetime. In such a case, it would be optimal to *reduce* long-term recall at the end of the lifetime; that is, memory loss would be optimal.

We can now start to understand the age-related phases of optimal memory that we saw in Figure 3-1. The memory lift at the beginning of the lifetime occurs because the young age of the organism limits its recall. We call this a “constraint deviation” in Figure 3-4d, because the young organism is constrained by its age to remember less than $m_H(x)$. (The other constraint deviation is for organisms in low state where $m_H(x)$ is predicted to be infinite, but the organism must have finite memory). The memory plateau begins at ages larger than $m_H(x)$ and gently slopes towards the end of the lifetime. The deviations here we call “signature deviations” because the signature of the final fitness function is echoed in the optimal memory at the end of the organism’s lifetime, where Φ functions that are convex relative to $H(x)$ lift final optimal memory and relatively concave Φ functions (e.g., like that used to produce Figure 3-4d) depress final optimal memory.

Figure 3-4. (a) The final fitness function $\Phi(x)=100(x/X)^4$ is given in grey and $\Phi(x)=100(x/X)^{0.25}$ is given in black. These functions are normalized to range from 0 to 1. (b) The corresponding renormalized $F(x,t,T)$ functions at time $t=20$, generated through backwards iteration of equation [1] (with $X=50$, $N=10$, $T=40$, $\delta=0.1$, $\pi_i=5$, $\pi_u=10$, $\pi_c=15$, $\kappa=12$, and $\rho=0.1$) starting with $F(x,T,T)=100(x/X)^4$ (in grey) and $F(x,T,T)=100(x/X)^{0.25}$ (in black). (c) The corresponding renormalized $F(x,t,T)$ functions at $t=5$. Both functions are approaching an equilibrium function $H(x)$ given in black dots in (a)-(c). (d) Using $H(x)$ in place of $F(x,t+1,T)$ we calculate $m_H(x,t)$ by equation [7] for all x and t values. Here we show the difference between m_H and m^* to show the appropriateness of the m_H approximation. We see two “types” of deviations between m_H and m^* . The constraint deviations occur where the m_H approximation is outside of the allowed range for memory (e.g. $m_H(x,t)>t$) and the signature deviations occur when occur when m_H is near attainable memory values, but other memory values are optimal (these deviations are due to the fact that $H(x)$ is an inappropriate approximation for $F(x,t,T)$). In the case shown, the signature deviations (m_H-m^*) are positive, suggesting that there is optimal decline in memory for large values of t (note that we are dealing



with the concave final fitness function shown in black in (a)).

Discussion

Using a dynamic programming approach, we have suggested that optimal retrieval of memories should not be expected to be constant, but rather should change with the age and state of the organism. Thus, this model proposes that specific changes in memory (e.g., loss of memory retrieval towards the end of the lifetime) could be adaptive. Of course, these findings do not deny that there may be other non-adaptive factors or constraints that affect memory retrieval, such as the general deterioration of memory for older individuals due to inevitable break-down of neural function. Rather, this model simply states that it is not necessary to invoke such constraints and that an organism that maximizes its future reproductive output is expected to alter its retrieval of memories under certain circumstances.

We should note that, even within the context of our model, there are changes in memory retrieval that result from constraints. For instance, an organism is forced to have shorter memory windows when it is younger because it cannot have windows that exceed its age. Mathematically, we could iterate recursion [3.1] without constraining the memory window, m , to be less than the current age of the organism, t . If m can assume any positive value, we have a situation that an organism can “remember” events before it was born. This exercise, although biologically unrealistic, reveals that our convergence approximation, m_H , is a good approximation of the optimal m value all the way to the

youngest age, thus demonstrating that constraints keep the memory window small early in the lifetime.

However, this same mathematical exercise does not affect results from later in the lifetime. Thus, it is not constraints on m that lead to memory decline or increase towards the end of the lifetime. Why is optimal memory changing here? Within our model, memory generates a variance in payoffs. This means that the concavity of the maximal expected future reproductive success function, F , will be important. Specifically, the spread in payoffs are translated positively for convex F and negatively for concave F . Of course, F changes with age. Specifically, for the oldest age, F is given by Φ , the final fitness function, and as age decreases, F converges to another function, which we've called H . If Φ is more convex than H , then memory tends to increase towards the end of the lifetime. If Φ is more concave than H , then memory tends to decrease towards the end of the lifetime. When H does not have much curvature (e.g., Figure 3-4), then strongly convex or concave final fitness functions will tend to leave "signatures" on the optimal memory surface (gradual increase or gradual decline, respectively). The question remains whether a convex or concave final fitness function is more biologically realistic. If fitness returns consistently diminish (as opposed to accelerate) with energy reserves, then a concave function would be apt. In such a case, our model predicts that optimal memory should decline with age, an effect due to the signature of the fitness function rather than constraints on memory.

Many theoretical treatments on the evolutionary advantages of learning center on environmental variability (Arnold, 1978b; Bergman and Feldman, 1995; Cohen, 1991; Dukas, 1998; Feldman et al., 1996; Johnston, 1982; Kerr and Feldman, 2003; Mangel,

1990; McNamara and Houston, 1987; Shettleworth, 1998; Stephens, 1987; Stephens, 1991). Consistent with other models, our model shows that as stimuli in the environment become less reliable, optimal recall (at any age or state) tends to decrease. However, the rate of stimulus change, ρ , will also have a strong effect on the shape of H , the function to which the maximal expected future reproductive success function converges. As the concavity of H changes, this will affect the differences in concavity between Φ and H and thus will change the potential for optimal memory decline or increase (see above).

When the organism is in a low state (and the cost of living is high, which we assume in our model), maximal memory is optimal. Consistent with this prediction, Pravosudov & Clayton (2001) found that mountain chickadees maintained on a limited and unpredictable food supply made fewer errors on a cache recovery task than well-fed birds (see also Friedrich et al., 2004; and Orsini et al., 2004). Of course, the motivation of an organism with low energy levels and high costs of living might be very different than that of an organism with high energy levels and low costs of living, and cognitive performance may vary with motivation.

While this model provides some novel insights, there are a number of potential extensions that could be explored. First, reproduction in the model occurs only at the end of the lifetime (a form of semelparity). It is not difficult to incorporate reproduction throughout the lifetime (Houston and McNamara, 1999; Mangel and Clark, 1988). We explore an iteroparous extension to our model in Appendix 5D, where we show that optimal memory can still decay with age despite repeated individual bouts of reproduction. Thus, adaptive memory decay does not depend on an assumption of semelparity. Second, we assumed that the cost of living, κ , was constant. However, it has

been proposed that memory itself may be explicitly costly (Dukas, 1999; Kuhl et al., 2007; Mery and Kawecki, 2005), and memory performance is sensitive to imposed costs (Balda and Bednekoff 1997; Laughlin and Mendl 2004). We could easily include cost by writing κ as an increasing function of memory ($\kappa(m)$). We consider this model extension in Appendix 5E. While explicit costs to memory depress optimal values, we see that optimal memory can still decay with age. With explicit memory costs, we can also reorder the payoffs, such that using memory is always more valuable than sampling from scratch: $\pi_u < \pi_i < \pi_c$. Such an ordering could apply to the case where stimulus change leads to only a slight decrease in payoff for a remembered behavior. Again, we see that optimal memory can still decrease with age in such a case (Appendix 5E). We have found that our basic results are robust to alterations in several of our model assumptions (Appendices 5D and 5E).

Our model does offer a new perspective on the evolutionary reasons for deterioration of memory with age. The classic theories of senescence (Medawar, 1952; Williams, 1957) suggest that deterioration with age occurs either due to selection for genes that are beneficial early in life but detrimental later in life (antagonistic pleiotropy) or due to the inefficacy of selection to weed out deleterious mutations late in life when reproductive output decreases (mutation accumulation). These evolutionary explanations generally posit that, all else being equal, decay at any age is selected against (and thus senescence is either a side effect or a selectively elusive phenomenon). Some of the classic explanations of memory loss (e.g., synaptic decay) fit well within this classic theoretical framework. In contrast, our model claims that memory loss can in and of itself be

adaptive. We expect this to be the case when condition is poor, the individual is near the end of its lifetime, and fitness returns consistently diminish with energy reserves.

Furthermore, our model makes several specific predictions about changes in memory. First, and in agreement with other models, recall should decrease as the environment becomes less reliable (Kerr and Feldman, 2003; Mangel, 1990; McNamara and Houston, 1987). Second, retrieval should roughly increase with the cost of living (see Appendix 5B). Third, cost of living should interact with state of the organism. If the cost of living is high, organisms in a very poor state should have high memory; whereas if the cost of living is low, organisms in a very good state should have little to no memory. Fourth, (and unlike general theories of senescence) our model does not predict inevitable decline of memory. The behavior of memory length with age depends on the relationship between the state of the organism and its final fitness. However, if fitness shows diminishing returns with state, memory should decline gently with age for intermediate levels of stimulus reliability.

The validity of this model can be addressed empirically. For organisms at reproductive age, the relationship between energy (or fat) reserves and number of offspring is an empirical issue. The predicted changes in memory could be checked (perhaps within a comparative framework) against the observed concavity of this relationship. Individuals with high metabolic demands of living tend to demonstrate superior memory capability while in poor states (e.g. Friedrich et al., 2004; Orsini et al., 2004; Pravosudov and Clayton, 2001). It would be interesting to repeat such experiments with organisms that possess a low cost of living (where our model predicts the effect will be less dramatic). Given the rich history of empirical work on memory, models such as this one will be

important in both reconciling seemingly disparate observations and suggesting new avenues of research in the behavioral ecology of memory. The essential point underscored by this model is that learning is a dynamic process and critical components of that process, such as memory length, may change in optimal ways within the lifetime of the learner.

Chapter 4 – Choice in a variable environment: an experimental test of tracking, sampling, and memory length³

Introduction

The world is a variable place. Foraging animals need information to make decisions about what and where to forage, but animals are not omniscient; they face the classic foraging problem of making decisions with incomplete information. An additional complication to this problem is that once an animal gathers enough information to make a decision, the environment can change: so the best place to forage right now is not necessarily the best place four hours from now. Making good decisions in a variable environment depend upon a suite of cognitive capabilities. In a changing world, a forager needs to sample from possible patches or food sources and track the changes that are occurring. Gathering this information requires learning, and the later use of that information requires memory and making choices: the forager must decide among the potential options at any given point in the future.

Models predict cost/benefit tradeoffs to sampling and acquiring new information: there can be costs to the forager who does not sample the environment frequently enough, but there are also costs to foragers who sample too often or over-track the environment

³ This project was approved by the Institutional Animal Care and Use Committee of the University of Minnesota.

(Stephens, 1987) . A key aspect of when to acquire information is how frequently the environment changes. At one extreme, if the environment is completely persistent, there is nothing to track: the information currently held is accurate and will be accurate in the future. At the other extreme, if the environment changes randomly, tracking is pointless and random choice is best. At rates of change in between these extremes, however, the broad prediction is that animals should sample and acquire information more frequently as change increases: there are more changes to track. Some studies have looked at the effect of differing food amounts in options being tracked (e.g. Inman, 1990; Shettleworth et al., 1988; Tamm, 1987) , but the aspect of variability has not been as well-tested empirically.

Once they acquire information, animals must often make decisions about that information at varying time points after the time of the last sample. The passage of time is an important variable in decision making, but most investigators chose to ignore changing memory lengths. A few models of memory have tied choice to variable environments. The classic approaches use exponential weighting of past experiences, linear operators, and often some application of Bayesian analysis (e.g. Harley, 1981; Houston et al., 1982; Kacelnik et al., 1987; Killeen, 1981; McNamara and Houston, 1987; Valone, 2006). In a more variable environment, animals should choose the most recently encountered stimulus, and similarly in a more variable environment, animals should forget more quickly (Mangel, 1990; McNamara and Houston, 1985; McNamara and Houston, 1987). Kerr and Feldman came to similar conclusions in a model learning system focusing in heterogeneous environments (Kerr and Feldman, 2003). A related approach to memory length explicitly considers how foragers balance recent and past

information and choose appropriately (e.g. Devenport and Devenport, 1994). How does environment change affect the balance between recent and past information?

In this paper, we approach sampling and memory in varying environments through model and experiment. We first modeled a very simple two-stimulus task, where choosing one stimulus is rewarded with food, while choosing the second results in no food and a longer wait until the next choice. We created environmental change by reversing which stimulus is rewarded. We then tested our predictions using blue jays in an operant setting. By varying the rates of reversal of stimulus reward, we test hypotheses about how the jays should sample and learn in the different environments and how they should apply the information they have gathered. We hypothesize that 1) birds will sample more in environments which change more frequently, 2) birds in more variable environments will more closely track those changes, and 3) rates of change will interact with the passage of time when birds are tested on after differing retention intervals, with birds in more variable environments forgetting more quickly.

Model

At any given trial, the subject can select stimulus A or stimulus B. Only two situations are possible: A correct and B wrong, or A wrong, B correct. We call these the A correct and B correct states. We assume that a symmetric persistence process governs the transition between these two states, with q giving the probability of remaining in the same state (and $1-q$ the probability of switching correct states). It would seem that there is nothing to track in this situation because an animal can immediately recognize which state is true. For instance, if the subject samples stimulus A and fails to obtain a reward, then it should immediately switch to choosing stimulus B (because only one stimulus can be

correct at a time). It is a simple, but likely unrealistic, result that assumes animals can perfectly recognize the current state (A or B correct). Instead, we distinguish between the true state and the recognized state. We do this using an error rate, ε , which we define as the probability that the recognized state differs from the true state.

We solved this problem using dynamic optimization, first determining the optimal behavior in the final step. Let p be the subjective probability of the A correct state. In the final step, then, the optimal behavior is to choose A if $p \geq 1/2$, with a payoff of p , and to choose B if $p < 1/2$, with a payoff of $1-p$). The optimal payoff for the last step is a function of p , which we'll call $V_o(p)$. This takes the form of a simple V shape.

Solving the second to last step is a bit more complicated. The payoffs are the same (choose A, get p ; choose B, get $1-p$). Regarding the state variable, p , four things can happen (Table 4-1):

Event	Probability
A true and A recognized	$p(1 - \varepsilon)$
A true and B recognized	$p\varepsilon$
B true and A recognized	$(1 - p)\varepsilon$
B true and B recognized	$(1 - p)(1 - \varepsilon)$

Table 4-1. Possible outcomes concerning the state variable.

The animal, however can only observe two events: recognizing A or recognizing B, and these occur regardless of the actual choice the subject makes (Table 4-2).

Event	Probability
A recognized	$p(1 - \varepsilon) + (1 - p)\varepsilon$
B recognized	$p\varepsilon + (1 - p)(1 - \varepsilon)$

Table 4-2. Events observable by the subject.

If A is recognized, we can find the probability that A is true by Bayes Theorem:

$$P(A_{correct}|A_{recognized}) = \frac{P(A_{rec}|A_{cor})P(A_{cor})}{P(A_{rec}|A_{cor})P(A_{cor}) + P(A_{rec}|B_{cor})P(B_{cor})} \quad [4.1]$$

$$p' = \frac{(1-\varepsilon)p}{(1-\varepsilon)p + \varepsilon(1-p)} \quad [4.2]$$

Similarly

$$P(A_{correct}|B_{recognized}) = \frac{\varepsilon p}{\varepsilon p + (1-\varepsilon)(1-p)} \quad [4.3]$$

Now we add change. If the animal knows p , the probability that A is correct in the next step is:

$$Q(p) = p'' = q \cdot p' + (1 - q)(1 - p') \quad [4.4]$$

Let $V_A(p)$ be the value of choosing A when p is true and let $V_B(p)$ be the value of choosing B when A is true. So:

$$V_A(p) = p + [(1 - \varepsilon)p + \varepsilon(1 - p)] \cdot V_0 \left[Q \left(\frac{(1-\varepsilon)p}{(1-\varepsilon)p + \varepsilon(1-p)} \right) \right] + [p\varepsilon + (1 - p)(1 - \varepsilon)] \cdot V_0 \left[Q \left(\frac{\varepsilon p}{\varepsilon p + (1-\varepsilon)(1-p)} \right) \right] \quad [4.5]$$

$$V_B(p) = (1 - p) + [(1 - \varepsilon)p + \varepsilon(1 - p)] \cdot V_0 \left[Q \left(\frac{(1-\varepsilon)p}{(1-\varepsilon)p + \varepsilon(1-p)} \right) \right] + [p\varepsilon + (1 - p)(1 - \varepsilon)] \cdot V_0 \left[Q \left(\frac{\varepsilon p}{\varepsilon p + (1-\varepsilon)(1-p)} \right) \right] \quad [4.6]$$

We give $V_1(p)$ as the optimal behavior when two steps remain, and

$$V_1(p) = \max[V_A(p), V_B(p)] \quad [4.7]$$

We use V_0 to calculate V_1 , and since the logic for the second to last step applies generally, we can calculate $V_n(p)$ from $V_{n-1}(p)$ in the same way. We solved the problem numerically by finding $V_i(p)$ for a list of p values, and then found $V_{i+1}(p)$ using the previously calculated values of $V_i(p)$.

Figure 4-1 represents a typical result from the model. In general, memory loses value more quickly when persistence, q , is low (and the value starts out at a lower point). With changes in error rate, we see few changes in overall effects, except when error rates are extremely high (e.g. $\varepsilon=0.5$) and the world is completely persistent ($q=1.0$). In this case,

the surface becomes a much more pronounced V-shape across all of the time steps. When we look at the rate of change (Fig. 4-2), the value divided by time step, as time steps increase, the rate converges upon the persistence value, q , as long as the error rate, ε , is relatively low. This simple two-choice model, where the value of one item implies the value of the second item, qualitatively agrees with the memory models of the past: increasing change decreases the value of memory across time.

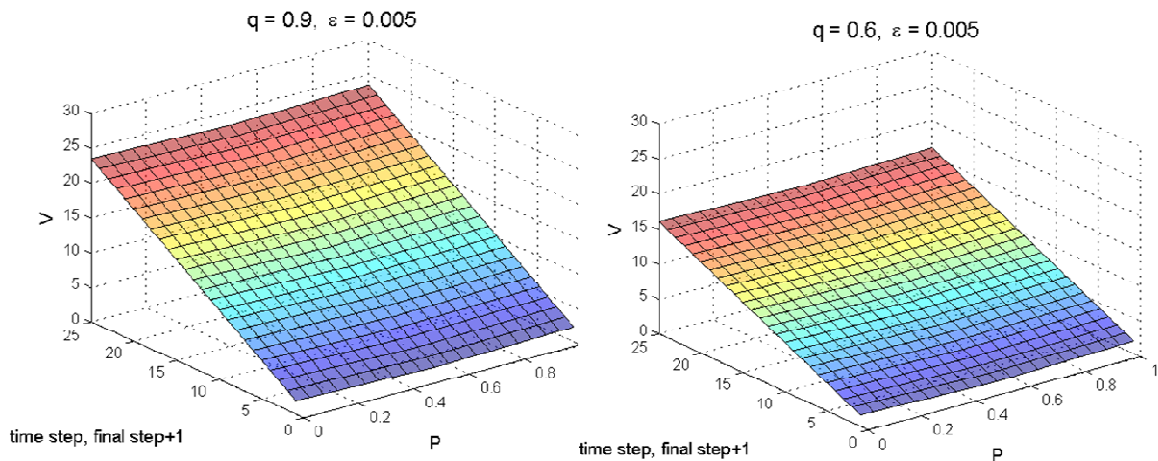


Figure 4-1. The effect of increasing change (lower values of q) is to depress the value of memory.

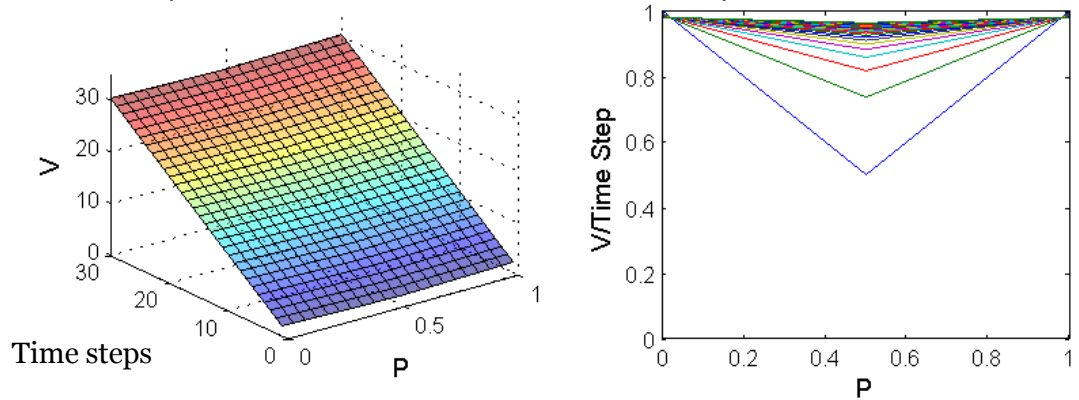


Figure 4-2. Results for $q=0.98186$ and $\varepsilon=0.01$. We chose the q value based upon the rate of change in the experiment in this paper. In 2b, the sharpest V-shape represents the beginning time step, with each of the next time steps following in order with more and more shallow V-shapes.

Materials & Methods

Subjects & Experiment Apparatus

Eight adult blue jays (*Cyanocitta cristata*) of unknown sexes and mixed experimental histories served as subjects. These birds were captured from the wild as nestlings and were hand reared in the laboratory. All birds were between three and ten years old. We tested the birds in operant boxes (Fig 4-3). Each box was equipped with two stimulus projectors in the front of the box and one stimulus projector in the rear. We placed a hinged perch attached to a micro switch under each stimulus projector so that we could determine when the subject hopped to each perch. Choice of a stimulus was registered as a hop to the perch beneath the projector. We attached a pellet feeder to each box, and food was delivered to a cup in the front of each box, located between the two stimulus

projectors. A small light in the food cup (a magazine light) flashed with each food delivery. A computer controlled all the equipment, ran the programmed the experimental contingencies and collected the resulting data. Prior to the experiment, we trained the birds to perform in the operant boxes using shaping techniques.

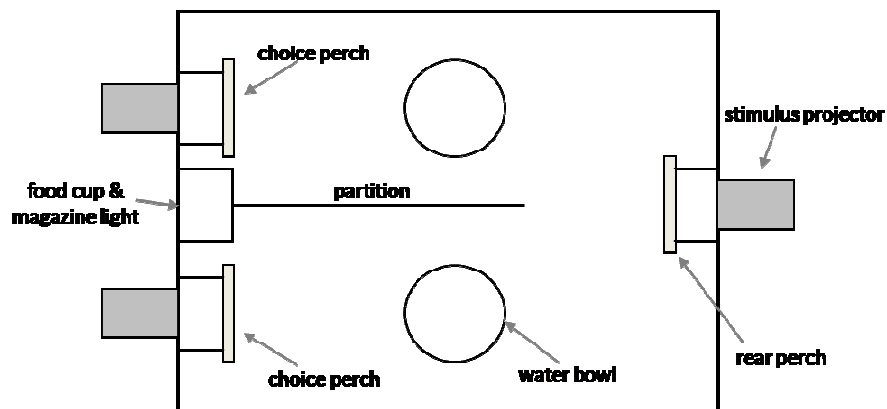


Figure 4-3. The experimental apparatus.

We used a modified closed economy system, in which birds received all of their daily food within the experimental trials. This allows each bird to experience the consequences of their choices: good choices resulted in more food, poor choices in less. Birds performed in the experiment for 8 hours each day, and remained in their boxes for 23 hours each day. We removed the birds for one hour each afternoon to clean the boxes, save the data, and weigh the bird and check its health. To ensure that each bird stayed in good health, we set a minimum amount of food a bird would receive each day. If a bird did not achieve the minimum on any day, we fed the bird the difference between the amount gained in the box during trials and the minimum allowance.

Basic Experimental Design

We first designed an environment in which we can vary how frequently the relationship between two stimuli changes. We did this using a two-choice reversal learning task. A choice of one stimulus, say stimulus A, is rewarded with food, while a choice of the second stimulus, stimulus B, results in no food and a longer time interval to wait until the next trial. The relative values of these stimuli switched according to a prior probability, e.g. stimulus A is correct for x minutes, then stimulus B becomes correct. Around this basic framework of a varying environment, we built experimental contingencies to test additional hypotheses concerning tracking and decision making over different time intervals. To do this, birds experienced three phases of trials (Fig. 4-4). In the first phase, the birds performed in the reversal learning task, where the two variability treatments were enacted by changing the rate at which the reversals occurred. After experiencing this variability, birds entered the second phase, during which we tested tracking behavior. Here, birds learned a final reversal of stimuli values to a set criterion. Upon reaching the criterion, the third phase of trials began, during which we tested decision-making. During these trials, the birds performed a simple light-following task, into which we inserted single probe trials to test the choices the birds made about the stimuli.

We randomized the order of the high and low treatments, such that the order was balanced across the eight birds. Within each change treatment, we randomized the retention intervals tested in four blocks of five intervals each (one of each interval in each block). We did this to ensure that any order effects within the probes tested would be balanced across the retention intervals.

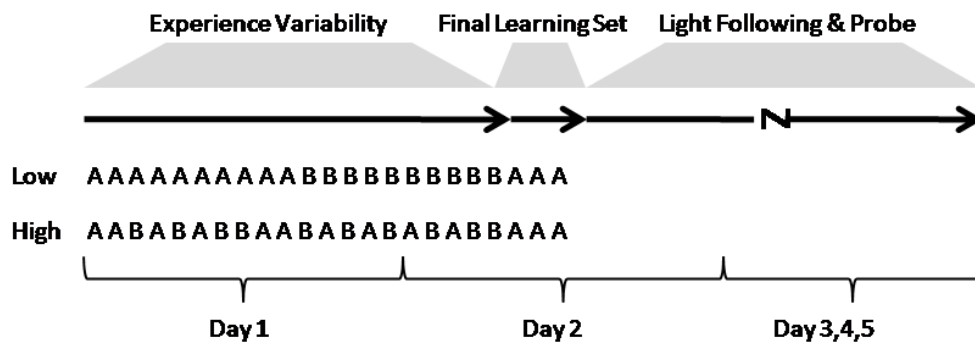


Figure 4-4. Design of the experimental trials. Birds experience variability in the form of stimulus reward reversals. In the low change treatment, birds experience stimulus A correct for five hours, then stimulus B is correct for the next five elapsed hours of trials. In the high change treatment, stimulus A and B reverse reward on average every 30 min. Stimuli rewards reverse for a final time in both treatments during the Day 2, and this final learning set continues until the bird reaches a criterion of accuracy. Birds then began a simple task following white lights and we inserted a single probe trial to test one of 5 retention intervals. At the end of the experimental day of the probe trial, the bird began the process again with a new stimuli pair.

Testing a Single Retention Interval

Each stimulus was a combination of a color (red, blue, green, or yellow), and a shape. We always presented the stimuli as pairs, with the computer randomly determining the side each stimulus appeared on in each trial. Each learning trial would begin with the rear light flashing, upon which, the bird would hop to the rear perch. Once the bird was on the rear perch, the computer would present the stimuli flashing on the front projectors. A choice would be registered by a hop to the perch below the stimulus, at which point both

stimulus lights would extinguish and the bird would either be fed one pellet, or fed no pellets. A correct choice would be followed by an intertrial interval (ITI) of 40 seconds, while an incorrect choice would be followed by a longer ITI of 120 seconds. At the end of the ITI, the rear light would flash, beginning a new trial. A trial would abort and begin anew following any 20 minute period without hops to perches.

The rewarded stimulus was chosen by the computer, and remained the same until changed according to the assigned treatment. In the low change treatment, the rewarded stimulus changed only once, six hours into the first day of each retention interval test. In the high change treatment, the time between changes was pulled from a distribution with a mean of 60 min, with a new time chosen after each change interval. The birds performed learning trials for eight hours during the first day of an interval test and then during the morning of the next day.

Following 2.5 hours of learning trials on the second day, the rewarded stimulus would switch for a final time for both treatments. The rewarded stimulus would then remain the same for the duration of the retention interval test. We then tested the birds for learning the new stimulus values to a criterion of five correct rewarded trials in a row. We chose this criterion based on pilot data. Immediately upon meeting the criterion for the final set of learning, the computer would switch the bird to a simple light following task. In this task, a trial would begin with the rear white light flashing. The bird would hop to the rear perch and then only one of the front white lights would flash. Birds hopping to the perch in front of the flashing light would be rewarded. This task would continue for eight hours each day until the retention interval was reached. At the retention interval, the bird would be presented with a single probe trial, with rewards

equal to the final learning set they reached criterion on. Following the probe trial, birds would continue in the simple light following task until the end of the experimental day.

On the following day the bird would begin the learning trials for the next retention interval. To avoid interference between tests of retention intervals, we rewired the stimulus lights to present a new pair following the completion of each retention interval test. Series of pairs were randomly assigned before the experiment, such that each box would cycle through fourteen possible combinations before repeating the sequence.

Dependent Measures

In this simple two stimulus system, we defined sampling as choosing the stimulus that is not currently rewarded. To address our hypothesis on sampling, we assess choice of the non-rewarded option during the 10 hours of trials during which subjects experience variability for each iteration of a retention interval test. To successfully track the environment, an individual must quickly switch to more profitable new behaviors and abandon old behaviors. With a stimulus switch, a bird must choose the newly rewarded stimulus and stop choosing the old stimulus, which no longer provides a reward. To analyze tracking behavior, we focused on the second day of learning each new stimulus pair. Our measure of tracking is how quickly the subjects acquired the last stimulus switch during the final learning set. We measure acquisition by the number of trials required to reach the criterion described above. Finally, to assess memory and decision making across the five retention intervals we tested, we compare the choice made in each probe trial to the correct choice as per the last experience of the final learning set.

Results

Sampling the Environment

Our first analyses ask how the birds sampled the experimental environments. To do this, we examined choice behavior during the first day of learning trials for each new stimulus pair. We predicted that birds in the high change treatment sample the non-rewarding stimulus more frequently than birds in the low change treatment. At the largest scale of analysis, we can look across each entire first learning day at how often birds sampled the unrewarded stimulus. Birds sampled more frequently in the high change treatment (unrewarded choices/day $\bar{X} \pm SD$, high change= 103.60 ± 32.60 , low change = 46.36 ± 20.68 unrewarded choices/day; Paired t-test: $t_7 = -8.987$, $P < 0.0001$).

Birds could be sampling, but they could also be making overrun errors, which occur when stimulus values change and an individual continues to choose the now unrewarding stimulus. Our next analysis, therefore, takes a more restrictive view by considering sampling choices made after initial learning has occurred. We compared sampling during periods in which no stimulus changes occurred. To control for time of day effects, we chose the closest 20-trial interval with no stimulus value change occurring at around 1400 on the first day of learning each new stimulus pairing. Even with the more stringent criteria, we still found a statistically significant difference between the environmental variability treatments, with birds sampling more frequently in the high change treatment than in the low change treatment (errors in 20 trial sample, $\bar{X} \pm SD$, high change= 2.833 ± 1.08 , low change= 0.532 ± 0.383 ; Paired t-test: $t_7 = -7.30$, $P < 0.0002$).

Tracking Change in the Environment With Learning

To assess how well the birds tracked the environment, we compared how many trials birds required to reach the learning criterion during the final learning set for each stimulus pairing. A quick adjustment to this final change indicated “good tracking,” while a slower adjustment indicated less effective tracking. As predicted, birds reached the learning criterion more quickly during the high change treatment (mean trials to criterion $\bar{X} \pm SD$, high change = 17.218 ± 7.620 , low change = 31.696 ± 16.534 ; Paired t-test: $t_7 = 3.16$, $P < 0.016$).

Birds in the high change treatment tracked the environment more efficiently, requiring a smaller number of trials, but across the 10 previous hours of experience they also had many more changes to track. In a highly variable environment, like our high change treatment, foragers must track to maintain an adequate intake rate. Did our birds track the environment well enough to keep up with their rate of food intake under the low change environment? We measured the average rate of food gain for each bird during the first day of learning trials under each of the environmental change treatments. Birds obtained a higher rate of food gain under the low change treatment (g/min $\bar{X} \pm SD$, high change = 0.0126 ± 0.0026 , low change = 0.0179 ± 0.0013 ; Paired t-test: $t_7 = 5.899$, $P = 0.0006$). A higher level of tracking did not completely offset the losses incurred due to the changes in the variable environment (theoretical max is 0.0279 g/min, both low and high significantly differ from this maximum, $P < 0.05$).

Memory and Choice of the Last Experienced Stimulus

Once birds have sampled and tracked the environment, and acquired the information about the environment, they must make decisions using that information at some point

in the future. We assessed their decisions in probe trials across the five retention intervals. We first analyzed how birds chose according to their most recent experience. A correct choice according to their most recent experience would be to choose the same stimulus that was rewarded during their final learning set. Figure 4-5 shows the mean choices for each treatment and retention interval. Birds in the low change treatment chose no differently after 1 hour than they did after 4 hours; however birds chose much less accurately at 24 hrs. In the high change treatment, birds made more accurate choices after 1 hr, less accurate choices after 4 hrs, and their best choices of all following 24 hrs. In both treatments, we saw similar choices after 48 and 72 hrs, an increase in performance from 24 hrs for birds in the low change treatment, but a decrease in performance from 24 hrs for birds in the high change treatment.

We analyzed these choices using a repeated measures ANOVA and found a significant interaction between the main effects of treatment and interval ($F_{(4,28)} = 2.77, P < 0.048$). However, there were no significant effects of treatment ($F_{(1,7)} = 2.60, P = 0.15$) or interval ($F_{(4,28)} = 0.65, P = 0.63$) alone. We analyzed this further using contrasts. Within the high change treatment the dramatic change from 4 hrs to 24 hrs is statistically significant ($F_{(1,7)} = 14.00, P = 0.0072$). The next change, from 24 hrs to 48 hrs, was statistically significant ($F_{(1,7)} = 7.06, P = 0.032$). The most striking difference was between the performance of birds in the low and in the high change treatments after 24 hrs. This difference was significant, with birds choosing correctly in the high change treatment, but incorrectly in the low change treatment ($F_{(1,7)} = 11.67, P = 0.011$). Following a full day, and importantly, and presumably overnight consolidation of the previous day's memories, birds in the high change treatment chose the most recent stimulus now more frequently. But, birds in the low change treatment chose the opposite. They chose the

stimulus, which in their case had not been rewarded most recently, but which had been rewarded for the longer period of time that previous day: the first 2.5 hrs of that second day of learning those stimuli, before the final learning set began, as well as the afternoon of the first day of learning (see figure 4-4).

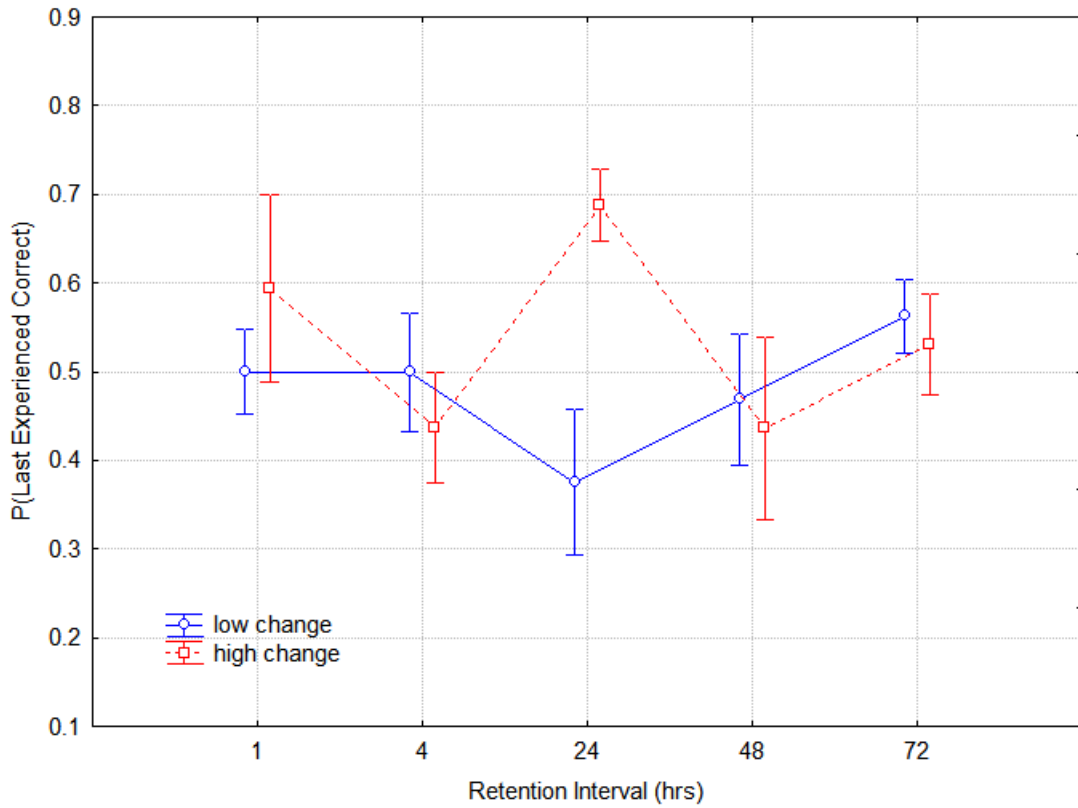


Figure 4-5. The mean performance of birds when making choices relative to the last stimulus they experienced as correct, over five retention intervals. Each set of birds was exposed to different rates of environmental change (low or high). Significant interactions were observed between the low and high change treatments at 24 hours, and within the high change treatment between 4 and 24 hours. The bars reflect the standard deviations.

Balancing Recent with Past Information

The previous analysis scored the probe choice as correct when it was the same as the most recently correct stimulus from that environment. However, birds may be making choices based on a different criterion. Birds may place more weight on past information. What is past information in this experiment? Across the entire experience variability phase of trials, the birds received equal time during trials when each stimulus was correct. However one stimulus remained “correct” during the entire overnight period between the two learning days, and we used this stimulus as the past information. The assumption behind this choice is based on a mechanistic explanation that information that has been consolidated overnight might hold more weight in a choice than the most recently encountered information. And indeed, the most striking change in choice in our first analysis was at 24 hrs.

What will birds choose when this past information conflicts with recent information? We analyzed this by looking at a deviation from expected choices if the birds were following the past information. To obtain these data, we gathered all instances of choice where the past information (the overnight correct stimulus) and recent information (most recently correct stimulus) conflicted. We counted the number of times each bird chose according to the past information correct, when the past and recent options conflicted, and then divided those totals by what we would expect by chance for the number of conflicted trials available in each case. The variable we analyzed was the difference between the observed choices and the expected choices, divided by the expected choices.

Environmental change interacted significantly with the interval at which the choice was

made (Repeated measures ANOVA: $F_{(4,28)} = 3.68$, $P = 0.0157$). Neither main effect was significant (change: $F_{(1,7)} = 0.207$, $P > 0.66$; interval: $F_{(4,28)} = 1.42$, $P = 0.254$).

Figure 4-6 shows these data. After one hour, birds in the low change chose the past/overnight stimulus as we would expect by chance, while birds in the high change treatment chose the past/overnight stimulus more than we would expect by chance. The difference between the two treatments at this interval was significant ($F_{(1,7)} = 7.716$, $P = 0.0273$). In a highly variable environment, when the most recent experience conflicted with the experience from the past (i.e. information that was consolidated over the previous night) birds appeared to rely on that past information. After 4 hrs, birds in both treatments appear to be choosing between the contradicting information in roughly the same way, with a slight bias towards the stimulus that was correct during the more distant past.

At 24 hrs the results mirrored the basic results in figure 4-5: when in the high change environment, birds after 24 hrs chose the stimulus that was rewarded most recently, and they did so even when this conflicted with what was the best stimulus from the past (the overnight correct stimulus before the second day of learning). But, when in the low change treatment, birds chose the previously rewarded stimulus rather than the most recently rewarded correct stimulus, and they only did this after 24 hrs of waiting. The two environmental change treatments did not quite differ significantly at 24 hrs ($F_{(1,7)} = 3.91$, $P = 0.088$). The shift between 1 and 24 hrs was not statistically significant for either treatment (low change: $F_{(1,7)} = 4.2$, $P = 0.0796$; high change $F_{(1,7)} = 4.16$, $P = 0.0807$).

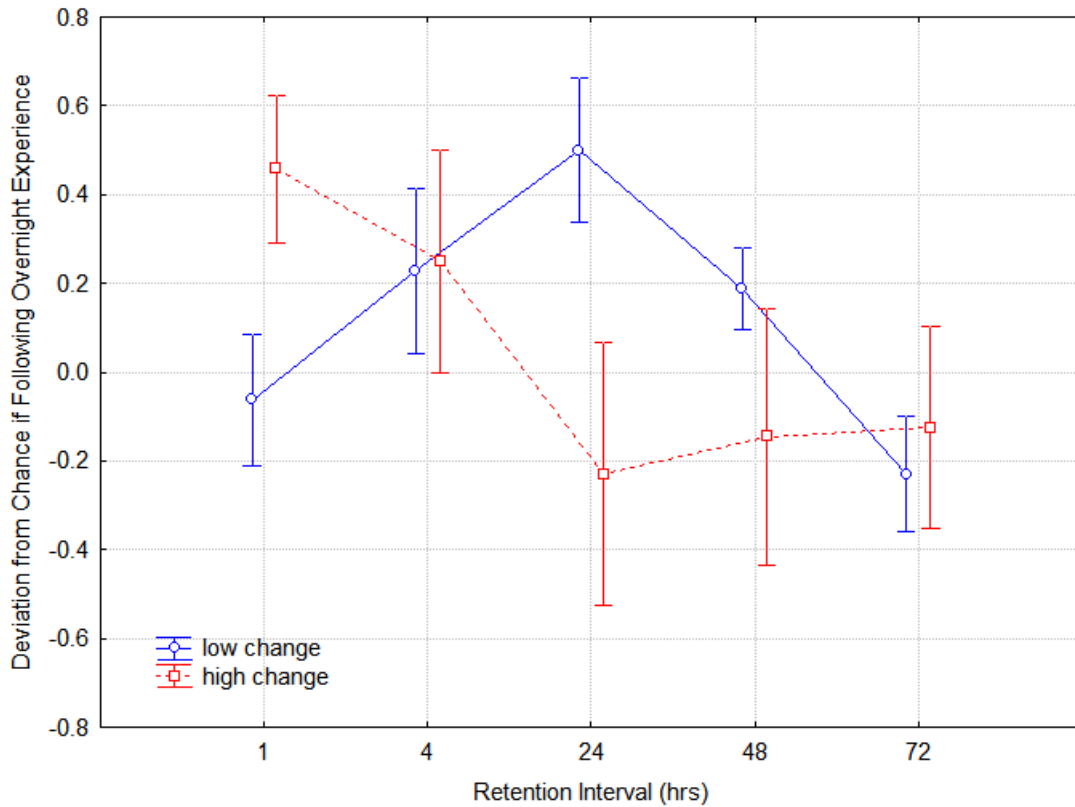


Figure 4-6. The deviation from expected for birds choosing according to the stimulus that was correct after the first day of learning, when that stimulus conflicted with the most recently correct stimulus. Values above zero on the Y-axis are tend towards choosing the overnight stimulus (past information) more than expected by chance, while negative values are moving towards choosing the most recently-rewarded stimulus (recent information) more than expected by chance.

Discussion

Birds' choices in a low change environment differed from their choices in an environment that changed much more frequently, and these differences came in a number of forms. First, birds sampled the unrewarded option more frequently when under a high rate of change compared to when under a low rates of change. Second, we found significant differences in birds tracking the environment: birds learned the new values of reversed set of stimuli more quickly in the high change treatment than in the low change treatment. However, while much theoretic work suggests an orderly decay of memory, with birds making more accurate choices under low change than under high change, this is not what we found. Instead we found significant interactions between the rates of change of the environment and the birds' choices after the five different retention intervals.

As predicted, birds sampled more under high rates of change. We expected birds exposed to high rates of change to choose the alternate stimulus more frequently, simply because the environment was changing and sampling the alternate versus making an error immediately following a change were indistinguishable. However we also found a significant difference between the high and low change treatments when we looked at choice in groups of trials between the reversals, after learning had been acquired. When the stimulus value changed frequently, birds had more to track: they had to learn those changes and switch their behavior. While a few studies have tested the Stephens (1987) tracking model, only one tested more than a single level of environmental change. Hummingbirds also sampled more in more varying environments than in less varying environments (Tamm, 1987).

The experimental system described in this paper differs from the traditional tracking models because we had only two states: good and bad, and the current value of the alternate stimulus could always be inferred from the value of the currently chosen stimulus. Birds however, still demonstrated sampling behavior, which is what one might expect when the current value of a stimulus does not perfectly predict its future value. From a purely psychological point of view, finding a difference in learning speed ability—what we have called tracking—should be expected. From an interference point of view, birds under a low rate a change should have a more difficult time learning any given reversal because of a greater susceptibility to proactive interference: the associative strength built up by many trials with the original stimulus values can interfere with quickly learning reversed stimuli values. Animals faced with serial reversals in learning become quicker at learning those reversals (e.g. Mackintosh, 1974). Not finding a difference between our two rates of change would have been rather surprising from this point of view. However, giving animals reversal learning tasks continuously in a closed-economy system, rather than in one-trial per day, or short sessions, is not a common approach. This type of interference, which would occur for a constantly foraging animal, is probably quite biologically relevant.

Our memory results were not what we expected. We generally expect animals to choose according to their most recent experience, because the most recent state of the environment is the best predictor of the current state of the environment. However, we found that choice across time interacted with the experienced rate of change. The most dramatic difference was at 24hrs, with both a significant difference between treatments and a significant positive increase between 4hrs and 24 hrs when birds were under the high rate of change. The results for 48 and 72 hours were more along our predictions:

birds performing marginally better under 48 hours with no difference under 72 hours (pilot testing indicated no retention after 72 hours under no change). Because of the striking difference at 24 hours, the birds were likely weighting past experience differently in these two treatments. At 24 hrs, birds in the low change treatment chose towards the opposite of their most recent experience, which was the stimulus that was correct for the rest of the second day and the second overnight. A temporal weighted average, such as that described by Devenport and Devenport (1994; 1997) qualitatively predicts such a switch, however, according to their model, it should have occurred within a few dozen of trials, rather than a full days of trials later.

Clearly, the jays balance past and recent information in some way. When the most recent experience conflicts with what was correct overnight, one can easily predict that the variability within that recent experience affects what the birds eventually choose. And that is what we found. When birds were in a minimally variable environment, at 1 hr following their most recent experience, neither that stimulus nor the overnight stimulus was preferred. This preference changed after 24 hrs, when the birds switched from preferring the stimulus that was correct during their first overnight to choosing the most recently correct stimulus, which at that point had an overnight memory consolidation period. The switch at 24 hours probably reflected the same pattern from the first analysis of choice (according to the most recently correct). Birds here averaged across a longer period of time. Their most recent experience was relatively brief compared to the previous 5 hours of trials contradicting that experience. And across time, the choice became more evident after 24 hrs.

We saw a different pattern for birds in a highly variable environment. At 1hr following their most recent experience, these birds more heavily relied upon their overnight memories. However, after 24 hrs, and another night, birds in the high change treatment tended to choose according to their most recent experience. The rates of change in this treatment occurred on average every 60 min. Following any period of multiple hours, the best strategy is to choose randomly because it is impossible to predict the state of the environment. The birds, however, did not choose randomly and choice did interact with time. A possible explanation is that in a variable world, the stimulus memory with more strength is what will be chosen. In short retention intervals, the birds relied on their past experience of what was correct overnight. Unlike birds in the low change treatment, the birds in the high change treatment did not disregard the most recent information when it came to 24 hr intervals: this information presumably gained more strength as time progressed, and another overnight interval occurred.

While time may degrade the value of remembered information in a linear fashion, as theory predicts, the birds in our experiment did not respond as if this were the case. Although we can only speculate as to the mechanisms, processing of the memories within the brain occurring overnight, i.e.. consolidation, likely had a strong effect on the differential weighting of memories and this interacted with the passage of time as we tested with our retention intervals. This type of interaction should be more closely investigated in how information from past events is weighted as animals make decisions.

Literature Cited

- Anderson MC, Ochsner KN, Kuhl B, Cooper J, Robertson E, Gabrieli SW, Glover GH, Gabrieli JDE, 2004. Neural systems underlying the suppression of unwanted memories. *Science* 303:232-235.
- Anderson RW, 1995. Learning and evolution: a quantitative genetics approach. *Journal of Theoretical Biology* 175:89-101.
- Arnold SJ, 1978a. The evolution of a special class of modifiable behaviors in relation to environmental pattern. *American Naturalist* 112:415-427.
- Arnold SJ, 1978b. The evolution of a special class of modifiable behaviors in relation to environmental pattern. *The American Naturalist* 112:415-427.
- Bergman A, Feldman MW, 1995. On the evolution of learning: Representation of a stochastic environment. *Theoretical Population Biology* 48:251-276.
- Bolles RC, 1970. Species-specific defense reactions and avoidance learning. *Psychological Review* 77:32-48.
- Borenstein E, Feldman MW, Aoki K, 2008. Evolution of learning in fluctuating environments: when selection favors both social and exploratory learning. *Evolution* 62:586-602.
- Bouton ME, 1994. Conditioning, Remembering, and Forgetting. *Journal of Experimental Psychology-Animal Behavior Processes* 20:219-231.
- Bouton ME, Moody EW, 2004. Memory processes in classical conditioning. *Neuroscience and Biobehavioral Reviews* 28:663-674.
- Brembs B, Heisenberg M, 2001. Conditioning with compound stimuli in *Drosophila melanogaster* in the flight simulator. *Journal of Experimental Biology* 204:2849-2859.
- Brembs B, Wiener J, 2006. Context and occasion setting in *Drosophila* visual learning. *Learning & Memory* 13:618-628.
- Clark CW, 1993. Dynamic models of behavior: an extension of life history theory. *Trends in Ecology & Evolution* 8:205-209.
- Cohen D, 1991. The equilibrium distribution of optimal search and sampling effort of foraging animals in patchy environments. In: *Adaptation in Stochastic Environments* (Yoshimura J, Clark CW, eds). Berlin: Springer-Verlag; 173-191.
- Cohen G, 1996. Memory and learning in normal aging. In: *Handbook of the Clinical Psychology of Ageing* (Woods RT, ed). London: Wiley; 43-58.

- Cook M, Mineka S, 1990. Selective associations in the observational conditioning of fear in rhesus monkeys. *Journal of Experimental Psychology* 16:372-389.
- Cook M, Mineka S, Wolkenstein B, Laitsch K, 1985. Observational conditioning of snake fear in unrelated rhesus monkeys. *Journal of Abnormal Psychology* 94:591-610.
- Craik FIM, Jennings JM, 1992. Human memory. In: *The handbook of aging and cognition* (Craik FIM, Salthouse TA, eds). Hillsdale NJ: Erlbaum; 51-110.
- Davis HP, Small SA, Stern Y, Mayeux R, Feldstein SN, Keller FR, 2003. Acquisition, recall, and forgetting of verbal information in long-term memory by young, middle-aged, and elderly individuals. *Cortex* 39:1063-1091.
- Davis HP, Trussell LH, Klebe KJ, 2001. A ten-year longitudinal examination of repetition priming, incidental recall, free recall, and recognition in young and elderly. *Brain and Cognition* 46:99-104.
- Devenport LD, Devenport JA, 1994. Time-dependent averaging of foraging information in least chipmunks and golden-mantled ground squirrels. *Animal Behaviour* 47:787-802.
- Devenport LD, Hill T, Wilson M, Ogden E, 1997. Tracking and averaging in variable environments: a transition rule. *Journal of Experimental Psychology: Animal Behavior Processes* 23:450-460.
- Dobrzecka C, Szwejkowska G, Konorski J, 1966. Qualitative versus directional cues in two forms of differentiation. *Science* 153:87-89.
- Dukas R, 1998. Evolutionary ecology of learning. In: *Cognitive ecology: the evolutionary ecology of information processing and decision making* (Dukas R, ed). Chicago: University of Chicago Press; 129-164.
- Dukas R, 1999. Costs of memory: Ideas and predictions. *Journal of Theoretical Biology* 197:41-50.
- Feldman MW, Aoki K, Kumm J, 1996. Individual versus social learning: Evolutionary analysis in a fluctuating environment. *Anthropological Science* 104:209-231.
- Friedrich A, Thomas U, Muller U, 2004. Learning at different satiation levels reveals parallel functions for the cAMP-protein kinase A cascade in formation of long-term memory. *Journal of Neuroscience* 24:4460-4468.
- Garcia J, Koelling RA, 1966. Relation of cue to consequence in avoidance learning. *Psychonomic Science* 4:123-124.
- Guilford T, Dawkins MS, 1991. Receiver psychology and the evolution of animal signals. *Animal Behaviour* 42:1-14.

- Guo A, Li L, Xia SZ, Feng CH, Wolf R, Heisenberg M, 1996. Conditioned visual flight orientation in *Drosophila*: Dependence on age, practice, and diet. *Learning & Memory* 3:49-59.
- Harley CB, 1981. Learning the evolutionarily stable strategy. *Journal of Theoretical Biology* 89:611-633.
- Hirvonen H, Ranta E, Rita H, Peuhkuri N, 1999. Significance of memory properties in prey choice decisions. *Ecological Modelling* 115:177-189.
- Houston AI, Kacelnik A, McNamara JM, 1982. Some learning rules for acquiring information. In: *Functional ontogeny* (McFarland D, ed). Boston: Pitman; 140-191.
- Houston AI, McNamara JM, 1999. *Models of Adaptive Behavior: An Approach Based on State*. Cambridge: Cambridge University Press.
- Inman A, 1990. Foraging decisions: the effects of conspecifics and environmental stochasticity (Ph.D. dissertation): Oxford University.
- Johnston TD, 1982. Selective costs and benefits in the evolution of learning. *Advances in the Study of Behavior* 12:65-106.
- Kacelnik A, Krebs JR, Ens B, 1987. Foraging in a changing environment: an experiment with starlings (*Sturnus vulgaris*). In: *Foraging* (Commons ML, Kacelnik A, Shettleworth SJ, eds). Hillsdale, New Jersey: Lawrence Erlbaum Associates; 63-87.
- Karlin S, Lieberman U, 1974. Random temporal variation in selection intensities: case of large population size. *Theoretical Population Biology* 6:355-382.
- Karlin S, Lieberman U, 1975. Random temporal variation in selection intensities: one-locus two-allele model. *J Math Biol* 2:1-17.
- Kerr B, Feldman MW, 2003. Carving the cognitive niche: Optimal learning strategies in homogeneous and heterogeneous environments. *Journal of Theoretical Biology* 220:169-188.
- Killeen PR, 1981. Averaging theory. In: *Quantification of steady-state operant behavior* (Bradshaw CM, Szabodi E, Lowe CF, eds). Amsterdam: Elsevier; 21-34.
- Kraemer PJ, Golding JM, 1997. Adaptive forgetting in animals. *Psychonomic Bulletin & Review* 4:480-491.
- Kuhl BA, Dudukovic NM, Kahn I, Wagner AD, 2007. Decreased demands on cognitive control reveal the neural processing benefits of forgetting. *Nature Neuroscience* 10:908-914.

- Kulahci IG, Dornhaus A, Papaj DR, 2008. Multimodal signals enhance decision making in foraging bumble-bees. *Proc R Soc Lond B* 275:797-802.
- Lewis AC, 1986. Memory constraints and flower choice in *Pieris-Rapae*. *Science* 232:863-865.
- Loftus EF, Loftus GR, 1980. On the permanence of stored information in the brain. *American Psychologist* 35:409-420.
- Logue AW, 1979. Taste aversion and the generality of the laws of learning. *Psychological Bulletin* 86:276-296.
- LoLordo VM, 1979. Selective Associations. In: *Mechanisms of learning and motivation: a memorial volume to Jerzy Konorski* (Dickinson A, Boakes RA, eds). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.; 367-398.
- Lupien SJ, de Leon M, De Santi S, Convit A, Tarshish C, Nair NPV, Thakur M, McEwan BS, Hauger RL, Meaney MJ, 1998. Cortisol levels during human aging predict hippocampal atrophy and memory deficits. *Nature Neuroscience* 1:69-73.
- Mackintosh NJ, 1974. *The Psychology of Animal Learning*. New York: Academic Press.
- Mangel M, 1990. Dynamic information in uncertain and changing worlds. *Journal of Theoretical Biology* 146:317-332.
- Mangel M, Clark CW, 1988. *Dynamic Modeling in Behavioral Ecology*. Princeton: Princeton University Press.
- McEwen BS, Sapolsky RM, 1995. Stress and Cognitive Function. *Current Opinion in Neurobiology* 5:205-216.
- McGaugh JL, Roozendaal B, 2002. Role of adrenal stress hormones in forming lasting memories in the brain. *Current Opinion in Neurobiology* 12:205-210.
- McNamara JM, Houston AI, 1985. Optimal foraging and learning. *Journal of Theoretical Biology* 117:231-249.
- McNamara JM, Houston AI, 1987. Memory and the efficient use of information. *Journal of Theoretical Biology* 125:385-395.
- Medawar PB, 1952. *An unsolved problem of biology*. London: H. K. Lewis.
- Mery F, Kawecki TJ, 2002. Experimental evolution of learning ability in fruit flies. *Proc Natl Acad Sci U S A* 99:14274-14279.
- Mery F, Kawecki TJ, 2004a. The effect of learning in experimental evolution of resource preference in *Drosophila melanogaster*. *Evolution* 58:757-767.

- Mery F, Kawecki TJ, 2004b. An operating cost of learning in *Drosophila melanogaster*. *Animal Behaviour* 68:589-598.
- Mery F, Kawecki TJ, 2005. A cost of long-term memory in *Drosophila*. *Science* 308:1148.
- Neissen U, 2004. Memory development: new questions and old. *Developmental Review* 24:154-158.
- Orsini C, Buchini F, Conversi D, Cabib S, 2004. Selective improvement of strain-dependent performances of cognitive tasks by food restriction. *Neurobiology of Learning and Memory* 81:96-99.
- Papaj DR, 1994. Optimizing learning and its effect on evolutionary change in behavior. In: *Behavioral Mechanisms in Evolutionary Ecology* (Real LA, ed). Chicago: University of Chicago Press; 133-153.
- Papaj DR, Prokopy RJ, 1989. Ecological and evolutionary aspects of learning in phytophagous insects. *Annual Review of Entomology* 34:315-350.
- Plotkin HC, Odling-Smee FJ, 1979. Learning, change, and evolution: an enquiry into the teleonomy of learning. *Advances in the Study of Behavior* 10:1-41.
- Pravosudov VV, Clayton NS, 2001. Effects of demanding foraging conditions on cache retrieval accuracy in food-caching mountain chickadees (*Poecile gambeli*). *Proceedings of the Royal Society of London Series B-Biological Sciences* 268:363-368.
- Punzo F, Chavez S, 2003. Effect of aging on spatial learning and running speed in the shrew (*Cyrtotis parva*). *Journal of Mammalogy* 84:1112-1120.
- Rowe C, 1999. Receiver psychology and the evolution of multicomponent signals. *Animal Behaviour* 58:921-931.
- Rozin R, Kalat JW, 1971. Specific hungers and poison avoidance as adaptive specializations of learning. *Psychological Review* 78:459-486.
- Salthouse TA, 1996. The processing-speed theory of adult age differences in verbal tasks. *Psychological Review* 103:403-428.
- Schacter DL, 1999. The seven sins of memory: insights from psychology and cognitive neuroscience. *American Psychologist* 54:182-203.
- Seeman TE, McEwan BS, Singer BS, Albert MS, Rowe JW, 1997. Increase in urinary cortisol excretion and memory declines: MacArthur studies of successful aging. *Journal of Clinical Endocrinology and Metabolism* 82:2458-2465.
- Seligman MEP, 1970. On the generality of the laws of learning. *Psychological Review* 77:406-418.

- Shafir S, Roughgarden J, 1996. The effect of memory length on individual fitness in a lizard. In: Adaptive Individuals in Evolving Populations: Models and Algorithms (Belew RK, Mitchell M, eds). Reading, MA: Addison-Wesley; 173-181.
- Shettleworth SJ, 1972. Constraints on learning. *Advances in the Study of Behavior* 4:1-68.
- Shettleworth SJ, 1979. Constraints on conditioning in the writings of Konorski. In: Mechanisms of Learning and Memory: A Memorial Volume to Jerzy Konorski (Dickinson A, Boakes RA, eds). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.
- Shettleworth SJ, 1998. *Cognition, Evolution, and Behavior*. Oxford: Oxford University Press.
- Shettleworth SJ, Krebs JR, Stephens DW, Gibbon J, 1988. Tracking a fluctuating environment: a study of sampling. *Animal Behaviour* 36:87-105.
- Shukitt-Hale B, McEwen JJ, Szprengiel A, Joseph JA, 2004. Effect of age on the radial arm water maze - a test of spatial learning and memory. *Neurobiology of Aging* 25:223-229.
- Slobodkin LB, Rapoport A, 1974. A optimal strategy of evolution. *Quarterly Review of Biology* 49:181-200.
- Spear NE, 1973. Retrieval of memory in animals. *Psychological Review* 80:163-194.
- Spreng M, Rossier J, Schenk F, 2002. Spaced training facilitates long-term retention of place navigation in adult but not in adolescent rats. *Behavioural Brain Research* 128:103-108.
- Squire LR, 1989. On the course of forgetting in very long term memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 15:241-245.
- Staddon JER, 1983. *Adaptive behavior and learning*. New York: Cambridge University Press.
- Stephens DW, 1987. On economically tracking a variable environment. *Theoretical Population Biology* 32:15-25.
- Stephens DW, 1989. Variance and the value of information. *The American Naturalist* 134:128-140.
- Stephens DW, 1991. Change, regularity, and value in the evolution of animal learning. *Behavioral Ecology* 2:77-89.
- Stephens DW, 1993. Learning and behavioral ecology: incomplete information and environmental predictability. In: *Insect Learning: Ecological and Evolutionary Perspectives* (Papaj DR, Lewis A, eds). New York: Chapman and Hall; 195-217.

- Stephens DW, Krebs JR, 1986. Foraging Theory. Princeton: Princeton University Press.
- Tamm S, 1987. Tracking varying environments: sampling by hummingbirds. *Animal Behaviour* 35:1725-1734.
- Thorndike EL, 1932. The fundamentals of learning. New York: Teachers College Press.
- Thorpe WH, 1963. Learning and instinct in animals, 2nd ed. London: Methuen.
- Tomsic D, Romano A, Maldonado H, 1998. Behavioral and mechanistic bases of long-term habituation in the crab *Chasmagnathus*. In: *Molecular and Cellular Mechanisms of Neuronal Plasticity*; 17-35.
- Valone TJ, 2006. Are animals capable of Bayesian updating? An empirical review. *Oikos* 112:252-259.
- Williams GC, 1957. Pleiotropy, natural selection, and the evolution of senescence. *Evolution* 11:398-411.
- Wylie GR, Foxe JJ, Taylor TL, 2008. Forgetting as an active process: an fMRI investigation of item-method-directed forgetting. *Cerebral Cortex* 18:670-682.
- Xia SZ, Liu L, Feng CH, Guo AK, 1997. Nutritional effects on operant visual learning in *Drosophila melanogaster*. *Physiology & Behavior* 62:263-271.
- Yanai S, Okaichi Y, Okaichi H, 2004. Long-term dietary restriction causes negative effects on cognitive functions in rats. *Neurobiology of Aging* 25:325-332.
- Zelinski EM, Burnight KP, 1997. Sixteen-year longitudinal and time lag changes in memory and cognition in older adults. *Psychology and Aging* 12:503-513.

Appendix 1 — Modeling the Evolution of Learning versus Non-Learning⁴

To begin, we introduce some notation that simplifies our presentation. First, we use A and B to represent the two types of media (orange and pineapple in our experiment). Second, we use the notation QA to mean that quinine is paired with media type A in the experience phase, similarly QB means that we paired quinine with media type B in the consequence phase. Third, we use the notation A^* to mean that the experimenter rears eggs from media type A in the consequence phase; similarly we use B^* to mean that we rear eggs from media type B . To simplify the terminology, we say that the A^* is the ‘A-best’ condition, similarly B^* is the B-best condition.

Next, we use this notation to define parameters that represent the fixity of the stimulus-consequence and action-consequence relationships. Let q measure the fixity of the stimulus-consequence relationship, specifically $P(A^*|QB)=P(B^*|QA)=q$. In words, q is the conditional probability that pairing with quinine in the experience phase predicts the media that flies should avoid in the consequence phase.

We use p to represent the fixity of action-consequence relationship. Specifically, let p be the probability that the A-best condition applies for any given realization of the consequence phase. For example, if $p=1.0$ it is always best to lay eggs of media type A , whereas if $p=0.5$ the best place to lay varied unpredictably from one generation to the next. We remark we lose no generality by defining p in terms of the A-best condition,

⁴ This appendix provides the mathematical details for the model described in Chapter 1.

because types A and B are arbitrary. In practice, this means that we define type A to be the type that is ‘best’ most frequently, i.e. $P(A^*) \geq P(B^*)$, implying that $P(A^*) \geq 0.5$.

Now we consider two types of flies: a non-learner who always tries to oviposit on A and a learner who oviposits on A if quinine was paired with B in the experience phase, but oviposits on B if quinine was paired with A in the experience phase. We assume that a female lays N eggs in the consequence phase. In addition, we assume that a female makes some oviposition errors so that she cannot lay 100% of her eggs in her preferred media. Instead she lays $1-\varepsilon$ of her eggs in the media she “prefers” and ε in the media she “intends” to avoid; so ε is the error rate. Finally we assume that a proportion r of the eggs a female lays in the “best” media survive to reproduce, while none survive to reproduce when they are laid on the “worst” media. Within a generation, there are four possible events, as shown in Table A1-1.

Event			Non-Learner		Learner	
Experience Phase	Consequence Phase	Probability	Behavior	Fitness	Behavior	Fitness
QA	A*	$p(1-q)$	Prefer A	$r(1-\varepsilon)N$	Prefer B	$r\varepsilon N$
QA	B*	$(1-p)q$	Prefer A	$r\varepsilon N$	Prefer B	$r(1-\varepsilon)N$
QB	A*	pq	Prefer A	$r(1-\varepsilon)N$	Prefer A	$r(1-\varepsilon)N$
QB	B*	$(1-p)(1-q)$	Prefer A	$r\varepsilon N$	Prefer A	$r\varepsilon N$

Table A1-1. Contingencies and fitnesses for learners and non-learners.

From this we can calculate the fitnesses of the two types. When fitness varies temporally (from one generation to the next), we calculate the geometric mean fitness of the two types (Karlin and Lieberman, 1974; Karlin and Lieberman, 1975). The fitness of the non-learner is

$$p(1-q) \ln[r(1-\varepsilon)N] + (1-p)q \ln(r\varepsilon N) + pq \ln[r(1-\varepsilon)N] + (1-p)(1-q) \ln(r\varepsilon N) \quad [\text{A1.1}]$$

Which simplifies to

$$P \ln[r(1-\varepsilon)N] + (1-p) \ln(r\varepsilon N) \quad [\text{A1.2}]$$

The reliability term, q , cancels out because the non-learner ignores the pairing with quinine. Similarly, the fitness of the learner is

$$p(1-q) \ln(r\varepsilon N) + (1-p)q \ln[r(1-\varepsilon)N] + pq \ln[r(1-\varepsilon)N] + (1-p)(1-q) \ln(r\varepsilon N) \quad [\text{A1.3}]$$

Which simplifies to

$$Q \ln[r(1-\varepsilon)N] + (1-q) \ln(r\varepsilon N) \quad [\text{A1.4}]$$

Here, the frequency of the A -best state cancels out because the learner's fitness depends on whether the quinine cue reliably predicts the best media. The only difference between the two simplified expressions is the presence of p or of q , thus learning should be favored whenever $q > p$. The learning and non-learning traits will be neutral whenever $q = p$. This includes the so-called absolute fixity case— $q = p = 1$ —and the completely random case— $q = p = 1/2$.

Appendix 2 — Modeling Prepared Learning to Color Versus Odor⁵

This model is logically similar to the model described in Appendix 1. However, here, flies experience pairings of two modalities of stimuli: odor, which we will denote using roman letters, and color, denoted by greek letters. There are two odors and two colors. To make this explicitly the same as the experiment, we can say that for the odors stimulus A is amyl acetate and stimulus B is benzaldehyde, and that for the colors, α is aqua, and β is blue.

As with the previous model, we imagine a learning scenario where individuals use quinine (Q), as an aversive unconditioned stimulus and the odor and color as potential conditioned stimuli. As with the experiment described in chapter 1, flies experience two phases: an experience phase and a consequence phase.

There are four possible pairings with quinine in the experience phase:

$$Q+A\alpha \text{ or } Q+B\beta \text{ or } Q+A\beta \text{ or } Q+B\alpha$$

The probabilities of quinine being placed with A or B, or α or β , are independent. We can define these probabilities as $P(Q+A) = a$, and $P(Q+B) = 1-a$; and $P(Q+\alpha) = b$, and $P(Q+\beta) = 1-b$.

Now we define the possible outcomes in the consequence phase (Table A2-1).

⁵ This appendix provides the mathematical details for the model described in Chapter 2.

Roman/Color Best		Greek/Odor Best
$P(A \text{ best} Q+A\alpha)$	Neither Predict	$P(\alpha \text{ best} Q+A\alpha)$
$P(B \text{ best} Q+B\beta)$		$P(\beta \text{ best} Q+B\beta)$
$P(B \text{ best} Q+A\alpha)$	Both Predict	$P(\beta \text{ best} Q+A\alpha)$
$P(A \text{ best} Q+B\beta)$		$P(\alpha \text{ best} Q+B\beta)$
$P(B \text{ best} Q+A\beta)$	Color (greek) predicts, but odor (roman) does not	$P(\alpha \text{ best} Q+A\beta)$
$P(A \text{ best} Q+B\alpha)$		$P(\beta \text{ best} Q+B\alpha)$
$P(A \text{ best} Q+A\beta)$	Odor predicts, but color does not	$P(\beta \text{ best} Q+A\beta)$
$P(B \text{ best} Q+B\alpha)$		$P(\alpha \text{ best} Q+B\alpha)$

Table A2-1. Table of all possible outcomes in the consequence phase.

These conveniently simplify into 2 independent probabilities. First, the probability that olfactory stimuli predict the best environment in the consequence phase, given any previous pairing in the experience phase, $P(\text{olfactory best} | \text{any pairing})$. And then the probability that color stimuli predict the best environment, $P(\text{color best} | \text{any pairing})$.

We can assign a probability to the reliability of each stimulus with regards to the quinine pairing predicting the best environment, where O is the probability odor predicts the best environment, and C is the probability color predicts the best environment. As with

the model for the evolution of learning versus nonlearning (in appendix 1), we define the number of eggs oviposited by the female in the consequence phase as N , the error rate for the female as ε (with $1-\varepsilon$ being the eggs she lays in the substrate she “prefers”), and r being the proportion of eggs which survive to adulthood (with zero surviving on the “bad” media).

There are 16 different combinations possible of experience phase pairings and consequence phase outcome (Table A2-2).

Experience Phase	Probability of Pairings	Consequence Phase	Probability of "Best"	Selective Learning to Roman	Selective Learning to Greek
Q+A α	ab	Roman only predicts	$O(1-C)$	$r(1-\varepsilon)N$	$r\varepsilon N$
Q+B β	$(1-a)(1-b)$				
Q+A β	$a(1-b)$				
Q+B α	$(1-a)b$				
Q+A α	ab	Greek only predicts	$C(1-O)$	$r\varepsilon N$	$r(1-\varepsilon)N$
Q+B β	$(1-a)(1-b)$				
Q+A β	$a(1-b)$				
Q+B α	$(1-a)b$				
Q+A α	ab	Both Predict	CO	$r(1-\varepsilon)N$	$r(1-\varepsilon)N$
Q+B β	$(1-a)(1-b)$				
Q+A β	$a(1-b)$				
Q+B α	$(1-a)b$				
Q+A α	ab	Neither Predict	$(1-C)(1-O)$	$r\varepsilon N$	$r\varepsilon N$
Q+B β	$(1-a)(1-b)$				
Q+A β	$a(1-b)$				
Q+B α	$(1-a)b$				

Table A2-2. Possible pairings and outcomes in the two stimulus modality system.

We can now calculate the fitnesses for each of the two types of learners. As in appendix 1, we do this using the geometric mean. The fitness for the roman only learner (learning about odor only) is

$$\begin{aligned}
& ab O(1-C)\ln r(1-\varepsilon)N + (1-a)(1-b) O(1-C)\ln r(1-\varepsilon)N + a(1-b) O(1-C)\ln r(1-\varepsilon)N + (1- \\
& a)b O(1-C)\ln r(1-\varepsilon)N + ab C(1-O)\ln r\varepsilon N + (1-a)(1-b) C(1-O)\ln r\varepsilon N + a(1-b) C(1- \\
& O)\ln r\varepsilon N + (1-a)b C(1-O)\ln r\varepsilon N + ab CO \ln r(1-\varepsilon)N + (1-a)(1-b) CO \ln r(1-\varepsilon)N + \\
& a(1-b) CO \ln r(1-\varepsilon)N + (1-a)b CO \ln r(1-\varepsilon)N + ab (1-C)(1-O)\ln r\varepsilon N + (1-a)(1-b) (1- \\
& C)(1-O)\ln r\varepsilon N + a(1-b) (1-C)(1-O)\ln r\varepsilon N + (1-a)b(1-C)(1-O)\ln r\varepsilon N
\end{aligned}$$

The fitness for the greek-only learner (learning about color only) is

$$\begin{aligned}
& ab O(1-C)\ln r\varepsilon N + (1-a)(1-b) O(1-C)\ln r\varepsilon N + a(1-b) O(1-C)\ln r\varepsilon N + (1-a)b O(1- \\
& C)\ln r\varepsilon N + ab C(1-O)\ln r(1-\varepsilon)N + (1-a)(1-b) C(1-O)\ln r(1-\varepsilon)N + a(1-b) C(1-O)\ln \\
& r(1-\varepsilon)N + (1-a)b C(1-O)\ln r(1-\varepsilon)N + ab CO \ln r(1-\varepsilon)N + (1-a)(1-b) CO \ln r(1-\varepsilon)N + \\
& a(1-b) CO \ln r(1-\varepsilon)N + (1-a)b CO \ln r(1-\varepsilon)N + ab (1-C)(1-O)\ln r\varepsilon N + (1-a)(1-b) (1- \\
& C)(1-O)\ln r\varepsilon N + a(1-b) (1-C)(1-O)\ln r\varepsilon N + (1-a)b(1-C)(1-O)\ln r\varepsilon N
\end{aligned}$$

Appendix 3 – Derivation of the Dynamic

Programming Recursion⁶

In order to derive recursion [3.1], we list out the possible events that can occur at time step t to an organism in state x . Let **U** be the event that our organism is unfamiliar with the stimulus it experiences. Let **C** be the event that our organism is familiar with the stimulus and employs the correct behavior (i.e., the correct behavior to the stimulus has not changed since its most recent recording in the organism’s memory). Let **I** be the event that our organism is familiar and employs an incorrect behavior (i.e., the correct behavior to the stimulus has changed since its most recent recording in the organism’s memory). We let the dummy event variable, **A**, stand for one of the stimulus encounter events; that is $\mathbf{A} \in \{\mathbf{U}, \mathbf{C}, \mathbf{I}\}$. Let \mathbf{X}_x be the event that the organism’s state is x , \mathbf{T}_t be the event that an organisms age is t and \mathbf{M}_m be the event that an organism’s memory window is m .

Consider that our organism in state x at age t has memory size m , where this memory window may not be optimal. After this time step we assume the organism adjusts its memory window optimally (i.e., we assume $m(x, \tau) = m^*(x, \tau)$ for all $\tau \geq t+1$). We call this organism the “nearly optimal” learner. Let the expected future reproductive success of the nearly optimal learner be given by $F_m(x, t, T)$. In order to calculate F_m , we let $EF[\mathbf{A} | \mathbf{X}_x \cap \mathbf{T}_t]$ be the expected future reproductive success for the nearly optimal learner if event **A** occurs (given that the organism’s state is x and its age is t). We let $\Pr\{\mathbf{A} | \mathbf{M}_m\}$

⁶ This appendix describes the derivation of the dynamic programming recursion in Chapter 3.

be the probability of event \mathbf{A} , given our organism has a memory window of size m . By the law of total expectation, we must have:

$$F_m(x, t, T) = \sum_{\mathbf{A} \in \{\mathbf{U}, \mathbf{C}, \mathbf{I}\}} EF[\mathbf{A} | \mathbf{X}_x \cap \mathbf{T}_t] \Pr\{\mathbf{A} | \mathbf{M}_m\}, \quad [\text{A3.1}]$$

Our nearly optimal learner becomes a “completely optimal” learner if it picks the memory window size m (where $0 \leq m \leq t$) that maximizes $F_m(x, t, T)$ (that is, if it chooses $m = m^*(x, t)$). Thus, we can express the maximal expected future reproductive success of our completely optimal learner in state x at age t as

$$F(x, t, T) = \max_{\substack{m \\ 0 \leq m \leq t}} F_m(x, t, T). \quad [\text{A3.2}]$$

Now we simply need to derive the expectations and the probabilities in equation [A1]. Since we assume that the nearly optimal learner behaves optimally from time $t+1$ onwards, the expectations are simply

$$EF[\mathbf{U} | \mathbf{X}_x \cap \mathbf{T}_t] = (1 - \delta)F(u(x), t + 1, T), \quad [\text{A3.3}]$$

$$EF[\mathbf{C} | \mathbf{X}_x \cap \mathbf{T}_t] = (1 - \delta)F(c(x), t + 1, T), \quad [\text{A3.4}]$$

$$EF[\mathbf{I} | \mathbf{X}_x \cap \mathbf{T}_t] = (1 - \delta)F(i(x), t + 1, T). \quad [\text{A3.5}]$$

For example, if the stimulus is unfamiliar, the organism receives a payoff of π_u and pays a metabolic cost of κ (with the constraint that the next state is between 0 and X). Thus, if the organism survives from age t to age $t+1$ (which occurs with probability $1 - \delta$), the

organism changes from state x at age t to state $u(x)=chop(x+\pi_t-\kappa;0,X)$ at age $t+1$. Since the organism is assumed to remember optimally from $t+1$ onwards, the expected future reproductive success at $t+1$ is the maximal expected future reproductive success, which is given by $F(u(x),t+1,T)$. Thus, given that a dead organism has zero future reproductive success (and given that death occurs with probability δ) the expected reproductive success is $EF[\mathbf{U} | \mathbf{X}_x \cap \mathbf{T}_t] = (1-\delta)F(u(x),t+1,T) + \delta(0) = (1-\delta)F(u(x),t+1,T)$. The other EF values are derived similarly.

Now we turn to the probabilities in equation [A3.1]. We assume that every time step a stimulus is chosen randomly from the N stimuli, so the probability that the currently experienced stimulus does not occur in any one of the last m time steps must be $(N-1)/N$. Since we assume that each stimulus occurs independently of previous stimuli, the probability that the current stimulus does not occur in all the previous m time steps is

$$\Pr\{\mathbf{U} | \mathbf{M}_m\} = \left(\frac{N-1}{N}\right)^m. \quad [\text{A3.6}]$$

Equation [A3.6] gives the probability that the organism does not remember the current stimulus (i.e., event \mathbf{U} occurs) given that the organism has a memory window of size m .

We know that without any memory, the stimulus cannot be familiar and therefore $\Pr\{\mathbf{C} | \mathbf{M}_0\} = \Pr\{\mathbf{I} | \mathbf{M}_0\} = 0$. To calculate $\Pr\{\mathbf{C} | \mathbf{M}_m\}$ and $\Pr\{\mathbf{I} | \mathbf{M}_m\}$ when $m > 0$, it helps to condition on other events. Let \mathbf{F} be the event that the stimulus is familiar. Since \mathbf{F} is the complement of \mathbf{U} , we must have

$$\Pr\{\mathbf{F} | \mathbf{M}_m\} = 1 - \left(\frac{N-1}{N}\right)^m. \quad [\text{A3.7}]$$

Let \mathbf{S}_r be the event that the stimulus was most recently recorded in memory r time steps before the current time step (thus, \mathbf{S}_r is only defined for $r \leq m$). Given that an organism is dealing with a familiar stimulus and has a memory size of m , the probability that the most recently recorded encounter with the current stimulus happened r time steps back is

$$\Pr\{\mathbf{S}_r | \mathbf{F} \cap \mathbf{M}_m\} = \frac{N^{m-r}(N-1)^{r-1}}{N^m - (N-1)^m}. \quad [\text{A3.8}]$$

To derive equation [A3.8], let us consider the memory window of size m as being constructed of m “slots,” ordered from 1 (the most recent memory) to m (the most distant memory). There are a total of $N^m - (N-1)^m$ sequences of stimuli in memory where the current stimulus occurs in at least one of the m slots. How many of those sequences have the current stimulus occurring most recently in slot r ? For this to occur, we must have any stimulus except for the current stimulus in each of the slots from slot 1 to slot $r-1$ (a total of $(N-1)^{r-1}$ combinations for these slots), whereas we can have any stimulus in each of the slots from slot $r+1$ to slot m (a total of N^{m-r} combinations for these slots). Note we must have the current stimulus in slot r (a single “combination” for this slot). Thus, the total number of combinations for the current familiar stimulus to be located most recently in the r^{th} slot is $(N-1)^{r-1}N^{m-r}$ and equation [A3.8] follows. Given that our organism with memory length m experienced a familiar stimulus most recently r time steps back, the probabilities that the behavior employed is correct is

$$\Pr\{\mathbf{C} | \mathbf{S}_r \cap \mathbf{F} \cap \mathbf{M}_m\} = (1 - \rho)^r, \quad [\text{A3.9}]$$

since the probability of stimulus change every time step is ρ and change occurs independently each time step. The conditional probability giving an incorrect behavior is calculated easily by noting that $\mathbf{I} | \mathbf{S}_r \cap \mathbf{F} \cap \mathbf{M}_m$ is the complement of $\mathbf{C} | \mathbf{S}_r \cap \mathbf{F} \cap \mathbf{M}_m$ and therefore

$$\Pr\{\mathbf{I} | \mathbf{S}_r \cap \mathbf{F} \cap \mathbf{M}_m\} = 1 - (1 - \rho)^r. \quad [\text{A3.10}]$$

Putting everything together, we have

$$\Pr\{\mathbf{C} | \mathbf{M}_m\} = \sum_{r=1}^m \Pr\{\mathbf{C} | \mathbf{S}_r \cap \mathbf{F} \cap \mathbf{M}_m\} \Pr\{\mathbf{S}_r | \mathbf{F} \cap \mathbf{M}_m\} \Pr\{\mathbf{F} | \mathbf{M}_m\}, \quad [\text{A3.11}]$$

$$\Pr\{\mathbf{I} | \mathbf{M}_m\} = \sum_{r=1}^m \Pr\{\mathbf{I} | \mathbf{S}_r \cap \mathbf{F} \cap \mathbf{M}_m\} \Pr\{\mathbf{S}_r | \mathbf{F} \cap \mathbf{M}_m\} \Pr\{\mathbf{F} | \mathbf{M}_m\}. \quad [\text{A3.12}]$$

Substituting equations [A3.7], [A3.8] and [A3.9] into equation [A3.11] and substituting equations [A3.7], [A3.8], and [A3.10] into equation [A3.12] and then simplifying gives

$$\Pr\{\mathbf{C} | \mathbf{M}_m\} = \frac{1 - \rho}{1 + (N - 1)\rho} \left[1 - \left(\frac{(N - 1)(1 - \rho)}{N} \right)^m \right], \quad [\text{A3.13}]$$

$$\Pr\{\mathbf{I} | \mathbf{M}_m\} = 1 - \left(\frac{N - 1}{N} \right)^m - \frac{1 - \rho}{1 + (N - 1)\rho} \left[1 - \left(\frac{(N - 1)(1 - \rho)}{N} \right)^m \right]. \quad [\text{A3.14}]$$

Now, plugging equations [A3.3], [A3.4], [A3.5], [A3.6], [A3.13] and [A3.14] into equation [A3.1] and then plugging equation [A3.1] into equation [A3.2] yields equation [3.1], our backwards recursion.

Appendix 4 – The Proxy for Optimal Memory⁷

If we differentiate equation [3.6] with respect to m , we have

$$\frac{\partial G}{\partial m} = \left\{ \begin{array}{l} (\ln \alpha) \alpha^m \{F(u(x), t+1, T) - F(i(x), t+1, T)\} \\ - \omega (\ln \beta) \beta^m \{F(c(x), t+1, T) - F(i(x), t+1, T)\} \end{array} \right\}. \quad [\text{A4.1}]$$

Solving $\partial G / \partial m = 0$ yields a single critical point given by [3.7], which we label m^{**} . The second derivative of G is

$$\frac{\partial^2 G}{\partial m^2} = \left\{ \begin{array}{l} (\ln \alpha)^2 \alpha^m \{F(u(x), t+1, T) - F(i(x), t+1, T)\} \\ - \omega (\ln \beta)^2 \beta^m \{F(c(x), t+1, T) - F(i(x), t+1, T)\} \end{array} \right\}, \quad [\text{A4.2}]$$

The second derivative is negative when

$$m < \frac{\ln\left(\frac{\ln \beta}{\ln \alpha}\right) + \ln\left(\frac{\omega \ln \beta \{F(u(x), t+1, T) - F(i(x), t+1, T)\}}{\ln \alpha \{F(c(x), t+1, T) - F(i(x), t+1, T)\}}\right)}{\ln\left(\frac{1}{1-\rho}\right)}, \quad [\text{A4.3}]$$

but from equation [3.7] we know

$$m^{**} = \frac{\ln\left(\frac{\omega \ln \beta \{F(c(x), t+1, T) - F(i(x), t+1, T)\}}{\ln \alpha \{F(u(x), t+1, T) - F(i(x), t+1, T)\}}\right)}{\ln\left(\frac{1}{1-\rho}\right)}. \quad [\text{A4.4}]$$

⁷ This appendix describes the calculations for the proxy for optimal memory in Chapter 3.

Since

$$\frac{\ln\left(\frac{\ln \beta}{\ln \alpha}\right)}{\ln\left(\frac{1}{1-\rho}\right)} > 0, \quad [\text{A4.5}]$$

we know

$$\left. \frac{\partial^2 G}{\partial m^2} \right|_{m=m^{**}} < 0, \quad [\text{A4.6}]$$

and therefore m^{**} is a local maximum. Since m^{**} is the only critical point and G is a continuous and continuously differentiable function, m^{**} is also the global maximum. If $0 < m^{**} < t$, then the integer directly above ($\lceil m^{**} \rceil$) or the integer directly below ($\lfloor m^{**} \rfloor$) gives the optimal memory window m^* .

In general, when $m < m^{**}$, we have

$$m < \frac{\ln\left(\frac{\omega \ln \beta \{F(c(x), t+1, T) - F(i(x), t+1, T)\}}{\ln \alpha \{F(u(x), t+1, T) - F(i(x), t+1, T)\}}\right)}{\ln\left(\frac{1}{1-\rho}\right)}, \quad [\text{A4.7}]$$

$$\ln\left(\frac{\alpha}{\beta}\right)^m < \ln\left(\frac{\omega \ln \beta \{F(c(x), t+1, T) - F(i(x), t+1, T)\}}{\ln \alpha \{F(u(x), t+1, T) - F(i(x), t+1, T)\}}\right), \quad [\text{A4.8}]$$

$$\alpha^m \ln \alpha \{F(u(x), t+1, T) - F(i(x), t+1, T)\} > \beta^m \omega \ln \beta \{F(c(x), t+1, T) - F(i(x), t+1, T)\} \quad [\text{A4.9}]$$

which means

$$\frac{\partial G}{\partial m} > 0 \Leftrightarrow m < m^{**} . \quad [\text{A4.10}]$$

Similarly, we can show,

$$\frac{\partial G}{\partial m} < 0 \Leftrightarrow m > m^{**} . \quad [\text{A4.11}]$$

That is, G always slopes upward below m^{**} and always slopes downward above m^{**} . Since $0 \leq m^*(x, t) \leq t$, this means that if $m^{**} \leq 0$, then $m^* = 0$ and if $m^{**} \geq t$, then $m^* = t$.

Appendix 5 — Extensions and Explanations of the Dynamic Programming Model for Optimal Memory⁸

5A: The Effect of Stimulus Change on Optimal Memory

Differentiating the proxy for optimal memory size [3.7] with respect to ρ gives

$$\frac{\partial m^{**}}{\partial \rho} = -\frac{\frac{N}{1+(N-1)\rho} + \frac{1}{\ln \beta} + m^{**}}{(1-\rho)\ln\left(\frac{1}{1-\rho}\right)}. \quad [\text{A5.1}]$$

For $m^{**} > 0$, if

$$\frac{N}{1+(N-1)\rho} > \frac{1}{\ln(1/\beta)}, \quad [\text{A5.2}]$$

then we must have $\partial m^{**}/\partial \rho < 0$. However, [A5.2] is equivalent to

$$\ln(1/\beta) > 1 - \beta \quad [\text{A5.3}]$$

or with $y=1/\beta$,

$$\ln y > 1 - \frac{1}{y}, \quad [\text{A5.4}]$$

⁸ This appendix describes a series of extensions and explorations, as referred to in Chapter 3. As mentioned in the footnote to Chapter 3, this work (and especially this appendix) was completed in collaboration with Benjamin Kerr.

However, [A5.4] is true since $y > 1$ and

$$\ln 1 = 1 - \frac{1}{1} = 0, \quad [\text{A5.5}]$$

and

$$\frac{d \ln y}{dy} = \frac{1}{y} > \frac{1}{y^2} = \frac{d[1 - 1/y]}{dy}. \quad [\text{A5.6}]$$

for all $y > 1$. Therefore, $\partial m^{**} / \partial \rho < 0$.

5B: Exploration of the Cost of Living

With regards to the cost of living (κ), there are really only two interesting orderings:

$\pi_i < \kappa < \pi_u < \pi_c$ and $\pi_i < \pi_u < \kappa < \pi_c$. In the first ordering, memory is not required to increase the value of the state variable x over a time step. In the second ordering, memory is essential to increase the value of x (we assume this second ordering in the paper). If the cost of living is low (memory is not required), many memory windows are equally optimal. We set the optimal memory to the lowest size if there are “ties”—zero memory in the case shown in Figure 3-3. In this case, optimal memory only increases from zero at the end of the lifetime. If the cost of living is low, note that throughout most of the early lifetime, the optimal memory is zero (without ties) for low states (see below). If the cost of living is high (memory is required), the optimal memory is non-zero and for low states is the maximum window allowed (see Figure 3-1).

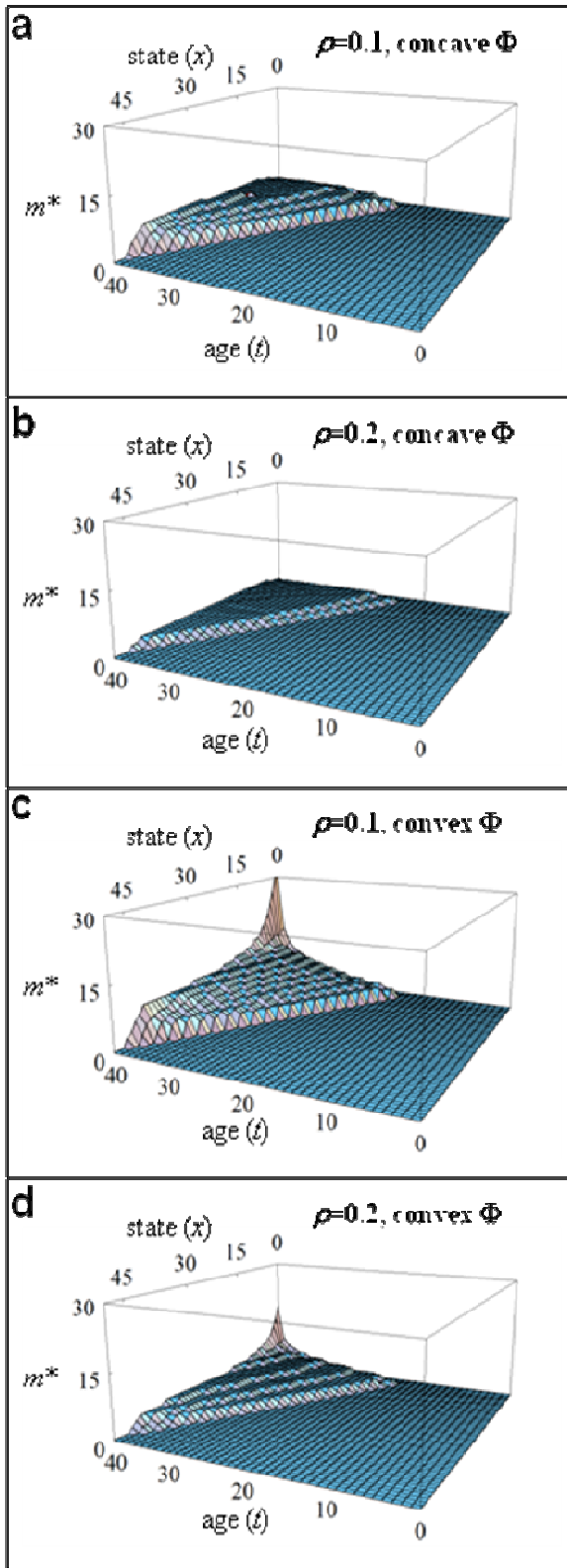


Figure A5-1. This figure corresponds to Figure 1 except $\kappa=8$ here (and $X=50$, $N=10$, $T=40$, $\delta=0.1$, $\pi_i=5$, $\pi_u=10$, and $\pi_c=15$). Each surface gives optimal memory $m^*(x,t)$ and corresponds to a different value of the rate of stimulus change (ρ) and either a convex final fitness function $\Phi(x)=100(x/X)^4$ or a concave final fitness function $\Phi(x)=100(x/X)^{0.25}$.

To understand these patterns, we focus on the action that helps an organism in a low state survive. If the cost of living is low, then an organism in a low state may “play it safe” by not remembering a stimulus, thus avoiding the possibility of an incorrect response. That is, if $\kappa < \pi_u$, then the organism can improve its state without memory. However, if the cost of living is high (i.e., $\kappa > \pi_u$), then an organism can no longer depend on the payoff of an unfamiliar stimulus to ensure survival. In such a case, remembering a stimulus and employing a correct behavior (even if probabilistically difficult) may be the only way for a low-state organism to survive.

We can illustrate some further results using the graphical approach from the Results section. Assume that the current state of the organism is x' . If the state of the organism is very high and if the cost of living is low (i.e., $\kappa < \pi_u$), then both $c(x')$ and $u(x')$ will be X (see Figure A5-2a, where the green dotted and blue dashed arrows point to X). Since $c(x')=u(x')=X$, this means that $B(x',t)=0$ and $v=0$. When $v=0$, we can show that $m^{**}<0$. Thus, when cost of living is low, organisms in a very high state should possess no memory (see Figure 3-3 –first graph in this section). The idea here is that employing a correct response to a familiar stimulus is identical to responding to an unfamiliar stimulus, therefore memory can only serve to open the door for incorrect responses. If the costs of living are higher (i.e., $\kappa > \pi_u$), then $u(x')<X$ for an organism in a very high state and optimal memory is no longer necessarily zero memory (see Figure A5-2b).

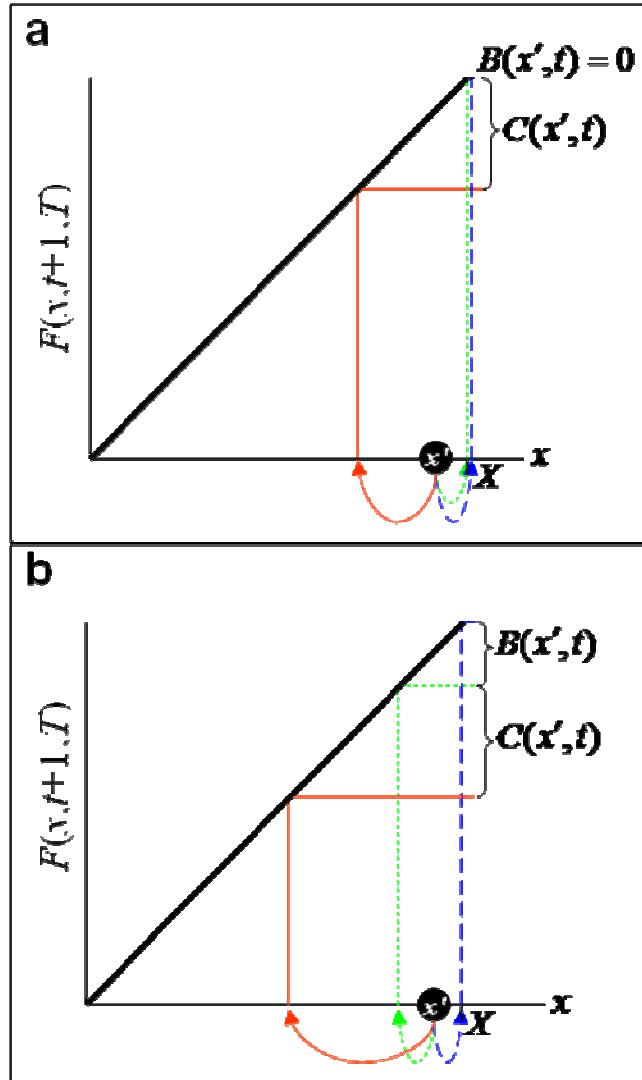


Figure A5-2. The arrows and terminology in this figure correspond to Figure 3. (a) If the state of the organism is very high and the cost of living is low ($\kappa < \pi_w$), then $u(x') = c(x') = X$ while $i(x') < X$. Consequently $C > 0$ and $B = 0$; that is, there are no benefits to memory, whereas there are costs—thus, the organism should have no memory. (b) If the cost of living is high ($\kappa > \pi_w$), then $C > 0$ and $B > 0$ and some memory may be favored.

When the cost of living is low, as we move back in time, it turns out that the function F converges to

$$H(x) = \begin{cases} 0 & \text{if } x = 0 \\ 1 & \text{if } x > 0 \end{cases} \quad [\text{A5.7}]$$

Remember that we are considering the sequence of renormalized F functions converging to $H(x)$. Under [A5.7], most energy reserve states would yield zero cost and zero benefit for memory ($C(x,t)=B(x,t)=0$) and consequently, every length of memory is equivalent (this accounts for the occurrence of “ties” above). However, for low states (i.e., x near 0), there is the possibility of positive cost of memory (since $i(x)=0$ and $u(x)\neq 0$), while there is still no benefit to memory (since $u(x)=c(x)$). Thus, for low states and at early ages, zero memory is optimal (see Figure A5-1).

5C: Concavity and Jensen’s Inequality

There are two important components to the payoff of any stimulus: (1) its expected value and (2) its variance. The expected value of an unfamiliar stimulus is π_u , whereas the expected value of a familiar stimulus (call it π_f) is some weighted average of π_i and π_c (i.e., $\pi_i \leq \pi_f \leq \pi_c$). The unfamiliar stimulus has zero variance in payoff (it always gives π_u). When the environmental stimuli are of intermediate reliability, a familiar stimulus gives a non-zero variance in potential payoffs (because a response to a familiar stimulus can either be correct or incorrect). Payoffs are not the same as maximal future expected

reproductive success, that is, these payoffs need to be filtered through the function $F(x, t+1, T)$ (see Figure 3-3). If F is convex and we treat it as a continuous function, Jensen's inequality guarantees that

$$E [F(x_f, t+1, T)] > F(E[x_f], t+1, T), \quad [\text{A5.8}]$$

where x_f is either $c(x)$ or $i(x)$. By spreading out payoffs, the maximal expected future reproductive value of a familiar stimulus is elevated over that predicted from its expected payoff. In this case, the variance in payoffs generated by memory is a good thing. And if increases in memory window do not lower π_f too substantially, then higher values of memory can be favored. If F is concave, we have

$$E [F(x_f, t+1, T)] < F(E[x_f], t+1, T), \quad [\text{A5.9}]$$

In this case, the future reproductive value of a familiar stimulus is lower than that predicted from its expected payoff. In this case, the variance in payoffs generated by memory is a bad thing. And if decreases in memory window do not raise π_f too substantially, then lower values of memory can be favored.

5D: Iteroparity

The model given by equation [3.1] refers to a semelparous organism. If an organism has multiple reproductive bouts during its lifetime, then some part of its reproductive value (F) can come from offspring produced before its final time step. This can be modeled through a slight adjustment to equation [3.1]:

$$F(x, t, T) = (1 - \delta) \max_{0 \leq m < t} \left\{ \begin{array}{l} \alpha^m F(u(x), t+1, T) \\ + \omega [1 - \beta^m] F(c(x), t+1, T) \\ + (1 - \alpha^m - \omega [1 - \beta^m]) F(i(x), t+1, T) \\ + \phi(x, t) \end{array} \right\} \quad [\text{A5.10}]$$

where all parameters and functions have the same meaning as before, and the immediate reproductive output at time step t , $\phi(x, t)$, is given by

$$\phi(x, t) = \begin{cases} 0 & \text{if } t < t_{RM} \\ \Phi(x) & \text{if } t \geq t_{RM} \end{cases} \quad [\text{A5.11}]$$

The parameter t_{RM} is the age of reproductive maturity. At this age and thereafter, the organism realizes immediate reproductive output that depends on its state. We assume that the previously defined “final fitness function” $\Phi(x)$ now gives the state-dependent reproductive output at all ages of reproductive maturity. For simplicity, we do not weight offspring born earlier with higher value here. In Figure A5-3, we see that optimal memory can still decline with age even if organisms are iteroparous. Note that the function $\Phi(x)$ is concave in the figure.

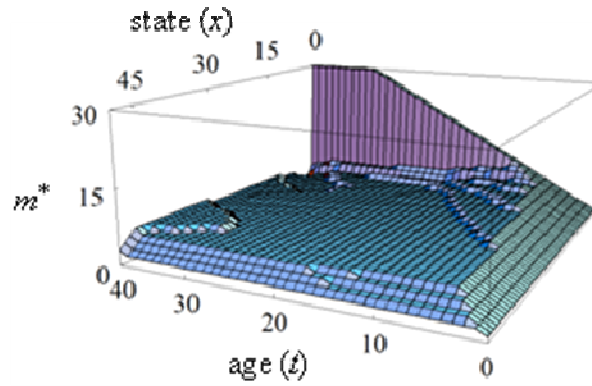


Figure A5-3. This surface gives optimal memory $m^*(x,t)$ as a function of age (t) and state (x) under the iteroparous extension of Appendix 5D. The other parameters are $X=50$, $N=10$, $T=40$, $\delta=0.1$, $\pi_i=5$, $\pi_u=10$, $\pi_c=15$, $\kappa=12$, $\rho=0.1$, and $t_{RM}=20$. The fitness function is concave: $\Phi(x)=100(x/X)^{0.25}$.

5E : Explicit Costs to Memory

In our model, costs to memory are implicit. However, memory may carry explicit costs. For instance, flexibility in retrieval may be costly. Indeed, it would be an interesting exercise to determine the strength of selection for variable retrieval over constant options within our model framework (see McNamara 1996; Houston & McNamara 1999). Here, though, we focus on costs to memory length. For instance, the cost of living could increase linearly with memory length:

$$\kappa(m) = \kappa_0 + \kappa_1 m . \quad [A5.12]$$

Without explicit costs to memory, it was important that the payoff for an incorrect response from memory was less than sampling ($\pi_i < \pi_u$). It was from this inequality that the implicit costs of memory emerged. If $\pi_u < \pi_i < \pi_c$, optimal memory would be maximal memory.

With explicit costs to memory, we now consider the case where sampling always yields lower payoffs than remembering responses; namely, $\pi_u < \pi_i < \pi_c$. In Figure , we see that optimal memory can still decline with age with reordered payoffs when there are explicit costs to memory.

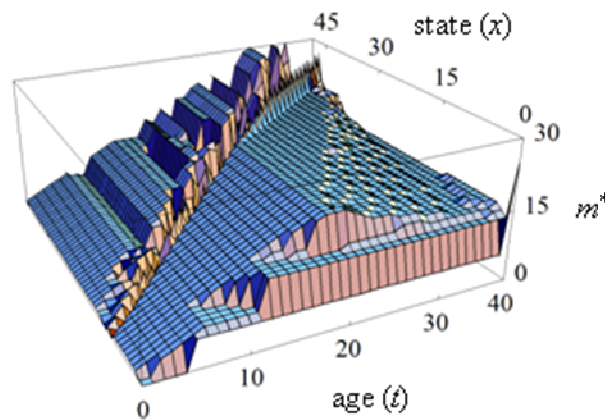


Figure A5-4. This surface gives optimal memory $m^*(x,t)$ as a function of age (t) and state (x) when memory is explicitly costly as laid out above. The other parameters are $X=50$, $N=10$, $T=40$, $\delta=0.1$, $\pi_i=10$, $\pi_u=5$, $\pi_c=15$, $\rho=0.1$, $\kappa_o=4$ and $\kappa_i=0.15$. The final fitness function is concave: $\Phi(x)=100(x/X)^{0.25}$.