

UNIVERSITÉ DU QUÉBEC À MONTRÉAL

DECISION MAKING BY LEARNING RULES IN A FREQUENCY  
DEPENDENT TASK

THESIS  
PRESENTED  
AS PARTIAL REQUIREMENT  
OF THE DOCTORATE OF BIOLOGY

BY  
MOHAMMAD AFSHAR

JULY 2015

UNIVERSITÉ DU QUÉBEC À MONTRÉAL  
Service des bibliothèques

Avertissement

La diffusion de cette thèse se fait dans le respect des droits de son auteur, qui a signé le formulaire *Autorisation de reproduire et de diffuser un travail de recherche de cycles supérieurs* (SDU-522 – Rév.01-2006). Cette autorisation stipule que «conformément à l'article 11 du Règlement no 8 des études de cycles supérieurs, [l'auteur] concède à l'Université du Québec à Montréal une licence non exclusive d'utilisation et de publication de la totalité ou d'une partie importante de [son] travail de recherche pour des fins pédagogiques et non commerciales. Plus précisément, [l'auteur] autorise l'Université du Québec à Montréal à reproduire, diffuser, prêter, distribuer ou vendre des copies de [son] travail de recherche à des fins non commerciales sur quelque support que ce soit, y compris l'Internet. Cette licence et cette autorisation n'entraînent pas une renonciation de [la] part [de l'auteur] à [ses] droits moraux ni à [ses] droits de propriété intellectuelle. Sauf entente contraire, [l'auteur] conserve la liberté de diffuser et de commercialiser ou non ce travail dont [il] possède un exemplaire.»

UNIVERSITÉ DU QUÉBEC À MONTRÉAL

LA PRISE DE DÉCISION PAR RÈGLE D'APPRENTISSAGE DANS UN  
CONTEXTE FRÉQUENCE-DÉPENDANT

THÈSE  
PRÉSENTÉE  
COMME EXIGENCE PARTIELLE  
DU DOCTORAT EN BIOLOGIE

PAR  
MOHAMMAD AFSHAR

JUILLET 2014

[Cette page a été laissée intentionnellement blanche]

DEDICATED TO

My mother, my sister, and my wife,  
three women to whom I owe all my achievements.

## FOREWORD

In this doctoral thesis, three chapters are presented as scientific articles. The first, Chapter I, is published in *Animal Behaviour* (Vol. 96, 2014, pp. 167-176) under the title “A unified modelling approach for producer-scrounger games in complex ecological conditions” and is co-authored by Mohammad Afshar and Luc-Alain Giraldeau. The second, Chapter II, is published in *Animal Behaviour* (Vol. 105, 2015, 181-186) under the title “Zebra finches scrounge more when patches vary in quality: Experimental support of the linear operator learning rule” and is co-authored by Mohammad Afshar, Carolyn L. Hall, and Luc-Alain Giraldeau. The third, Chapter III, is entitled “The evolution of learning and non-learning rules in dynamic environments: A test of behavioural gambit”, is co-authored by Mohammad Afshar and Luc-Alain Giraldeau and will be submitted shortly to *Proceedings of the Royal Society B: Biological Sciences or Behavioral Ecology*.

During this study, Mohammad Afshar was supported financially by a Natural Sciences and Engineering Research Council Discovery Grant (number 303112) to Luc-Alain Giraldeau. In addition, Mohammad Afshar was financially supported through a UQAM FARE Scholarship, a foreign fee remission from UQAM's Faculté des Sciences and a Doctoral Scholarship from Group For Research In Decision Analysis (GERAD).

## ACKNOWLEDGEMENTS

I have been fortunate throughout my life to encounter brilliant people who generously shared their invaluable knowledge and wisdom with me, and one of these people was Luc-Alain Giraldeau. More than being a research director, Luc-Alain taught me how to think and how to ask questions, a priceless skill that I believe will help me in the next 100 years of my life. I learned many things from Luc-Alain and doing science was only one of them. Coming from a different background, I could not catch up with behavioral ecology and finish my study without Luc-Alain's exceptional patience. The financial support that he kindly provided during that past four and half years was more than I deserved.

I am grateful to Carolyn L. Hall for her enormous help with the writing of the articles and statistical analysis of the experiments as well as all the administrative tasks in Université du Québec à Montréal (UQAM). I have no doubt that everything would have taken longer if Carolyn had not been a part of the lab. We had great discussions along the way and she kindly provided much needed expertise.

I would also like to thank my lab-mates for their valuable advice, friendly help, and discussions during this study: Morgan David, Marie-Pierre Bonin, Simone Tomassina, Mewen Le Hô, Steven Hamblin, and Mélanie Dappen-Couture.

I would also like to thank Denis Réale for accepting to provide extensive help and guidance for statistical analyses as well as comments about the experiments in this study. I am grateful also to William Vickery and Frédérique Dubois for the great discussions we had and the comments they provided on the early versions of the articles. During the first year of my study, Pedro R. Peres-Neto generously allowed

me to use the computational resources in his lab that were provided through a Canada Foundation for Innovation (CFI) Leaders Opportunity Fund.

Special thanks to the management team and technical staffs in animal care facilities of UQAM and the Centre National de Biologie Expérimentale at INRS–Institut Armand-Frappier Research Centre that provided amazing conditions and collaborations during the course of my experiments: Manon St-Germain, Normand Lapierre, Angèle Sansoucy, Sandie Papineau, Luc Gladu, Louise Arvisais, Hakim Slimani and Marie-Andrée Blais.

My sincere thanks to all my colleagues and professors in Groupe de Recherche en Écologie Comportementale et Animale (GRECA) and Département des sciences biologiques at UQAM as well as Société Québécoise pour l'Étude Biologique du Comportement (SQEBC). Through their discussions, talks and seminars, I learned a lot about biology and behavioral ecology.

Throughout my Ph.D., I was supported financially by a Natural Science and Engineering Research Council Discovery Grant (number 303112) to Luc-Alain Giraldeau. In addition I was financially supported through a UQAM FARE Scholarship and a Faculty of Science Foreign Fee Remission as well as a doctoral student's Scholarship from Group For Research In Decision Analysis (GERAD).

Finally, my wife, Zahra Najarzadeh, who supported me in all the ups and downs of my journey and encouraged me to finish what I had never dared to start on my own. You are the best supporter anyone could ever wish to have. I must be the luckiest man who has ever lived, and I thank you for that.

## TABLE OF CONTENTS

LIST OF FIGURES .....	xiii
LIST OF TABLES .....	xvii
RÉSUMÉ .....	xix
ABSTRACT .....	xxi
INTRODUCTION .....	1
0.1 Behavioral ecology .....	1
0.1.1 Phenotypic gambit .....	2
0.1.2 Behavioral gambit .....	2
0.1.3 Evolution and learning .....	3
0.2 Learning can be as powerful as evolution to predict some animal behavior .....	4
0.3 State of knowledge .....	6
0.3.1 Frequency-dependency, tactic and strategy .....	6
0.3.2 Evolutionary game theory .....	7
0.3.3 Learning rule .....	12
0.3.4 On the extent of matching .....	17
0.4 Thesis overview .....	18
CHAPTER I	
A UNIFIED MODELLING APPROACH FOR PRODUCER-SCROUNGER GAMES IN COMPLEX ECOLOGICAL CONDITIONS .....	21
1.1 Abstract .....	21
1.2 Introduction .....	22
1.3 Producer-scrouter game .....	28
1.4 Linear operator model .....	29
1.4.1 Results .....	33
1.5 The effect of variable patch quality on joining .....	40

1.5.1 Results .....	40
1.6 Discussion .....	44
1.7 Future directions.....	46
1.8 Acknowledgments.....	49
CHAPTER II	
ZEBRA FINCHES SCROUNGE MORE WHEN PATCHES VARY IN QUALITY: EXPERIMENTAL SUPPORT OF THE LINEAR OPERATOR LEARNING RULE .....	
2.1 Abstract .....	51
2.2 Introduction.....	52
2.3 Methods.....	54
2.3.1 Study subjects .....	54
2.3.2 Experimental Procedure.....	55
2.3.3 Statistical Analysis.....	57
2.4 Results .....	58
2.5 Discussion .....	64
2.6 Acknowledgements.....	67
CHAPTER III	
THE EVOLUTION OF LEARNING AND NON-LEARNING RULES IN DYNAMIC ENVIRONMENTS: A TEST OF THE BEHAVIOURAL GAMBIT ....	
3.1 Abstract .....	69
3.2 Introduction.....	70
3.3 Simulation model .....	73
3.3.1 Static environment .....	80
3.3.2 Dynamic environment .....	80
3.4 Results .....	81
3.4.1 Static environment .....	81
3.4.2 Dynamic environment .....	90
3.5 Discussion .....	93
3.6 Acknowledgements .....	97
CHAPTER IV	
GENERAL DISCUSSION AND FUTURE DIRECTIONS .....	
	99

4.1 Future directions .....	100
APPENDIX A	
FINAL PROPORTION OF PURE, MIXED, AND LINEAR OPERATOR STRATEGIES WHEN COMPETING WITH EACH OTHER IN PRODUCER- SCROUNGER GAME IN DYNAMIC ENVIRONMENTS .....	103
BIBLIOGRAPHY .....	109

[Cette page a été laissée intentionnellement blanche]

## LIST OF FIGURES

Figure	Page
1.1 Evidence of temporal adjustment of the producer-scrouter strategy for a sample group of eight individuals engaged in choosing tactics using a linear operator learning rule. Each line shows the proportion of join tactic choices for one individual. The thick dashed line shows the average for the group. This shows, that by the end of the simulation, individuals do not necessarily adopt a pure strategy (all find, or all search), which means that, in terms of expressed tactics, the population ends in a strong polymorphism of tactics.....	35
1.2 Proportion of join tactic at equilibrium as a function of changes in eight ecological parameters (group size, metabolic requirement, finder's advantage, food intake rate, cost of searching, cost of joining, patch encounter probability and patch richness) for group of eight individuals while foraging in a producer-scrouter game and choosing tactics based on a linear operator learning rule. Means $\pm$ SD are shown (N=30). The general parameter settings for all simulations are shown in Table 1.1.....	36
1.3 Average temporal change in the proportion of join tactic in groups of eight individuals using the linear operator learning rule in a producer-scrouter game when the average patch richness ( $\mu_F$ ) = 5 and patch richness variance ( $\sigma_F$ ) = 0 or 5. Means $\pm$ SD are shown (N = 30). To keep the graph readable, only the first 100 time steps (out of 500) are shown. ....	42
1.4 Effect of patch richness (mean and variance) on the proportion of join tactic. Means $\pm$ SD are shown (N = 30). To compare the results for different $\mu_F$ , we show the results based on the patch richness coefficient variance ( $\sigma_F \mu_F$ ). ....	43
2.1 Mean $\pm$ SE proportion of patches found by joining in four flocks of five zebra finches, <i>Taeniopygia guttata</i> , foraging in nonvariable and variable patch richness treatments on day 1 (N = 19) and days 2-5 (N = 20). The curves show the predicted proportion of patches by join for each treatment from the GLMM model shown in Table 2.1 and 2.2 .....	62

2.2	Strategy consistency across the two treatments (nonvariable and variable patch richness). A scatter plot of the average proportion of patches found by joining in the variable treatment and the nonvariable treatment for individuals in four flocks of five zebra finches, <i>Taeniopygia guttata</i> . ....	63
3.1	Evolutionarily Stable Strategy plot for producer–scrounger game for groups of 90 individuals engaged in choosing tactics using pure strategy, mixed strategy, or linear operator learning rule in a static environment, $\Delta\mu, W = 0\%$ and $\Delta\mu, B = 0\%$ . Each point shows the lifetime average payoff of each tactic (per time step) for one individual vs. its lifetime proportion of join tactic. The general parameter settings for all simulations are shown in Table 3.1 and Table 3.2. ....	84
3.2	Payoff of search and join tactics in producer–scrounger game for groups of 90 individuals engaged in choosing tactics using pure strategy, mixed strategy, or linear operator learning rule in a static environment during 500 generations. Means $\pm$ SE are shown ( $N=20$ ). The general parameter settings for all simulations are shown in Table 3.1 and Table 3.2. To keep the graph readable, only the results for every five generations is shown. ....	85
3.3	The histogram of frequency of different mixed strategy genotypes for individuals that play producer–scrounger game using only mixed strategy in a static environment during 500 generations. Mixed strategy genotype is the threshold to randomly choose between search or join tactic. Despite the initially uniform distribution, the mixed strategy genotype finally converged to either close to one (only join) or close to zero (only search). We can therefore say that even populations of mixed strategists evolved to behave similar to pure strategists. Frequencies are shown for the last 100 generations when the populations have become stable. The general parameter settings for all simulations are shown in Table 3.1 and Table 3.2. 20 simulations were performed and each colour represents the result of one of the 20 simulations. ....	88
3.4	The evolution of different strategies in a producer–scrounger game for groups of 90 individuals engaged in choosing tactics using pure strategy, mixed strategy, or a linear operator learning rule in a static environment during 500 generations ( $N = 20$ ). After initial changes in the distribution of strategies, the pure strategy remains dominant in the populations for the remainder of the simulation. The linear operator learning strategy thus has no advantage over inflexible strategies, either pure or mixed strategies. The general parameter settings for all simulations are shown in Table 3.1 and Table 3.2. ....	89

- 3.5 Final proportions of a) pure, b) mixed, and c) linear operator strategies when competing with each other in populations of 90 individuals playing a producer–scrounger game in a dynamic environment with different levels of within-generation change ( $\Delta\mu, W$ ) and between-generation change ( $\Delta\mu, B$ ) during 500 generations. The final proportions were calculated by averaging over the last 100 generations ( $N = 20$ ). The general parameter settings for all simulations are shown in Table 3.1 and Table 3.2. To keep the graph readable, axes of within-generation and between-generation were reversed for (a). The results for populations of sizes 10 and 40 are available in the Appendix..... 92
- A.1 Final proportions of a) pure, b) mixed, and c) linear operator strategies when competing with each other in populations of 10 individuals playing a producer–scrounger game in a dynamic environment with different levels of within-generation change ( $\Delta\mu, W$ ) and between-generation change ( $\Delta\mu, B$ ) during 500 generations. The final proportions were calculated by averaging over the last 100 generations ( $N = 20$ ). The general parameter settings for all simulations are shown in Table 3.1 and Table 3.2. To keep the graph readable, axes of within-generation and between-generation were reversed for (a)..... 106
- A.2 Final proportions of a) pure, b) mixed, and c) linear operator strategies when competing with each other in populations of 40 individuals playing a producer–scrounger game in a dynamic environment with different levels of within-generation change ( $\Delta\mu, W$ ) and between-generation change ( $\Delta\mu, B$ ) during 500 generations. The final proportions were calculated by averaging over the last 100 generations ( $N = 20$ ). The general parameter settings for all simulations are shown in Table 3.1 and Table 3.2. To keep the graph readable, axes of within-generation and between-generation were reversed for (a)..... 108

[Cette page a été laissée intentionnellement blanche]

## LIST OF TABLES

Table	Page
0.1 Different ways of having ESS in a population proposed by: a) Maynard smith (1982), b) Krebs & Davies (1993), c) Gross (1996) .....	10
0.2 Examples of having an ESS in a population .....	11
0.3 Selected important learning rules in behavioral ecology. Here, only the updating part is mentioned. Depending on the experiment, either matching or maximizing was used for decision making. $T_i(t)$ is the estimated value of tactic $i$ at time $t$ . $P_i(t)$ is the observed payoff for tactic $i$ at time $t$ . $x$ is the memory factor ( $1 - x$ is the learning rate) used to adjust the estimation with new experience. ....	14
1.1 Parameter settings for the linear operator model in the producer-scrounger game. ....	32
1.2 Comparison of the predictions made by the Linear Operator agent-based simulation model with predictions of previous models and results obtained from experiments. Only studies that provided predictions or results concerning the parameters that we investigated were considered.....	38
2.1 Variance explained by different random-effect structures of the GLMM on strategy use, using the proportion of patches found by joining as the dependent variable .....	60
2.2 Fixed effects of the GLMM on strategy use, using the proportion of patches found by joining as the dependent variable .....	61
3.1 Genetic algorithm chromosome description and initial values.....	76
3.2 General parameter settings for the producer-scrounger (PS) game simulations .....	78
3.3 Final proportion of join tactic (Means $\pm$ SE) for groups of individuals playing a producer-scrounger game using different types of strategies in a static environment, $\Delta\mu, W = 0\%$ and $\Delta\mu, B = 0\%$ , after 500	

	generations. The final proportion was calculated by averaging over the last 100 generations ( $N = 20$ ) .....	83
3.4	Final frequency of different strategies (Means $\pm$ SE) when competing with each other in producer–scrounger game in a static environment during 500 generations. The final frequency was calculated by averaging over the last 100 generations ( $N = 20$ ). .....	87
3.5	Final proportions of pure, mixed, and linear operator strategies (Means $\pm$ SE) when competing with each other in populations of 90 individuals playing a producer–scrounger game in a dynamic environment with different levels of within-generation change ( $\Delta\mu, W$ ) and between-generation change ( $\Delta\mu, B$ ) during 500 generations. The final proportion was calculated by averaging over the last 100 generations ( $N = 20$ ). The general parameter settings for all simulations are shown in Table 3.1 and Table 3.2. The results for the populations of sizes 10 and 40 are available in the Appendix. ....	91
A.1	Final proportions of pure, mixed, and linear operator strategies (Means $\pm$ SE) when competing with each other in populations of 10 individuals playing a producer–scrounger game in a dynamic environment with different levels of within-generation change ( $\Delta\mu, W$ ) and between-generation change ( $\Delta\mu, B$ ) during 500 generations. The final proportion was calculated by averaging over the last 100 generations ( $N = 20$ ). The general parameter settings for all simulations are shown in Table 3.1 and Table 3.2. ....	105
A.2	Final proportions of pure, mixed, and linear operator strategies (Means $\pm$ SE) when competing with each other in populations of 40 individuals playing a producer–scrounger game in a dynamic environment with different levels of within-generation change ( $\Delta\mu, W$ ) and between-generation change ( $\Delta\mu, B$ ) during 500 generations. The final proportion was calculated by averaging over the last 100 generations ( $N = 20$ ). The general parameter settings for all simulations are shown in Table 3.1 and Table 3.2. ....	107

## RÉSUMÉ

Les écologistes du comportement se sont satisfaits de l'approche évolutive pour prédire et expliquer le comportement animal. Lorsqu'on étudie les comportements flexibles l'approche évolutive en générale, et l'approche des jeux évolutifs en particulier, supposent que le mécanisme psychologique comportemental sous-jacent e.g. l'apprentissage, n'empêche pas l'expression de comportements adaptés; une supposition connue sous l'intitulée du «pari comportemental». Ce pari, cependant, pourrait être invalide, surtout si le mécanisme en question est soumis aux pressions découlant des conséquences de différents systèmes comportementaux. Des études récentes prétendent que l'étude des mécanismes psychologiques peut générer de meilleures compréhensions et prédictions que celles des modèles évolutifs. Je crois en conséquence que dans certaines circonstances l'apprentissage peut être autant, voire plus puissant, que l'évolution pour prédire le comportement animal. J'ai adopté trois mesures pour tester cette hypothèse: 1) je propose un modèle d'apprentissage pour un jeu producteur chapardeur, un jeu d'approvisionnement social qui a été bien étudié en écologie comportementale. Les résultats de ce modèle recourent presque tous les résultats empiriques et les prédictions théoriques qui ont été publiés pour le jeu producteur chapardeur à ce jour de sorte qu'il fournit un outil prédictif plus général qu'aucun des modèles évolutifs existant pour des conditions de fréquence dépendance. De plus, le modèle d'apprentissage génère une prédiction inédite à l'effet que la fréquence de chapardeur devrait s'accroître avec la variance dans la qualité des parcelles. 2) J'ai testé cette nouvelle prédiction avec des groupes captifs de diamants mandarins (*Taeniopygia guttata*) avec des conditions de parcelles de qualité fixe ou variable. J'ai trouvé qu'à mesure que les individus accumulent de l'expérience dans la condition variable, la fréquence de chapardeur augmente. Ce résultat appuie le modèle. 3) J'ai exploré les conditions sous lesquelles on s'attend à ce que l'apprentissage évolue dans un contexte fréquence dépendant afin de permettre aux écologistes du comportement de comprendre les situations où l'apprentissage sera présent. En faisant cela j'ai pu étudier les propriétés évolutives des règles d'apprentissage tout en palliant à deux limitations d'études précédentes: en confrontant les règles d'apprentissages à des stratégies fixes et en évaluant les effets de variations environnementales intra et intergénérationnelles. J'ai trouvé que malgré les bénéfices marginaux de l'apprentissage comparé aux stratégies fixes, l'apprentissage n'arrive pas à se répandre à l'ensemble de la population dans des environnements dynamiques. Ma thèse démontre donc que lorsqu'on s'intéresse au comportement au niveau de la population dans un environnement statique, les

écologistes du comportement peuvent continuer faire le pari comportemental et supposer que le mécanisme psychologique qui sous-tend le comportement produira le même résultat que celui attendu par le processus évolutif. Cependant, si on s'intéresse au comportement individuel et qu'il s'agit d'un environnement fluctuant, alors le pari peut être risqué et il serait mieux de tenir compte du mécanisme d'apprentissage.

**Mots-clés:** écologie comportementale, pari comportemental, prise de décision, théorie des jeux évolutive, sélection fréquence-dépendante, règle d'apprentissage, opérateur linéaire, stratégie mixte, polymorphisme, producteur-chapardeur, stratégie pure, sensibilité au risque, approvisionnement social, diamants mandarins (*Taeniopygia guttata*).

## ABSTRACT

Behavioral ecologists have been comfortable predicting and explaining animal behavior from an evolutionary approach. When studying flexible behaviors, evolutionary approaches in general and evolutionary game theory in particular implicitly assume that the underlying psychological mechanism, e.g. learning, does not constrain the expression of the adaptive behavior: an assumption known as the behavioral gambit. The gambit, however, may not be valid, especially if the mechanism is under selection pressure from the outcomes of multiple behavior systems. Recent studies claim that focusing instead on the psychological mechanism can provide new insights and predictions that could not have been made from evolutionary models. I thus believe that in some contexts learning is as powerful as, or perhaps even more powerful than, evolution in predicting animal behavior. I have taken three steps to explore this hypothesis: 1) I proposed a learning based model for producer-scrouter game, a social foraging game that has been extensively studied in behavioral ecology. The results replicate almost every producer-scrouter prediction and experimental result published to date such that it provides a more general tool than any single game-theoretic model to predict behavior under frequency-dependent conditions. This model, furthermore, provided a novel prediction: an increase in scrounging behavior when patch quality varies. 2) I tested this new prediction using captive foraging flocks of male zebra finches (*Taeniopygia guttata*) with variable and nonvariable patch quality. I found that as individuals gain experience in the variable environment, they increase their scrounging behavior. This supports the model. 3) I explored the conditions under which learning is expected to evolve in a frequency-dependent context to guide behavioral ecologists as to where one should expect learning. In doing so, I studied the evolutionary properties of learning rules and addressed both limitations of earlier studies by pitting fixed against learning agents while exploring the effects within and between-generation fluctuations. I found that despite the marginal benefits of learning over fixed strategies in dynamic environments, learning never goes to fixation. The thesis shows that when the research interest is in the population level behavior in a static environment, behavioral ecologist can continue using the behavioral gambit and assume that the mechanism in a frequency-dependent situations will produce outcomes that are the same as those from an evolutionary process. However, if the interest is in individual level behavior or involves environmental fluctuations, the gambit may be incomplete and learning should be considered.

**Keywords:** behavioral ecology, behavioral gambit, decision making, evolutionary game theory, frequency-dependent selection, learning rule, linear operator, mixed strategy, polymorphism, producer-scrounger, pure strategy, risk sensitivity, social foraging, zebra finches (*Taeniopygia guttata*)

## INTRODUCTION

### 0.1 Behavioral ecology

You see a pigeon, in the middle of a flock on the ground, standing and watching other birds while they are walking and looking at the ground to find food. Why does the animal do this? Tinbergen (1963) suggested four types of answers to this question: immediate causation, development (ontogeny), adaptive value (survival function), and evolutionary history. Immediate causation explains the behavior based on the underlying physiological mechanisms that are involved in performing the behavior. Development explains the behavior in relation to early experiences in the animal's life. Adaptive value, however, explains the behavior based on the utility of the behavior in relation to the animal's natural surroundings. Finally, evolutionary history explains the behavior of this animal based on the behaviors of its ancestors. Focusing on the adaptive value of a behavior, behavioral ecology has emerged from ethology to study the evolutionary basis for animal behavior due to past ecological pressures (Danchin, Giraldeau, & Cézilly, 2008)

Evolution has traditionally been the fundamental theory in behavioral ecology (Davies, Krebs, & West, 2012). When behavioral ecologists want to explain or predict an animal's behavior, they usually come up with a model that shows that animal does X (and not Y) because X maximizes its fitness, overall survival and reproduction. By considering costs and benefits associated with alternative options, we have been able for the past 30-40 years to make detailed predictions of animal behavior of which many have been supported empirically (Danchin et al., 2008; Davies et al., 2012; Stephens & Krebs, 1986).

### 0.1.1 Phenotypic gambit

Applying an economical view to animal behavior problems requires two implicit assumptions. First, the genetic architecture of an organism does not constrain which phenotype can evolve, known as the phenotypic gambit (Fawcett, Hamblin, & Giraldeau, 2013; Grafen, 1984). For example, if a clutch size of six eggs is optimal then we should expect to see a similar clutch size in nature, regardless of what the evolutionary pathway to reaching this optimal solution has been. In behavioral ecology, we do not often question the phenotypic gambit and when the prediction of the model fails, the first thing we verify are the details of the model to see if its predictions could be improved, for example by adding a factor to make it more realistic.

Accepting the phenotypic gambit has proven to be a good way to simplify problems and come up with powerful models and predictions, especially about fixed behaviors, such as clutch size. Adaptation, however, occurs not only through genetic evolution but also through changes that take place within an individual's lifetime via psychological mechanisms such as learning (Shettleworth, 2010). There is thus an extension of the phenotypic gambit for flexible behaviors.

### 0.1.2 Behavioral gambit

The second implicit assumption resulting from applying an economical view to animal behavior problems is that for flexible behaviors, i.e. where the behavior can be adjusted in response to the behavior of others (Fawcett et al., 2013), the psychological mechanism does not constrain the expression of the adaptive behavior, known as behavioral gambit (Giraldeau & Dubois, 2008). Whatever the optimal outcome expected from an evolutionary perspective, we can thus assume the animal will express it, even for behaviors generated by adaptive decision making. The behavioral gambit simplifies the problem by neglecting the mechanism and focusing on the function.

However, we can think of at least two reasons that behavioral ecologists should be cautious with the behavioral gambit. First, the behavioral gambit may fail; that is the mechanism may constrain the optimal behavior. For example, if three different behaviors, such as foraging, mate choice, and nest building, operate according to a common underlying mechanism such as learning, the overall outcome of these three behaviors applies selection pressure on the mechanism. The mechanism, however, may not maximize each behavior independently because different behaviors may require different or conflicting optimization mechanisms. For example, foraging behavior may favor an exploratory trait while such a trait may not be optimal for mate choice. Natural selection thus favors a mechanism that finds an optimal compromise for these behaviors and maximizes the overall outcome. In this case, an evolutionary model that accepts the behavioral gambit may be handicapped in predicting or explaining the optimality of the animal's foraging behavior.

Second, by focusing on the psychological mechanism, we can obtain new insights about the behavior and we can come with predictions that we could not have made had we not focused on the mechanism. The first part of this study demonstrates a model based on learning that is shown to be more effective than evolutionary models in predicting the social foraging behaviors of ground feeding birds (Chapter I). Considering these two reasons it is thus important to examine the extent and ecological conditions that allow the behavioral gambit to be accepted, that is to continue ignoring the underlying mechanism for the flexible adaptive behavior.

### 0.1.3 Evolution and learning

Based on the speed of adaptation, different types of adaptation can be generally categorized into physiological adaptation, learning, and evolution. Unlike the functional role of a physiological adaptation, its underlying mechanisms, such as chemical and hormonal reactions, are of no particular interest to behavioral ecology.

Learning that happens during the lifetime of an individual, however, has a complex relation with evolution.

Learning can be broadly defined as “a change in state due to experience. [...] The changes in state referred to as learning seem to involve a change in cognitive state, not just behavioral potential” (Shettleworth, 2010). There is much debate about the relationship between evolution and learning, and their roles in animals adaptation (Baldwin, 1896; Bateson, 2004; Carroll, Hendry, Reznick, & Fox, 2007; Duckworth, 2008; Hinton & Nowlan, 1987; Kawecki, 2009; Levey et al., 2009; Luniak, 2004; Nolfi & Floreano, 1999; Suzuki & Arita, 2004). Learning can inhibit the selection pressure and hence evolution by allowing an animal to adopt an adaptive response to a completely novel situation that had never been encountered in the evolutionary history of the species. Learning is thus thought to generally slow down the rate of evolutionary change because behavioral plasticity can shield organisms from strong directional selection (Duckworth, 2008).

On the other hand, learning is also argued to accelerate evolution because it allows individuals to survive and hence maintains genetic diversity within the population. The increased diversity increases the likelihood of evolution if a new adaptive unlearned phenotype occurs. This mutant with unlearned behavior will spread in the population, replacing the learning phenotype by virtue of the saved costs that are required of learning: a process known as genetic assimilation (Bateson, 2004). Selection of the best learners, therefore, would eventually result in the appearance of fixed behavior occurring entirely in the absence of learning (Price, Qvarnstrom, & Irwin, 2003; Shimada, Ishii, & Shibao, 2010; Tebbich, Sterelny, & Teschke, 2010). Learning can thus enhance the path of evolution (Suzuki & Arita, 2004).

0.2 Learning can be as powerful as evolution to predict some animal behavior

Despite the complex relation between learning and evolution, the mechanism has remained neglected in behavioral ecology. Moreover, behavioral ecologists do not

commonly distinguish between evolutionary and behavioral decisions. Mechanisms of decision, hence, remain ill explored, except perhaps for the occasional attention devoted to learning rules: a mathematical expression for behavioral decisions (Aoki & Feldman, 2014; Arbilly, Motro, Feldman, & Lotem, 2011; Dridi & Lehmann, 2014; Dubois, Morand-Ferron, & Giraldeau, 2010; Fawcett et al., 2014; Groß et al., 2008; Hamblin & Giraldeau, 2009; Katsnelson, Motro, Feldman, & Lotem, 2012; Kolodny, Edelman, & Lotem, 2014; Lotem & Biran-Yoeli, 2014; Rendell et al., 2010; Trimmer, McNamara, Houston, & Marshall, 2012). Learning rules are often composed of information updating and decision making components and for several decades each has been studied separately and sometimes even misinterpreted. Moreover, little attention has been devoted to exploring the ecological conditions under which decision should be based on learning.

I however believe that in some context learning is as powerful as, perhaps even more powerful, than evolution in predicting animal behavior. To explore this hypothesis, I have taken three steps that I will explain in more details in the following chapters:

1. I proposed a learning-based model for producer-scrouter game, a social foraging behavior that has been extensively studied in behavioral ecology. I then compared the results of the model with previously published models, experiments, and observation. This model, furthermore, provided novel prediction concerning an increase in scrounging behavior when patch quality varies.
2. I tested this new prediction using captive foraging flocks of male zebra finches (*Taeniopygia guttata*) in two conditions with variable and non-variable patch quality. The results show that as individuals gain experience in

the variable environment, they increase their scrounging behaviour. This support of the model emphasizes the power of the learning rule approach in studying animal behavior.

3. To identify situations where one should look for learning, I explored the conditions under which learning is expected to evolve in a frequency-dependent context like a producer-scrounger-game. In doing so, I studied the evolutionary properties of learning rules and addressed both limitations of earlier studies by pitting fixed against learning agents while exploring the effects of no fluctuations, within or between-generation fluctuations.

In the remaining of this chapter, I review earlier studies in behavioral ecology concerning learning and decision rule.

### 0.3 State of knowledge

#### 0.3.1 Frequency-dependency, tactic and strategy

One circumstance in which learning could play a role as important as evolution, is when individuals adjust in a frequency-dependent context, where the payoff of a tactic depends on the combined frequency of all tactics used in the group simultaneously (Maynard Smith, 1982). Tactics are known as the behavioral components of a strategy (Krebs & Davies, 1993), e.g. wander vs. defend, or the alternative phenotypes generated by a strategy or decision rule (Gross, 1996; Tomkins & Hazel, 2007), such as fight with probability 0.3 and sneak with probability 0.7. The word “strategy” has been used differently over time. For example, Krebs & Davies (1993) defined strategy as what animals do when competing for a scarce resource (Krebs & Davies, 1993), e.g. if young, wander; if old, defend a territory. Houston & McNamara (2005), however, defined strategy as a

rule for choosing a tactic as a function of state and possibly time. Strategy sometimes has been defined as a genetically based decision rule that results in the allocation of the somatic and reproductive effort of an organism, such as energy and development, among alternative phenotypes (Gross, 1996; Tomkins & Hazel, 2007), e.g. fight if larger than X and sneak if smaller than X, or fight with prob. 0.3 and sneak with prob. 0.7. Throughout this document, tactic and action refer to the same concept. I use the definition by Krebs & Davies (1993) for tactic and Houston & McNamara's (2005) definition for strategy:

- Tactic (Action): The behavioral components of a strategy (e.g. search, join, eat)
- Strategy: A rule for choosing a tactic as a function of state and possibly time (e.g. join early in the morning and search other times, search all the time)

What becomes a tactic or a strategy can depend on how and at which level we are asking a question about the behavior. If one asks, for instance, *how does a ground feeding bird find food*, then either “search” for a new patch or “join” a patch already found by others become possible tactics. In this case, possible strategies can be “use only one tactic (pure)” or “join with probability of 0.7 (mixed)”. If, however, the question is: *how should an individual forage in a dynamic environment*, then the tactics can be either pure or mixed. In such a situation the strategies can become “*Always use mixed*” or “*use mixed only if patch richness varies*”. Hence, depending on the question, the exact definition of expressions such as pure or mixed can be either tactic or strategy.

### 0.3.2 Evolutionary game theory

When different tactics are not simultaneously compatible in a single individual, but are simultaneously compatible in a population, there will be a stable equilibrium

number of individuals playing each tactic. Maynard Smith (1982) says this equilibrium can be reached through pure strategies or mixed strategies. In a pure strategy, individuals choose one action because, as a result of small size, lack of experience, etc., they cannot effectively perform any other. In a mixed strategy, the payoffs of actions depend on the frequencies with which they are performed (Table 0.1a). He also introduces the concept of Evolutionarily Stable Strategy (ESS) which is a strategy (or a combination of tactics) which, once fixed in a population, cannot be invaded by any other strategy (or combination of tactics).

Krebs & Davies (1993) redefined the pure strategy (as a polymorphism), mixed strategy, and also add behavioral assessment as the third type of solution (Table 0.1b). In a polymorphism, different individuals play different strategies and the frequencies of the morphs will be fixed by frequency dependent selection. A polymorphism may arise because of genetic differences or environmental differences. In a mixed ESS, individuals play a combination of different strategies in a proportion that satisfies the ESS. In behavioral assessment, instead of having a fixed population-wide rule, every individual is flexible and bases its decision on an assessment of what others in the population are doing. Examples of having ESS in a population are shown in Table 0.2.

Gross (1996) considered differences in individual state, e.g. energy level, and proposed a new "conditional strategy" that switches tactics depending on individual state (Table 0.1c). In this classification, alternative and mixed strategies are very similar to the pure and mixed strategies in Krebs & Davies (1993). In the conditional strategy an individual changes its behavior in response to changes in its internal state, which can be frequency dependent or independent. Here, the tactic related to each state is genetically determined. Throughout this document, I will not consider differences in individual state and will use the terminology of Krebs & Davies (1993).

For both pure strategies (polymorphism) and mixed strategies, selection pressure is on the decision, i.e. the equilibrium or switching point is determined by natural selection. The strategy thus remains the same during an individual's lifetime. In behavioral assessment, however, an individual changes its behavior flexibly during its lifetime by using a decision rule, and selection pressure acts on the decision rule.

**Table 0.1** Different ways of having ESS in a population proposed by: a) Maynard smith (1982), b) Krebs & Davies (1993), c) Gross (1996)

(a)		
Pure strategy	Mixed strategy	
<ul style="list-style-type: none"><li>- Fixed behavior</li><li>- Genetic polymorphism selects one tactic because others are not possible</li></ul>	<ul style="list-style-type: none"><li>- Fixed behavior</li><li>- Frequency dependent selection acts to equalize the payoffs for different tactics</li></ul>	
(b)		
Polymorphism	Mixed strategy	Behavioral assessment
<ul style="list-style-type: none"><li>- Fixed behavior</li><li>- Frequency dependent selection</li><li>- Two or more strategies with unequal payoffs because of genetic differences or environmental differences</li></ul>	<ul style="list-style-type: none"><li>- Fixed behavior</li><li>- Frequency dependent selection</li><li>- One strategy: different tactics in a proportion that satisfies the mixed ESS with equal payoffs</li></ul>	<ul style="list-style-type: none"><li>- Flexible behavior</li><li>- Frequency dependent selection</li><li>- Two or more strategies with equal payoffs</li></ul>
(c)		
Alternative strategies	Mixed strategy (with alternative tactics)	Conditional strategy (with alternative tactics)
<ul style="list-style-type: none"><li>- Genetic polymorphism</li><li>- Frequency dependent selection</li><li>- Two or more strategies with equal average fitness</li><li>- Evolutionarily Stable State Frequency</li></ul>	<ul style="list-style-type: none"><li>- Genetic monomorphism</li><li>- Frequency dependent selection</li><li>- One strategy: tactics with equal average fitness</li><li>- Evolutionarily Stable Strategy Frequency</li></ul>	<ul style="list-style-type: none"><li>- Genetic monomorphism</li><li>- Status dependent selection (with/out frequency dependent selection)</li><li>- One strategy: tactics with unequal average fitness</li><li>- Evolutionarily Stable Strategy Switch point</li></ul>

**Table 0.2      Examples of having an ESS in a population**

Strategy equilibrium by	Example
<b>Polymorphism</b> (Pure strategies / Alternatives)	Sex: male vs. female Male Coho salmon: jack vs. hooknose Male fig wasp: fight vs. disperse Male stickleback throat: bright red vs. dull
<b>Mixed strategy</b>	Reproduction in hermaphrodite organisms Female digger wasp: dig vs. enter
<b>Behavioral assessment</b>	Finch & pigeon: producer vs. scrounger Male bullfrog: caller vs. satellite

### 0.3.3 Learning rule

Decision rules assign the probability of displaying different tactics. If the probability of playing a tactic is based on previous experience, it is called a learning rule (Harley, 1981; Kacelnik & Krebs, 1985; Stephens, 1991). Learning rules typically have two components, one that forms an estimate of the value of alternative tactics from experience, and another that decides on a tactic based on these estimates. The approach can be traced to Estes (1950) who proposed the first statistical theory of learning. In a series of papers, Bush and Mosteller developed the linear operator model for learning that is based on previous experiences (Bush & Mosteller, 1951a, 1951b, 1953, 1955; Mosteller, 1958). In these primary models, only the occurrence of an event, i.e. experiencing a reward or a punishment, is important and the size of the reward or punishment is not considered. This means that receiving a big reward is not different from receiving a small reward.

The linear operator model has been modified (Table 0.3) and attention has specifically been directed to the decision component (Killeen, 1984; Krebs, Kacelnik, & Taylor, 1978; Lea & Dow, 1984; Lester, 1984; McNamara & Houston, 1985; Myerson & Miezin, 1980; Ollason, 1980; Pulliam & Dunford, 1980) in an effort to reconcile learning rules with phenomena such as ideal free distribution (Krebs & Davies, 1993), matching ratio (Herrnstein, 1961, 1970; Herrnstein & Heyman, 1979), evolutionarily stable strategy (ESS), evolutionarily stable (ES) learning rule, and rule for ESS (Hines & Bishop, 1983; Maynard Smith, 1982). Looking for the ES learning rule, Harley (1981) modified the linear operator by adding residuals to the estimation component (Table 0.3). These residuals simulate innate values for different options that provide an absolute floor below which the estimates of alternatives cannot go. The residuals according to him insure that all options get sampled at least periodically and he claimed that the decision component should be based on the matching ratio, i.e. allocating responses to alternative in proportion to their relative estimated payoffs. He claimed that this rule, called the relative payoff sum, is an evolutionary

stable learning rule. This incited a debate between Harley and Houston (Harley, 1983, 1987; Houston, 1983; Houston & Sumida, 1987) on the optimality of relative payoff sum, and more specifically the optimality of the matching ratio, when compared to other learning rules such as linear operator and perfect memory (Table 0.3).

**Table 0.3** Selected important learning rules in behavioral ecology. Here, only the updating part is mentioned. Depending on the experiment, either matching or maximizing was used for decision making.  $T_i(t)$  is the estimated value of tactic  $i$  at time  $t$ .  $P_i(t)$  is the observed payoff for tactic  $i$  at time  $t$ .  $x$  is the memory factor ( $1 - x$  is the learning rate) used to adjust the estimation with new experience.

Learning rule	Formula	Description
Pulliam & Dunford's linear operator (Pulliam & Dunford, 1980)	$T_i(t) = x.T_i(t-1) + (1-x).P_i(t)$	
Relative Payoff Sum (Harley, 1981)	$T_i(t) = x.T_i(t-1) + (1-x).r_i + P_i(t)$	$r_i$ is the residual value for tactic $i$ representing the innate value for that tactic.
Lester's linear operator (Lester, 1984)	$T_i(t) = P_i(t).T_i(t-1)$	
Dual process (Kacelnik & Krebs, 1985)	$T_i(t) = x.T_i(t-1) + (1-x).r_i + P_i(t)$ $C_i(t) = C_i(t-1).(1-T_i(t-1))^n$ $A_i(t) = T_i(t).C_i(t)$	$n$ is the time passed since the last time tactic $i$ gave payoff. $C_i(t)$ is a mechanism to decrease preference to tactic $i$ if it does not pay back for a while. $A_i(t)$ is the total estimated value for tactic $i$ at time $t$ used in decision making.
Perfect memory (Houston & Sumida, 1987)	$T_i(t) = \frac{\alpha + \sum_t P_i(t)}{\beta + N_i(t)}$	$\alpha$ and $\beta$ are the parameter of learning rules and same values are used for different tactics. $N_i(t)$ is the number of times tactic $i$ is used until time $t$ .

Regelmann (1984) used the relative payoff sum in simulations of the ideal free distribution and ideal despotic distribution, where some animals are good competitors while the others are poor in competition and good in defending a food source. He extended relative payoff sum with and without travelling cost between food patches and used equal values for all residuals. The distribution of individuals among different patches was very close to that predicted by the matching ratio. Experiments by Milinski (1984) show that stickleback behavior in a foraging context is consistent with the predictions of Harley's and Regelmann's relative payoff sum. Based on the experiment and the reported results, however, linear operator, which was not considered by Milinski, could also explain the results.

Lester (1984) proposed a simplified version of linear operator (Table 0.3) based on the matching ratio where the probability of displaying each behavior is proportional to its relative profitability, i.e. prey/time. Kacelnik & Krebs (1985) compared Harley's relative payoff sum, Lester's linear operator, and a proposed dual-process decision rule similar to linear operator (Table 0.3) in three series of experiments with sticklebacks, goldfish, and starlings all in foraging contexts. They claim that the relative payoff sum and Lester's linear operator are not good when there is rapid patch depletion, and that models based on simple averages of the past and present experience are insufficient to describe some important features of the way changes in food availability are detected by experimental animals. This paper was followed by a survey of learning rules (Kacelnik, Krebs, & Ens, 1987) comparing Pulliam & Dunford (1980)'s linear operator (Table 0.3), Harley's relative payoff sum, Lester's linear operator, the dual-process decision rule, and Killeen (1984)'s incentive theory where memory factor is also updated by linear operator. They compared the results of simulations with the results from experiments with three adult starlings in random ratio and fixed ratio feeders. In both conditions, the feeder that gave the higher rate of food dropped rapidly to zero and the starlings decided to forage on the other one. The trend of switching to other feeder is called the drop to zero phenomena which, of the

models tested, only the dual-process model predicted it well. Results seem to be as expected from a partial reinforcement extinction effect that is the situation where an animal stops responding sooner to a less variable option than a more variable option, when both stop providing a reward. They did not however consider Bayesian approach either in the models or in the analysis. But they claimed that under some constraints Bayesian approaches may be more powerful than linear operator to describe the results. The Bayesian statistical approach is known as a competitor to linear operator based learning rules. To study the distinction between linear operator and statistical approaches, McNamara & Houston (1987) show how linear operator comes from a statistical approach by choosing memory factor properly as a function of variance and mean in reward.

Because information has a central role in learning, Stephens (1989) used an analytical approach to approximate the value of information. His results suggest that two factors are important in determining its value: first, the optimal behavior that would be chosen if the decision maker knew which subtype of the resource it faced; second, the cost of small deviations from the optimal behavior. He shows that the value of information is approximately equal to the product of the mean cost of small deviations from the subtype optima and the variance of a modified distribution of the optimal behaviors. It is interesting to note that in this paper Stephens shows that a linear operator offers a greedy approach: maximization of immediate reward (Cormen, 2001). However, greedy algorithms are not necessarily optimal such that the linear operator could be suboptimal under some ecological conditions (Sutton & Barto, 1998). Linear operator, also known as the Rescorla-Wagner rule, was shown to be favored by natural selection in a simulation study (Trimmer et al., 2012) and provides the basic structure for many more complicated forms of reinforcement learning rules (Sutton & Barto, 1998).

#### 0.3.4 On the extent of matching

It has long been claimed that an optimal decision maker must adopt alternative actions in proportion to their observed or estimated payoffs, called the matching ratio (matching law, law of effect) (Herrnstein, 1961, 1970; Herrnstein & Heyman, 1979). Matching allows an individual to sample all of the options, while still exploiting the best option most often. This is in contrast to maximizing, where an individual exploits only the best option. Matching vs. maximizing have been frequently tested with both models (Lea & Dow, 1984; Lester, 1984; McNamara & Houston, 1985; Myerson & Miezin, 1980; Staddon, 1977; Tracy & Seaman, 1995) and experiments (Krebs et al., 1978; Milinski, 1979; Plowright & Plowright, 1987). Several experiments and models have shown that in some condition matching is not an optimal behavior (Staddon, Hinson, & Kram, 1981) and animals behave differently from matching (Herrnstein & Heyman, 1979). Houston (Harley, 1983, 1987; Houston, 1983; Houston & Sumida, 1987) therefore was highly critical of the use of matching in Harley's relative payoff sum.

Instead of pure matching or pure maximizing, other methods of balancing sampling and exploitation were also proposed. In the  $\epsilon$ -greedy strategy (Sutton & Barto, 1998), the decision maker samples all options equally with probability  $\epsilon$ , and exploits the best known option with probability  $1-\epsilon$ . When it enters a new environment, it can have large  $\epsilon$  during the early trials allowing it to decay gradually. So, the decision-maker samples more at the beginning of trials to gather a coarse estimation of the payoffs. Then as  $\epsilon$  moves closer to zero, it gradually fine tunes its knowledge. In the failure strategy (Thuijsman, Peleg, Amitai, & Shmida, 1995), an animal leaves option  $A$  after  $p_A$  failed trials, and leave  $B$  after  $p_B$  failed trials. It updates  $p_A$  and  $p_B$  with new experiences. When it selects a new option, it selects each option relative to its  $p$ . Thuijsman et al. (1995) compared  $\epsilon$ -greedy and failure strategies along with linear operator in simulations of an ideal free distribution problem and the matching law in bees. They find that the optimal balance between sampling (exploration)-exploitation

is sensitive to time horizon. The results show that bees in groups have a matching behavior that is optimal and successfully converge to the ideal free distribution. Matching behavior is however not optimal when they forage solitarily.

#### 0.4 Thesis overview

The following chapters explore the hypothesis that in some contexts learning is as powerful as evolutionary models in predicting animal behavior:

Chapter I. A unified modelling approach for producer-scrouter games in complex ecological conditions

Chapter II. Zebra finches scrounge more when patches vary in quality: Experimental support of the linear operator learning rule

Chapter III. The evolution of learning and non-learning rules in dynamic environments: A test of the behavioural gambit

Chapter IV. General discussion and future directions

In chapter I, I propose an agent-based model using a linear operator learning rule as the decision mechanism for producer-scrouter game, a social foraging behavior that has been extensively studied in behavioral ecology. The model provides a unified framework from which to predict the effects on the expected equilibrium of producers and scroungers of group size, metabolic requirement, finder's advantage, food intake rate, cost of searching, cost of joining, patch encounter probability, and patch richness. I then compared the results of the model with previously published models, experiments, and observation. The simulation results replicate almost every producer-scrouter prediction and experimental result published to date such that it provides a more general tool than any single game-theoretic model to predict behaviour under frequency-dependent conditions. The model furthermore allows us to develop a novel

prediction about the behavior in a more realistic environment of variable patch richness.

In chapter II, I test Chapter I's novel prediction that increased patch quality variances calls for high equilibrium frequency of scrounging. In doing so, we used captive foraging flocks of male zebra finches in two conditions with variable and non-variable patch quality. The results show that as individuals gain experience in the variable environment, they increase their scrounging behavior. This supports the prediction of the linear operator model, emphasizing the power of the learning rule approach in studying animal behavior.

In chapter III, I explored the conditions under which a linear operator type of learning is expected to evolve in a frequency-dependent context like a producer-scrounger game. In doing so, I addressed both limitations of earlier studies by putting fixed against learning agents while exploring the effects of no fluctuations, as well as within or between-generation fluctuations. The results show that at least within the context of a producer-scrounger game, we can accept the behavioral gambit if we are only interested in the population level equilibrium in static environment, because learning generates a combination of strategy use that corresponds exactly to what would be expected from evolutionary model based on selection acting on genetic alternatives. The results however emphasize that we need to consider learning if there is environmental changes or if we want to predict the behavior at individual-level.

Finally, chapter IV provides a general discussion of the earlier chapters and their relation to my original hypothesis. I identified the key contributions of this thesis and discussed the directions for future research.

[Cette page a été laissée intentionnellement blanche]

## CHAPTER I

### A UNIFIED MODELLING APPROACH FOR PRODUCER-SCROUNGER GAMES IN COMPLEX ECOLOGICAL CONDITIONS

Afshar, M., & Giraldeau, L.-A. (2014). A unified modelling approach for producer–scrounger games in complex ecological conditions. *Animal Behaviour*, 96, 167–176. doi:10.1016/j.anbehav.2014.07.022

#### 1.1 Abstract

Animal decision making in frequency-dependent situations, where the payoff of an action depends on the actions of others, has gained prominence in behavioural ecology and in social foraging in particular. One such situation involves cases where an animal can search for a new resource (produce) or join what others have already found (scrounge). A number of game-theoretic models have been proposed to predict the equilibrium combination of producer and scrounger strategists based on the evolutionarily stable strategy. However, each game model can only handle a few environmental parameters at a time and none address the flexible use of tactics that allows individuals to respond quickly and adaptively to changes in payoffs. In this study we propose an agent-based model using a linear operator learning rule as the decision mechanism. The model provides a unified framework from which to predict the effects on the expected equilibrium of producers and scroungers of group size, metabolic requirement, finder's advantage, food intake rate, cost of searching, cost of joining, patch encounter probability and patch richness. The simulation results replicate almost every producer–scrounger prediction and experimental result

published to date such that the simulation provides a more general tool than any single game-theoretic model to predict behaviour under frequency-dependent conditions. The model furthermore allows us to develop a novel prediction about foraging behaviour in a more realistic environment of variable patch richness. By modelling the operation of a plausible decision rule, we can explore the validity of the behavioural gambit, the assumption that the unspecified decision mechanisms of game-theoretic models faithfully reproduce outcomes expected of natural selection operating over generations on fixed alternatives. We suggest that this simulation model can provide a tool for others to explore and predict the effect of more complex and hence realistic foraging conditions on individual levels of producer and scrounger use.

**Keywords:** decision making, evolutionary game theory, frequency-dependent selection, learning rule, producer–scrounger, risk sensitivity, social foraging

## 1.2 Introduction

Animals of the same species sometimes use different tactics when engaged in foraging or mate choice (Davies et al., 2012). The choice of tactic is generally made as a function of state and or time, and the rule that governs this choice is referred to as a strategy (e.g. for the tactics of wander or defend, the strategy might be ‘if young, wander; if old, defend’; Gross, 1996; Houston & McNamara, 2005; Tomkins & Hazel, 2007). Behavioural ecologists are interested in why different tactics coexist in the same population under the same ecological conditions. One possible explanation is that none of the tactics can reach fixation because their fitnesses are negatively correlated with their frequencies in the population: a negative frequency dependence that predicts stable mixtures of tactics (Barnard & Sibly, 1981; Beauchamp, Bélisle, & Giraldeau, 1997; Coolen, Giraldeau, & Lavoie, 2001; Giraldeau & Livoreil, 2000; Giraldeau, Soos, & Beauchamp, 1994; Maynard Smith, 1982; Morand-Ferron, Giraldeau, & Lefebvre, 2007; Mottley & Giraldeau, 2000; Wu & Giraldeau, 2005).

For more than three decades, evolutionary game theory has been the main tool used in frequency dependent situations to predict an equilibrium mixture of tactics based on evolutionarily stable strategy (ESS) theory. ESS theory assumes that strategies (and so the equilibrium mixture of tactics) are determined by natural selection acting on genetic alternatives (Maynard Smith, 1982). It can predict the equilibrium when there are either pure or mixed strategies in the population. A pure strategy is when each individual always uses either one tactic or the other, and so does not alternate between them (Gross, 1996). It can arise because of genetic differences between individuals or environmental differences (Davies et al., 2012) or when the frequency of each strategy is set over generations by frequency-dependent selection and the population then appears dimorphic (Bergstrom & Godfrey-Smith, 1998). A mixed strategy is when each individual plays a fixed combination of different tactics so that the ESS is satisfied by the overall combination of the population (Gross, 1996; Tomkins & Hazel, 2007).

ESS theory only predicts the distributions of tactics. It however remains silent about the ways in which these distributions are realized and makes no explicit prediction as to the distribution of tactics in an infinite population accommodating anything from monomorphism, everyone uses the same mixture of tactics, to extreme polymorphism, where each individual adopts a unique combination of tactics (Bergstrom & Godfrey-Smith, 1998). This leaves some important questions about how complexity and diversity are realized within individuals and across a population. These questions have been addressed in Bergstrom and Godfrey-Smith (1998), Maynard Smith (1988) and Vickery (1987, 1988). In a finite population, only pure strategies are expected from purely deterministic dynamics and no mixed strategy can be an ESS (Vickery, 1987, 1988). However, if the frequency of one pure strategist is low at ESS, stochastic environmental fluctuations may cause its extinction. In this case, mixed strategies are favoured and can invade the population (Bergstrom &

Godfrey-Smith, 1998; Maynard Smith, 1988; Orzack & Hines, 2005). Deterministic and stochastic methods can thus have different opposing predictions.

ESS theory also accommodates situations where the fixed combination of each tactic is acquired by learning. In this case it assumes that an evolutionary stable (ES) learning rule exists whose characteristics lead the population within a generation to the same equilibrium combination as would frequency-dependent selection over generations (Maynard Smith, 1982). It is unclear whether a learned equilibrium should generate a monomorphic, dimorphic or polymorphic population. Whatever the type of strategy expected at the ESS, be it pure, mixed or learned, the ESS theory approach in the context of animal behaviour suffers three important limitations that we consider in turn.

The first is that, if selection is acting on genetic alternatives, individual behaviour must be genetically fixed. Yet, in many frequency dependent tasks, such as the producer-scrouter (PS) game, prisoner's dilemma and ideal free distribution, behaviour does not appear genetically fixed but rather a learned ESS (Belmaker, Motro, Feldman, & Lotem, 2012; Lendvai, Barta, & Liker, 2004; Morand-Ferron & Giraldeau, 2010; Morand-Ferron, Varennes, & Giraldeau, 2011; Morand-Ferron, Wu, & Giraldeau, 2011). Furthermore, many behavioural games are probably played several times and under different conditions within a generation (Stephens & Clements, 1998). The optimality of any tactic thus changes during an individual's lifetime. Here, any genetically fixed mechanism may be handicapped to respond to the changes in the structure of the games. Therefore, it seems unlikely that the optimality of such a game would be fixed in merely a genetic mechanism during evolutionary timescale. The expectation that a learned ESS will always generate a combination of strategy use that corresponds to that expected from selection acting on genetic alternatives remains an untested assumption (Harley, 1981; Harley & Maynard Smith, 1983; Hines & Bishop, 1983; Maynard Smith, 1982; Tracy &

Seaman, 1995), and its acceptance, what Giraldeau and Dubois (2008) described as accepting the 'behavioural gambit', may be incorrect (Fawcett et al., 2013; McNamara & Houston, 2009).

The second limitation of the ESS approach is its inability to deal effectively with complicated situations (Fawcett et al., 2014). For example, the possibility that ecological conditions (the frequency-dependent nature of many games and the specificities of the decision mechanism) may constrain the equilibrium that a population can be expected to reach is also largely ignored (Fawcett et al., 2013; Grodzinski, Dor, & Lotem, 2011; McNamara & Houston, 2009). Because of the analytical complexity of solving ESS models in more realistic situations, most of these models consider conditions in which the organism is omniscient and the environment is oversimplified and often static. The predictions from these models then encourage experimentalists to test their predictions in equally simplistic and hence unrealistic conditions. This oversimplification of conditions in which frequency dependence is considered has been justly criticized for being too far removed from the world in which organisms have evolved (Fawcett et al., 2014).

Finally, behaviourists who wish to predict how individuals behave in a population cannot rely on ESS theory. This is because ESS theory only predicts the average stable mixtures of strategies across populations, not the tactic combination for each individual within a population. ESS theory remains silent concerning whether all individuals should adopt the same proportion of each tactic (monomorphism), or whether some individuals should use higher proportions of one tactic while others use lower proportions (dimorphism). If we are to predict individual behaviour it would be important, therefore, to develop a way of predicting it in situations of frequency dependence that rely on learning. Such an approach should allow us to predict individual behaviour and explore the effects of more realistic environmental conditions not only on individual tactic use, but on its expected variation among

individuals within populations. A behavioural mechanism based on learning not only responds to a repeated game with changing conditions during an individual's lifetime (Stephens & Clements, 1998), it can furthermore be used in multiple games affecting different equilibriums simultaneously.

To this end, in this study we address the first and the second limitations of ESS theory by proposing a simulation model that predicts the same population equilibrium as an ESS. This model incorporates a commonly used learning rule as the individual's cognitive decision mechanism (Arbilly & Laland, 2014; Beauchamp et al., 1997; Beauchamp, 2000; Dridi & Lehmann, 2014; Dubois, Morand-Ferron, & Giraldeau, 2010; Hamblin & Giraldeau, 2009; Katsnelson, Motro, Feldman, & Lotem, 2012; Kurvers, Hamblin, & Giraldeau, 2012). We then tackle the second limitation by adding complexity to the environment beyond what has been possible with game-theoretic PS models and the experimental studies they generated. We use the model to bring a novel prediction about the expected effect of patch quality variance on the equilibrium use of search and join tactics.

In our simulation we assume that agents use their experience of payoffs to decide on the tactic they adopt: we assume learning. We acknowledge that the evolution and use of learning in the context of such frequency-dependent games is not straightforward. For instance, Dubois et al. (2010) showed that although learning individuals experience an initial advantage in a producer–scrounger game, they never evolve to fixation. The population is thus made up of a mixture of fixed and flexible (learning) players. Moreover, by explicitly modelling learning in a producer–scrounger game, Katsnelson et al. (2012) showed that learning can evolve to fixation but only when there is sufficient environmental fluctuations or when there is asymmetry in individual's phenotype. Notwithstanding these difficulties concerning the origins of learning in groups, we choose to assume that our agents used learning rules to decide on tactic use, an assumption that is supported by a number of empirical studies.

Learning rules, for instance, have been used in psychology, ethology and behavioural ecology for many years to model animal behaviour individually or in groups (Arbilly & Laland, 2014; Bush & Mosteller, 1951a; Dridi & Lehmann, 2014; Estes, 1950; Hamblin & Giraldeau, 2009; Harley, 1981; Herrnstein, 1961; Krebs et al., 1978; McNamara & Houston, 1985; Stephens, 1989). Moreover, in a series of producer–scrounger experiments with nutmeg mannikins, *Lonchura punctulata*, Morand-Ferron and Giraldeau (2010) found that the group-level proportion of the join tactic adjusts to different environmental conditions by learning the payoffs associated with each tactic. Birds trained in a low-scrounging condition joined less than those trained in a high-scrounging condition and continued to do so even when subsequently tested in the same conditions. Similarly, by direct manipulation of the success rate experienced by adult house sparrows, *Passer domesticus*, Belmaker et al. (2012) could change the birds' subsequent strategy choice. Birds more frequently used the strategy with which they were more successful, and the differences in tactic use were correlated with differences in success rate.

The learning rule approach, coupled with an agent-based simulation, could specify more clearly the sensitivity and the attributes of the equilibria to the animal decision making mechanism. In this case, the population equilibrium comes from the interactions of individuals with each other and with the environment at an ecological timescale rather than from the action of natural selection over generations. Furthermore, it is now possible to study the behaviour of each animal individually, and ask, for instance, how the parameters of the learning rule (e.g. its memory window, sampling frequency) and the decision rule (matching versus maximizing) influence not only the equilibrium but the extent of individual variation in strategy use at equilibrium.

We use an agent-based simulation model operating in a PS game context for which there already exists a number of published models and empirical results. This allows

us to test the validity of our simulation by comparing its results against the predictions of several published PS models as well as the results reported from a number of experimental studies. Our simulation model uses a linear operator (LO) learning rule, the most commonly used learning rule in behavioural ecology. The LO is simple and has fewer parameters compared with alternative learning rules (e.g. relative payoff sum: Harley, 1981; perfect memory: Houston & Sumida, 1987). Linear operator, also known as the Rescorla–Wagner rule, was shown to be favoured by natural selection in a simulation study (Trimmer et al., 2012) and provides the basic structure for many more complicated forms of reinforcement learning rules (Sutton & Barto, 1998). LO has been previously used in the PS game in competition with other learning rules (Beauchamp, 2000a; Hamblin & Giraldeau, 2009), alone (Arbilly, Motro, Feldman, & Lotem, 2010) and with nonlearning strategies (Arbilly & Laland, 2014; Katsnelson et al., 2012). The interest of these earlier studies on learning rules resided in the evolution of the learning rule itself and in determining the performance of the rules in reaching an equilibrium, namely their evolutionary stability. Here, instead, we simply assume the LO is the stable learning rule and use it to explore the expected effects of a number of different ecological parameters on the equilibrium of expressed tactics in a PS game.

### 1.3 Producer-scrounger game

When foraging in a group, an individual can obtain food by searching for a new food patch or by joining patches already found by others (Barnard & Sibly, 1981). In some cases, such as some ground-feeding birds, the two alternatives are temporally incompatible because the animal cannot look at the ground to search for food (head down) and watch others in the group to join them when they find food (head up) simultaneously (Coolen et al., 2001). Individuals may adopt pure strategies, a situation where they use only one tactic: either search only (pure producer) or join only (pure scrounger). They can have different combinations of search and join tactics. Thus, to be more accurate, we define a ‘producer’ more broadly as an

individual that searches more than it joins, and a 'scrounger' as an individual that joins more than it searches. This means there is variation in the search-join combination within this broadly defined producer, so a 'producer' does not represent a single strategy but a continuum of unique strategies with different proportions of join tactic. The same holds for the 'scrounger'. ESS theory predicts a specific evolutionary stable frequency of join where the payoffs of both search and join tactics are equal (Giraldeau & Beauchamp, 1999). The general question in a producer-scrounger game is to find this equilibrium and explain how changes in environmental parameters affect it.

#### 1.4 Linear operator model

In this model of social foraging, a given group of individuals forages in a virtual environment, represented by a matrix. Each individual is a separate learner and tries to maximize the amount of food gained: its payoff. It uses a linear operator learning rule to estimate the value of each tactic (search or join) by the following equation:  $V_i(t) = V_i(t-1) + x(P_i(t) - V_i(t-1))$ ,  $V_i(0) = \text{unif}(0,1)$  where  $V_i(t)$  and  $P_i(t)$  are the estimated value and the observed payoff, respectively, for tactic  $i$  at time  $t$ .  $x$  is the learning rate ( $1 - x$  is the memory factor). The first term,  $V_i(t-1)$ , is the previous estimation, and the second term,  $P_i(t) - V_i(t-1)$ , is the difference between the observed payoff and the previous estimated value. This process updates the estimated value of the most recent tactic. For the other tactic that has not been sampled at time  $t$ , the estimated value remains unchanged ( $V_i(t) = V_i(t-1)$ ).  $V_i(0)$  is the initial condition, which is chosen at random from a uniform distribution between zero and one. This causes the average initial preference of the population to be 50% for the search tactic and 50% for the join tactic.

So, in the linear operator, with each new experience, the estimated value moves towards the recent observed payoff by a rate of  $x$ . At each time step, the learner decides which of the two tactics to express by adopting alternative tactics in

proportion to their estimated values, called matching ratio. This choice mechanism allows the individual to sample both tactics, while still exploiting the best one proportionally more often.

In the simulation model, time is divided into discrete steps. Each time step can be interpreted as one sampling event in the real world. In each time step, an individual adopts a tactic. If it searches, then whether it encounters a patch or not depends on the patch encounter probability. Simultaneous discoveries by different individuals are possible when they concurrently use the search tactic. Individuals that join, see the location of all discovered patches and they join the nearest patch. The patch may be depleted just before the individual arrives because of other individuals already present at the patch.

Upon encountering a food patch, the finder receives a fraction of the food in the patch: the finder's advantage. This represents the amount of food the finder would be able to eat before the joiners arrive. In subsequent time steps, all individuals at the patch, the finder and those that have joined, receive food at the same food intake rate. Normally, due to competition, the rate of food intake is affected by the number of scroungers at the patch, but for simplicity, we kept intake rate fixed. The individuals at the patch continue eating until the patch is depleted, at which point each individual chooses a tactic using the LO. We assume that in each time step individuals require a minimum amount of energy to survive, a parameter we termed metabolic requirement. If an individual does not meet its metabolic requirement, it survives but gets a negative payoff to ensure that there is value in exploiting the better tactic more often. In the literature, terms such as energy reserve or energy requirement (Koops & Giraldeau, 1996; Lendvai et al., 2004; Wu & Giraldeau, 2005) have also been used for the same concept. We however use metabolic requirement as it is well defined and easier to understand. Individuals must also pay a cost of searching or joining. These

costs are absolute values and are incurred in each time step, depending on the tactic played.

Each simulation involves 500 time steps. Because of the learning mechanism and the random initial condition, the proportion of individuals using the join tactic changes rapidly at the beginning of each simulation, then gradually converges to a stable point. To determine the predicted combination of tactics, we consider only the second half of the simulation, when values have stabilized.

The general parameter settings for all simulations are shown in Table 1.1. We chose parameter values to be as similar as possible to empirical simulation and modelling studies to which they will be related. The range of parameters was chosen to allow maximum exploration of their effect while keeping the simulation computationally possible. To study the effect of an ecological parameter on the expected PS equilibrium, we conducted a set of 30 simulations for each value of that parameter, while all the other general parameters remained unchanged. Thus, the results are averaged over the 30 simulations. All simulations were conducted in MatLab (R2013a) for Windows (Mathworks, Natick, MA, U.S.A.).

**Table 1.1** Parameter settings for the linear operator model in the producer-scrounger game.

Parameter	Symbol	Value	Range
Grid size	$G$	15 x 15	-
Maximum simulation time step	$T_{max}$	500	-
Group size	$N$	8	5-50
Metabolic requirement	$C_{MR}$	0.125	0-2
Finder's advantage	$a$	2.5	1-20
Food intake rate	$b$	1	0.5-2.5
Cost of searching	$C_{search}$	0.125	0-2
Cost of joining	$C_{join}$	0.125	0-2
Patch encounter probability	$p$	0.2	0.1-1
Patch richness (mean)	$\mu_F$	20	5-35
Learning rate	$x$	0.3	-

To see the effect of each parameter, we kept all parameters equal to their values and changed only one parameter at a time based on the mentioned range.

To test the simulation results with published predictions and observations, we collected predictions from published models and empirical results for PS games by conducting a literature search using the Scopus bibliographic database. We searched for ‘producer scrounger’, without quotation marks and as two separate words, in the article title, abstract and keyword fields for papers published in and before 2013. We located 121 publications, of which we excluded three because they were not English and two because they were duplicate entries. Another 29 papers were excluded because they used the concept of ‘producer scrounger’ in other domains, such as group search optimizers in computer science, and did not provide results in the area of foraging behaviour. We expanded our search by following up the reference sections of these papers, which provided an additional 11 publications that had not appeared in our original Scopus search. In the end, we used the models or experimental results reported in 98 publications.

We chose to explore the effect of the following eight ecological parameters on the expected equilibrium mixture of tactics: group size, metabolic requirement, finder’s advantage (the amount of food from a patch that goes to the exclusive use of the finder), food intake rate, cost of searching, cost of joining, patch encounter probability and patch richness.

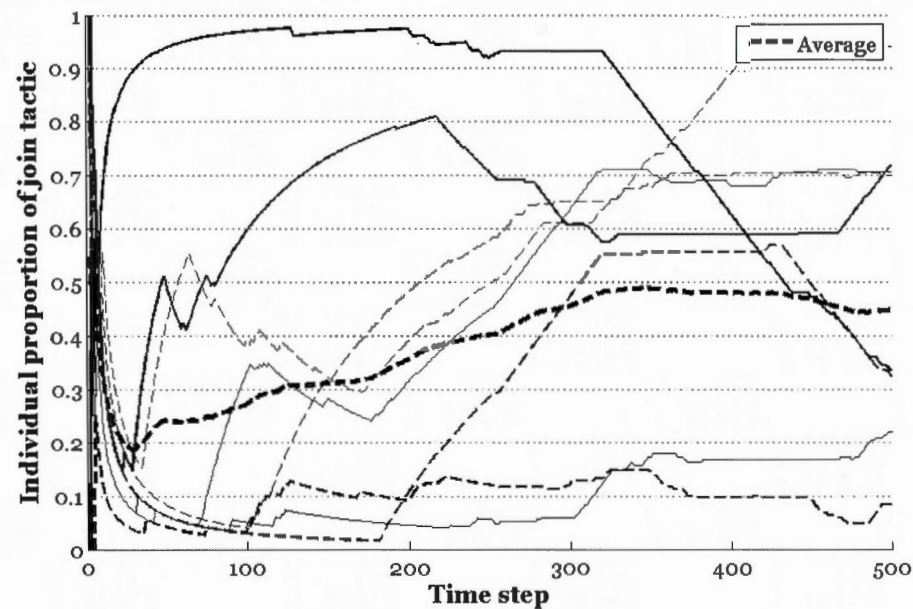
#### 1.4.1 Results

Our results show that an individual’s proportion of join tactic (its strategy) changes over time and so responds to its experience of the environment (Figure 1.1). Although each individual changes its use of joining tactic, it seems to respond idiosyncratically to its own experience. If one considers only the last half of the simulation, when the use of tactics by the group is most likely to have reached an equilibrium, we can observe that an individual’s final strategy is often not pure; individuals obtain their food from a combination of both finding and joining. This means that the expected

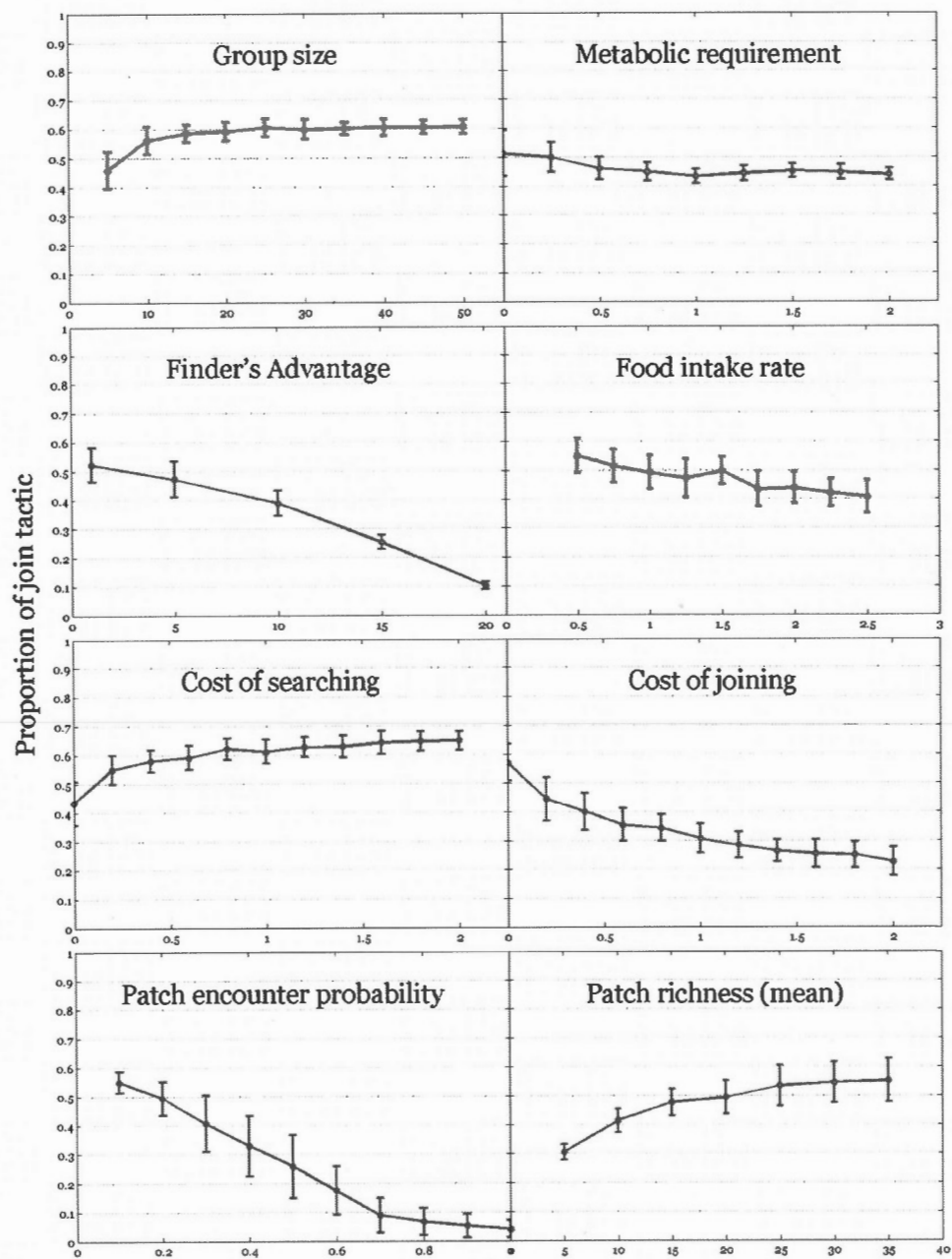
equilibrium state of the population is highly polymorphic, with each individual adopting a distinct combination of finding and joining. Furthermore, the result shows that the average of the population does converge towards an unchanging value, consistent with an equilibrium. However, individuals that try to maximize their own payoff by choosing the better tactic based on their experience do not converge to this mean and continue to change their strategy over time. Instead, the population is strongly polymorphic in its use of the alternative foraging tactics.

The simulation predicts that the equilibrium level of joining changes in response to each of the eight parameters we explored (Figure 1.2). It is noteworthy that the weakest effects are expected for group size, metabolic requirement and food intake rate. The parameters that exert the strongest effects are patch encounter probability, patch richness and the finder's advantage.

These simulation results match the predictions of published game-theoretic models as well as all but two published empirical results (Table 1.2). One of the mismatches concerns the effect of energy reserves or metabolic requirement. In this case, the empirical results reported by Lendvai et al. (2004) are opposite to those expected by our simulation. The other mismatch concerns the observation by Koops and Giraldeau (1996) that joining declines with declining likelihood of patch encounter, a result that also contradicts the results of Beauchamp and Giraldeau's (1996) simulation model.



**Figure 1.1** Evidence of temporal adjustment of the producer-scrouter strategy for a sample group of eight individuals engaged in choosing tactics using a linear operator learning rule. Each line shows the proportion of join tactic choices for one individual. The thick dashed line shows the average for the group. This shows, that by the end of the simulation, individuals do not necessarily adopt a pure strategy (all find, or all search), which means that, in terms of expressed tactics, the population ends in a strong polymorphism of tactics.



**Figure 1.2** Proportion of join tactic at equilibrium as a function of changes in eight ecological parameters (group size, metabolic requirement, finder's advantage,

food intake rate, cost of searching, cost of joining, patch encounter probability and patch richness) for group of eight individuals while foraging in a producer–scrounger game and choosing tactics based on a linear operator learning rule. Means  $\pm$  SD are shown (N=30). The general parameter settings for all simulations are shown in Table 1.1.

**Table 1.2 Comparison of the predictions made by the Linear Operator agent-based simulation model with predictions of previous models and results obtained from experiments. Only studies that provided predictions or results concerning the parameters that we investigated were considered.**

Parameter Reference (M: model, E: experiment): Effect	Predicted qualitatively by simulation
<b>Group size</b>	
Caraco & Giraldeau (1991; M): Larger group size, more scroungers	✓
Vickery et al. (1991; M): Larger group size, more scroungers	✓
Beauchamp & Giraldeau (1996; M): Larger group size, more scroungers	✓
Coolen (2002; E): Larger group size, more scroungers	✓
<b>Metabolic requirement</b>	
Koops & Giraldeau (1996; E): Higher energy requirement (lower energy reserve), no significant effect on scrounging	✓
Lendvai et al. (2004; E): Higher energy requirement (lower energy reserve), more scroungers	✗
Wu & Giraldeau (2005; E): Higher energy requirement (lower energy reserve), small non-significant effect of fewer scrounging	✓
<b>Finder's advantage</b>	
Caraco & Giraldeau (1991; M): Larger finder's share, fewer scroungers	✓
Vickery et al. (1991; M): Larger finder's share, fewer scroungers	✓
Giraldeau & Livoireil (2000; E): Larger finder's share, fewer scroungers	✓
Hamilton (2002; M): Larger ownership advantage, fewer scroungers	✓
<b>Food intake rate</b>	
Caraco & Giraldeau (1991; M): Higher food intake, less scroungers (if probability of finding food by producing is high. Otherwise no change in scroungers)	✓
<b>Cost of searching (cost of finding food by producing)</b>	
Giraldeau et al. (1994; E): Higher cost of producing, more scroungers	✓
Kameda & Nakanishi (2002; M+E): Lower cost of searching, fewer scroungers	✓
Morand-Ferron et al. (2007; E): Higher cost of producing, more scroungers	✓
<b>Cost of joining (cost of finding food by scrounging)</b>	
Caraco & Giraldeau (1991; M): Higher cost of scrounging, fewer scroungers	✓
Morand-Ferron et al. (2007; E): Higher cost of scrounging, fewer scroungers	✓
<b>Patch encounter probability (probability of finding food by producing)</b>	
Koops & Giraldeau (1996; E): Lower probability of finding food by producing,	✗

fewer scroungers	
Beauchamp & Giraldeau (1996; M): Higher probability of patch encounter, fewer scroungers	✓
Hamilton (2002; M): Higher rate of finding food, fewer scroungers	✓
Barrette & Giraldeau (2006, E): Cryptic food (more difficult to find), more scroungers	✓
Beauchamp (2008, M): Lower probability of finding food by producing, more scroungers	✓
<b>Patch richness (average number of seeds in each patch)</b>	
Vickery et al. (1991; M): Larger patch richness, more scroungers	✓
Beauchamp & Giraldeau (1996; M): Larger patch richness, more scroungers	✓
Beauchamp (2008, M): Larger patch richness, more scroungers	✓

### 1.5 The effect of variable patch quality on joining

None of the models that have been generated so far consider the potential effect of variance in patch quality on the use of PS alternatives. All published accounts of PS games, whether simulations or experimental tests, have involved collections of identical food patches. This is surprising given that, in the field, foragers probably encounter foraging patches of various richness. Our simulation model however allows us to address whether adding variance in patch richness has repercussions on the equilibrium proportion of joining.

To analyse the effect of patch richness variance on the expected frequency of the join tactic, we extended the LO model by adding variance to the patch richness. This allowed the patches to contain different quantities of food. We used the same parameter settings as the value in Table 1.1 with the exception of patch richness mean ( $\mu_F$ ) and patch richness standard deviation ( $\sigma_F$ ). We studied the effect of  $\sigma_F$  for three habitats with different mean patch richness,  $\mu_F = 5, 10$  and  $15$ . For each  $\mu_F$ , we varied patch quality variance by altering the standard deviations of patch qualities and changing  $\sigma_F$  from  $0$  to  $\mu_F$ . A set of 30 simulations was conducted for each combination of  $\sigma_F$  and  $\mu_F$ , and the results averaged over the 30 simulations.

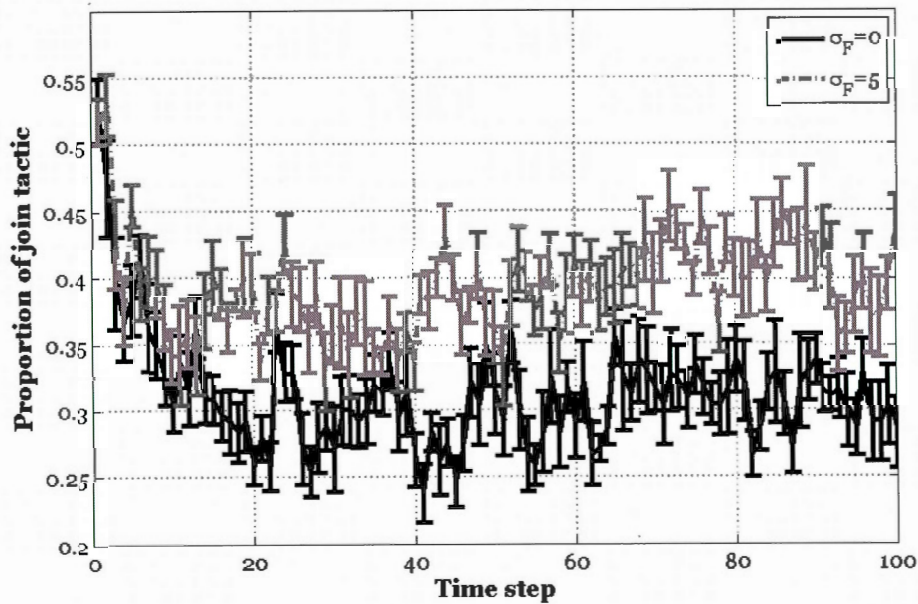
#### 1.5.1 Results

The simulated individuals rapidly changed their use of the join tactic at the beginning of each simulation because of the learning mechanism and the random initial condition (Figure 1.3). The level of joining then gradually converged to a stable point. To determine the predicted combination of tactics, we used only the second half of the simulation, when the values had stabilized.

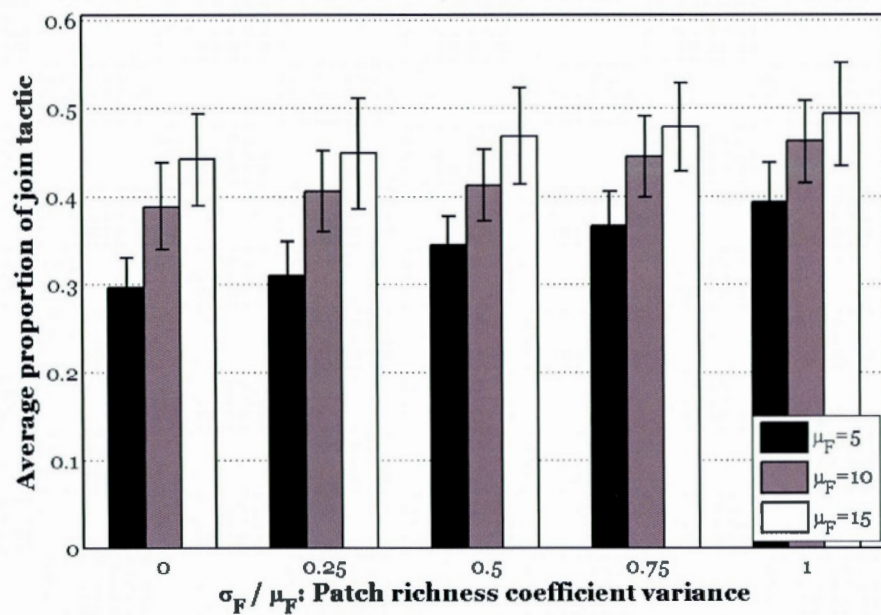
Similar to previous models (Beauchamp, 2008; Beauchamp & Giraldeau, 1996; Vickery et al., 1991), the results of our simulation (Figure 1.4) showed that a higher mean patch richness ( $\mu_F$ ) predicts a higher proportion of joining. However, our model

goes further because for each  $\mu_F$ , increasing the variance of patch richness  $\sigma_F$  also increases the proportion of joining within the population. This is an entirely new prediction that no earlier model or empirical study has yet explored. Moreover, as  $\sigma_F/\mu_F$  increases, the average proportion of the join tactic for different  $\mu_F$  values converge. This means that the effect of patch richness variance should decrease as the mean patch richness of a habitat increases.

The proportion of joining was higher when the variance of patch richness was higher (Figure 1.3). Because the initial condition ( $Vi(0)$ ) was set randomly, both conditions started at a 50% joining proportion (Figure 1.3). The difference in the proportion of joining between conditions was therefore small initially and then increased with more experience until it stabilized and remained so until the end of the simulations (time step = 500). Both conditions caused individuals to decrease their use of joining below the initial 50%, but the decline was more pronounced when the variance of patch richness was lower (Figure 1.3).



**Figure 1.3** Average temporal change in the proportion of join tactic in groups of eight individuals using the linear operator learning rule in a producer–scrounger game when the average patch richness ( $\mu_F$ ) = 5 and patch richness variance ( $\sigma_F$ ) = 0 or 5. Means  $\pm$  SD are shown ( $N = 30$ ). To keep the graph readable, only the first 100 time steps (out of 500) are shown.



**Figure 1.4** Effect of patch richness (mean and variance) on the proportion of join tactic. Means  $\pm$  SD are shown ( $N = 30$ ). To compare the results for different  $\mu_F$ , we show the results based on the patch richness coefficient variance ( $\sigma_F / \mu_F$ ).

## 1.6 Discussion

The agent-based simulation we provide here represents a realistic and effective tool to predict producer–scrounger behaviour under a number of different foraging situations. The model is trustworthy to the extent that it has succeeded in duplicating, at least qualitatively, almost every prediction made by different published producer–scrounger models. It is worth noting that no single game-theoretic PS model published to date is capable of making all of our model's predictions. So, this simulation model predicts PS behaviour under a greater range of situations than any single game-theoretic PS model.

We noted two instances where predictions of our simulation model did not match published observations. One of these refers to the effect of energy requirement on the observed level of joining. The effect of metabolic requirement on a PS game has been studied in laboratory flocks of European starlings, *Sturnus vulgaris* (Koops & Giraldeau, 1996), nutmeg mannikins (Wu & Giraldeau, 2005) and house sparrows (Lendvai et al., 2004). Lendvai et al. (2004) reported that lower energy reserves increased the use of joining during the first feed of the day. Koops and Giraldeau (1996) and Wu and Giraldeau (2005) however found small (nonsignificant) effects on the proportion of join tactic, which were similar to the nonsignificant effects (SD intervals overlaps for all values of metabolic requirement in Figure 1.2) that we observed from the LO model. The significant effect observed in Lendvai et al. (2004) might be because it was limited only to the first feed of the day and the birds were not allowed to adapt to food deprivation.

The other instance where our simulation model diverged from observation concerned Koops and Giraldeau's (1996) report that the expected level of joining declines when the probability of finding food declines. Our simulation instead predicted (Figure 1.2) that the proportion of join tactic should increase when the probability of finding food declines. This prediction is compatible with the predictions of Beauchamp and

Giraldeau (1996), Hamilton (2002), Beauchamp (2008), and an empirical study with nutmeg mannikins reported by Barrette and Giraldeau (2006). Koops & Giraldeau argued that, based on variance-sensitive game-theoretic models (Caraco & Giraldeau, 1991), when the probability of encountering a patch by search increases, the intake for all foragers will increase in a given foraging period. Consequently, individuals that are variance averse should reduce their use of the variance-prone search tactic and increase their use of variance-averse join tactic. This interpretation, however, depends on the costs of searching and joining. If the cost of joining is high relative to the cost of searching, say because animals must appropriate the resource aggressively, then the search tactic is less variance prone when the probability of patch encounter is higher. An extreme example of this would be if patch encounter probability was 1, such that individuals find instantaneously upon choosing to search, making joining a totally unnecessary and hence irrational option, especially if joining is more costly than searching. Defendable, indivisible and very poor food patches are all examples where increasing patch encounter probability increases the payoff for the search tactic but not for the join tactic.

Unlike game-theoretic models, our simulation predicts explicitly that foragers are unlikely to specialize in only searching or joining and hence that populations should appear polymorphic in terms of tactic use. Most published accounts of individual use of tactics are compatible with our model's expectation (Barnard & Sibly, 1981; Beauchamp, 2001; Bicca-Marques & Garber, 2005; Coolen, 2002; Coolen et al., 2001; Giraldeau, Hogan, & Clinchy, 1990; Ha & Ha, 2003; Ilan, Katsnelson, Motro, Feldman, & Lotem, 2013; Koops & Giraldeau, 1996; Mathot & Giraldeau, 2008; Mónus & Barta, 2008; Morand-Ferron, Varennes, et al., 2011; Morand-Ferron, Wu, et al., 2011). In a series of experiments with house sparrows, Barnard and Sibly (1981) reported that producers obtained 19.8–38.3% of their food by joining and that scroungers obtained 38.5–48.8% of their food by searching. As Barnard and Sibly (1981) emphasized, the flexible use of tactics allows individuals to respond quickly

and adaptively to changes in payoffs. The only published instance of pure tactic use was from a study that placed animals in situations that constrained them to specialize (Mottley & Giraldeau, 2000).

In the context of the producer–scrounger game, producing is a more uncertain foraging strategy than scrounging (Wu & Giraldeau, 2005). Producers rely on their own effort and the patch richness variance directly affects their performance (Rita & Ranta, 1998). However, as has been shown in Koops and Giraldeau (1996) and Lendvai et al. (2004), scrounging is a variance-averse strategy. By exploiting the effort of multiple individuals, scrounging is more resistant to changes in patch richness variance and, therefore, is only affected by the average patch richness (Flynn & Giraldeau, 2001). The stable solution in a more uncertain environment is therefore to scrounge more (Lendvai et al., 2004; Rita & Ranta, 1998). However, this is true only if the average patch richness is large enough to provide the minimum amount of energy required by each individual. Based on the shortfall minimizing model of variance-sensitive foraging theory, if average food intake is less than the minimum survival requirement, individuals might favour the variance-prone strategy since it offers at least a chance of survival (McNamara & Houston, 1992; Stephens & Krebs, 1986).

### 1.7 Future directions

An important limitation of analytical evolutionary approaches (including ESS theory) is the excessive simplification of individuals and the environment in order to determine an expected stable outcome (Fawcett et al., 2014). This is not as true for agent-based simulations, making it a particularly powerful means of exploring the expected outcomes for more complex and hence realistic conditions. Our results demonstrate that the simulation approach is entirely capable of replicating results obtained from highly simplified conditions. This is reassuring and suggests that it can provide a valuable basis for generating predictions in more complex situations. We

used the model to predict foraging behaviour when patches were of variable quality. Here, we explore some of the more likely situations where the simulation should be used to generate testable predictions for future enquiry.

Based on previous studies (Belmaker et al., 2012; Morand-Ferron & Giraldeau, 2010; Stephens, 1991), we assumed that learning can evolve and fix in the population. For simplicity, we used the model to predict foraging behaviour when patches were of variable richness but the environment was not changing and individuals were identical. Katsnelson et al. (2012), however, showed that learning is more likely to evolve when there is asymmetry in phenotype and environmental fluctuation. It will thus be important to verify the model and reanalyse the effect of these parameters on PS behaviour when phenotypic asymmetry or environmental fluctuations exist. Real-life individuals are rarely identical. For instance, Mathot et al. (2009) reported that individuals differ in their basal metabolic rates (BMR) and that their preferences for the search or join tactic depends on their BMR; high BMR individuals join more. Several other studies also investigated the effect of individual differences in PS games (Barta & Giraldeau, 1998; Hamilton, 2002; Jolles, Ostojić, & Clayton, 2013; Marchetti & Drent, 2000; Mathot & Giraldeau, 2010; McCormack, Jablonski, & Brown, 2007; Morand-Ferron, Varennes, et al., 2011; Morand-Ferron, Wu, et al., 2011; Rita, Ranta, & Peuhkuri, 1997). For instance, individuals within groups occupy different spatial positions, and simulation and experimental studies have demonstrated that spatial position affects an individual's preferred tactic (Barta, Flynn, & Giraldeau, 1997; Flynn & Giraldeau, 2001; King, Isaac, & Cowlshaw, 2009; Mõnus & Barta, 2008). A number of recent studies have shown considerable individual differences in the speed of learning or in the propensity to learn (Katsnelson et al., 2012), but we have little idea concerning how variation in learning affects tactic use and the extent of tactic polymorphism within a group. Furthermore, individuals show differences in their impulsiveness or exploratory tendencies as well as a whole slew of other behaviours that have been considered under the theme of

animal personalities (Réale & Dingemanse, 2012). What consequences might individual differences have on PS outcomes (Dubois, Giraldeau, & Réale, 2012)? These are all examples where the agent-based linear operator model, unlike evolutionary game-theoretic models, can be used to explore the consequences for social foraging. In addition to explore the effect of individual differences, we suggest that the agent-based simulation could be used to explore the effects of more complex environments. Examples include, but are not limited to, predation risk, patch defense and group size. One could also address the effect of PS behaviour on expected group size.

The simulation predicted that the equilibrium level of joining varied in response to each of the eight parameters we explored (Figure 1.2). For example, group size, metabolic requirement and food intake rate had the weakest effects, while patch encounter probability, patch richness and the finder's advantage exerted the strongest effects. An additional step, not yet taken, would be a detailed experimental examination of such differences. Knowing why some parameters have stronger effects may lead to a better understanding of animal behaviour and decision making in frequency-dependent tasks. We cannot however relate the slope of increase from our simulation results to real life experiments as we cannot relate the time steps in the simulation to seconds, minutes or trials in the real world. We thus need to verify and measure the effects experimentally. Then, we can compare the effects of different parameters and study why some effects are stronger than others.

In conclusion, we argue that learning rule models like the one we present here are more effective tools, in term of generality, to study and predict behaviour under frequency-dependent conditions. We model the decision rule as a mechanism to determine foraging behaviour directly and explicitly as suggested by Arbilly et al. (2010) and Grodzinski et al. (2011). The approach allows us to explore behaviour without indulging in a behavioural gambit, but the correspondence between

predictions from our model and those of game-theoretic models provide some reassurance that the gambit may be acceptable. Ultimately, however, the validity of the gambit needs to be established experimentally. Our laboratory is already engaged in exploring some of these effects, particularly the effect of patch variance on the use of the joining tactic. We encourage other researchers to use our approach to extend PS games to a greater range of diverse, complex and hence more realistic foraging situations.

### 1.8 Acknowledgments

This research, as well as M.A., was supported by a Natural Sciences and Engineering Research Council of Canada discovery grant to L.-A.G (grant number 303112). In addition, M.A. was supported through a Fund Accessibility and Academic Success (FARE) Scholarship from the University of Quebec in Montreal and a foreign fee remission. We thank Pedro Peres-Neto for access to computational resources, Carolyn Hall and Denis Réale for extensive help in statistical analysis and comments, and William L. Vickery and Frédérique Dubois for commenting on earlier drafts of the manuscript, and two anonymous referees for their constructive comments.

[Cette page a été laissée intentionnellement blanche]

## CHAPTER II

### ZEBRA FINCHES SCROUNGE MORE WHEN PATCHES VARY IN QUALITY: EXPERIMENTAL SUPPORT OF THE LINEAR OPERATOR LEARNING RULE

Afshar, M., Carolyn L. Hall, & Giraldeau, L.-A. (2014). Zebra finches scrounge more when patches vary in quality: Experimental support of linear operator learning rule. *Animal Behaviour*, 105, 2015, 181-186. doi: 10.1016/j.anbehav.2015.04.016

#### 2.1 Abstract

The use of evolutionary game theory and the concept of the evolutionarily stable strategy (ESS) have come under criticism lately because game-theoretic models are often constrained to overly simplistic situations. Furthermore, game-theoretic models commit the behavioural gambit, that is, they assume that individuals have some unspecified decision mechanism that duplicates the outcomes that would be expected from natural selection acting on genetic alternatives. The producer–scrounger game is an ideal illustration of the criticisms aimed at the game-theoretic approach; it has generated a number of game-theoretic models that make specific predictions for highly simplified foraging conditions. Simulation models of the producer–scrounger game that incorporate the linear operator learning rule, however, have been more successful in replicating the empirical results. In these models, groups of animals forage for patches of food using a linear operator learning rule. The linear operator model furthermore provides a new prediction about the effect of variation in patch quality within a producer–scrounger game. Current models based on evolutionarily stable strategies and empirical investigations have always assumed or used patches of

uniform quality. Yet patches vary in quality and some environments are potentially more variable than others. The linear operator model predicts an increase in scrounging behaviour when patch quality varies. We tested this prediction using captive foraging flocks of male zebra finches, *Taeniopygia guttata*, in two treatments with variable and nonvariable patch quality. The results show that as individuals have more experience in the variable environment, they increase their scrounging behaviour. Linear operator models may therefore provide a general and empirically valid means of exploring the outcome of animal decisions in social foraging situations too complex to be addressed by evolutionary game models.

**Keywords:** decision making, evolutionary game theory, frequency-dependent selection, learning rule, linear operator, producer–scrounger, risk sensitivity, social foraging, zebra finches (*Taeniopygia guttata*)

## 2.2 Introduction

Situations where the outcome of an individual's behaviour depends on the behaviour used by others commonly call for evolutionary game theory as an analytical tool (Giraldeau & Caraco, 2000; Maynard Smith, 1982). The use of evolutionary game theory and the concept of the evolutionarily stable strategy (ESS), however, have come under criticism because game-theoretic models are often constrained to overly simplistic situations that bear little resemblance to actual conditions under which the behavioural decision has evolved (Afshar & Giraldeau, 2014; Fawcett et al., 2014; Grodzinski et al., 2011). As a result, tests of these models' predictions are often conducted under equally unrealistic and oversimplified situations.

Another critique of evolutionary game theory, and of optimality in general, has been that it assumes the operation within individuals of some unspecified decision mechanism that duplicates the outcomes that would be expected from natural selection acting on genetic alternatives. It therefore commits the behavioural gambit by assuming that psychological mechanisms do not constrain the behavioural

phenotypes that can evolve (Fawcett et al., 2013; Giraldeau & Dubois, 2008; McNamara & Houston, 2009). Several authors have thus argued that it would be more useful to explore the outcomes of decision mechanisms in game situations directly, rather than assume their operation and outcome (Arbilly et al., 2011; Fawcett et al., 2013; Katsnelson et al., 2012).

In a recent study, Afshar and Giraldeau (2014) developed a simulation model that explored the outcomes of groups of individuals using an explicit learning rule in a producer–scrounger (PS) game (Barnard & Sibly, 1981). In a PS game, individuals can obtain food by searching for a new food patch (producing) or by joining patches already found by others (scrounging). The game assumes that these two foraging alternatives are temporally incompatible, as is the case for animals like ground-feeding birds (Coolen et al., 2001). When the scrounger strategy is rare, it receives a higher payoff than the producer tactic. Increasing the frequency of scrounging, however, decreases everyone's payoffs. The payoff for both alternatives is therefore affected by the frequency of the scrounger strategy. ESS game-theoretic models predict a specific evolutionarily stable frequency of scrounging where the payoffs of both alternatives are equal (Giraldeau & Beauchamp, 1999). The principal question, therefore, in PS research has been to explore how environmental and individual state parameters affect this equilibrium. The PS game is an ideal illustration of the criticisms aimed at the game-theoretic approach; it has generated a number of game-theoretic models that make specific predictions for highly simplified foraging conditions (Afshar & Giraldeau, 2014). In Afshar and Giraldeau (2014), the individual chooses which tactic to express based on the matching ratio, adopting the tactics in proportion to their estimated values, at each time step. This agent-based simulation was remarkably successful in replicating almost every published prediction and experimental result concerning PS games (Afshar & Giraldeau, 2014). Given the success of their simulation, they argued that the explicit decision model was a more effective means of predicting the effect of environmental factors on PS

behaviour than ESS models. They backed this claim by generating an entirely novel prediction for a situation that was too complex for the outcome to be anticipated using a conventional ESS approach. Every previous ESS model and test of PS behaviour had considered only environments of homogeneous patch quality; ignoring the effect that a mixture of patch qualities would have on PS behaviour. The Afshar and Giraldeau (2014) simulation model showed, however, that the variance in patch quality should affect the expected level of scrounging, and it predicted that increasing variance should lead to increased scrounging. Furthermore, the simulation model showed that, because individuals use learning and hence their experience to adjust their behaviour, any changes in the amount of scrounging is expected to happen gradually as individuals gain experience and the population slowly converges to a stable level of scrounging after sufficient time.

In this study we provide the first experimental test of the new expected level of scrounging to test whether this linear operator learning model represents an empirically valid way to predict the outcomes of individual decisions in complex frequency-dependent foraging conditions. We test this novel prediction using captive flocks of zebra finches, *Taeniopygia guttata*. Specifically, we ask whether increasing the degree of variance in patch quality, while keeping its mean constant, will affect the frequency of scrounging in flocks of foraging captive zebra finches.

## 2.3 Methods

### 2.3.1 Study subjects

We used zebra finches purchased from a local supplier, L'oisellerie de l'Estrée, QC, Canada. These social granivorous birds have been widely used to test predictions of PS games (Beauchamp, 2000b, 2001, 2006; Biondolillo, Stamp, Woods, & Smith, 1997; David, Cézilly, & Giraldeau, 2011; David & Giraldeau, 2012; Giraldeau et al., 1990; Mathot & Giraldeau, 2008, 2010; Mathot et al., 2009). We randomly selected 20 males from our colony of 32 males.

Before the experiments, the birds were kept in groups of three to four individuals in small home cages ( $29 \times 52 \times 38$  cm) with ad libitum access to water, a mixture of vitamin-supplemented seeds, and regular access to vegetables and egg mixture. Room temperature was maintained at  $24 \pm 1$  °C on a 12:12 h light:dark cycle. For identification, each individual had a coloured leg ring. All the birds had previous experience with the PS game, but none had been used in an experiment for at least 1 year. We randomly assigned the birds to four flocks of five birds.

During the experiments, each bird was temporarily marked with a dab of nontoxic acrylic paint on top of the head. The colour was randomly chosen from orange, yellow, green, blue and purple with the constraint that no two birds in the same flock had the same colour. The experiment lasted 15 days for each flock, and two flocks were tested concurrently in neighbouring aviaries. The experiments were conducted from 9 July to 4 August 2013 within the animal care facility at the University du Quebec à Montreal. All experimental procedures complied with the guidelines from Canadian Council for Animal Care and were approved under protocol 0313-C1-712-0314 by the UQAM committee for animal care.

During the experiment, each flock was kept in a large indoor aviary ( $1.5 \times 3.8 \times 2.3$  m high) with a 12:12 h light:dark cycle (0700–1900 hours). Birds had ad libitum access to water at all times and access to a bath 1 day per week. A  $1.1 \times 1.15$  m piece of plywood grid in which 100 wells (1.3 cm diameter and 1 cm depth) had been drilled in a grid pattern ( $10 \times 10$  cm with 10 cm central distance) was placed on the floor of each aviary. A video camera hung from the ceiling 1.9 m from floor, directly above the center of the grid so that the image captured the whole grid surface.

### 2.3.2 Experimental Procedure

The flock was allowed two consecutive days within the aviary to familiarize themselves with their surroundings. During this time, two feeders filled with millet seeds were placed on the grid. All food was removed from the aviaries at 1800 hours

on day 2. The next morning five trials were conducted for one flock at 30 min intervals starting at 0800 hours and ending at 1000 hours. Trials for the second flock were started with a 15 min delay at 0815 hours, using the same procedure.

At the beginning of a trial, the experimenter (M.A.) entered the aviary, turned on the camera, and filled six randomly chosen patches with white millet seeds. To prevent the birds from recognizing or remembering the location of the full patches on the grid, the experimenter filled six experimental patches and mimed the addition of seeds at six empty patches, visiting the 12 patches in random order. He then covered the grid with a sheet of opaque white corrugated plastic (Coroplast) before exiting the aviary. From outside the aviary, the experimenter pulled a string to slide the opaque covering off the grid. The birds then landed and searched for food on the grid. Each trial continued for 9 min, after which the experimenter entered the aviary, cleaned the grid of all food and turned off the camera.

After the last trial of the day the birds were provided with ad libitum access to food until 1800 hours, when all food was removed. This procedure was repeated from day 3 to day 8 and then again from day 10 to day 15. During day 9, the birds had no trials and were provided with ad libitum food, water and bath. Days 3 and 10, the first day of each treatment, were considered training days and no data were collected. At the end of the 15th day we returned the birds to their holding colony.

Each flock was tested in two treatments, variable and nonvariable patch richness. In the nonvariable treatment, all six patches in each trial contained exactly 10 seeds. In the variable treatment, there were two patches for each of three richness values: 5, 10 and 15 seeds. This changed the patch richness variance from zero (in the nonvariable treatment) to 20 (standard deviation of 4.47). Two flocks were randomly selected to experience the nonvariable treatment first and then the variable; the other two flocks experienced the reverse order.

In each trial, one bird was randomly selected as the focal bird, with the constraint that each bird had to be observed once per day (excluding the training day) and once in each time slot over the course of the week. After the experiment was completed for all flocks, the videos were viewed in random order by the experimenter, who was blind to the treatment type.

Behavioural observations started when the first bird landed on the grid and continued for 2 min. We counted find and join events. We defined a 'find event' when the focal bird pecked seeds from a patch that was unoccupied at the time the bird arrived at the patch. A 'join event' was noted when the focal bird pecked seeds from a patch where at least one other bird was feeding. If the focal bird 're-found' an abandoned patch or 're-joined' a previously joined patch, we also counted the new event. We then calculated the proportion of patches joined by the focal bird in each trial as the number of its join events divided by the sum of its join and find events. Similar to previous studies (Beauchamp, Giraldeau, & Ennis, 1997; Coolen, 2002; Mathot et al., 2009; Morand-Ferron & Giraldeau, 2010; Morand-Ferron, Varennes, & Giraldeau, 2011; Wu & Giraldeau, 2005), we used the proportion of patches found by joining as an index of the level of scrounging for the focal individual. We used the proportion of patches found by joining rather than the proportion of seeds found by joining because it was difficult to obtain a reliable count of the number of seeds from the videos, and because Barrette and Giraldeau (2006) have shown that the two give similar results. All video analysis was conducted using JWatcher (Blumstein & Daniel, 2007). The data were then transferred to R (R Development Core Team, 2008) for statistical analysis (Field, Miles, & Field, 2012; Zuur, Ieno, Walker, Saveliev, & Smith, 2009).

### 2.3.3 Statistical Analysis

As the dependent variable was the proportion of patches discovered by joining, we used a generalized linear mixed-effect model with binomial distribution (Bolker et al.,

2009; McDonald, 2009; Nakagawa & Schielzeth, 2010). For the model construction and selection, we followed the protocol in Zuur et al. (2009). First, we started with a model that contained all of our fixed effects: patch richness variance (nonvariable = 0, variable = 1), day of the week, and their interaction as the initial fixed effects. Rather than testing the effect of patch richness variance using an intercept of a nonexistent 'day 0', we were interested in testing the effect on day 5, once the birds had time to learn about the treatment. We thus transformed the day of the week to vary from -4 to 0, rather than from 1 to 5, so that the intercept of the model (transformed value of 0) was on the fifth day of the experiment. We then tested the random effects by comparing different random-effect structures, keeping a stable fixed-effect structure. We included week (order of the nonvariable and variable treatment), flock identity and bird identity as possible random effects. As the fixed-effects structure did not change at this step, we selected the model with the highest conditional variance ( $R^2_{\text{GLMM(c)}}$ ) (Bolker et al., 2009; Nakagawa & Schielzeth, 2013; Schielzeth & Nakagawa, 2013). We then kept the selected random structure, and tested for the optimal fixed-effect structure using maximum likelihood estimation by Akaike's information criterion (AIC).

## 2.4 Results

One subject became ill during day 3 of the second series and was removed. To keep flock size constant we replaced it with a bird from the flocks that were used in the first series the previous week. This bird was not used as a focal bird, though, and its behaviour was not recorded for the experiment this time. We thus ended up with 19 data series from 19 focal birds. Each of the 19 birds was used as a focal bird in 10 trials. In only one trial (on day 1), the focal bird did not land on the grid during the 2 min in the nonvariable treatment. We thus had a total of 189 trials with valid data.

Of the random effects, bird identity explained by far the largest amount of variance (Table 2.1) and was the only random effect retained in the model. The best model to

explain the proportion of patches found by joining included treatment, day and their interaction. The proportion of patches found through joining was significantly higher in the variable treatment than in the nonvariable treatment (Table 2.2). The proportion of patches found through joining decreased slightly but nonsignificantly overall with day (Table 2.2). There was also a significant interaction between patch richness variance and day (Table 2.2), such that the proportion of patches that the birds found by joining increased over time in the variable treatment and decreased over time in the nonvariable treatment (Figure 2.1). Individuals were consistent in their strategy use across both treatments; those that scrounged more in the variable patch treatment also scrounged more in the nonvariable patch (Pearson correlation coefficient:  $r = 0.83$ ,  $P < 0.001$ ; Figure 2.2).

**Table 2.1** Variance explained by different random-effect structures of the GLMM on strategy use, using the proportion of patches found by joining as the dependent variable

Random effect(s) in the model	$R^2_{\text{GLMM(m)}}$	$R^2_{\text{GLMM(c)}}$	AIC
Week <sup>a</sup>	1.31%	1.31%	714.2
Flock identity	1.40%	2.50%	711.9
Bird identity	2.14%	42.25%	509.2
Week* flock identity* bird identity	2.05%	42.91%	512.1

Different random-effect structures were compared using the conditional variance ( $R^2_{\text{GLMM(c)}}$ ). Marginal variance explained ( $R^2_{\text{GLMM(m)}}$ ) and Akaike's information criterion (AIC) are also shown. Only bird identity was retained as a random effect in the final model.

<sup>a</sup> Order of the nonvariable and variable treatment.

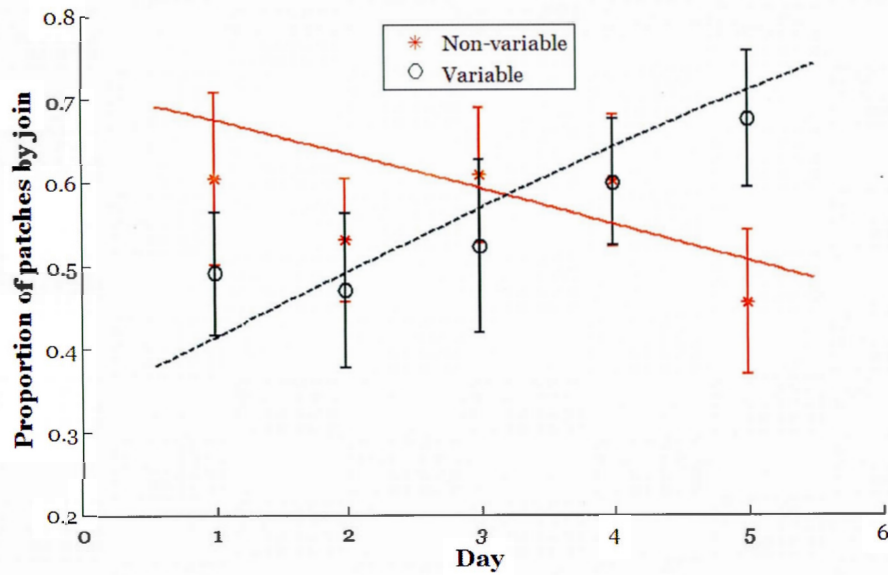
**Table 2.2** Fixed effects of the GLMM on strategy use, using the proportion of patches found by joining as the dependent variable

	Estimate	SE	Z score	P
(Intercept)	0.017	0.429	0.039	0.969
<b>Treatment<sup>a</sup></b>	0.831	0.309	2.688	<b>0.007</b>
<b>Day<sup>b</sup></b>	-0.171	0.089	-1.932	0.054
<b>Treatment*day</b>	0.465	0.127	3.647	<b>&lt;0.001</b>

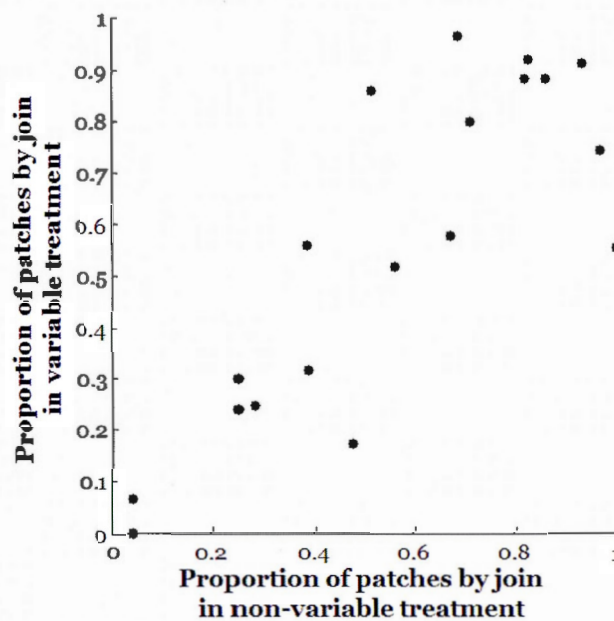
Significant effects are shown in bold.

<sup>a</sup> Nonvariable patch richness = 0; variable patch richness = 1.

<sup>b</sup> Days 1 to 5 were transformed to -4 to 0 (see text for explanation).



**Figure 2.1** Mean  $\pm$  SE proportion of patches found by joining in four flocks of five zebra finches, *Taeniopygia guttata*, foraging in nonvariable and variable patch richness treatments on day 1 ( $N = 19$ ) and days 2–5 ( $N = 20$ ). The curves show the predicted proportion of patches by join for each treatment from the GLMM model shown in Table 2.1 and 2.2



**Figure 2.2** Strategy consistency across the two treatments (nonvariable and variable patch richness). A scatter plot of the average proportion of patches found by joining in the variable treatment and the nonvariable treatment for individuals in four flocks of five zebra finches, *Taeniopygia guttata*.

## 2.5 Discussion

Our experimental results show that the linear operator learning model (Afshar & Giraldeau, 2014) represents an empirically valid means to predict the outcomes of individual decisions in complex social foraging situations with frequency dependence that cannot be analysed using conventional evolutionary game theory. We conclude this because our results confirm a totally novel prediction of the simulation model to the effect that scrounging increases with the variance in patch quality. This prediction has not previously been made by any of the game-theoretic models of social foraging. We discuss the implications for both social foraging and the prediction of behaviour in complex social conditions.

The significant effect of patch richness variance and its significant interaction with day (Table 2.2), supports the linear operator model. As predicted, the birds gradually increased their level of scrounging over time in the variable treatment and decreased it over time in the nonvariable treatment (Figure 2.1). As a consequence, on the last day of each treatment, the birds scrounged significantly more in the variable treatment than they did in the nonvariable treatment. In fact, this gradual change over time is exactly what we would expect if the birds were using a learning rule to adjust their scrounging behaviour, and is similar to what we see in the simulation results of Afshar and Giraldeau (2014). Consequently, had the birds been given more time, the difference between the proportion of patches scrounged in the variable and nonvariable patch treatments would likely have become even more prominent.

This increase in scrounger use with increasing variance in patch quality may be consistent with risk-sensitive foraging behaviour (Giraldeau & Caraco, 2000). Wu and Giraldeau (2005) argued that producing was a more uncertain foraging strategy than scrounging because producers rely on their own effort and have a more variable payoff than scroungers (Rita & Ranta, 1998). Scrounging, therefore, is thought to be a variance-averse strategy (Koops & Giraldeau, 1996; Lendvai et al., 2004) because

scroungers can exploit the effort of several individuals and are thus less sensitive to changes in patch quality variance than they are to the average patch richness (Flynn & Giraldeau, 2001). Scrounging more under increased patch variance may therefore offer a stable solution to a less certain environment (Afshar & Giraldeau, 2014; Lendvai et al., 2004; Rita & Ranta, 1998). This is even more likely when the average patch richness is large enough to consistently provide the minimum amount of energy for survival, as was the case in our experiment.

We found that individuals were consistent in their strategy use across both treatments (Fig. 2). Similar consistent individual differences in strategy use during the PS game have been reported when patch quality is constant within treatments but differs between treatments (Morand-Ferron, Varennes, & Giraldeau, 2011; Morand-Ferron, Wu, & Giraldeau, 2011). This suggests that individuals are not identical in their strategy use. Individual differences, such as differences in the speed of learning, may affect the way that individuals learn. When individuals in a population have different speeds of learning, fast individuals respond more quickly to spatiotemporal fluctuations. For example, when the payoff of a strategy drops temporarily, fast individuals will sample alternative options and stop expressing that strategy more quickly. In a frequency-dependent situation, this quick response can recover the payoff of the strategy. It thus removes the need for other individuals to sample alternative behaviours. In our experiment, for example, when individuals were introduced to an environment that favoured a higher level of scrounging, quick increases in the level of scrounging by some individuals would have reduced the need for others to change their strategies. This might explain why some individuals were so consistent in their strategy use regardless of the treatment, while others were more flexible (Figure 2.2).

Incorporating nonidentical individuals would allow us to explore questions about which type of individuals produce more and why. We thus argue that integrating

individual differences into the model would be an important next step towards increasing the realism of situations for which we need behavioural predictions. Kurvers, Hamblin, and Giraldeau (2012), for example, modelled a PS game where individuals had different levels of exploratory behaviour. They found that environments with a high patch density favoured fast-exploring individuals. Another way to integrate individual differences would be to use our linear operator learning simulation and allow for individuals to be endowed with different learning rates or different levels of prior experience with the environment.

Afshar and Giraldeau (2014) have already demonstrated that the linear operator model regenerates the vast majority of results and predictions from previous producer-scrouter studies and does so in a single united framework. It also made a new prediction that was experimentally supported by the results of the experiment presented here. This prediction would have been rather difficult to make using conventional game-theoretic models and in fact no model had yet addressed the question of the effect of patch variance in a PS game. As Fawcett et al. (2014) argued, studies on decision making and behaviour generally suffer from oversimplification, while trying to study the behaviour in a complex heterogeneous environment. We therefore argue that learning rule models are better tools, in terms of both simplicity of the models and their capability to handle more complex situations, to study and predict animal behaviour in frequency-dependent systems than is the ESS approach.

We acknowledge that an optimal learning rule is both a consequence and an outcome of an evolutionary process and so is not divorced from it. Such a learning rule, be it linear operator or any other, may be what Maynard Smith called evolutionary stable learning rule, a learning rule whose characteristics lead the population within a generation to the same equilibrium combination as would frequency-dependent selection over generations (Maynard Smith, 1982). It is however important to

emphasize the difference between the two approaches, evolutionary game theory versus learning (Afshar & Giraldeau, 2014). In addition to dealing with experimental realism, the learning approach is particularly powerful in predicting the individual behaviour for which evolutionary game models remain entirely silent. More importantly, evolutionary game theory models cannot deal well with multiple ecological parameters that match more realistic environments. Agent-based models using learning rules therefore provide a convenient alternative to study the outcome of individual decisions in complex social situations. Recent growing interest in the behaviour of individuals within behavioural ecology (Réale & Dingemanse, 2012) raises the importance of having more powerful predictive tools capable of generating expectations at the individual level. The linear operator learning simulation approach (Afshar & Giraldeau, 2014) therefore appears to be an empirically valid means to generate the expected outcomes within groups of individuals engaged in complex games with frequency-dependent payoffs.

## 2.6 Acknowledgements

M.A. and this research were supported by a Natural Sciences and Engineering Research Council of Canada discovery grant (number 303112) to L.-A.G. In addition M.A. was financially supported through a Fund Accessibility and Academic Success (FARE) Scholarship and a foreign fee remission from University of Québec in Montréal and a doctoral student's financial aid from Group for Research in Decision Analysis (GERAD). We thank Pedro Peres-Neto for access to computational resources, Denis Réale for extensive help in statistical analysis and comments, Mélanie Dappen-Couture for the help with experiment design and William L. Vickery for commenting on earlier drafts of the manuscript.

[Cette page a été laissée intentionnellement blanche]

## CHAPTER III

### THE EVOLUTION OF LEARNING AND NON-LEARNING RULES IN DYNAMIC ENVIRONMENTS: A TEST OF THE BEHAVIOURAL GAMBIT

Afshar, M., & Giraldeau, L.-A. (2014). The evolution of learning and non-learning rules in dynamic environments: A test of the behavioural gambit. For submission to *Proceedings of the Royal Society B: Biological Sciences*.

#### 3.1 Abstract

Learning, a way to adapt to environmental changes, has been studied for decades by behavioural ecologists using models based on learning rules performing within a social foraging context. Most of these studies, however, had two important shortcomings: learning rules were rarely vetted against non-learning rules and the environmental variance required for learning was rarely present within an individual's lifetime. Furthermore, we still do not understand the conditions necessary for learning to evolve initially. In this study, we explore the conditions under which a linear operator type of learning is expected to evolve in a frequency-dependent context like a producer-scrouter (PS) game, where producer search for food and scrounger obtain food by joining others. In doing so, we address both limitations of earlier studies by allowing fixed strategies to compete against learning strategies in environments that are stable, vary, within generations or vary between generations. We found that all populations with similar group size converged to a similar proportion of join tactic, regardless of whether individuals used a pure strategy, mixed strategy, or linear operator learning strategy. Our results show that the

pure strategy is superior to both the mixed strategy and the linear operator learning strategy across a majority of the environmental conditions. Depending on the type and magnitude of environmental fluctuations, learning may either be dominated by fixed rules or it can coexist with pure strategies. This result suggests that, at least within the context of producer-scrounger game, learning generates a combination of strategy use that corresponds to that expected from selection acting on genetic alternatives and we can therefore accept the behavioural gambit when studying population level equilibrium.

**Keywords:** producer-scrounger, learning rule, decision making, frequency-dependent selection, Evolutionary game theory, social foraging, pure strategy, mixed strategy, polymorphism, linear operator, behavioural gambit

### 3.2 Introduction

Learning can be broadly defined as “a change in state due to experience. [...] The changes in state referred to as learning seem to involve a change in cognitive state, not just behavioral potential” (Shettleworth, 2010, pp. 98-99). The underlying mechanism for learning has been studied for decades by psychologists (Shettleworth, 2010) and more recently by behavioural ecologists (Dugatkin & Reeve, 2000; Dunlap & Stephens, 2009; Harley, 1981; Henly et al., 2008; Kacelnik & Krebs, 1985; McNamara & Houston, 1985; Niv, Joel, Meilijson, & Ruppin, 2002; Stephens, 1991). In behavioural ecology, a popular tradition has been to model the process as a “learning rule”, a mathematical expression that combines previous experiences and then uses this information to assign a probability to each possible behaviour of being expressed at a given time (Aoki & Feldman, 2014; Arbilly et al., 2011; Dridi & Lehmann, 2014; Dubois et al., 2010; Fawcett et al., 2014; Groß et al., 2008; Hamblin & Giraldeau, 2009; Katsnelson et al., 2012; Kolodny et al., 2014; Lotem & Biran-Yoeli, 2014; Rendell et al., 2010; Trimmer et al., 2012). Several studies have observed learning rules performing within a social foraging context, especially for the

producer-scrourer game (Afshar & Giraldeau, 2014; Arbilly et al., 2010, 2011; Beauchamp, 2000a; Dridi & Lehmann, 2014; Hamblin & Giraldeau, 2009; Katsnelson et al., 2012), but most of these studies have been plagued with two important shortcomings: learning rules are rarely vetted against non-learning rules and the environmental variance required for learning to be of any value is rarely present within an individual's lifetime. Furthermore, we do not know what conditions are necessary for learning to evolve in the first place, and accepting that it will always generate a combination of strategies that corresponds to that expected from selection acting on genetic alternatives (what Giraldeau and Dubois 2008 describe as accepting the Behavioural Gambit) may be incorrect (Fawcett et al., 2013; McNamara & Houston, 2009).

Several studies investigated the design of an optimal learning rule (Dridi & Lehmann, 2014; Fawcett et al., 2014; Harley, 1981; Houston & Sumida, 1987; Kacelnik & Krebs, 1985; Kacelnik et al., 1987; Lea & Dow, 1984; March, 1996; Trimmer et al., 2012). Beauchamp (2000) was one of the first studies to compare the optimality of different learning rules by letting them compete against each other in a simulated foraging game. He compared three rules: linear operator, perfect memory, and relative payoff sum within the context of two social foraging games: producer-scrourer and ideal free distribution. The linear operator rule yielded consistently higher payoffs than other learning rules across various contexts, while the relative payoff sum rule performed poorly, especially in the ideal free distribution. Beauchamp (2000), however, selected the parameters of the learning rules arbitrarily. We therefore cannot be sure whether the success of the linear operator over the others was due to the parameters chosen or to the form of the rule itself. Furthermore, the study never pitted learning rules against fixed players and so it is not clear whether learning was actually the best solution in these games.

Hamblin and Giraldeau (2009) addressed the issue of arbitrary parameter choice by using genetic algorithm simulations to first find the best parameter combinations for each rule in a producer-scrounger game. Once the optimal parameter values for each rule were established, each rule played against the others in a round-robin tournament seeking the rule that would evolve to fixation. The relative payoff sum performed better than either the linear operator or perfect memory rules, but as in Beauchamp (2000) these rules were never tested against fixed players.

By omitting non-learning fixed players, both the studies of Beauchamp (2000) and Hamblin and Giraldeau (2009) implicitly assumed that learning was superior to fixed behaviour in social foraging games. This is surprising given that several studies have shown that learning only outperforms fixed individuals when there are environmental fluctuations (Dubois et al., 2010; Dunlap & Stephens, 2009; Groß et al., 2008; Katsnelson et al., 2012; Stephens, 1991). If the environment is stable, as was the case in Hamblin and Giraldeau (2009), the fixed individuals may very well have outperformed the learning individuals. In fact, the optimal parameters of the best relative payoff sum in the Hamblin and Giraldeau (2009) study paradoxically generated almost fixed behaviour in its players.

Some studies, however, have investigated the value of learning over fixed non-learning strategies. For instance, Dubois, Morand-Ferron, and Giraldeau (2010) found that flexible learning individuals can invade population of inflexible individuals in a producer-scrounger game, but learning individuals can never evolve to fixation. Compared to fixed pure producers or pure scroungers, learners can outperform one but not both fixed strategies. However, the study simply assumes that learners are omniscient and know at no cost the payoff of both search and join tactic instantaneously and so does not consider alternative learning rules. Katsnelson et al. (2012) compared the linear operator rule with pure and mixed strategies while explicitly modeling the learning rule. They found that the linear operator rule can

invade a population of non-learners and even evolve to fixation if there is either sufficient environmental fluctuations or phenotypic asymmetries among individuals. Although they compared the linear operator rule with either pure or mixed strategies, they never pitted all three together. Furthermore, Katsnelson et al. (2012), as well as Hamblin and Giraldeau (2009), Arbilly et al. (2010), and Arbilly et al. (2011), only consider between-generation environmental fluctuations, ignoring within-generation fluctuations. This is surprising, given that learning can evolve when there is within generation changes, although very highly variable environment can discourage learning (Stephens, 1991).

Previous studies (Arbilly et al., 2010, 2011; Dubois et al., 2010; Hamblin & Giraldeau, 2009; Katsnelson et al., 2012; Trimmer et al., 2012) were interested in determining which strategy becomes fixed in the population. However, the main question in our opinion should be when to expect the stable coexistence of different strategies in the same population. It is therefore also necessary to determine the conditions under which no strategy is expected to reach fixation, and to monitor the average frequency of each strategy in the population.

In this study, we explore the conditions under which a linear operator type of learning rule could evolve in a frequency-dependent context like a producer-scrourer game. In doing so, we address the limitations of earlier studies by pitting fixed strategies against learning strategies in environments that are stable, vary within-generation or vary between-generations. We also measure the frequency of each strategy during the process rather than look only for fixation so that we can investigate when multiple strategies are expected to coexist simultaneously.

### 3.3 Simulation model

We chose the linear operator as the base structure for the learning rule in this study because it was found to be favoured by natural selection in a simulation study (Trimmer et al., 2012) and a recent agent-based simulation based on a linear operator

learning strategy was remarkably successful in replicating almost every published prediction and empirical result from the producer-scrouter game (Afshar & Giraldeau, 2014). The linear operator is simple and has fewer parameters than alternative learning rules (e.g. relative payoff sum; Harley 1981; perfect memory; Houston & Sumida 1987). It also provides the basic structure for many more complicated forms of reinforcement learning rules (Sutton & Barto, 1998). Moreover, it has been used previously in producer-scrouter game in competition with other learning rules (Beauchamp, 2000a; Hamblin & Giraldeau, 2009), with non-learning strategies (Arbilly & Laland, 2014; Katsnelson et al., 2012), and alone (Arbilly et al., 2010).

The simulation starts with a generation of individuals initialized using the parameters of Table 3.1. An individual can have one of the three strategies, pure, mixed, or linear operator. The type of the strategy is determined by  $L_1$ . For a pure strategy ( $L_1 = 1$ ), the individual's genes fix the tactic and depending on the pure strategy genotype ( $L_2$ ), the individual either always joins or always searches throughout its lifetime. For a mixed strategy ( $L_1 = 2$ ), an individual demonstrates both tactics with a fixed probability throughout its life. The probability of join tactic (1 - probability of search tactic) is fixed by the individual's genes ( $L_3$ ) and does not change during its lifetime. For a linear operator strategy ( $L_1 = 3$ ), the individual has flexible behaviour. It uses the linear operator learning rule to learn the current payoffs to each tactic, and so can respond to environmental changes that affect these payoffs (Afshar & Giraldeau, 2014). The individual adjusts its behaviour to try to maximize the amount of food gained, deciding whether to search or join by using the following linear operator rule:  $V_i(t) = V_i(t - 1) + x (P_i(t) - V_i(t - 1))$ ,  $V_i(0) = \text{unif}(0,1)$ ,  $0 < x < 1$  where  $V_i(t)$  and  $P_i(t)$  are the estimated value and observed payoff for tactic  $i$  at time  $t$  respectively.  $x$  is the learning rate and is determined by  $L_4$  ( $1 - x$  is the memory factor). The previous estimation was  $V_i(t - 1)$ , and  $P_i(t) - V_i(t - 1)$  is therefore the difference between the observed payoff and the previous estimated value. The update is conducted only for

the estimated value of the most recently used tactic. The estimated value remains unchanged,  $V_i(t) = V_i(t - 1)$ , for the tactic that has not been sampled at time  $t$ . The individual then decides between the two tactics using either maximizing or matching ratio, depending on  $L_5$ . When maximizing, the individual chooses the tactic with the highest estimated payoff. When matching, the individual performs the tactics in proportion to their relative estimated payoff.

**Table 3.1 Genetic algorithm chromosome description and initial values**

<b>Locus</b>	<b>Parameter</b>	<b>Range</b>	<b>Mutation bound</b>
<i>L</i> <sub>1</sub>	Strategy: Pure (1), Mixed (2), Linear operator (3)	1, 2, 3	1, 2, 3
<i>L</i> <sub>2</sub>	Pure strategy gene: Search tactic (1), Join tactic (2)	1, 2	1, 2
<i>L</i> <sub>3</sub>	Mixed strategy gene: Probability of using join tactic	[0, 1]	±0.1
<i>L</i> <sub>4</sub>	Linear operator memory factor	[0, 1]	±0.1
<i>L</i> <sub>5</sub>	Linear operator decision rule: Maximizing (0), Matching (1)	0, 1	0, 1

In the simulation model, a group of individuals forage in a  $G \times G$  grid, representing a virtual environment. Time is divided into discrete steps, each of which can be thought of as one sampling event in the real world. An individual who searches, encounters patches according to the patch encounter probability. Simultaneous discoveries by different individuals are possible when they concurrently use the search tactic. An individual who joins, sees the location of all discovered patches and joins the nearest patch by moving one grid cell per time step. The patch may be depleted before the individual arrives because of other individuals are already present at the patch.

Upon encountering a food patch the finder receives a fraction of the food in the patch called the finder's advantage. This represents the amount of food the finder would be able to eat before the joiners arrive. In the subsequent time steps all individuals at the patch, the finder and those that have joined, receive food at the same food intake rate. Competition would normally cause the rate of food intake to be affected by the number of scroungers at the patch, but for simplicity we kept intake rate fixed. The rate of depletion is thus based on the intake rate and the number of individuals.

The individuals at the patch continue eating until the patch is depleted, at which point each individual either joins or searches again, depending on its strategy. We assume that at each time step individuals require a minimum amount of energy to survive, a parameter we termed metabolic requirement. If an individual does not meet its metabolic requirement, it survives but gets a negative payoff to ensure that there is value in exploiting the better tactic more often. This is similar to the concepts of energy reserve or energy requirement (Koops & Giraldeau, 1996; Lendvai et al., 2004; Wu & Giraldeau, 2005). Individuals must also pay a cost of searching or a cost of joining. These costs have absolute values and are incurred in each time step, depending on the tactic they are playing. After the initialization, the population plays producer-scrounger game for 500 time steps. The general parameter settings for the game are shown in Table 3.2.

**Table 3.2** General parameter settings for the producer-scrounger (PS) game simulations

Parameter	Symbol	Value(s)
Grid size	$G$	20 x 20
Maximum simulation time step	$T_{max}$	500
Group size	$N$	10, 40, 90
Metabolic requirement	$C_{MR}$	0.125
Finder's advantage	$a$	1
Food intake rate	$b$	1
Cost of searching	$C_{search}$	0.125
Cost of joining	$C_{join}$	0.125
Patch encounter probability	$p$	0.1
Patch richness (mean)	$\mu_{F,static}$	20
Within-generation change	$\Delta_{\mu,W}$	0%, 25%, 50%, 75%
Between-generation change	$\Delta_{\mu,B}$	0%, 25%, 50%, 75%

At the end of the producer-scrourer game we use a process similar to Hamblin and Giraldeau (2009) to generate the next generation of foragers. We rank individuals based on their fitness, measured as the lifetime cumulative amount of food gained. Individuals with the lowest 10% fitness die. These individuals are then replaced with other individuals in the population in proportion to their fitness. Individuals who have higher fitness thus have a higher chance of replacing dead individuals. The resulting population will have the same number of individuals and will be the parents of the subsequent generation. Parents randomly pair with a uniform distribution. Each of the five loci (Table 3.1) may then be exchanged between paired parents with 50% probability, as a crossover operation. Finally, each locus of each gene may mutate with the probability of 0.002. The mutation happens by changing the locus by an amount within the bounds shown in Table 3.1. The resulting genes make up the individuals of the next generation. This process is repeated for 500 generations. We measured the frequency of each strategy during the last 100 generations.

We used static as well as dynamic environments where different levels of within-generation and between-generation changes were present. For each configuration, a set of 20 genetic algorithm simulations was conducted in Matlab (R2013a) for Windows (Mathworks, Natick, MA, U.S.A.). All the results are thus averaged over the 20 simulations.

We replicated all simulations with three different group sizes (10, 40, and 90). Several studies have shown that the evolutionarily stable proportion of join tactic changes with group size (Afshar & Giraldeau, 2014; Beauchamp & Giraldeau, 1996; Caraco & Giraldeau, 1991; Coolen, 2002; Vickery et al., 1991). Using different group sizes thus ensured that the equilibrium proportion of join tactic changed and that our results were not affected by the value of the equilibrium.

### 3.3.1 Static environment

We first tested the model in a static environment, where each patch had a fixed value of 20 energy units (Table 3.2), with five different strategy configurations, pure strategy only ( $L_1 = 1$ ), mixed strategy only ( $L_1 = 2$ ), pure and mixed strategies ( $L_1 \in \{1,2\}$ ), linear operator only ( $L_1 = 3$ ), and all three strategies together ( $L_1 \in \{1,2,3\}$ ). For each strategy configuration we tested three different group sizes (10, 40, and 90).

### 3.3.2 Dynamic environment

We studied both types of environmental fluctuation: within-generation change and between-generation change. To simulate within-generation change, individuals experienced three different conditions during their lifetime, an intermediate condition similar to the static environment,  $\mu_F = \mu_{F,static}$ , a condition with smaller patch richness,  $\mu_F * (1 - \Delta_{\mu,W})$ , and a condition with larger patch richness,  $\mu_F * (1 + \Delta_{\mu,W})$ , compared to the intermediate condition. The patch richness in the intermediate condition ( $\mu_{F,static}$ ) was 20 units. We tested four different values of  $\Delta_{\mu,W}$ , 0%, 25%, 50%, and 75% to see the effect of the amount of within-generation change. Within each condition, though, all patches had the same value.

In each simulation, all individuals first encountered the condition with low patch richness for one third of the simulation. Several models predict that animals should join less when patch richness is lower (Afshar & Giraldeau, 2014; Beauchamp & Giraldeau, 1996; Ruxton & Beauchamp, 2008; Vickery et al., 1991). We therefore expected to see the highest amount of searching in this environment. Individuals then encountered the intermediate patch richness condition for a third of the time, and then the higher patch richness condition for the last third of the time.

With between-generation change, the patch richness changes from generation to generation. During each generation, the group encountered one of the three

environments with different patch richness ( $\mu_{F,static}, \mu_{F,static} * (1 - \Delta_{\mu,B}), \mu_{F,static} * (1 + \Delta_{\mu,B})$ ) chosen at random, with a uniform distribution. We used four amounts of between-generation changes,  $\Delta_{\mu,B} = 0\%, 25\%, 50\%, \text{ and } 75\%$ .

In our study we had different combinations of within-generation change and between-generation change. Considering the four levels of each one, we ended up with 16 different configurations. The case where both  $\Delta_{\mu,W}$  and  $\Delta_{\mu,B}$  are 0% is equal to a static environment. The case with  $\Delta_{\mu,W} > 0\%$  and  $\Delta_{\mu,B} = 0\%$  represents an environment with only within-generation change. The case with  $\Delta_{\mu,W} = 0\%$  and  $\Delta_{\mu,B} > 0\%$  represents an environment with only between-generation change. For each configuration, we tested three different group sizes (10, 40, and 90) with all three strategies competing against each other.

### 3.4 Results

#### 3.4.1 Static environment

We found that regardless of the strategy configuration, the populations converged to similar stable proportions of join tactic for each group size of 40 or 90 (Table 3.3). For small groups of 10 individuals, however, the effect of genetic drift was strong and the groups ended at different proportions of join tactic.

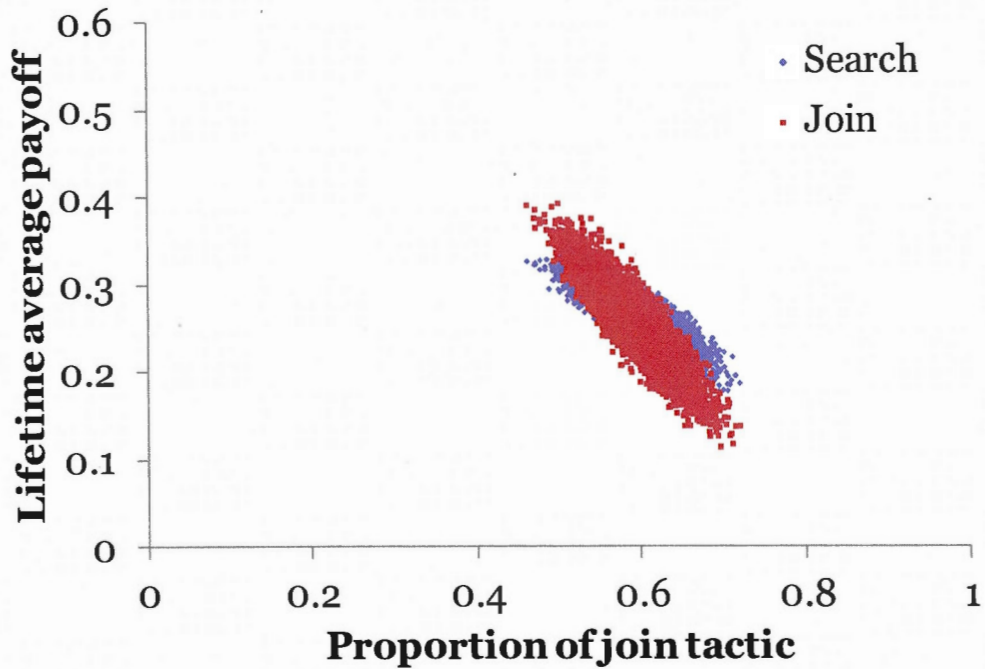
For all strategy configurations, we found higher proportions of join tactic for larger group sizes, with the exception of group size 10 with mixed strategy, likely due to genetic drift (Table 3.3). The increase in the proportion of join tactic for larger group sizes is consistent with previous models and empirical studies (Afshar & Giraldeau, 2014; Beauchamp & Giraldeau, 1996; Caraco & Giraldeau, 1991; Vickery et al., 1991).

Individuals' lifetime average payoff of each tactic vs. individual's lifetime proportion of join tactic shows that when the individual used a higher proportion of join, the

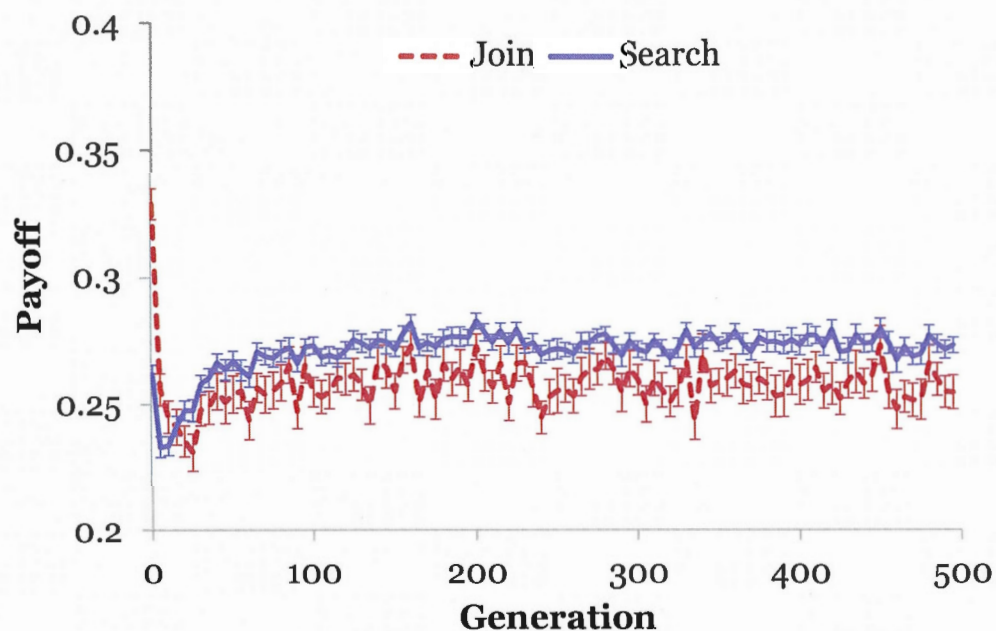
payoff of both tactics decreased (Figure 3.1) as expected in a producer-scrounger game. Although the payoff of join was higher for lower proportions of join tactic, it decreased as the proportion of join increased and eventually fell below the payoff of the search tactic. We found that for all configurations, the populations evolved to a combination of search and join where both tactics had equal payoffs (Figure 3.2). The variance in the payoff for the join tactic, however, was larger than that for the search tactic.

**Table 3.3** Final proportion of join tactic (Means  $\pm$  SE) for groups of individuals playing a producer–scrounger game using different types of strategies in a static environment,  $\Delta_{\mu,W} = 0\%$  and  $\Delta_{\mu,B} = 0\%$ , after 500 generations. The final proportion was calculated by averaging over the last 100 generations (N = 20)

Possible strategies	Group size		
	10	40	90
Pure strategy only	0.15 $\pm$ 0.06	0.39 $\pm$ 0.05	0.58 $\pm$ 0.03
Mixed strategy only	0.49 $\pm$ 0.34	0.41 $\pm$ 0.05	0.60 $\pm$ 0.03
Pure and mixed strategies	0.15 $\pm$ 0.07	0.40 $\pm$ 0.05	0.59 $\pm$ 0.03
Linear operator	0.33 $\pm$ 0.22	0.44 $\pm$ 0.06	0.60 $\pm$ 0.03
Pure, mixed, and linear operator strategies	0.14 $\pm$ 0.07	0.40 $\pm$ 0.05	0.59 $\pm$ 0.03



**Figure 3.1** Evolutionarily Stable Strategy plot for producer–scrounger game for groups of 90 individuals engaged in choosing tactics using pure strategy, mixed strategy, or linear operator learning rule in a static environment,  $\Delta_{\mu,W} = 0\%$  and  $\Delta_{\mu,B} = 0\%$ . Each point shows the lifetime average payoff of each tactic (per time step) for one individual vs. its lifetime proportion of join tactic. The general parameter settings for all simulations are shown in Table 3.1 and Table 3.2.

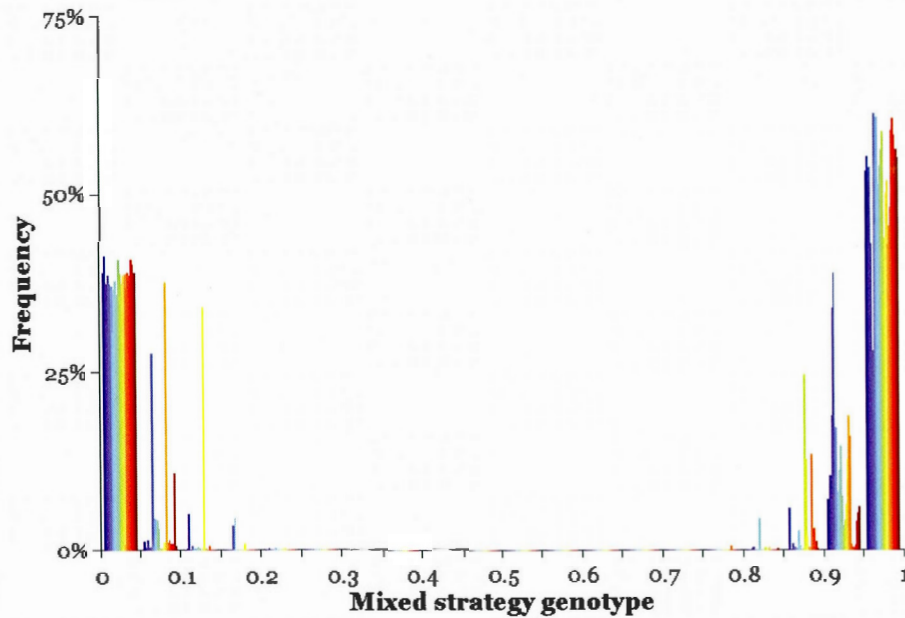


**Figure 3.2** Payoff of search and join tactics in producer–scrounger game for groups of 90 individuals engaged in choosing tactics using pure strategy, mixed strategy, or linear operator learning rule in a static environment during 500 generations. Means  $\pm$  SE are shown (N=20). The general parameter settings for all simulations are shown in Table 3.1 and Table 3.2. To keep the graph readable, only the results for every five generations is shown.

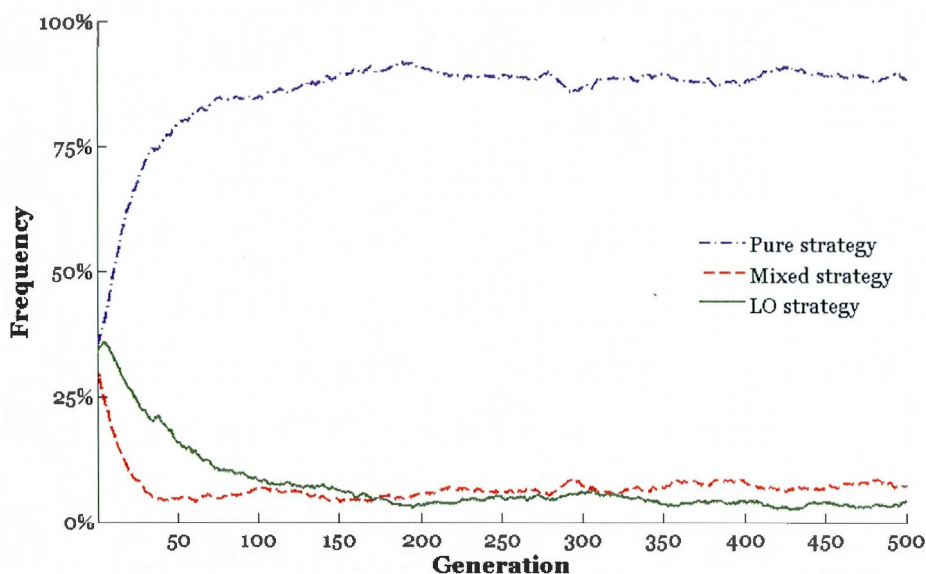
When both pure strategies and mixed strategies compete with each other in the same population in a static environment, the final combination is close to fixation with majority of the population using pure strategies (Table 3.4). Even the small proportion of mixed strategies behaved like a pure strategy, by having genotypes close to one (being similar to pure joiners) or close to zero (being similar to pure searchers). Similarly, when all individuals used a mixed strategy, the final frequency distribution of mixed strategy genotypes is made up of genotypes that are close to either one or zero, and therefore similar to pure strategies (Figure 3.3). Even when we added the linear operator, the pure strategy was still dominant (Table 3.4 and Figure 3.4). The frequency of individuals with the linear operator learning rule strategy increased with group size, but remained insignificant compared to pure and mixed strategies.

**Table 3.4** Final frequency of different strategies (Means  $\pm$  SE) when competing with each other in producer–scrounger game in a static environment during 500 generations. The final frequency was calculated by averaging over the last 100 generations (N = 20).

Possible strategies		Group size		
		10	40	90
<b>Pure strategy and mixed strategy</b>	Pure strategy	99% $\pm$ 1%	97% $\pm$ 1%	83% $\pm$ 3%
	Mixed strategy	1% $\pm$ 1%	3% $\pm$ 1%	17% $\pm$ 3%
<b>All three strategies</b>	Pure strategy	98% $\pm$ 1%	96% $\pm$ 1%	89% $\pm$ 2%
	Mixed strategy	1% $\pm$ 1%	2% $\pm$ 1%	7% $\pm$ 1%
	Liner operator	1% $\pm$ 0%	2% $\pm$ 1%	3% $\pm$ 1%



**Figure 3.3** The histogram of frequency of different mixed strategy genotypes for individuals that play producer-scrouter game using only mixed strategy in a static environment during 500 generations. Mixed strategy genotype is the threshold to randomly choose between search or join tactic. Despite the initially uniform distribution, the mixed strategy genotype finally converged to either close to one (only join) or close to zero (only search). We can therefore say that even populations of mixed strategists evolved to behave similar to pure strategists. Frequencies are shown for the last 100 generations when the populations have become stable. The general parameter settings for all simulations are shown in Table 3.1 and Table 3.2. 20 simulations were performed and each colour represents the result of one of the 20 simulations.



**Figure 3.4** The evolution of different strategies in a producer–scrounger game for groups of 90 individuals engaged in choosing tactics using pure strategy, mixed strategy, or a linear operator learning rule in a static environment during 500 generations ( $N = 20$ ). After initial changes in the distribution of strategies, the pure strategy remains dominant in the populations for the remainder of the simulation. The linear operator learning strategy thus has no advantage over inflexible strategies, either pure or mixed strategies. The general parameter settings for all simulations are shown in Table 3.1 and Table 3.2.

### 3.4.2 Dynamic environment

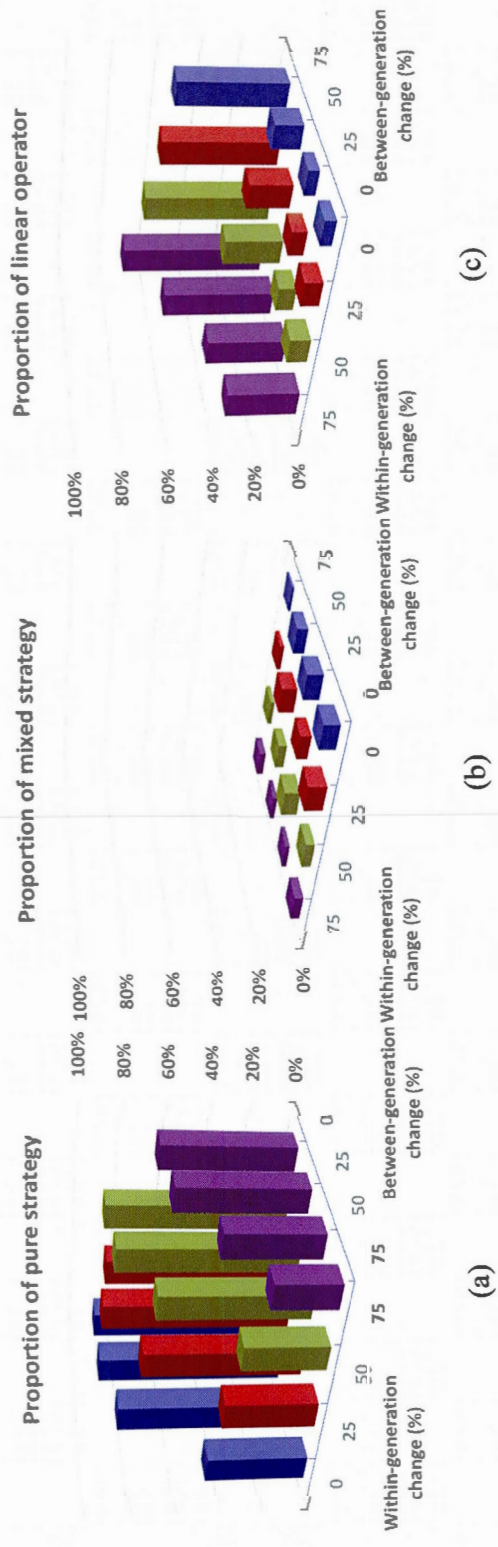
We found that both within-generation change ( $\Delta_{\mu,W}$ ) and between-generation change ( $\Delta_{\mu,B}$ ) can increase the percentage of the linear operator strategy in the population (Table 3.5, Figure 3.5a, and Appendix A). As the amount of environmental change, either  $\Delta_{\mu,W}$  or  $\Delta_{\mu,B}$ , increases, this effect becomes stronger. The population, however, remains dominated by pure strategy.

Between-generation change had stronger effects than within-generation change and the proportion of the linear operator strategy was often larger than for within-generation change (Table 3.5, Figure 3.5a, and Appendix A). Out of the 16 configurations, individuals with linear operator represented the majority of the population in only five cases, four of which was when the between-generation change was 75% (Table 3.5). We, however, did not find any situation where the linear operator reached fixation.

**Table 3.5** Final proportions of pure, mixed, and linear operator strategies (Means  $\pm$  SE) when competing with each other in populations of 90 individuals playing a producer-scrouter game in a dynamic environment with different levels of within-generation change ( $\Delta_{\mu,W}$ ) and between-generation change ( $\Delta_{\mu,B}$ ) during 500 generations. The final proportion was calculated by averaging over the last 100 generations ( $N = 20$ ). The general parameter settings for all simulations are shown in Table 3.1 and Table 3.2. The results for the populations of sizes 10 and 40 are available in the Appendix.

	Pure strategy				Mixed strategy				Linear operator			
	Within-generation change ( $\Delta_{\mu,W}$ )				Within-generation change ( $\Delta_{\mu,W}$ )				Within-generation change ( $\Delta_{\mu,W}$ )			
	0%	25%	50%	75%	0%	25%	50%	75%	0%	25%	50%	75%
0%	88% $\pm$ 2%	84% $\pm$ 3%	86% $\pm$ 2%	64% $\pm$ 4%	7% $\pm$ 2%	8% $\pm$ 3%	4% $\pm$ 1%	4% $\pm$ 1%	5% $\pm$ 1%	8% $\pm$ 2%	9% $\pm$ 2%	32% $\pm$ 4%
25%	88% $\pm$ 3%	88% $\pm$ 2%	84% $\pm$ 3%	61% $\pm$ 3%	7% $\pm$ 2%	5% $\pm$ 1%	7% $\pm$ 2%	2% $\pm$ 0%	5% $\pm$ 1%	7% $\pm$ 1%	8% $\pm$ 1%	38% $\pm$ 4%
50%	82% $\pm$ 2%	73% $\pm$ 2%	69% $\pm$ 2%	44% $\pm$ 3%	5% $\pm$ 1%	7% $\pm$ 1%	4% $\pm$ 1%	2% $\pm$ 0%	13% $\pm$ 2%	20% $\pm$ 2%	27% $\pm$ 2%	54% $\pm$ 3%
75%	46% $\pm$ 2%	41% $\pm$ 3%	36% $\pm$ 1%	28% $\pm$ 2%	1% $\pm$ 0%	2% $\pm$ 0%	2% $\pm$ 1%	3% $\pm$ 1%	53% $\pm$ 2%	57% $\pm$ 3%	62% $\pm$ 2%	70% $\pm$ 1%

Between-generation  
change ( $\Delta_{\mu,B}$ )



**Figure 3.5** Final proportions of a) pure, b) mixed, and c) linear operator strategies when competing with each other in populations of 90 individuals playing a producer–scrounger game in a dynamic environment with different levels of within-generation change ( $\Delta_{\mu,W}$ ) and between-generation change ( $\Delta_{\mu,B}$ ) during 500 generations. The final proportions were calculated by averaging over the last 100 generations ( $N = 20$ ). The general parameter settings for all simulations are shown in Table 3.1 and Table 3.2. To keep the graph readable, axes of within-generation and between-generation were reversed for (a). The results for populations of sizes 10 and 40 are available in the Appendix.

### 3.5 Discussion

We found that, depending on the type and the intensity of environmental variability, learning may either be completely displaced by fixed rules or, at best, it can coexist with pure strategies. We discuss the conditions under which each of these situations apply and then relate this to the validity of accepting the behavioural gambit.

For all configurations of the simulation, the payoffs obtained from each strategy behaved as assumed by a producer-scrourer Game. Consequently the populations evolved to an equilibrium where search and join tactics obtained equal payoffs (Figure 3.2). Furthermore, when an individual joined less than this equilibrium, the join tactic received a higher payoff than the search tactic, which should cause them to join more if they use linear operator strategy, but when an individual joined more than the equilibrium the join tactic had a lower payoff than the search tactic, which should cause them to join less if they use linear operator strategy (Figure 3.1). We can therefore argue that the populations had converged to the evolutionarily stable strategy (ESS) equilibrium which, once reached, cannot be invaded by alternative combinations of search and join tactics. It is worth noting that our simulation provides the first numerical demonstration of the evolutionarily stable strategy payoff plot based on the outcome of an actual agent-based producer-scrourer game model (Figure 3.1). The functions obtained appear to match those measured in empirical tests of the producer-scrourer game with house sparrows (*Passer domesticus*, Barnard and Sibly 1981) and with nutmeg mannikins (*Lonchura punctulata*, Mottley and Giraldeau 2000).

Katsnelson et al. (2012) report that a mixed strategy can invade a population of pure strategies and evolve to fixation. Our results contradict this because in the static environments the pure strategy dominated the mixed strategy (Table 3.4). Moreover, the small proportion of individuals that use a mixed strategy acted like pure strategists by having mixed strategy genotypes close to one or zero (Figure 3.3). This

difference in conclusion between our study and Katsnelson et al. (2012) is likely attributable to two distinguishing features. Katsnelson et al. (2012) set the producer-scrounger game parameters such that the optimal solution is an equal proportion of join and search tactics. The mixed strategy individuals, therefore, already had the optimal behaviour even before the evolutionary process began and so enjoyed an unfair advantage in competition with the other fixed strategies. This was not so in our study. In addition in Katsnelson et al. (2012) the mixed strategies were constrained to always use the join tactic 50% of the time. No change was therefore possible within the mixed strategy. In our study, however, the proportion of join tactic used by the mixed strategy genotype ( $L_3$ ) could evolve along with other alleles, including the strategy itself ( $L_1$ ). We argue that in a genetic algorithm study in order to draw robust conclusions it is important to allow the parameters of all strategies to change and adapt without placing strong constraints on any one of them.

Our finding that the pure strategy completely dominates the mixed strategy in a static environment (Table 3.4) supports the predictions of earlier game models (Vickery 1987, 1988) that in a finite population only pure strategies should be expected from purely deterministic dynamics and that no mixed strategy can be an ESS. On the other hand, our result contradicts arguments based on stochastic dynamics that when the frequency of one pure strategist is low at the ESS, stochastic environmental fluctuations may cause its extinction and hence lead to the fixation of mixed strategies (Bergstrom & Godfrey-Smith, 1998; Maynard Smith, 1988; Orzack & Hines, 2005). These contradictory predictions highlight the importance of the deterministic or stochastic nature of the dynamics and our results suggest that the producer-scrounger game may favour the deterministic assumptions of Vickery (1987, 1988) rather than the stochastic assumptions. This result appears robust given that we included random between-generation fluctuations and random initial values in our simulated producer-scrounger game.

We found that in a static environment the pure strategy dominates both the linear operator and the mixed strategies (Table 3.4 and Figure 3.4). This is not surprising because many models have already predicted that learning is of no use when there is no environmental change (Dubois et al., 2010; Dunlap & Stephens, 2009; Groß et al., 2008; Katsnelson et al., 2012; Stephens, 1991). In this case, innate behaviour can outperform flexible behaviours based on learning (Stephens, 1991). In the absence of between-generation change, the frequency of the learning strategy remains low even when we introduced within-generation environmental changes, with the exception of extreme environmental changes ( $\Delta_{\mu,W} = 75\%$ ) with the large group size (Table 3.5 and Figure 3.5c). In our model, the fitness of an individual is measured based on the lifetime food intake. Lifetime food intake includes all three periods of different environmental conditions ( $\mu_F, \mu_F * (1 - \Delta_{\mu,W}), \mu_F * (1 + \Delta_{\mu,W})$ ). When the environment changes during an individual's lifetime but the same amount of changes repeats for several generations, innate behaviour can outperform learning. This is especially true, when a small portion of population uses the learning strategy (Table 3.5 and Figure 3.5c). This portion of the population can act as a buffer and compensate for individuals that consistently show a higher or lower level proportion of join tactic (Dubois et al., 2010). It therefore buffers the inflexible strategies from the force of natural selection.

Between-generation change had stronger effects on strategy type than within-generation change and the proportion of the linear operator strategy was often higher than for within-generation change (Table 3.5 and Figure 3.5a). We found the highest proportion of linear operator strategies when there was highest amount of between-generation changes and highest amount of within-generation change (Table 3.5 and Figure 3.5c). In this case, innate behaviour that is optimal for one generation might not be useful for the following generation. It is therefore important to adapt and find the optimal proportion of join tactic in each generation. This can be done only with a flexible strategy such as the linear operator. However, the intrinsic frequency-

dependency property of the producer-scrourer game implies that when one individual changes its behaviour it also affects the performance of other individuals in the population creating, once again, a buffer effect by adjusting their behaviour until the payoffs of both tactics are equal. When this happens there is no advantage to flexibility for the remaining fixed individuals and so, as was the case in Dubois et al. (2010) the linear operator cannot evolve to fixation in the population (Table 3.5 and Figure 3.5c). Although Katsnelson et al. (2012) also found that the frequency of learning individuals in the population increases with the magnitude of the between-generation environmental change they observe the learning strategy evolving to fixation. There are several possible reasons for the difference between our results and those of Katsnelson et al. (2012). One has already been addressed and concerns the constrained parameters of the mixed strategy. But they also kept a fixed memory factor for the linear operator strategy at one of three values: 0.5, 0.85, and 0.95. Instead we allowed the memory factor ( $L_5$ ) to evolve freely. Moreover, Katsnelson et al. (2012) studied the linear operator along with either the pure strategy or the mixed strategy but there was no situation where all three competed together. As Katsnelson et al. (2012) suggested, the benefit of the linear operator may come in part by demonstrating a mixed combination of both tactics, similar to a mixed strategy. This may give it an advantage when competing with a pure strategy. However, when all three strategies play together, as in our study, the linear operator strategy may be less likely to evolve to fixation. We can think of the game involving the strategies as a Rock-Scissor-Paper game, where each strategy is better than one other strategy, but none is better than all. Our results underscore the importance of playing all strategies together in order to obtain robust conclusions.

We found that, regardless of the possible strategies, all populations with similar group size converged to a comparable proportion of join tactics (Table 3.3), suggesting that the ESS is independent of the mechanisms that governs the decision-making of individuals. This fact provides us with two important results. First, regardless of

whether the population uses pure, mixed, or learning strategies, we can expect the effects of environmental parameters on the population proportion of join tactic to be similar to those predicted by Afshar and Giraldeau (2014) where all individuals in the population used linear operator to choose between search and join tactics.

Second, when the research interest is in the population level behavior in a static environment, behavioral ecologists can continue using the behavioral gambit and assume that the mechanism in a frequency-dependent situations will produce outcomes that are the same as those from an evolutionary process. However, if the interest is in individual level behavior or involves environmental fluctuations, the gambit may be wrong and learning should be considered.

### 3.6 Acknowledgements

This research and M.A. were supported by a Natural Science and Engineering Research Council of Canada Discovery Grant (number 303112) to L.-A.G. In addition M.A. was financially supported through a Fund Accessibility and Academic Success (FARE) Scholarship and a foreign fee remission from University of Québec in Montréal and a doctoral Scholarship from Group For Research In Decision Analysis (GERAD).

[Cette page a été laissée intentionnellement blanche]

## CHAPTER IV

### GENERAL DISCUSSION AND FUTURE DIRECTIONS

My primary goal in this thesis was to explore the hypothesis that learning can be as powerful as, or in some contexts perhaps even more powerful than, evolution for predicting animal behavior. To do this I proposed a learning based model for a producer-scrouter game and compared it with previous published studies, conducted an experiment to test the novel prediction of my model using captive foraging flocks of male zebra finches (*Taeniopygia guttata*), and developed second simulation model to explore the conditions under which learning is expected to evolve in a frequency-dependent context, each of which were explained in detail in the previous chapters. In this chapter, I discuss my contributions to the field of behavioral ecology and the implications of my findings for the study of animal behavior.

My results demonstrate that adopting a learning approach when studying a frequency-dependent behaviour is more effective than adopting an evolutionary game-theoretic approach. The learning approach was particularly powerful in two situations: i) when more complex, and hence more realistic, models are required and ii) when individual behaviour needs to be predicted. Evolutionary game models cannot easily deal with the complexity required of realistic social foraging situations. As such, the evolutionary game theory models used by behavioral ecologists for the past 30-40 years have required an oversimplification of conditions that has recently been criticised. More importantly, evolutionary game theory models cannot deal well with multiple ecological parameters that likely correspond to more realistic environments.

Agent-based models using learning rules, similar to those I used in this thesis, provide a convenient alternative to study complex behaviors based on simple individual level models. Furthermore, when we need to dig deeper than the population level behavior and predict individual level behavior the learning approach is again more successful than evolutionary game models. Individual behaviour is of particular interest to behavioural ecologists who have now discovered the importance of persistent individual differences often referred to as “personality” (Réale & Dingemanse, 2012). I have shown that different populations with similar population level behavior can indeed have individuals with different behaviours. Some of the individuals can have fixed behaviours while others have flexible behaviours. If one studies individual differences, it is important to consider flexible behaviour in general, and learning in particular, as a potential source of variation. Neglecting the mechanism, i.e. learning, thus means neglecting the behaviour of what can be a significant portion of the population when there are high levels of environmental fluctuations. Moreover, this means that models developed for studying behaviour in static environments may not be appropriate for dynamic environments.

We can thus argue that two of the early questions a cautious behavioural ecologist should ask are first, whether the environment is static and, second, whether he or she is interested in behaviour only at the population level. If the answer to either of these questions is “no”, then he or she needs to consider learning as a potential part of the model. We should thus be careful when using the results of previous or future studies that failed to take this into account and, for example, used a model entirely based on fixed behaviours in a changing environment.

#### 4.1 Future directions

The majority of the models on learning in behavioural ecology are based on the assumption that either individuals are identical or their individual differences does not affect the outcome of the learning process. Researchers later verify these predictions

assumption in their experiments. That this assumption is so often made is surprising, given that individual differences are well documented in the literature (Réale & Dingemanse, 2012). For example, in Chapter II I showed that individuals varied consistently in their use of tactics across both non-variable and variable patch conditions (Figure 2.2). Furthermore, the simulations I present in Chapter I, show that individual differences can arise because of different early experiences (Figure 1.1), even when individuals use similar learning mechanisms. It is therefore important to be cautious about accepting this assumption, and to investigate the contexts under which it may or may not hold using both models and experiments.

In a non-frequency dependent task, the optimal solution is equivalent for all identical individuals. We therefore might expect identical individuals to express similar behaviour even if they have different learning rules. I can think of only one case where we might expect persistent individual differences in a non-frequency dependent task. Several studies have shown that learning is costly. Animals may need to spend time sampling less profitable choices, for example. In extreme cases, individuals may choose not to sample all choices and may thus end up with an inaccurate estimate of the different options. In such an environment, individuals may express different behaviour, with some individuals making what appear to be suboptimal choices. Based on Stephens (1989), we can expect these differences to persist if the cost of deviation from the optimal behaviour is smaller than the cost of obtaining the information necessary to make the optimal choice.

In a frequency-dependent task, on the other hand, each individual may experience a different environment due to the behaviour of other individuals. The optimal behaviour for each individual may therefore also differ, and we can expect to observe individual differences. My simulations show that when individuals use same learning rule in a frequency-dependent task, different early experiences lead to temporal, and not persistent, individual differences over time (Figure 1.1). The next step would be

to see what happens when individuals differ intrinsically, having either different learning rules or the same learning rule with different parameters.

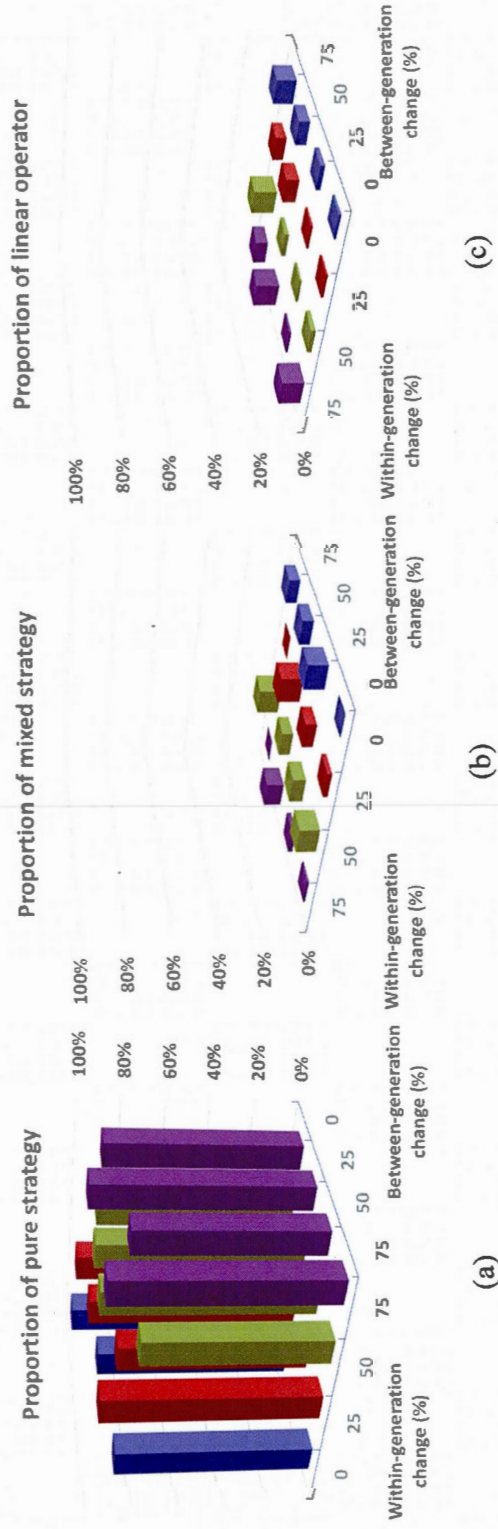
[Cette page a été laissée intentionnellement blanche]

## APPENDIX A

### FINAL PROPORTION OF PURE, MIXED, AND LINEAR OPERATOR STRATEGIES WHEN COMPETING WITH EACH OTHER IN PRODUCER- SCROUNGER GAME IN DYNAMIC ENVIRONMENTS

**Table A.1** Final proportions of pure, mixed, and linear operator strategies (Means  $\pm$  SE) when competing with each other in populations of 10 individuals playing a producer–scrounger game in a dynamic environment with different levels of within-generation change ( $\Delta_{\mu,W}$ ) and between-generation change ( $\Delta_{\mu,B}$ ) during 500 generations. The final proportion was calculated by averaging over the last 100 generations ( $N = 20$ ). The general parameter settings for all simulations are shown in Table 3.1 and Table 3.2.

Between-generation change ( $\Delta_{\mu,B}$ )	Pure strategy				Mixed strategy				Linear operator			
	Within-generation change ( $\Delta_{\mu,W}$ )				Within-generation change ( $\Delta_{\mu,W}$ )				Within-generation change ( $\Delta_{\mu,W}$ )			
	0%	25%	50%	75%	0%	25%	50%	75%	0%	25%	50%	75%
0%	99% $\pm$ 0%	98% $\pm$ 1%	90% $\pm$ 5%	89% $\pm$ 7%	0% $\pm$ 0%	2% $\pm$ 1%	9% $\pm$ 5%	1% $\pm$ 1%	1% $\pm$ 0%	1% $\pm$ 0%	2% $\pm$ 1%	10% $\pm$ 7%
25%	89% $\pm$ 7%	94% $\pm$ 5%	93% $\pm$ 5%	97% $\pm$ 1%	9% $\pm$ 6%	5% $\pm$ 5%	6% $\pm$ 5%	2% $\pm$ 1%	2% $\pm$ 2%	1% $\pm$ 1%	1% $\pm$ 0%	1% $\pm$ 1%
50%	89% $\pm$ 7%	84% $\pm$ 8%	93% $\pm$ 5%	82% $\pm$ 8%	5% $\pm$ 5%	10% $\pm$ 7%	5% $\pm$ 5%	8% $\pm$ 5%	5% $\pm$ 5%	6% $\pm$ 5%	2% $\pm$ 1%	11% $\pm$ 7%
75%	86% $\pm$ 7%	94% $\pm$ 5%	79% $\pm$ 9%	94% $\pm$ 5%	5% $\pm$ 5%	0% $\pm$ 0%	10% $\pm$ 7%	0% $\pm$ 0%	9% $\pm$ 6%	5% $\pm$ 5%	11% $\pm$ 7%	6% $\pm$ 5%

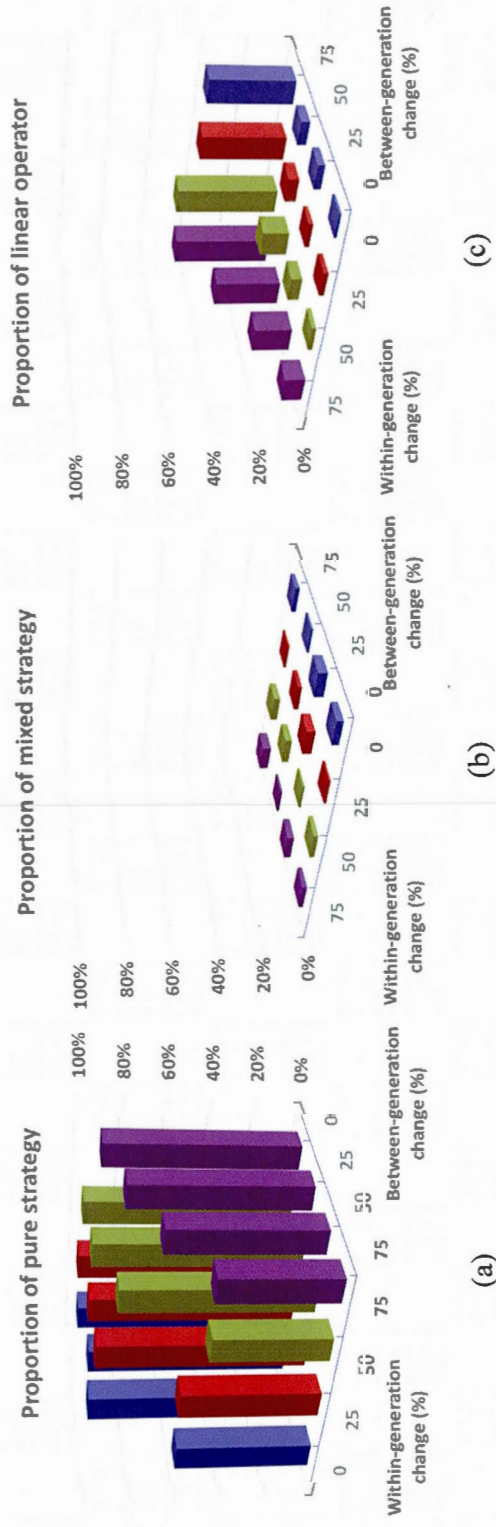


**Figure A.1** Final proportions of a) pure, b) mixed, and c) linear operator strategies when competing with each other in populations of 10 individuals playing a producer-scrounger game in a dynamic environment with different levels of within-generation change ( $\Delta_{\mu_W}$ ) and between-generation change ( $\Delta_{\mu_B}$ ) during 500 generations. The final proportions were calculated by averaging over the last 100 generations ( $N = 20$ ). The general parameter settings for all simulations are shown in Table 3.1 and Table 3.2. To keep the graph readable, axes of within-generation and between-generation were reversed for (a).

**Table A.2** Final proportions of pure, mixed, and linear operator strategies (Means  $\pm$  SE) when competing with each other in populations of 40 individuals playing a producer–scrounger game in a dynamic environment with different levels of within-generation change ( $\Delta_{\mu,w}$ ) and between-generation change ( $\Delta_{\mu,B}$ ) during 500 generations. The final proportion was calculated by averaging over the last 100 generations ( $N = 20$ ). The general parameter settings for all simulations are shown in Table 3.1 and Table 3.2.

	Pure strategy				Mixed strategy				Linear operator			
	Within-generation change ( $\Delta_{\mu,w}$ )				Within-generation change ( $\Delta_{\mu,w}$ )				Within-generation change ( $\Delta_{\mu,w}$ )			
	0%	25%	50%	75%	0%	25%	50%	75%	0%	25%	50%	75%
0%	96% $\pm$ 1%	97% $\pm$ 1%	96% $\pm$ 1%	89% $\pm$ 3%	3% $\pm$ 1%	1% $\pm$ 0%	2% $\pm$ 1%	2% $\pm$ 1%	1% $\pm$ 0%	2% $\pm$ 0%	2% $\pm$ 1%	9% $\pm$ 3%
25%	93% $\pm$ 2%	94% $\pm$ 3%	94% $\pm$ 2%	81% $\pm$ 3%	4% $\pm$ 2%	4% $\pm$ 2%	1% $\pm$ 0%	3% $\pm$ 2%	3% $\pm$ 1%	2% $\pm$ 1%	5% $\pm$ 2%	17% $\pm$ 3%
50%	95% $\pm$ 2%	93% $\pm$ 2%	85% $\pm$ 3%	68% $\pm$ 4%	1% $\pm$ 0%	2% $\pm$ 1%	3% $\pm$ 1%	1% $\pm$ 1%	4% $\pm$ 2%	5% $\pm$ 2%	12% $\pm$ 3%	30% $\pm$ 4%
75%	59% $\pm$ 4%	60% $\pm$ 3%	50% $\pm$ 5%	51% $\pm$ 3%	2% $\pm$ 1%	1% $\pm$ 0%	3% $\pm$ 2%	4% $\pm$ 2%	39% $\pm$ 4%	39% $\pm$ 3%	47% $\pm$ 4%	45% $\pm$ 2%

Between-generation change ( $\Delta_{\mu,B}$ )



**Figure A.2** Final proportions of a) pure, b) mixed, and c) linear operator strategies when competing with each other in populations of 40 individuals playing a producer-scrouter game in a dynamic environment with different levels of within-generation change ( $\Delta_{\mu_W}$ ) and between-generation change ( $\Delta_{\mu_B}$ ) during 500 generations. The final proportions were calculated by averaging over the last 100 generations ( $N = 20$ ). The general parameter settings for all simulations are shown in Table 3.1 and Table 3.2. To keep the graph readable, axes of within-generation and between-generation were reversed for (a).

## BIBLIOGRAPHY

- Afshar, M., & Giraldeau, L.-A. (2014). A unified modelling approach for producer–scrounger games in complex ecological conditions. *Animal Behaviour*, 96, 167–176. <http://doi.org/10.1016/j.anbehav.2014.07.022>
- Aoki, K., & Feldman, M. W. (2014). Evolution of learning strategies in temporally and spatially variable environments: A review of theory. *Theoretical Population Biology*, 91, 3–19. <http://doi.org/10.1016/j.tpb.2013.10.004>
- Arbilly, M., & Laland, K. N. (2014). The local enhancement conundrum: in search of the adaptive value of a social learning mechanism. *Theoretical Population Biology*, 91, 50–57. <http://doi.org/10.1016/j.tpb.2013.09.006>
- Arbilly, M., Motro, U., Feldman, M. W., & Lotem, A. (2010). Co-evolution of learning complexity and social foraging strategies. *Journal of Theoretical Biology*, 267(4), 573–581. <http://doi.org/10.1016/j.jtbi.2010.09.026>
- Arbilly, M., Motro, U., Feldman, M. W., & Lotem, A. (2011). Evolution of social learning when high expected payoffs are associated with high risk of failure. *Journal of the Royal Society Interface*, 8(64), 1604–1615. <http://doi.org/10.1098/rsif.2011.0138>
- Baldwin, J. M. (1896). A New Factor in Evolution. *The American Naturalist*, 30(354), 441–451.

- Barnard, C. J., & Sibly, R. M. (1981). Producers and scroungers: A general model and its application to captive flocks of house sparrows. *Animal Behaviour*, 29(2), 543–550.  
[http://doi.org/10.1016/S0003-3472\(81\)80117-0](http://doi.org/10.1016/S0003-3472(81)80117-0)
- Barrette, M., & Giraldeau, L.-A. (2006). Prey crypticity reduces the proportion of group members searching for food. *Animal Behaviour*, 71(5), 1183–1189.  
<http://doi.org/10.1016/j.anbehav.2005.10.008>
- Barta, Z., Flynn, R., & Giraldeau, L.-A. (1997). Geometry for a Selfish Foraging Group: A Genetic Algorithm Approach. *Proceedings: Biological Sciences*, 264(1385), 1233–1238.
- Barta, Z., & Giraldeau, L.-A. (1998). The effect of dominance hierarchy on the use of alternative foraging tactics: a phenotype-limited producing-scrounging game. *Behavioral Ecology and Sociobiology*, 42(3), 217–223.  
<http://doi.org/10.1007/s002650050433>
- Bateson, P. (2004). The Active Role of Behaviour in Evolution. *Biology and Philosophy*, 19(2), 283–298. <http://doi.org/10.1023/B:BIPH.0000024468.12161.83>
- Beauchamp, G. (2000a). Learning Rules for Social Foragers: Implications for the Producer–Scrounger Game and Ideal Free Distribution Theory. *Journal of Theoretical Biology*, 207(1), 21–35. <http://doi.org/10.1006/jtbi.2000.2153>
- Beauchamp, G. (2000b). The effect of prior residence and pair bond on scrounging choices in flocks of zebra finches, *Taenopygia guttata*. *Behavioural Processes*, 52(2–3), 131–140. [http://doi.org/10.1016/S0376-6357\(00\)00132-7](http://doi.org/10.1016/S0376-6357(00)00132-7)

- Beauchamp, G. (2001). Consistency and flexibility in the scrounging behaviour of zebra finches. *Canadian Journal of Zoology*, 79(3), 540–544. <http://doi.org/10.1139/z01-008>
- Beauchamp, G. (2006). Phenotypic Correlates of Scrounging Behavior in Zebra Finches: Role of Foraging Efficiency and Dominance. *Ethology*, 112(9), 873–878. <http://doi.org/10.1111/j.1439-0310.2006.01241.x>
- Beauchamp, G. (2008). A spatial model of producing and scrounging. *Animal Behaviour*, 76(6), 1935–1942. <http://doi.org/10.1016/j.anbehav.2008.08.017>
- Beauchamp, G., Bélisle, M., & Giraldeau, L.-A. (1997). Influence of Conspecific Attraction on the Spatial Distribution of Learning Foragers in a Patchy Habitat. *Journal of Animal Ecology*, 66(5), 671–682. <http://doi.org/10.2307/5920>
- Beauchamp, G., & Giraldeau, L.-A. (1996). Group Foraging Revisited: Information Sharing or Producer-Scrounger Game? *The American Naturalist*, 148(4), 738–743.
- Beauchamp, G., Giraldeau, L.-A., & Ennis, N. (1997). Experimental evidence for the maintenance of foraging specializations by frequency-dependent choice in flocks of spice finches. *Ethology Ecology & Evolution*, 9(2). <http://doi.org/10.1080/08927014.1997.9522890>
- Belmaker, A., Motro, U., Feldman, M. W., & Lotem, A. (2012). Learning to Choose Among Social Foraging Strategies in Adult House Sparrows (*Passer domesticus*). *Ethology*, 118(11), 1111–1121. <http://doi.org/10.1111/eth.12013>
- Bergstrom, C. T., & Godfrey-Smith, P. (1998). On the Evolution of Behavioral Heterogeneity in Individuals and Populations. *Biology and Philosophy*, 13(2), 205–231. <http://doi.org/10.1023/A:1006588918909>

- Bicca-Marques, J. C., & Garber, P. A. (2005). Use of Social and Ecological Information in Tamarin Foraging Decisions. *International Journal of Primatology*, 26(6), 1321–1344. <http://doi.org/10.1007/s10764-005-8855-9>
- Biondolillo, K., Stamp, C., Woods, J., & Smith, R. (1997). Working and scrounging by zebra finches in an operant task. *Behavioural Processes*, 39(3), 263–269. [http://doi.org/10.1016/S0376-6357\(96\)00763-2](http://doi.org/10.1016/S0376-6357(96)00763-2)
- Blumstein, D. T., & Daniel, J. C. (2007). *Quantifying behavior the JWatcher way*. Sunderland, Mass.: Sinauer Associates.
- Bolker, B. M., Brooks, M. E., Clark, C. J., Geange, S. W., Poulsen, J. R., Stevens, M. H. H., & White, J.-S. S. (2009). Generalized linear mixed models: a practical guide for ecology and evolution. *Trends in Ecology & Evolution*, 24(3), 127–135. <http://doi.org/10.1016/j.tree.2008.10.008>
- Bush, R. R., & Mosteller, F. (1951a). A Mathematical Model for Simple Learning. In S. E. Fienberg & D. C. Hoaglin (Eds.), *Selected Papers of Frederick Mosteller* (pp. 221–234). New York, NY: Springer New York. Retrieved from <http://www.springerlink.com/content/u486mhu7pp33v081/>
- Bush, R. R., & Mosteller, F. (1951b). A Model for Stimulus Generalization and Discrimination. In S. E. Fienberg & D. C. Hoaglin (Eds.), *Selected Papers of Frederick Mosteller* (pp. 235–250). New York, NY: Springer New York. Retrieved from <http://www.springerlink.com/content/g1m11163787k40pj/>
- Bush, R. R., & Mosteller, F. (1953). A Stochastic Model with Applications to Learning. *The Annals of Mathematical Statistics*, 24(4), 559–585.
- Bush, R. R., & Mosteller, F. (1955). *Stochastic models for learning*. Wiley.

- Caraco, T., & Giraldeau, L.-A. (1991). Social foraging: Producing and scrounging in a stochastic environment. *Journal of Theoretical Biology*, 153(4), 559–583.  
[http://doi.org/10.1016/S0022-5193\(05\)80156-0](http://doi.org/10.1016/S0022-5193(05)80156-0)
- Carroll, S. P., Hendry, A. P., Reznick, D. N., & Fox, C. W. (2007). Evolution on ecological time-scales. *Functional Ecology*, 21(3), 387–393. <http://doi.org/10.1111/j.1365-2435.2007.01289.x>
- Coolen, I. (2002). Increasing Foraging Group Size Increases Scrounger Use and Reduces Searching Efficiency in Nutmeg Mannikins (*Lonchura punctulata*). *Behavioral Ecology and Sociobiology*, 52(3), 232–238.
- Coolen, I., Giraldeau, L.-A., & Lavoie, M. (2001). Head position as an indicator of producer and scrounger tactics in a ground-feeding bird. *Animal Behaviour*, 61(5), 895–903.  
<http://doi.org/10.1006/anbe.2000.1678>
- Cormen, T. H. (2001). *Introduction to algorithms*. MIT Press.
- Danchin, E., Giraldeau, L.-A., & Cézilly, F. (Eds.). (2008). *Behavioural Ecology: An Evolutionary Perspective on Behaviour* (1st ed.). Oxford University Press, USA.
- David, M., Cézilly, F., & Giraldeau, L.-A. (2011). Personality affects zebra finch feeding success in a producer–scrounger game. *Animal Behaviour*, 82(1), 61–67.  
<http://doi.org/10.1016/j.anbehav.2011.03.025>
- David, M., & Giraldeau, L.-A. (2012). Zebra finches in poor condition produce more and consume more food in a producer–scrounger game. *Behavioral Ecology*, 23(1), 174–180. <http://doi.org/10.1093/beheco/arr171>
- Davies, N. B., Krebs, J. R., & West, S. A. (2012). *An Introduction to Behavioural Ecology*. John Wiley & Sons.

- Dridi, S., & Lehmann, L. (2014). On learning dynamics underlying the evolution of learning rules. *Theoretical Population Biology*, 91, 20–36.  
<http://doi.org/10.1016/j.tpb.2013.09.003>
- Dubois, F., Giraldeau, L.-A., & Réale, D. (2012). Frequency-dependent payoffs and sequential decision-making favour consistent tactic use. *Proceedings of the Royal Society B: Biological Sciences*, 279(1735), 1977–1985.  
<http://doi.org/10.1098/rspb.2011.2342>
- Dubois, F., Morand-Ferron, J., & Giraldeau, L.-A. (2010). Learning in a game context: strategy choice by some keeps learning from evolving in others. *Proceedings of the Royal Society B: Biological Sciences*, rspb20100857.  
<http://doi.org/10.1098/rspb.2010.0857>
- Duckworth, R. A. (2008). The role of behavior in evolution: a search for mechanism. *Evolutionary Ecology*, 23(4), 513–531. <http://doi.org/10.1007/s10682-008-9252-6>
- Dugatkin, L. A., & Reeve, H. K. (2000). *Game Theory and Animal Behavior* (Reprint). Oxford University Press.
- Dunlap, A. S., & Stephens, D. W. (2009). Components of change in the evolution of learning and unlearned preference. *Proceedings of the Royal Society B: Biological Sciences*.  
<http://doi.org/10.1098/rspb.2009.0602>
- Estes, W. K. (1950). Toward a statistical theory of learning. *Psychological Review*, 57(2), 94.
- Fawcett, T. W., Fallenstein, B., Higginson, A. D., Houston, A. I., Mallpress, D. E. W., Trimmer, P. C., & McNamara, J. M. (2014). The evolution of decision rules in complex environments. *Trends in Cognitive Sciences*, 18(3), 153–161.  
<http://doi.org/10.1016/j.tics.2013.12.012>

- Fawcett, T. W., Hamblin, S., & Giraldeau, L.-A. (2013). Exposing the behavioral gambit: the evolution of learning and decision rules. *Behavioral Ecology*, 24(1), 2–11.  
<http://doi.org/10.1093/beheco/ars085>
- Field, A. P., Miles, J., & Field, Z. (2012). *Discovering statistics using R*. Los Angeles; London: SAGE.
- Flynn, R. E., & Giraldeau, L.-A. (2001). Producer–Scrounger Games in a Spatially Explicit World: Tactic Use Influences Flock Geometry of Spice Finches. *Ethology*, 107(3), 249–257. <http://doi.org/10.1046/j.1439-0310.2001.00657.x>
- Giraldeau, L.-A., & Beauchamp, G. (1999). Food exploitation: searching for the optimal joining policy. *Trends in Ecology & Evolution*, 14(3), 102–106.  
[http://doi.org/10.1016/S0169-5347\(98\)01542-0](http://doi.org/10.1016/S0169-5347(98)01542-0)
- Giraldeau, L.-A., & Caraco, T. (2000). *Social Foraging Theory*. Princeton University Press.
- Giraldeau, L.-A., & Dubois, F. (2008). Social Foraging and the Study of Exploitative Behavior (Vol. Volume 38, pp. 59–104). Academic Press.
- Giraldeau, L.-A., Hogan, J. A., & Clinchy, M. J. (1990). The Payoffs to Producing and Scrounging: What Happens when Patches are Divisible? *Ethology*, 85(2), 132–146.  
<http://doi.org/10.1111/j.1439-0310.1990.tb00393.x>
- Giraldeau, L.-A., & Livoreil, B. (2000). Game theory and social foraging. In *Game theory and animal behavior* (pp. 16–37). Oxford University Press.
- Giraldeau, L.-A., Soos, C., & Beauchamp, G. (1994). A test of the producer-scrounger foraging game in captive flocks of spice finches, *Loncbura punctulata*. *Behavioral Ecology and Sociobiology*, 34(4), 251–256. <http://doi.org/10.1007/BF00183475>

- Grafen, A. (1984). Natural selection, kin selection and group selection. In *Behavioural Ecology: An Evolutionary Approach* (2nd Edition, pp. 62–84). Oxford, UK: Blackwell Scientific Publications.
- Grodzinski, U., Dor, R., & Lotem, A. (2011). Begging for a better future: How far can behavioral ecologists go without specifying mechanisms? *Behavioral Ecology*, 22(5), 921–922.
- Gross, M. R. (1996). Alternative reproductive strategies and tactics: diversity within sexes. *Trends in Ecology & Evolution*, 11(2), 92–98. [http://doi.org/10.1016/0169-5347\(96\)81050-0](http://doi.org/10.1016/0169-5347(96)81050-0)
- Groß, R., Houston, A. I., Collins, E. J., McNamara, J. M., Dechaume-Moncharmont, F.-X., & Franks, N. R. (2008). Simple learning rules to cope with changing environments. *Journal of The Royal Society Interface*, 5(27), 1193–1202. <http://doi.org/10.1098/rsif.2007.1348>
- Hamblin, S., & Giraldeau, L.-A. (2009). Finding the evolutionarily stable learning rule for frequency-dependent foraging. *Animal Behaviour*, 78(6), 1343–1350. <http://doi.org/10.1016/j.anbehav.2009.09.001>
- Hamilton, I. M. (2002). Kleptoparasitism and the distribution of unequal competitors. *Behavioral Ecology*, 13(2), 260–267. <http://doi.org/10.1093/beheco/13.2.260>
- Harley, C. B. (1981). Learning the evolutionarily stable strategy. *Journal of Theoretical Biology*, 89(4), 611–633. [http://doi.org/10.1016/0022-5193\(81\)90032-1](http://doi.org/10.1016/0022-5193(81)90032-1)
- Harley, C. B. (1983). When do animals learn the evolutionarily stable strategy? *Journal of Theoretical Biology*, 105(1), 179–181.

- Harley, C. B. (1987). Learning rules, optimal behaviour, and evolutionary stability. *Journal of Theoretical Biology*, 127(3), 377–379. [http://doi.org/16/S0022-5193\(87\)80114-5](http://doi.org/16/S0022-5193(87)80114-5)
- Harley, C. B., & Maynard Smith, J. (1983). Learning -- an evolutionary approach. *Trends in Neurosciences*, 6, 204–208. [http://doi.org/16/0166-2236\(83\)90094-2](http://doi.org/16/0166-2236(83)90094-2)
- Ha, R. R., & Ha, J. C. (2003). Effects of ecology and prey characteristics on the use of alternative social foraging tactics in crows, *Corvus caurinus*. *Animal Behaviour*, 66(2), 309–316. <http://doi.org/10.1006/anbe.2003.2182>
- Henly, S. E., Ostdiek, A., Blackwell, E., Knutie, S., Dunlap, A. S., & Stephens, D. W. (2008). The discounting-by-interruptions hypothesis: model and experiment. *Behavioral Ecology*, 19(1), 154–162. <http://doi.org/10.1093/beheco/arm110>
- Herrnstein, R. J. (1961). Relative and absolute strength of response as a function of frequency of reinforcement. *Journal of the Experimental Analysis of Behavior*, 4(3), 267–272. <http://doi.org/10.1901/jeab.1961.4-267>
- Herrnstein, R. J. (1970). On the law of effect. *Journal of the Experimental Analysis of Behavior*, 13(2), 243–266.
- Herrnstein, R. J., & Heyman, G. M. (1979). Is matching compatible with reinforcement maximization on concurrent variable interval variable ratio? *Journal of the Experimental Analysis of Behavior*, 31(2), 209–223.
- Hines, W. G. S., & Bishop, D. T. (1983). On Learning and the Evolutionarily Stable Strategy. *Journal of Applied Probability*, 20(3), 689–695. <http://doi.org/10.2307/3213903>
- Hinton, G. E., & Nowlan, S. J. (1987, January 1). How Learning Can Guide Evolution [Journal (Paginated)]. Retrieved June 30, 2011, from <http://httpprints.yorku.ca/archive/00000172/>

- Houston, A. I. (1983). Comments on "Learning the evolutionarily stable strategy." *Journal of Theoretical Biology*, 105(1), 175–178.
- Houston, A. I., & McNamara, J. M. (2005). John Maynard Smith and the importance of consistency in evolutionary game theory. *Biology and Philosophy*, 20(5), 933–950.  
<http://doi.org/10.1007/s10539-005-9016-4>
- Houston, A. I., & Sumida, B. H. (1987). Learning rules, matching and frequency dependence. *Journal of Theoretical Biology*, 126(3), 289–308. [http://doi.org/10.1016/0022-5193\(87\)80236-9](http://doi.org/10.1016/0022-5193(87)80236-9)
- Ilan, T., Katsnelson, E., Motro, U., Feldman, M. W., & Lotem, A. (2013). The role of beginner's luck in learning to prefer risky patches by socially foraging house sparrows. *Behavioral Ecology*, 24(6), 1398–1406.  
<http://doi.org/10.1093/beheco/art079>
- Jolles, J. W., Ostojić, L., & Clayton, N. S. (2013). Dominance, pair bonds and boldness determine social-foraging tactics in rooks, *Corvus frugilegus*. *Animal Behaviour*, 85(6), 1261–1269. <http://doi.org/10.1016/j.anbehav.2013.03.013>
- Kacelnik, A., & Krebs, J. R. (1985). Learning to exploit patchily distributed food. In *Behavioural ecology: ecological implications of adaptive behaviour* (pp. 189–206). Oxford: Blackwell Scientific Publications.
- Kacelnik, A., Krebs, J. R., & Ens, B. (1987). Foraging in a changing environment: an experiment with starlings (*Sturnus vulgaris*). In *Quantitative analyses of behavior* (Vol. VI: Foraging, pp. 63–87). Lawrence Erlbaum and Associates.
- Kameda, T., & Nakanishi, D. (2002). Cost–benefit analysis of social/cultural learning in a nonstationary uncertain environment: An evolutionary simulation and an experiment

with human subjects. *Evolution and Human Behavior*, 23(5), 373–393.

[http://doi.org/10.1016/S1090-5138\(02\)00101-0](http://doi.org/10.1016/S1090-5138(02)00101-0)

- Katsnelson, E., Motro, U., Feldman, M. W., & Lotem, A. (2012). Evolution of learned strategy choice in a frequency-dependent game. *Proceedings of the Royal Society B: Biological Sciences*, 279(1731), 1176–1184. <http://doi.org/10.1098/rspb.2011.1734>
- Kawecki, T. J. (2009). Evolutionary ecology of learning: insights from fruit flies. *Population Ecology*, 52(1), 15–25. <http://doi.org/10.1007/s10144-009-0174-0>
- Killeen, P. R. (1984). Incentive theory III: Adaptive clocks. *Annals of the New York Academy of Sciences*, 423(1), 515–527.
- King, A. J., Isaac, N. J. B., & Cowlshaw, G. (2009). Ecological, social, and reproductive factors shape producer–scrounger dynamics in baboons. *Behavioral Ecology*, 20(5), 1039–1049. <http://doi.org/10.1093/beheco/arp095>
- Kolodny, O., Edelman, S., & Lotem, A. (2014). The evolution of continuous learning of the structure of the environment. *Journal of The Royal Society Interface*, 11(92), 20131091. <http://doi.org/10.1098/rsif.2013.1091>
- Koops, M. A., & Giraldeau, L.-A. (1996). Producer–scrounger foraging games in starlings: a test of rate-maximizing and risk-sensitive models. *Animal Behaviour*, 51(4), 773–783. <http://doi.org/10.1006/ambe.1996.0082>
- Krebs, J. R., & Davies, N. B. (1993). *An Introduction to Behavioural Ecology* (3rd ed.). Wiley-Blackwell.
- Krebs, J. R., Kacelnik, A., & Taylor, P. (1978). Test of optimal sampling by foraging great tits. *Nature*, 275(5675), 27–31.

- Kurvers, R. H. J. M., Hamblin, S., & Giraldeau, L.-A. (2012). The Effect of Exploration on the Use of Producer-Scrounger Tactics. *PLoS ONE*, 7(11), e49400.  
<http://doi.org/10.1371/journal.pone.0049400>
- Lea, S. E., & Dow, S. M. (1984). The integration of reinforcements over time. *Annals of the New York Academy of Sciences*, 423, 269–277.
- Lendvai, Á. Z., Barta, Z., & Liker, A. (2004). The effect of energy reserves on social foraging: hungry sparrows scrounge more. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, 271(1556), 2467–2472.  
<http://doi.org/10.1098/rspb.2004.2887>
- Lester, N. P. (1984). The “Feed:Feed” Decision: How Goldfish Solve the Patch Depletion Problem. *Behaviour*, 89(3/4), 175–199.
- Levey, D. J., Londoño, G. A., Ungvari-Martin, J., Hiersoux, M. R., Jankowski, J. E., Poulsen, J. R., ... Robinson, S. K. (2009). Urban mockingbirds quickly learn to identify individual humans. *Proceedings of the National Academy of Sciences*, 106(22), 8959.
- Lotem, A., & Biran-Yoeli, I. (2014). Evolution of learning and levels of selection: A lesson from avian parent–offspring communication. *Theoretical Population Biology*, 91, 58–74. <http://doi.org/10.1016/j.tpb.2013.10.003>
- Luniak, M. (2004). Synurbization—adaptation of animal wildlife to urban development. In *Proceedings of the 4th International Symposium on Urban Wildlife Conservation. Tuscon (AZ): University of Arizona* (pp. 50–55).

- Marchetti, C., & Drent, P. J. (2000). Individual differences in the use of social information in foraging by captive great tits. *Animal Behaviour*, 60(1), 131–140.  
<http://doi.org/10.1006/anbe.2000.1443>
- March, J. G. (1996). Learning to be risk averse. *Psychological Review*, 103(2), 309.
- Mathot, K. J., & Giraldeau, L.-A. (2008). Increasing vulnerability to predation increases preference for the scrounger foraging tactic. *Behavioral Ecology*, 19(1), 131–138.  
<http://doi.org/10.1093/beheco/arm114>
- Mathot, K. J., & Giraldeau, L.-A. (2010). Family-related differences in social foraging tactic use in the zebra finch (*Taeniopygia guttata*). *Behavioral Ecology and Sociobiology*, 64(11), 1805–1811. <http://doi.org/10.1007/s00265-010-0992-2>
- Mathot, K. J., Godde, S., Careau, V., Thomas, D. W., & Giraldeau, L.-A. (2009). Testing dynamic variance-sensitive foraging using individual differences in basal metabolic rates of zebra finches. *Oikos*, 118(4), 545–552. <http://doi.org/10.1111/j.1600-0706.2009.17357.x>
- Maynard Smith, J. (1982). *Evolution and the Theory of Games* (1st Edition). Cambridge University Press.
- Maynard Smith, J. (1988). Can a mixed strategy be stable in a finite population? *Journal of Theoretical Biology*, 130(2), 247–251. [http://doi.org/10.1016/S0022-5193\(88\)80100-0](http://doi.org/10.1016/S0022-5193(88)80100-0)
- McCormack, J. E., Jablonski, P. G., & Brown, J. L. (2007). Producer-Scrounger Roles and Joining Based on Dominance in a Free-Living Group of Mexican Jays (*Aphelocoma ultramarina*). *Behaviour*, 144(8), 967–982.
- McDonald, J. (2009). *Handbook of Biological Statistics*. Sparky House Publishing.

- McNamara, J. M., & Houston, A. I. (1985). Optimal foraging and learning. *Journal of Theoretical Biology*, 117(2), 231–249. [http://doi.org/10.1016/0022-5193\(85\)80219-8](http://doi.org/10.1016/0022-5193(85)80219-8)
- McNamara, J. M., & Houston, A. I. (1987). Memory and the efficient use of information. *Journal of Theoretical Biology*, 125(4), 385–395. [http://doi.org/10.1016/0022-5193\(87\)80209-6](http://doi.org/10.1016/0022-5193(87)80209-6)
- McNamara, J. M., & Houston, A. I. (1992). Risk-sensitive foraging: A review of the theory. *Bulletin of Mathematical Biology*, 54(2-3), 355–378. <http://doi.org/10.1007/BF02464838>
- McNamara, J. M., & Houston, A. I. (2009). Integrating function and mechanism. *Trends in Ecology & Evolution*, 24(12), 670–675. <http://doi.org/10.1016/j.tree.2009.05.011>
- Milinski, M. (1979). An Evolutionarily Stable Feeding Strategy in Sticklebacks<sup>1</sup>. *Zeitschrift Für Tierpsychologie*, 51(1), 36–40. <http://doi.org/10.1111/j.1439-0310.1979.tb00669.x>
- Milinski, M. (1984). Competitive resource sharing: An experimental test of a learning rule for ESSs. *Animal Behaviour*, 32(1), 233–242. [http://doi.org/10.1016/0003-3472\(84\)80342-5](http://doi.org/10.1016/0003-3472(84)80342-5)
- Mónus, F., & Barta, Z. (2008). The Effect of Within-Flock Spatial Position on the Use of Social Foraging Tactics in Free-Living Tree Sparrows. *Ethology*, 114(3), 215–222. <http://doi.org/10.1111/j.1439-0310.2007.01472.x>
- Morand-Ferron, J., & Giraldeau, L.-A. (2010). Learning behaviorally stable solutions to producer–scrounger games. *Behavioral Ecology*, 21(2), 343–348. <http://doi.org/10.1093/beheco/arp195>

- Morand-Ferron, J., Giraldeau, L.-A., & Lefebvre, L. (2007). Wild Carib grackles play a producer–scrounger game. *Behavioral Ecology*, 18(5), 916–921.  
<http://doi.org/10.1093/beheco/arm058>
- Morand-Ferron, J., Varennes, E., & Giraldeau, L.-A. (2011). Individual differences in plasticity and sampling when playing behavioural games. *Proceedings of the Royal Society B: Biological Sciences*, 278(1709), 1223–1230.  
<http://doi.org/10.1098/rspb.2010.1769>
- Morand-Ferron, J., Wu, G.-M., & Giraldeau, L.-A. (2011). Persistent individual differences in tactic use in a producer–scrounger game are group dependent. *Animal Behaviour*, 82(4), 811–816. <http://doi.org/10.1016/j.anbehav.2011.07.014>
- Mosteller, F. (1958). Stochastic Models for the Learning Process. *Proceedings of the American Philosophical Society*, 102(1), 53–59.
- Mottley, K., & Giraldeau, L.-A. (2000). Experimental evidence that group foragers can converge on predicted producer–scrounger equilibria. *Animal Behaviour*, 60(3), 341–350. <http://doi.org/10.1006/anbe.2000.1474>
- Myerson, J., & Miezin, F. M. (1980). The kinetics of choice: An operant systems analysis. *Psychological Review*, 87(2), 160.
- Nakagawa, S., & Schielzeth, H. (2010). Repeatability for Gaussian and non-Gaussian data: a practical guide for biologists. *Biological Reviews*, 85(4), 935–956.  
<http://doi.org/10.1111/j.1469-185X.2010.00141.x>
- Nakagawa, S., & Schielzeth, H. (2013). A general and simple method for obtaining R<sup>2</sup> from generalized linear mixed-effects models. *Methods in Ecology and Evolution*, 4(2), 133–142. <http://doi.org/10.1111/j.2041-210x.2012.00261.x>

- Niv, Y., Joel, D., Meilijson, I., & Ruppin, E. (2002). Evolution of reinforcement learning in foraging bees: A simple explanation for risk averse behavior. *Neurocomputing*, 44, 951–956.
- Nolfi, S., & Floreano, D. (1999). Learning and Evolution. *Autonomous Robots*, 7(1), 89–113.  
<http://doi.org/10.1023/A:1008973931182>
- Ollason, J. G. (1980). Learning to forage--optimally? *Theoretical Population Biology*, 18(1), 44–56.
- Orzack, S. H., & Hines, W. G. S. (2005). The Evolution of Strategy Variation: Will an ESS Evolve? *Evolution*, 59(6), 1183–1193.
- Plowright, C. M. S., & Plowright, R. C. (1987). Oversampling by great tits? A critique of Krebs, Kacelnik, and Taylor's (1978) "Test of optimal sampling by great tits." *Canadian Journal of Zoology*, 65(5), 1282–1283.
- Price, T. D., Qvarnstrom, A., & Irwin, D. E. (2003). The role of phenotypic plasticity in driving genetic evolution. *Proceedings of the Royal Society B: Biological Sciences*, 270(1523), 1433–1440. <http://doi.org/10.1098/rspb.2003.2372>
- Pulliam, H. R., & Dunford. (1980). The Cybernetic Lizard vs. The Toxic Ants. In *Programmed to Learn* (pp. 11–22). Columbia University Press.
- R Development Core Team. (2008). *R: A Language and Environment for Statistical Computing*. Retrieved from <http://www.R-project.org>
- Réale, D., & Dingemanse, N. J. (2012). Animal Personality. In John Wiley & Sons, Ltd (Ed.), *eLS*. Chichester, UK: John Wiley & Sons, Ltd. Retrieved from <http://www.els.net/WileyCDA/ElsArticle/refId-a0023570.html>

- Regelmann, K. (1984). Competitive resource sharing: A simulation model. *Animal Behaviour*, 32(1), 226–232. [http://doi.org/10.1016/0003-3472\(84\)80341-3](http://doi.org/10.1016/0003-3472(84)80341-3)
- Rendell, L., Boyd, R., Cownden, D., Enquist, M., Eriksson, K., Feldman, M. W., ... Laland, K. N. (2010). Why Copy Others? Insights from the Social Learning Strategies Tournament. *Science*, 328(5975), 208–213. <http://doi.org/10.1126/science.1184719>
- Rita, H., & Ranta, E. (1998). Group Foraging, Gain and the Finder's Advantage. *Oikos*, 82(3), 515–521. <http://doi.org/10.2307/3546372>
- Rita, H., Ranta, E., & Peuhkuri, N. (1997). Group foraging, patch exploitation time and the finder's advantage. *Behavioral Ecology and Sociobiology*, 40(1), 35–39. <http://doi.org/10.1007/s002650050313>
- Ruxton, G. D., & Beauchamp, G. (2008). The application of genetic algorithms in behavioural ecology, illustrated with a model of anti-predator vigilance. *Journal of Theoretical Biology*, 250(3), 435–448. <http://doi.org/10.1016/j.jtbi.2007.10.022>
- Schielzeth, H., & Nakagawa, S. (2013). Nested by design: model fitting and interpretation in a mixed model era. *Methods in Ecology and Evolution*, 4(1), 14–24. <http://doi.org/10.1111/j.2041-210x.2012.00251.x>
- Shettleworth, S. J. (2010). *Cognition, Evolution, and Behavior*. Oxford University Press US.
- Shimada, M., Ishii, Y., & Shibao, H. (2010). Rapid adaptation: a new dimension for evolutionary perspectives in ecology. *Population Ecology*, 52(1), 5–14. <http://doi.org/10.1007/s10144-009-0187-8>
- Staddon, J. E. R. (1977). On Herrnstein's equation and related forms. *Journal of the Experimental Analysis of Behavior*, 28(2), 163–170. <http://doi.org/10.1901/jeab.1977.28-163>

- Staddon, J. E. R., Hinson, J. M., & Kram, R. (1981). Optimal choice. *Journal of the Experimental Analysis of Behavior*, 35(3), 397–412.  
<http://doi.org/10.1901/jeab.1981.35-397>
- Stephens, D. W. (1989). Variance and the Value of Information. *The American Naturalist*, 134(1), 128–140.
- Stephens, D. W. (1991). Change, regularity, and value in the evolution of animal learning. *Behavioral Ecology*, 2(1), 77–89. <http://doi.org/10.1093/beheco/2.1.77>
- Stephens, D. W., & Clements, K. C. (1998). Game theory and learning. In *Game theory and animal behavior* (pp. 239–260). Oxford University Press, UK.
- Stephens, D. W., & Krebs, J. R. (1986). *Foraging theory*. Princeton, N.J.: Princeton University Press.
- Sutton, R. S., & Barto, A. G. (1998). *Reinforcement Learning: An Introduction*. The MIT Press.
- Suzuki, R., & Arita, T. (2004). Interactions between learning and evolution:: The outstanding strategy generated by the Baldwin effect. *Biosystems*, 77(1-3), 57–71.  
<http://doi.org/10.1016/j.biosystems.2004.04.002>
- Tebbich, S., Sterelny, K., & Teschke, I. (2010). The tale of the finch: adaptive radiation and behavioural flexibility. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 365(1543), 1099–1109. <http://doi.org/10.1098/rstb.2009.0291>
- Thuijsman, F., Peleg, B., Amitai, M., & Shmida, A. (1995). Automata, matching and foraging behavior of bees. *JOURNAL OF THEORETICAL BIOLOGY*, 175, 305–316.
- Tinbergen, N. (1963). On aims and methods of Ethology. *Zeitschrift Für Tierpsychologie*, 20(4), 410–433. <http://doi.org/10.1111/j.1439-0310.1963.tb01161.x>

- Tomkins, J., & Hazel, W. (2007). The status of the conditional evolutionarily stable strategy. *Trends in Ecology & Evolution*, 22(10), 522–528.  
<http://doi.org/10.1016/j.tree.2007.09.002>
- Tracy, N. D., & Seaman, J. (1995). Properties of Evolutionarily Stable Learning Rules. *Journal of Theoretical Biology*, 177(2), 193–198. <http://doi.org/10.1016/j.jtbi.1995.0238>
- Trimmer, P. C., McNamara, J. M., Houston, A. I., & Marshall, J. A. R. (2012). Does natural selection favour the Rescorla–Wagner rule? *Journal of Theoretical Biology*, 302, 39–52. <http://doi.org/10.1016/j.jtbi.2012.02.014>
- Vickery, W. L. (1987). How to cheat against a simple mixed strategy ESS. *Journal of Theoretical Biology*, 127(2), 133–139. [http://doi.org/10.1016/S0022-5193\(87\)80124-8](http://doi.org/10.1016/S0022-5193(87)80124-8)
- Vickery, W. L. (1988). Reply to Maynard Smith. *Journal of Theoretical Biology*, 132(3), 375–378. [http://doi.org/10.1016/S0022-5193\(88\)80222-4](http://doi.org/10.1016/S0022-5193(88)80222-4)
- Vickery, W. L., Giraldeau, L.-A., Templeton, J. J., Kramer, D. L., & Chapman, C. A. (1991). Producers, Scroungers, and Group Foraging. *The American Naturalist*, 137(6), 847–863.
- Wu, G.-M., & Giraldeau, L.-A. (2005). Risky decisions: a test of risk sensitivity in socially foraging flocks of *Lonchura punctulata*. *Behavioral Ecology*, 16(1), 8–14.  
<http://doi.org/10.1093/beheco/arh127>
- Zuur, A., Ieno, E. N., Walker, N., Saveliev, A. A., & Smith, G. M. (2009). *Mixed Effects Models and Extensions in Ecology with R*. Springer.