



This work is protected by copyright and other intellectual property rights and duplication or sale of all or part is not permitted, except that material may be duplicated by you for research, private study, criticism/review or educational purposes. Electronic or print copies are for your own personal, non-commercial use and shall not be passed to any other individual. No quotation may be published without proper acknowledgement. For any other use, or to quote extensively from the work, permission must be obtained from the copyright holder/s.

# The emergence and utility of social behaviour and social learning in artificial evolutionary systems

**James Martin Borg**

Submitted for the degree of

Doctor of Philosophy

March 2018

Keele University

# Abstract

The questions to be addressed here are all aimed at beginning to assess the emergence and utility of social behaviour and social learning in artificial evolutionary systems. Like any biological adaptation, the adaptation to process and use social information must lead to an overall increase in the long term reproductive capability of any population utilising such an adaptation - this increase in fecundity also being accompanied by increased survivability and therefore adaptability. In nature, social behaviours such as co-operation, teaching and agent aggregation, all seem to provide improved levels of fitness, resulting in an improved and more robust set of general behaviours - in the human case these social behaviours have led to cumulative culture and the ability to rapidly adapt to, and thrive in, an astonishing number of environments. In this thesis we begin to look at why the evolutionary adaptation to process and use social information, leading to social learning and social behaviour, proves to be such a useful adaptation, and under which circumstances we would expect to see this adaptation, and its resulting mechanisms and strategies, emerge.

We begin by asking these questions in two contexts; firstly what does social learning enable that incremental genetic evolution alone does not, and secondly what benefit does social learning provide in temporally variable environments. We go on to assess how differing social learning strategies affect the utility of social learning, and whether social information can be utilised by an evolutionary process without any accompanying within-lifetime learning processes (and whether the accommodation of social information results in any notable behavioural changes). By addressing the questions posed here in this way, we can begin to shed some light on the circumstances under which the adaptations for the accommodation and use of social information begin to emerge, and ultimately lead to the emergence of robust socially intelligent artificial agents.

## Acknowledgements

These acknowledgements have been more difficult to write than they appear to be, not just because they are the last thing I have tried to write after wrestling with an unwieldy thesis, but because it is so hard to fully express how deep my appreciation is to the people who have helped and supported me not only through this thesis but through my entire journey through education.

I suppose the best place to start is with my PhD supervisor Dr. Alastair Channon; not only is he the best dancer alive, but it also turns out that he was right about everything! Throughout my PhD Alastair has encouraged, guided and supported me, his advice has been consistently useful and his knowledge of artificial evolutionary systems is second to none. It has truly been a privilege to be supervised by him.

My unending thanks and gratitude also extend to the staff and my fellow PhD students in the School of Computing and Mathematics at Keele University. It hasn't been easy balancing teaching and a PhD, but the support I have received from my colleagues has been wonderful. Special thanks goes to Dr. Charles Day who was part of my thesis group in the early days of my PhD and has been a great source of advice and wisdom throughout my PhD studies, and Prof. Peter Andras who has arrived in my life just in time to help me make the most of the wonderful opportunities that await me as part of the wider Social Learning and Cultural Evolution community.

Adam Stanton, Nikki Williams, John Butcher and Liz Aston have been my brothers (and sisters) in arms over the course of my PhD studies. All have led the way for me to follow, with John also helping me along the way with a few shoves and kicks (more often than not physical rather than metaphorical)! Adam deserves more than a special mention here. I studied for my undergraduate degree with Adam, and now I have the honour and the privilege of graduating for a second time alongside him, now as a

Doctor of Philosophy. Adam has been more like a brother than a colleague over the last 7 years, and I am convinced that I could not have made it through this PhD without him.

I also owe a special mention to the lads from football: Neil Murphy, Steven De Cruz, John Butcher (again), Charles, Boyd, Ian, Dan, Mick, Richey, Thomas, and everyone else who has been kicked by me. Having a couple of hours a week to run off some steam in the company of some of the best people I have ever had the fortune of meeting was a god send. To my best mates from home; Bobolox, Liam and Peter, who didn't really do anything other than just be there for a drink and laugh - but that was more than enough.

Last but not least I owe special thanks to my amazing family, not just for supporting me through my education, but also for being a never ending source of love and encouragement. To Mom and Dad, Louise and Martin - ever since I was little you never questioned why I was interested in the things I was interested in, you were just happy to encourage me and let me enjoy asking questions about the world around me. That kind of unquestioning love and support has enabled me to achieve things I would never have thought of as being possible. I hope I have made you both proud. To my beautiful sisters Charlotte and Danielle - I am so proud of the amazing women you have both become, and I am forever thankful for the love and support you give me. Nanny Marg, I am sure I get these brains from you really! Grandad Andrew, I have never met a more resilient and determined man, your strength has been an inspiration to me. To Nana Doris, you are sorely missed every day - I hope I have made you proud. To Grandad Graham, gone but never forgotten, I am sorry I didn't follow you into the gas box business! To my multitude of aunties, uncles and cousins, having such a big family has been a blessing to me and I couldn't have asked to be related to more

wonderful group of people.

Finally, to my beautiful, intelligent, and brave partner Stephanie - I am so sorry that I not only put you through my undergraduate dissertation and my master thesis, but that I also dragged you all the way to Keele to have you suffer through over 7 years of a PhD! I know you have lived every achievement with me and felt my every disappointment. I am so thankful that I have such an amazing person there to support me day in, day out. You are my rock. I could not have done this without you. I cannot express in writing how much I love and appreciate you.

I would also like to extend my thanks to my examiners, Prof. Seth Bullock and Dr. John Bullinaria, who were fair, constructive and supportive in their comments and criticisms.

# Contents

<b>1</b>	<b>Introduction</b>	<b>1</b>
1.1	Aims and Organisation of Thesis . . . . .	4
1.2	Published Work and Attributions . . . . .	7
<b>2</b>	<b>Social Behaviour and Social Learning</b>	<b>15</b>
2.1	Social Behaviour and Social Learning in Nature . . . . .	15
2.1.1	Theories of Cultural Evolution and a Working Definition of Social Behaviour . . . . .	16
2.1.2	Social Information and Social Learning: Mechanisms and Strategies	21
2.2	Social Behaviour and Social Learning in Artificial Evolutionary Systems	25
<b>3</b>	<b>Research Questions</b>	<b>30</b>
3.0.1	Incremental Genetic Evolution and Social Learning . . . . .	30
3.0.2	Robustness in (Increasingly) Variable Environments . . . . .	31
3.0.3	Social Learning Strategies . . . . .	32
3.0.4	Social Information Use without Learning . . . . .	32
3.0.5	Social Information Use and Behavioural Change . . . . .	33
<b>4</b>	<b>Discovering and Maintaining Behaviours Inaccessible to Incremental Genetic Evolution</b>	<b>35</b>

4.1	Introduction . . . . .	35
4.1.1	Incremental Genetic Evolution . . . . .	36
4.1.2	Discovering and Maintaining Inaccessible Solutions: Transcription Errors and Imitation . . . . .	39
4.1.3	Neuroevolution of Deliberative Behaviours . . . . .	40
4.1.4	The RC+ Task . . . . .	41
4.2	The Model . . . . .	42
4.2.1	The Shunting Network . . . . .	42
4.2.2	The Decision Network . . . . .	44
4.2.3	Evolution of the Decision Network . . . . .	46
4.2.4	Learning in the Decision Network . . . . .	48
4.3	Experimentation . . . . .	49
4.4	Results . . . . .	51
4.5	Conclusions and Future Work . . . . .	55
<b>5</b>	<b>The Adoption of Social Learning in Increasingly Variable Environments</b>	<b>58</b>
5.1	Introduction . . . . .	58
5.1.1	The Variability Selection Hypothesis . . . . .	59
5.1.2	Social Learning . . . . .	60
5.1.3	Social Learning in Increasingly Variable Environments . . . . .	60
5.2	The Model . . . . .	62
5.2.1	Variable Environments . . . . .	63
5.2.2	Evolution and Learning . . . . .	64
5.3	Experimentation and Results . . . . .	66
5.3.1	Static Environments . . . . .	67



5.3.2	Consistently Variable Environments . . . . .	69
5.3.3	Environments of Increasing Variability . . . . .	71
5.3.4	Population Collapse in Variable Environments (Consistent and Increasing) . . . . .	73
5.4	Conclusions and Future Work . . . . .	77
<b>6</b>	<b>Social Learning Strategies</b>	<b>79</b>
6.1	Introduction . . . . .	79
6.1.1	Discovering and Maintaining Behaviours Inaccessible to Incre- mental Genetic Evolution Alone . . . . .	80
6.2	Experimentation . . . . .	83
6.3	Results . . . . .	85
6.4	Conclusions and Further Work . . . . .	90
<b>7</b>	<b>Evolutionary Adaptation to Social Information Use without Learning</b>	<b>92</b>
7.1	Environmental Set-up . . . . .	95
7.2	Neuroevolutionary Model . . . . .	96
7.2.1	The Shunting Network . . . . .	97
7.2.2	The Decision Network, Neuroevolution and Reproduction . . . .	100
7.2.3	Agent Actions and Action Energy Costs . . . . .	103
7.2.4	Social Information Strategies . . . . .	104
7.3	Results and Discussion . . . . .	106
7.3.1	Eat Action Performance . . . . .	107
7.3.2	Social Information Performance in Less Difficult Environments .	109
7.4	Conclusions and Further Work . . . . .	113

<b>8</b>	<b>The Effect of Social Information Use without Learning on the Evolution of Behaviour</b>	<b>115</b>
8.1	Social Information and Social Behaviour . . . . .	116
8.2	Experimentation and Results . . . . .	122
8.2.1	Action Profiles . . . . .	122
8.2.2	Reasons for Moving . . . . .	131
8.2.3	Behavioural Transitions . . . . .	136
8.3	Discussion and Conclusion . . . . .	142
<b>9</b>	<b>Conclusions</b>	<b>144</b>
9.1	Summary of Conclusions . . . . .	144
9.2	Contribution of this work . . . . .	147
9.3	Future Work . . . . .	149
	<b>Bibliography</b>	<b>151</b>

# List of Figures

2.1	Whiten and van Schaik’s cultural pyramid . . . . .	20
4.1	Mock fitness landscape with an inaccessible fitness peak . . . . .	38
4.2	RC+ task and activity landscapes . . . . .	47
4.3	RC+ task mean fitness in the best performing populations . . . . .	55
5.1	Fitness and reproduction rates in static environments . . . . .	68
5.2	Fitness and reproduction rates in consistently variable environments. (a) Mean fitness of each learning strategy, (b) Mean fitness of individual and social learning with the evolved learning rates, (c) Reproductive fitness of combined learning rates. . . . .	70
5.3	Fitness and reproduction rates in increasingly variable environments . .	72
5.4	Maximum learning rates . . . . .	74
6.1	Time taken for RC+ map completion for each social learning strategy .	90
7.1	Mock-up transition from agent visual field to shunting network activity landscape . . . . .	98
7.2	Difference between successful and unsuccessful eat actions . . . . .	107
7.3	Environment 1 box plots . . . . .	111
7.4	Environment 2 box plots . . . . .	112

8.1	Median agent action profiles for each social information strategy over all environments . . . . .	123
8.2	Median eat actions for each social information strategy over all environments . . . . .	126
8.3	Median move actions for each social information strategy over all environments . . . . .	127
8.4	Median wait actions for each social information strategy over all environments . . . . .	129
8.5	Environment 1 action box-plots . . . . .	130
8.6	Median move actions plotted against median agent views . . . . .	133
8.7	Median agent views for each social information strategy over all environments . . . . .	134
8.8	Median move actions for each social information strategy (barring non social) over all environments . . . . .	135
8.9	Transitions in movement behaviour in relation to eating action performance . . . . .	139
8.10	Transitions in eating behaviour in relation to eating action performance	140
8.11	Transitions in waiting behaviour in relation to eating action performance	141

# List of Tables

4.1	Non-Learners(1) strategy RC+ performance . . . . .	53
4.2	Non-Learners(2) strategy RC+ performance . . . . .	53
4.3	Learners strategy RC+ performance . . . . .	53
5.1	Population survival in consistently variable environments . . . . .	75
5.2	Population survival in increasingly variable environments . . . . .	75
5.3	Population survival in variable environments when individual learning is pre-evolved . . . . .	77
6.1	RC+ map completion for each social learning strategy . . . . .	86

# Chapter 1

## Introduction

This thesis is primarily concerned with the emergence and utility of social behaviour and social learning in artificial evolutionary system, including the evolutionary adaptation to the use of social information. The objective of this thesis is not just to contribute to our understanding of the emergence and utility and social behaviour and social learning in real biological systems, but also to provide a better understanding of the dynamics of emergent social behaviour and social learning, from which we may more effectively engineer social artificial systems.

In “On The Origin of Species” Charles Darwin [48] mused over the many adaptations and variations found in nature. Darwin concluded that for any adaptation to be maintained, greater fitness must be conferred by its existence resulting in an adaptive advantage to any individual exhibiting it. While it is true that some startling physiological variation first inspired Darwin, behavioural adaptations are just as apparent and varied in nature. Like physiological adaptation, the adaptation for social and cultural behaviour must confer some survival and reproductive advantage. It is the adaptive advantage of social behaviour via social learning and social information use that is to be investigated in this thesis.

Social learning can be defined as learning that is influenced by the interactions with, or observation of, other agents or the products of agent interactions with the environment [24, 60, 80]. Social learning primarily involves learning from conspecifics, though inter species social interactions are also being recognised as an increasingly important aspect of animal social learning [68]. Social learning is in contrast to individual or asocial learning which instead concerns itself with plasticity in agent private information as result of non-social interactions with an environment, and population level evolutionary change which is a process of long term adaptive change on a population level driven primarily by environmental change.

Social learning therefore relies on social information, which can be broadly defined as information derived from the behaviours, actions, cues or signals of other agents [91]. As social information necessarily involves the direct or indirect broadcasting of information in to the public domain, it is sometimes known as (or conflated with) public information [20]; public information generally being defined as inadvertently expressed social information about the performance or state of the information producer, or the quality of the environment [46].

From these definitions of social learning and social information, social behaviour may therefore be defined as any behaviour exhibited by an agent which results in an interaction with another agent, an interaction with the product of agent activity, or the expression or communication of private information.

In 1963 the Ethologist Nikko Tinbergen famously derived four questions one must ask when engaging in the “biological study of behaviour” [146]:

1. What causes a behaviour to be exhibited? (Causation)
2. What advantage, if any, does a behaviour confer on the exhibiting individual?  
(Survival Value)

3. What ‘cognitive machinery’ is necessary for a behaviour to be demonstrated?  
(Ontogeny)
4. How may a behaviour have evolved over time and over related species? (Evolution)

Tinbergen’s rationale for formalising these questions was firstly to unify the behavioural sciences (a goal that is still to be realised [108, 65]) and more importantly, to provide a framework for the behavioural sciences to work under.

Here Tinbergen’s four questions are reduced to two; the first primarily concerning the survival value of the adaptation for social behaviour and social learning. The second combining aspects of causation, ontogeny and evolution to investigate when social behaviour becomes adaptive and is therefore maintained as an evolutionary adaptation in artificial evolutionary systems. To investigate these two broad areas this work uses a number of Artificial Life [94] and simulated Evolutionary Robotics [156, 51] models of evolutionary systems. The artificial evolutionary systems used here are primarily grounded; that is to say simulated autonomous agents (Agents, Animats or Simulated Autonomous Robots) will be physically situated in a virtual environment. Agents in these grounded models will also be situated socially [96] and be controlled by artificial neural networks which will be subject to both evolution (by means of evolutionary algorithms [83]), and in some cases extra-genetic, within-lifetime learning (both individual and social). A non-grounded approach, making use of populations of binary strings evolved by means of evolutionary algorithms [83], but never being physically situated in a virtual environment, will also be applied here but less frequently than grounded models.



## 1.1 Aims and Organisation of Thesis

The purpose of this work is to use a variety of artificial evolutionary systems to address a series of fundamental questions regarding the emergence and utility of social behaviour and social learning. These aims, and the research questions they address (discussed in Chapter 3), are all motivated by observations from natural social systems. Specifically the work presented in this thesis aims to:

- Investigate whether the utilisation of social learning enables the access to behaviours that are inaccessible to incremental genetic evolution alone.
- Investigate the emergence of social learning in temporally variable environments.
- Explore the role different social learning strategies play in the adaptive value of social learning.
- Investigate whether social information is of an adaptive benefit when decoupled from within-lifetime learning.
- Explore the behavioural consequences of social information when decoupled from within-lifetime learning.

These aims have been achieved by:

- The extension of a simulation model known as the Rivercrossing (or RC) task (introduced by Robinson et al. [128] and expanded here in Chapter 4) which utilises an artificial evolutionary system comprised of grounded populations of neuroevolutionary artificial agents. This model will be extended to explore the utilisation of social learning when accessing behaviours which are inaccessible to incremental genetic evolution alone, and to explore the role different social learning strategies play in the adaptive value of social learning.

- The application of a non-grounded binary string evolutionary algorithm to investigate the emergence of social learning in temporally variable environments, whereby individual instantiations of binary strings with access to both individual and/or social learning are evaluated against their ability to match a temporally changing binary search string. The nature of temporal variability is adjusted to enable the testing of both consistently and increasingly variable environments.
- The creation of a simulation model utilising an artificial evolutionary system comprised of grounded populations of neuroevolutionary artificial agents. This simulation model will incorporate a simple food foraging task whereby populations of agents are tasked with maintaining above zero energy levels. Agents in this model will have no access to within-lifetime learning processes, and will therefore rely on the evolutionary process alone; populations of agents in this model will have access to a variety of different types of social information inadvertently expressed by conspecifics. The aims of this simulation model are to investigate whether social information alone is of an adaptive benefit when decoupled from within-lifetime learning and to explore the behavioural consequences of social information when decoupled from within-lifetime learning.
- All models utilised here will simulate non social populations alongside social populations in order to provide a baseline from which the utility of social behaviours may be gauged. A variety of statistical techniques including T-tests, Chi-squared tests, and Mann-Whitney U tests will be employed to ascertain the significance of the resulting differences between social and non social populations.

This thesis is organised into nine Chapters. The organisation of these nine Chapters is as follows. Chapters 4 to 7 report on previously published work - details of these published works, and their attributions may be found in Chapter 1.2.

- Chapter 1 (this Chapter) provides an introduction to this thesis by way of providing a brief overview of aims and objectives, along with the organisation of this thesis and details of the published work which contributes to this thesis.
- Chapter 2 provides an introductory overview to the topics of social behaviour and social learning, which provide the basis for the research questions explored in this thesis. This Chapter looks at social behaviour and social learning in nature and goes on to discuss how social behaviour and social learning has been explored in other artificial evolutionary systems.
- Chapter 3 provides a more detailed discussion of the research questions to be addressed by this thesis.
- Chapter 4 is the first Chapter that is directly concerned with the research questions and aims of this thesis. This Chapter discusses a task, known as the RC+ task, which is shown to be impossible to solve by incremental genetic evolution alone. The question of whether the inclusion of social learning enables access to the behaviours required is addressed, as per the research question addressed in Chapter 3.0.1.
- Chapter 5 discusses the role environmental variability plays in the adoption and utility of social learning, both when accompanied by individual learning and when utilised as the lone form of learning, as per the research question addressed in Chapter 3.0.2.
- Chapter 6 takes the task used in Chapter 4 and explores the utility of a variety of different social learning strategies, with the aim of exploring whether the ability of social learning to overcome the shortcomings of incremental genetic evolu-

tion alone extends to a variety of social learning strategies, as per the research questions addressed in Chapter 3.0.3.

- Chapter 7 looks at the evolutionary adaptation to social information use when not accompanied by within-lifetime learning. This Chapter explores the foundational benefits of social and public information use as per the research question addressed in Chapter 3.0.4.
- Chapter 8 explores the behavioural differences between the non social and social populations evaluated in Chapter 7. This Chapter addresses the question of emergent social behaviours as a result of social information use, as per the research question addressed in Chapter 3.0.5.
- Chapter 9 concludes this work by addressing the contributions made in this thesis. This Chapter goes on to reflect on the conclusions to each of the research questions addressed in this thesis, and proposes a number of possible avenues for future research.

## 1.2 Published Work and Attributions

A significant portion of the work presented in this thesis has already been published or is to be published in the near future. All published work has been through a thorough peer-review process. Each piece of published work is presented here as a Chapter, with Chapters being organised by date of publication. In order to accommodate for crossover between published works in regard to introductory materials and research methods, some of the content found in the published works has been re-organised. As some of the published works underpinning this thesis were published a number of years ago, a small amount of updating regarding the referenced literature has also

been undertaken. Despite these minor changes, the Chapters representing previously published work are still largely unchanged from the original peer reviewed publications. These previously published works, and the division of labour between co-authors, are briefly discussed below.

- Borg et al. [22]: “Discovering and maintaining behaviours inaccessible to incremental genetic evolution through transcription errors and cultural transmission”. Proceedings of the European Conference on Artificial Life 2011, MIT Press, 2011, 102-109

The work presented in Borg et al. [22] was presented as a poster at, and published as part of the proceedings for, the 2011 European Conference on Artificial Life, held in Paris, France. The work presented in Borg et al. [22] appears here in Chapter 4.

In Borg et al. [22] the question of whether the introduction of both transcription errors and cultural transmission, in the form of learning by imitation, can enable the evolution of behaviours inaccessible to incremental genetic evolution alone is assessed. To answer this question a neural network model using a hybrid of two different networks was implemented: one capable of demonstrating reactive qualities, the other controlling deliberative goal selecting behaviours. Animats using this model were evolved in an adaptation of the environment proposed by Robinson et al. [128] to solve increasingly difficult tasks. Simulations were run on populations with and without learning by imitation to assess the relative success of each strategy, leading to the conclusion that populations with learning by imitation can successfully demonstrate the most complex behaviour available to them, which was empirically found to be inaccessible to non-learning populations.

The work presented in Borg et al. [22] was co-authored by the author of this thesis along with Dr. Alastair Channon in his role as PhD supervisor, and Dr. Charles Day in his role as part of the original PhD supervisory team. The work was first authored by the author of this thesis. As a co-author of Robinson et al. [128], Dr. Alastair Channon provided the original C++ code for the simulation model used in Robinson et al. [128]. Coding to update this model for use in Borg et al. [22], including changes made to the structure of the simulation model (resulting in the RC+ task - discussed further in Chapter 4) plus the addition of transcription errors, cultural transmission and learning, was undertaken by the author of this thesis. The author of this thesis was also responsible for producing both the published work and resulting conference poster; both Dr. Channon and Dr. Day provided advice throughout, including proof-reading and commenting on a series of paper drafts, and contributing to the analysis of results.

- Borg and Channon [23]: “Testing the variability selection hypothesis - The adoption of social learning in increasingly variable environments”. ALIFE XIII: The Thirteenth International Conference on the Synthesis and Simulation of Living Systems, MIT Press, 2012, 317-324

The work presented in Borg and Channon [23] was selected for an oral presentation at, and published as part of the proceedings for, Artificial Life XIII: The Thirteenth International Conference on the Synthesis and Simulation of Living Systems, held in Michigan, USA. The work presented in Borg and Channon [23] appears here in Chapter 5.

The work published in Borg and Channon [23] primarily concerns itself with the variability selection hypothesis [119, 120, 121] which predicts the

adoption of versatile behaviours and survival strategies in response to increasingly variable environments. In hominin evolution the most apparent adaptation for versatility is the adoption of social learning. The hypothesis that social learning will be adopted over other learning strategies, such as individual learning, when individuals are faced with increasingly variable environments is tested here using a genetic algorithm with steady state selection and constant population size. Individuals, constituted of binary string genotypes and phenotypes, are evaluated on their ability to match a target binary string, nominally known as the environment, with success being measured by the Hamming distance between the phenotype and environment. The state of any given locus in the environment is determined by a sine wave, the frequency of which increases as the simulation progresses thus providing increasing environmental variability. Populations exhibiting combinations of genetic evolution, individual learning and social learning are tested, with the learning rates of both individual and social learning allowed to evolve. We show that increasingly variable environments are sufficient but not necessary to provide an evolutionary advantage to those populations exhibiting the extra-genetic learning strategies, with social learning being favoured over individual learning when populations are allowed to explore both strategies simultaneously. We also introduce a more biologically realistic model that allows for population collapse, and show that here the prior adoption of individual learning is a prerequisite for the successful adoption of social learning in increasingly variable environments.

The work presented in Borg and Channon [23] was co-authored by the author of this thesis and Dr. Alastair Channon in his role as PhD supervisor.

The author of this thesis wrote the simulation model used in Borg and Channon [23], analysed the results, and wrote the paper. Dr. Channon’s involvement in this work was primarily advisory, including the proof reading of paper drafts and advising on data analysis and simulation model design.

- Jolley et al. [87]: “Analysis of social learning strategies when discovering and maintaining behaviours inaccessible to incremental genetic evolution”: International Conference on Simulation of Adaptive behaviour, 2016, 293-304 (also presented at SLaCE at ALIFE XV)

The work presented in Jolley et al. [87] was selected for an oral presentation at, and published as part of the proceedings for, the 2016 International Conference on Simulation of Adaptive behaviour, held in Aberystwyth, UK. Aspects of this work were also accepted for presentation at the Social Learning and Cultural Evolution workshop at the Artificial Life XV conference in Cancun, Mexico. The author of this thesis was a co-organiser of the Social Learning and Cultural Evolution workshop at Artificial Life XV. The work presented in Jolley et al. [87] appears here in Chapter 6.

It had been previously demonstrated in Borg et al. [22] that social learning can enable agents to discover and maintain behaviours that are inaccessible to incremental genetic evolution alone. However, previous models investigating the ability of social learning to provide access to these inaccessible behaviours are often limited. Here we investigate teacher-learner social learning strategies. It is often the case that teachers in teacher-learner social learning models are restricted to one type of agent, be it a parent or some fit individual; here we broaden this exploration to include a variety of teachers to investigate whether these social learning strategies are also



able to demonstrate access to, and maintenance of, behaviours inaccessible to incremental genetic evolution. In this work new agents learn from either a parent, the fittest individual, the oldest individual, a random individual or another young agent. Agents are tasked with solving a river crossing task, with new agents learning from a teacher in mock evaluations. The behaviour necessary to successfully complete the most difficult version of the task has been shown to be inaccessible to incremental genetic evolution alone, but achievable using a combination of social learning and noise in the Genotype-Phenotype map. Here we show that this result is robust in all of the teacher-learner social learning strategies explored.

The work presented in Jolley et al. [87] was co-authored by Ben Jolley and Dr. Alastair Channon along with the author of this thesis, with Mr. Jolley being recognised as first author. As this work extended the work published by Borg et al. [22], the author of this thesis's role was initially to advise Mr. Jolley on the adaptation of the model used in Borg et al. [22] to include a larger variety of social learning strategies, with the initial research idea (to extend the Borg et al. [22] to incorporate some of the social learning strategies presented in Laland [93]) being that of the author of this thesis. As the work in Jolley et al. [87] progressed, the role of the author of this thesis increased, with a significant portion of the data analysis being conducted by the author of this thesis. The author of this thesis also undertook the vast majority of the work in writing and editing the publication associated with this work. Mr. Jolley's role as first author primarily comprised of coding the amended simulation model and producing and presenting data. Dr. Channon's involvement in this work was primarily advisory, including the

proof reading of paper drafts and advising on data analysis and simulation model design. The author of this thesis provided the basis for this research to be conducted, coded the original model which was amended here, led on the analysis of data, and wrote the majority of the publication associated with this work.

- Borg and Channon [21]: “Evolutionary adaptation to social information use without learning” European Conference on the Applications of Evolutionary Computation, Springer, 2017, 837-852

The work presented in Borg and Channon [21] was selected for presentation at, and publication as part of the proceedings for, the 2017 European Conference on the Applications of Evolutionary Computation, held in Amsterdam, Netherlands. This work was part of the conference’s Evolutionary Robots track. The work presented in Borg and Channon [21] appears here in Chapter 7

Social information can provide information about the presence, state and intentions of other agents; therefore it follows that the use of social information may be of some adaptive benefit. As with all information, social information must be interpretable and relatively accurate given the situation in which it is derived. In both nature and robotics, agents learn which social information is relevant and under which circumstances it may be relied upon to provide useful information about the current environmental state. However, it is not clear to what extent social information alone is beneficial when decoupled from a within-lifetime learning process, leaving evolution to determine whether social information provides any long term adaptive benefits. In this work we assess this question of the adaptive value of social

information when it is not accompanied by a within-lifetime learning process. The aim here is to begin to understand when social information, here expressed as a form of public information, is adaptive; the rationale being that any social information that is adaptive without learning will be a good base to allow the learning processes associated with social information to evolve and develop later. Here we show, using grounded neuroevolutionary artificial life simulations incorporating simulated agents, that social information can in certain circumstances provide an adaptive advantage to agents, and that social information that more accurately indicates success confers more reliable information to agents leading to improved success over less reliable sources of social information.

The work presented in Borg and Channon [21] was co-authored by the author of this thesis and Dr. Alastair Channon in his role as PhD supervisor. The author of this thesis wrote the simulation model used in Borg and Channon [21], analysed the results, and wrote the paper. Dr. Channon's involvement in this work was primarily advisory, including the proof reading of paper drafts and advising on data analysis and simulation model design.

- The Effect of Social Information Use without Learning on the Evolution of Behaviour - to be submitted to either the MIT Press Artificial Life journal, or a special issue on Social Learning and Cultural Evolution in the Cognitive Systems Research journal, alongside the work published in Borg and Channon [21]. As with Borg and Channon [21], the anticipated publication will be co-authored by the author of this thesis and Dr. Alastair Channon in his role as PhD supervisor. Dr. Channon's involvement in this work was advisory, including the proof reading of drafts and advising on data analysis and simulation model design.

# Chapter 2

## Social Behaviour and Social Learning

### 2.1 Social Behaviour and Social Learning in Nature

Social and cultural behaviours<sup>1</sup> have long fascinated scholars from fields ranging from Biology and Ethology to Anthropology, Sociology, Economics and Artificial Intelligence. Unlike non-social behaviour which may often be understood from a purely ecological and ontological view (that is to say that non-social behaviour is often coupled strongly with both environmental and biological factors and limitations) social and cultural behaviours are far less easy to study. In pursuit of the theoretical grounding of social and cultural evolution many theories of cultural evolution have arisen, many definitions of culture have been proposed, and social transmission and its benefits have been modelled mathematically and computationally.

---

<sup>1</sup>The terms social and cultural are often, but incorrectly, conflated to mean the same thing. Social behaviour simply requires a social interaction between two organisms. Cultural behaviour instead requires the social interaction to involve two organisms capable of exhibiting culture, with culture being defined by Whiten and Van Schaik as "the existence in the same species of multiple traditions forming unique local complexes"[164].

### 2.1.1 Theories of Cultural Evolution and a Working Definition of Social Behaviour

One of the first major theories to address the evolutionary basis for social and cultural behaviour was Sociobiology [165, 99]. Sociobiology offers what Boyd and Richerson [26] have termed the “argument from natural design”, placing the core driving force behind the evolution of social behaviour and culture on biological natural selection. In Rogers’ opinion [129] the argument from natural design implies that there are strong biological constraints on cultural evolution and limited scope for cultural transmission, that is to say that all aspects of culture may be explained as an extension of biological evolution thus precluding the possibility of two competing yet complimentary evolutionary systems (biological evolution and cultural evolution). The acultural nature of Sociobiology has led it to be widely rejected as a feasible theory of cultural evolution, with its scope instead limited to those non-social and social behavioural adaptations which can be seen to be primarily driven by ecological requirements. Sociobiology’s inability to fully accommodate cultural evolution in species such as humans led to a number of related theories of cultural evolution, the most prominent of these being Evolutionary Psychology [15]. Like Sociobiology, Evolutionary Psychology argues for strong biological constraints on the evolution of culture. Evolutionary Psychology argues that human psychological adaptations have evolved to solve recurrent problems faced by humans during our evolutionary history, the mind being broken down into functional modules akin to functional physiological adaptations. Evolutionary Psychology has been highly criticised for lacking the scope to incorporate other social animals due its focus on the evolutionary basis cultural behaviour in human, and making oversimplified assumptions about the nature of human behaviour (often described as ‘just so stories’) [44].

As opposed to Sociobiology and Evolutionary Psychology, which argue for strong biological constraints, there are a number of theories endorsing a stance of weaker biological constraints; the most notable of these theories being Memetics [48, 50, 58, 59, 17, 18, 81] and Dual Inheritance Theory (aka Gene/Culture co-evolution)[25, 26, 27, 78, 79, 75, 106]. Memetics as a theory came about through Richard Dawkins' musings on the universal nature of Darwinian evolution [47, 48]. Dawkins postulated that human culture could be viewed as an evolutionary system separate yet analogous to biological evolution. He suggested that culture contained all the necessary elements expected of an evolutionary system; discrete units of replication (dubbed memes), variation/mutation, competition between units, and methods of replication (cultural transmission mechanisms). As a theory of cultural evolution, Memetics has come under a huge amount of criticism (see Rose [130] for an early, but thorough overview of many of these criticisms). The weakness of the biological constraints in the Memetic view on culture, where the biological carrier of memes is perceived to be largely held hostage by the adaptive needs of the memes, is seen by many to entirely over-estimate the strength culture has to affect the biological vehicle, however the primary criticism is in regard to discrete replication. Though some simulation work has been done to demonstrate the feasibility of discrete cultural replication [58, 81], convincing empirical evidence is lacking [75]. The emergent social phenomenon often seen in artificial evolutionary systems making use of social agents controlled by neural networks or similar connectionist architectures [49, 42, 39, 5, 4, 3, 45, 2, 104, 154], also calls into question the need for discrete representation when considering social and cultural behaviour; behaviours emerging as a result of the evolution of a neural controller cannot be reduced down to a set of discrete units within the neural substrate. However, the use of discrete representations of culture or behaviour can be of use from a modelling perspective and

are therefore still widely used in artificial life models [81, 36, 37, 101, 103, 102].

As an alternative to the weak biological constraints and discrete replication endorsed by Memetics, the over-bearing biological constraints endorsed by Sociobiology and the limited scope for social and cultural learning, which results from Evolutionary Psychology, a theory of Gene/Culture co-evolution theory; more commonly known as Dual Inheritance Theory, was proposed by Robert Boyd and Peter J Richerson [26]. Dual Inheritance Theory endorses the view that two evolutionary systems; cultural and genetic, co-evolve resulting in both cultural (extra-genetic) and genetic inheritance. Replication of culture in Dual Inheritance Theory is proposed to be continuous; that is to say there is no necessity for discrete replication [75]. The Dual Inheritance Theory can be easily modelled by allowing genotypic and social information to evolve side by side, with both sources of information being expressed as part of an individual's phenotype and therefore having some affect on reproductive selection and long term survivability [103, 102].

In this thesis no single theory is binding, though Dual Inheritance Theory provides the most useful framework from a modelling perspective as it does not limit representation to discrete units and it does not limit the causation and survival value of social learning to purely genetic or environmental factors (whilst not excluding the importance of either). It is able to incorporate cumulative cultural evolution as well as the many potential evolutionary biases for the evolution of social adaptation.

To move forward with a robust theoretical grounding of cultural evolution and the adaptation for social behaviour, a working definition of culture must be formalised. Social behaviour has been observed in a large variety of wild animals [164, 123, 163], from mammals [92, 161, 124, 105, 100, 143, 144, 166] to birds [56, 115, 142] or even fish

[52, 151, 90, 12, 97, 13, 159]<sup>2</sup>. The problem with building a definition of culture is where to draw the line on what species may be deemed as cultural rather than merely social; on this problem there are two prevailing views. The first view is that culture should be reserved for traditions, taken in this case to be consistent behaviours maintained over a number of generations, transmitted via some mechanism of social information transfer; e.g., imitation or teaching [61]. The second view limits culture to behaviours that accumulate in complexity over time, thus restricting culture to behaviours subject to cumulative cultural evolution [95]. Given this later view culture may be limited to humans, though a growing body of work is attesting to the presence of (potentially cumulative) cultural evolution in chimpanzees [161, 105, 166]. To resolve the debate while taking into account the ever growing number of species demonstrating highly complex social behaviours, Andrew Whiten and Carel van Schaik devised a cultural pyramid (see Figure 2.1). The cultural pyramid accommodates low level social transfer of the kind exhibited by fish [12, 13] and high level cumulative culture as demonstrated by humans and some apes [161, 105, 166]. According to Whiten and van Schaik, for a species to be considered cultural they must exhibit and maintain multiple traditions; it is this definition of culture as the maintenance of multiple traditions that is used in this work. As the agents modelled in this work are by comparison to even the least complex animal very simple, we do not expect to find anything close to culture or cultural evolution in this work, with this definition of culture instead being used to provide a sense of scale to the social behaviours exhibited here.

---

<sup>2</sup>Reader and Biro [123] provides a thorough and detailed overview of social learning in wild animals



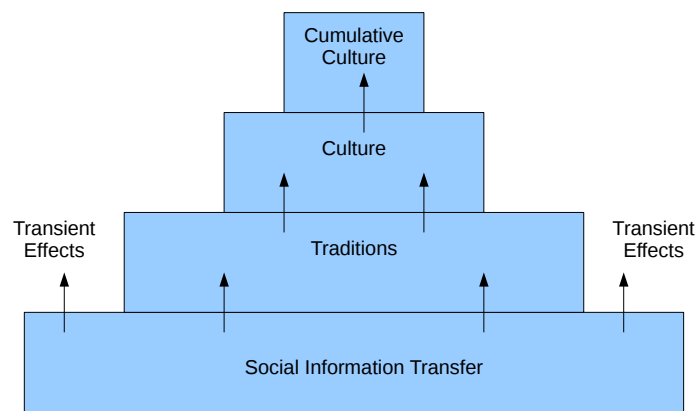


Figure 2.1: A reproduction of Whiten and van Schaik’s cultural pyramid [164], showing the variety and dependence of the emergent properties of social interaction and social behaviours. Social information transfer is considered to be the most common result of social interaction, with cumulative culture emerging only as a result of the most complex social and cultural interactions.

### 2.1.2 Social Information and Social Learning: Mechanisms and Strategies

A social learning or social information strategy is a preference exhibited by an agent or a population in regard to sources of social information. These are sometimes known as social biases, as they can manifest themselves as a bias towards a certain type of agent, an agent expressing a certain type of information, or an agent undertaking a certain action. Laland [93] groups social learning strategies into two broad categories: when strategies and who strategies. When strategies (also known as context dependent, state based, strategies [126]) may be simply described as rules defining when a social interaction may be undertaken, examples of such strategies being "copy when uncertain" [62] or "copy if current behaviour has a low pay-off" [131]. Who strategies (also known as context dependent, model based, strategies [126]) on the other hand are more concerned with the state or identity of conspecifics, examples of such strategies being "age-based copying" [52], "prestige-based copying" [76] or "kin-based copying" [77]. Rendell et al. [126] extends Laland's work [93] by considering a wider set of possible social learning strategies, including unbiased/random social learning, and also frequency dependent strategies such as "copy the majority" [26], which is also known as conformist bias, and copy based on "number of demonstrators" [19]. Importantly all of the social learning strategies discussed in Laland [93], Galef [62], and Rendell et al. [126] are apparent in nature, meaning they must all confer some adaptive benefit to the animals utilising them.

Alongside social learning strategies, a number of social learning mechanisms have been observed in nature. A social learning or social information mechanism is the method via which social information is transmitted and transferred - this does not necessarily mean how social information is physically communicated, but more how

social information is passed from one agent to another. A number of works have attempted to provide a set of useful categories of, or a taxonomy for, social learning mechanisms [162, 80, 169, 126, 84, 63]. In the taxonomy used by Galef [63] (based on the work of Whiten and Ham [162]) social learning mechanisms are split into two broad groups: social influence and social learning. Social influence mechanism are those which do not require any learning, instead allowing for social or public information to intentionally or unintentionally influence the behaviour of social agents. Examples of social influence include contagion, whereby the behaviour of one agent stimulates a similar behaviour to be expressed by others agents (e.g. yawning), and exposure, whereby one agent unintentionally exposes another to a new or novel environment or resource. Social learning mechanisms are those in which one agent learns something directly from the actions or state of another agent. Examples of social learning include imitation, goal emulation, observational conditioning, and even stimulus enhancement whereby one agent actively seeks to orient the attention of another agent towards a stimulus. Missing from the taxonomy used by Galef [63] is a discussion on local enhancement, which is often conflated with stimulus enhancement. Both Heyes [80] and Rendell et al. [126] discuss local enhancement as a social mechanism apart from, yet associated with, stimulus enhancement, with Heyes [80] considering both to be a form of social enhancement whereby social activity increases the likelihood of engagement with either a particular stimuli or a location or habitat. To further clarify this division Rendell et al. [126] (using the work of Hoppitt and Laland [84] as its basis) defines stimulus enhancement and local enhancement as follows:

- Stimulus Enhancement: "A demonstrator exposes an observer to a single stimulus, which leads to a change in the probability that the observer will respond to stimuli of that type"

- Local Enhancement: "A demonstrator attracts an observer to a specific location, which can lead to the observer learning about objects at that location"

Stimulus and local enhancement can be intentional or unintentional on the part of the demonstrator, and can both be used effectively alongside a variety of social learning strategies which may guide the circumstances under which these, and other, social mechanisms are applied.

It has been noted in recent work that cumulative culture, which is often seen as the ultimate expression of social behaviour, and has only been observed in humans and some apes [105, 166], may be achievable without the use of imitation [38] or teaching [170] - this opens up exciting possibilities regarding cultural behaviours in simple artificial life and evolutionary robotics models where both true imitation (which according to Thorpe [145] requires both self-consciousness and goal-directedness) and detailed teaching are extremely difficult to model. This echoes the work of Noble and Todd [113] where complex adaptive behaviours that seemed to be the result of imitative social learning were instead shown to be possible using simpler social learning mechanisms such as local enhancement.

Mechanisms of social information transfer and social learning strategies are central to this work. To properly simulate social behaviour and thus investigate the benefits and evolutionary necessities of social learning and social information use, social learning mechanisms must be simulated in a justifiable manner. In computational and mathematical simulation, imitation (or at least a mechanism described as being imitation) is by the far the most explored mechanism of social transmission [49, 81, 28, 39, 45, 104, 22], however there is a burgeoning set of simulation models exploring learning by emulation (learning to copy the results of a set of actions as opposed to imitating the precise set actions themselves) [98, 7].

In regard to social learning strategies in this thesis, we focus primarily on the utility of social learning and social information strategies in Chapters 6, 7 and 8. In Chapter 6 we consider a number of the who strategies alongside unbiased social learning to see whether the results presented in Chapter 4 are robust across a number of social learning strategies, we also consider this wider consideration of social learning strategies to be important in evolutionary robotics and artificial life research where a broader consideration of the adaptive implications of social learning strategies is often ignored. In Chapters 7 and 8 various social information strategies are considered regarding public and social information use in a simple food foraging task to investigate if social information without an accompanying within-lifetime learning process is of any adaptive benefit, and whether this adaptive benefit is dependent on the type of social information available to agents. Chapter 5 is less concerned with social learning strategies (or mechanisms), but can be said to incorporate unbiased social learning, as no information about the state of other agents (other than their existence in the population) is used to select social learning models.

Regarding social learning mechanisms, here we primarily concern ourselves with a form of teacher-learner imitation (in a simpler sense than intended by Thorpe [145]), stimulus enhancement, and local enhancement. In Chapters 4 and 6 social agents engage in rounds of teacher-learner imitation whereby a learner agent is guided around an environment in a manner akin to stimulus enhancement. Agents in these tests are supposed to imitate the reactive and deliberative behaviours exhibited by their teachers, before themselves being evaluated on the task. In Chapters 7 and 8 only social information use is considered whereby agents may decide to use the social information inadvertently expressed by others to affect their own behaviour in a food foraging task; this could be framed as a sort of social influence such as exposure or a form of non-learnt

local enhancement.

## 2.2 Social Behaviour and Social Learning in Artificial Evolutionary Systems

The artificial evolutionary systems employed here are all sub-sets of what may be described as Artificial Life systems. As a discipline Artificial Life is relatively young, especially in regard to simulating social systems where mathematical modelling is already well established and widely used [41], but it has vast amount of potential. According to one of the founders of Artificial Life, Chris Langton:

“Artificial Life [AL] is the study of man-made systems that exhibit behaviours characteristic of natural living systems. It complements the traditional biological sciences concerned with the analysis of living organisms by attempting to synthesize life-like behaviours within computers and other artificial media. By extending the empirical foundation upon which biology is based beyond the carbon-chain life that have evolved on Earth, Artificial Life can contribute to theoretical biology by locating life-as-we-know-it within the larger picture of life-as-it-could-be.” [94]

As a method of biological simulation, Artificial Life allows us to explore many aspects of biological systems from the bottom-up. By starting at the bottom many of the confounding factors that often stand in the way of the kind of clear synthesis of social behaviour sought by Tinbergen [146] may be avoided, thus allowing us to investigate which factors actually drive the evolutionary adaptation to social information use and social learning, resulting in adaptive social behaviour.

Building grounded (and non-grounded) Artificial Life models is often described as bottom-up due to behaviours and adaptations emerging as results of the simulation model rather than being built in or implicitly parametrised for from the outset. One of the many factors that affects the study of social behaviour in nature is that one must start from the middle; millions of years of evolution both physiologically and behaviourally had already occurred by the time human scientists decided to start studying cultural adaptation. When starting in the middle one must postulate, aided by fossil and geological evidence, on the potential evolutionary causation of any given cultural adaptation. For instance, why did the split between the chimpanzee and human evolutionary lines more than 2 million years ago cause one species to develop a highly complex culture and the other to remain relatively culturally naive? We cannot go back 2 million years and witness evolution in action despite the many fossils we now have at our disposal. Simulating from the bottom-up avoids this complication by instead simulating abstract animals (agents or animats) in abstract environments, we can then view the evolutionary histories and trajectories of our agents as well as investigating the many evolutionary permutations that may occur. While it is true that these agents and their virtual environments are not in fact close abstractions of real animals and ecosystems, the general purpose nature of such modelling techniques can provide an invaluable insight into the basic rules and necessities of grounded evolutionary systems. Generalising from the bottom-up also removes other confounding factors that those studying behaviour in nature cannot. In an abstract simulation we do not have to consider the issues surrounding sexual selection, or include species competing in the same niche or even subtle environmental factors such as rainy seasons, instead only factors we are interested in or those that seem to be immediately important need be included in bottom-up models. Added complications and complexity can always

be added later, once the basic factors and their effects have been sufficiently explored and understood. The work of Joanna Bryson is a prime example of the flexibility and importance of bottom-up grounded simulation to the many fields of natural science [30, 32, 122, 31].

One of the favoured methods of Artificial Life simulation is the use of grounded artificial agents situated in some simulated environment. Grounded agents of this sort are often referred to as animats [158] or agents, though this is not universal [104]. Animats are general purpose abstractions of biological organisms often composed of simple sensors, a means of locomotion, some decision making organ (artificial neural networks as a default) and a genotype coding for those aspects of the agent subject to evolution. Inspired by the Braitenberg Vehicle Architecture [29], animats take the core aspects of autonomous biological organisms as a starting point from which behaviour or physiology may be empirically explored. In this work populations of neural networks embodied in animats will be employed. There are a number of competing paradigms in artificial neural network (ANN) research from standard feed forward neural networks where weighted connections feed forward through layers of neural synapses to an eventual output, to echo state networks and reservoir computing. In many of the grounded simulations of social and cultural evolution to date, a simple feed forward approach has been taken [5, 4, 3, 45], though some subtle variations have also been employed [104].

In order for social behaviour to be simulated in a way that is not totally at odds with the natural world, a separation between the genetic encoding of an animat, and the phenotype of the animat is necessary [103, 102]. By separating the genotype and the phenotype, extra-genetic inheritance via social and non-social learning can take place with none of these within-lifetime changes affecting the underlying genotype. In simulations using ANNs, which are often described as neuroevolutionary models when



coupled with an evolutionary algorithm, genotypes are often direct encodings of neural network weights (sometimes with additional genetic information on network parameters and topology), with phenotypes initially being direct copies of the genotype which then diverge during an agent’s lifetime. This direct encoding approach is primarily used here, but other approaches to encoding neural networks where a direct mapping of genotype to weights is not employed are also widely used [138, 53]. Neuroevolution applies evolutionary principles to neural networks by coding the structure [45] and weighted connections [49, 39, 45, 4, 3, 128] that constitute the network into some genotype, these genotypes being subject to evolutionary pressures.

Evolutionary selection pressures in artificial evolutionary systems determine the fitness of an individual. In evolutionary computing, fitness has generally been judged by some absolute fitness measure in a mathematically predictable fitness landscape, often for the optimisation of some mathematical function. This however is not a biologically viable method of fitness assessment. In nature, the fitness of an individual is simply determined by the individual’s ability to survive in its environment to the point of successful reproduction, with fitter individuals having more offspring than less fit individuals. By judging an individual’s fitness by their ability to survive in their current environment, we are considering an individual to be environmentally and evolutionarily situated [132]. This measure of fitness, though biologically viable, does provide some challenges for social learning [117] especially when social learning reinforces sub-optimal behaviours that population level genetic adaptation is capable of escaping.

Alongside environmental situatedness, social or cultural situatedness [96] is also a pre-requisite for the emergence of adaptive social intelligences. If environmental situatedness implies that agent behaviour is ecologically embedded, social situatedness

is the idea that social intelligence also requires a level of social or cultural embedding [96]. To situate an agent socially in simulation, an agent must be able to obtain successful behaviours through social interaction, this is most effectively done when the adaptation for social learning is allowed to evolve rather than being simulated as an assumption. It may be the case that only a weak coupling of the agent and social factors is necessary, with overly conformist social learning potentially undermining the agent's ability to produce successful behaviours [129, 117, 14].

Here, Chapters 4, 6, 7 and 8 all report on work which involves grounded neuroevolutionary agents, which are environmentally, evolutionarily and socially situated to some extent. Though by the definition of evolutionary situatedness offered by Schut et al. [132], none of the models employed in this work are fully situated in an evolutionary sense as all of them involve some level of central authority when decided on reproduction, fitness evaluation, and death.

This work also includes, in Chapter 5, a non-grounded artificial evolutionary system. This system is neither environmentally situated in the traditional sense, nor does it incorporate animats or neuroevolutionary agents. Instead, a simple evolutionary algorithm is used whereby populations of individuals comprising only of binary strings are evaluated on their ability to match a search string, which changes periodically based some mathematical formulation of environmental variability. Non-grounded models such as these, whilst simple, have been demonstrated to be well suited to testing some aspects of learning (both social and non-social) and the evolution of plasticity [82, 81, 67, 71], though they do lack some of the behavioural richness seen in grounded neuroevolutionary and evolutionary robotics models.

# Chapter 3

## Research Questions

All models used here to address the following research questions employ what Marriott and Chebib [103] would describe as Evolutionary and Developmental Models, in which individual agents have a genotype and some phenotype, with any within-lifetime learning events altering the agent phenotype but leaving the genotype unchanged (as opposed to a Lamarckian model whereby within-lifetime changes to the phenotype are encoded in the genotype). In these models selection operates on phenotypic behaviour while reproduction operates on the genotype, thus creating a potential disconnect between the genotype and phenotype. Both grounded neuroevolutionary models and non-grounded models making use of binary string evolutionary algorithms, are employed to explore the research questions posed below.

### 3.0.1 Incremental Genetic Evolution and Social Learning

In order to be maintained by evolution, social learning should confer an evolutionary advantage over population level adaptation (via incremental genetic evolution alone) and individual learning for obvious reasons; it provides access to behaviours learned by

conspecifics without the need to engage in dangerous individual trial and error learning or to rely on the slow and incremental process of genetic adaptation, which is largely incapable of escaping sub-optimal peaks in the fitness landscape. These advantages, however, need to be tested empirically; grounded artificial life models provide the perfect environment for doing this.

- Q1: In grounded Artificial Life simulation, can extra-genetic inheritance lead to the emergence and maintenance of behaviours that are inaccessible to genetic and individual learning? (*When is social learning useful?*) (addressed in Chapter 4)

### 3.0.2 Robustness in (Increasingly) Variable Environments

Humans (*homo sapiens*) acquired their adaptations for complex social learning before the Out Of Africa Exodus [141], whereby increasing numbers of humans left Africa to settle in the Middle East, Asia and Europe, before going on to colonise a significant portion of planet's surface. This exodus would have provided changes to the human environmental situation that would have given ample advantage to those with a predisposition to cultural adaptation, however complex social learning was before the migration out of Africa occurred. Given the potential costs and drawbacks of cultural adaptation, such as the increased energy requirements of larger brains [85, 86] and the altricial state of children at birth [34, 35] stemming from increased risks in child birth owing to larger neonatal brain size, and the seemingly stable environmental situation pre-exodus humans found themselves in, why did the adaptation for cultural evolve? Assuming social learning is advantageous (see Research Question 1) under which temporal environmental circumstances will it evolve? Are there any environmental necessities for cultural adaptation that we may model using grounded Artificial Life simulation?

- Q2: In non-grounded Artificial Life simulation, under which temporally variable environmental conditions will a social learning mechanism evolve? (*When will culture evolve?*) (addressed in Chapter 5)

### 3.0.3 Social Learning Strategies

As discussed in Chapter 2.1.2, social learning and social behaviour in nature is associated with a large variety of social learning strategies. However, despite the large variety of social learning strategies seen in nature, it is not entirely clear whether all social learning strategies can be relied upon to provide an adaptive advantage over incremental genetic evolutionary processes alone. It is also not clear how different social learning strategies enable access to behaviours that are inaccessible to incremental genetic evolution alone. By exploring the adaptive advantages of different social learning strategies we may begin to better understand the role different social learning strategies play in natural social systems, and ultimately engineer artificial agents with more effective and robust social learning capabilities.

- Q3: In grounded Artificial Life simulation, can we demonstrate that behaviours inaccessible to incremental genetic evolution alone are still discovered, and maintained, when agents are permitted to learn via differing social learning strategies, access these incrementally inaccessible behaviours in differing ways? (*Are different social learning strategies adaptive, and how?*) (addressed in Chapter 6)

### 3.0.4 Social Information Use without Learning

Social information can provide information about the presence, state and intentions of other agents; therefore it follows that the use of social information may be of some adaptive benefit. As with all information, social information must be interpretable

and relatively accurate given the situation in which it is derived. In both nature and robotics, agents are required to learn which social information is relevant and under which circumstances it may be relied upon to provide useful information about the current environmental state. However, it is not clear to what extent social information alone is beneficial when decoupled from a within-lifetime learning process, leaving evolution to determine whether social information provides any long term adaptive benefits. By decoupling social information and learning, we may begin to understand the role social information alone plays in natural social systems, and in the future engineer artificial social agents with an improved understanding of the relative importance of different sources of social information.

- Q4: In grounded Artificial Life simulation, can the use of social information in populations comprised of simulated neuroevolutionary agents be shown to be adaptive when decoupled from within-lifetime learning processes? (*Is social information adaptive when decoupled from within-lifetime learning processes?*) (addressed in Chapter 7)

### 3.0.5 Social Information Use and Behavioural Change

Agents incorporating social information into their decision making processes should exhibit notable behavioural differences compared to non-social agents. However, it is not clear to what extent these social behaviours resulting from social information use are exhibited, and whether they persist when social information use is not necessarily adaptive. By exploring the behavioural differences observed in social populations, we may begin to better understand how many of the behaviours seen in nature are driven by the need to access social information, and ultimately engineer artificial social systems with a better understanding of the kinds of behaviours we would expect to

emerge in such systems.

- Q5: In grounded Artificial Life simulation, how effectively is social information used to increase agent adaptability, and do any behaviours resulting from social information use persist even when task performance does not outperform the performance seen when no social information is available? (*To what extent does social information affect agent behaviour; are any behavioural differences distinct from those exhibited by non-social populations; and do these behaviours persist when social information use is no longer adaptive?*) (addressed in Chapter 8)

## Chapter 4

# Discovering and Maintaining Behaviours Inaccessible to Incremental Genetic Evolution

### 4.1 Introduction

Here we present work showing animats in a virtual environment learning behaviours through imitation that are inaccessible to incremental genetic evolution alone. Learning by imitation is often considered to be a mechanism of social information transfer [41, 164], leading to what may be described as social or cultural learning. By combining social learning, which allows the transfer of information between different individuals in the population, and individual learning in the same evolutionary system it is possible to make use of both global and local search: global search through the underlying (multi-generational) genetic algorithm and local search through individual (lifetime) learning [82]. It has been demonstrated by Best [16] that by using social learning in



place of individual learning on a more challenging version of the Hinton and Nowlan [82] problem, it is possible to improve the speed at which a population of agents discover an adaptive goal. Social learning has also been shown to be effective when finding optima in narrow peaked and flat peaked search landscapes, as opposed to individual learning which struggles with narrow peaked search landscapes [6]. Social learning has the added advantage of allowing individuals to pass on learnt information to other members of the population, and so preserving extra-genetic information for the next generation. Beyond its uses in evolutionary optimisation and search, cultural and social learning is also a well known natural phenomenon with various species using social learning mechanisms such as imitation, emulation, teaching and the use of public information to produce adaptive behaviours in dynamic and challenging real world environments [46, 164, 123, 150, 159? ].

A number of studies have investigated the effect learning by imitation has on populations of evolving neural networks [16, 39, 5, 4, 45, 104]. In much of the literature these imitating neural networks are referred to as agents, with some, as is the case in this work, even taking on the role of animats or autonomous agents in virtual environments [104]. It is the aim of this work to investigate whether learning by imitation in a population of neural networks enables behaviours that are deemed to be inaccessible to incremental genetic evolution, to be learned and maintained. Here an increasingly complex virtual environment is used in which animats' behaviours are evaluated.

#### **4.1.1 Incremental Genetic Evolution**

Long-term incremental genetic evolution [43, 66] necessarily uses converged populations, which may be referred to as species (or quasi species [33]), as an initial starting point for evolution search or population genetic learning. In genetic algorithms (GAs)

this is referred to as the Species Adaptation Genetic Algorithm (SAGA) approach [73]. The SAGA approach impacts on the way populations evolve: recombination will have a far smaller effect on the motion of the population than in a standard GA, as each species is already genetically similar, leaving mutation as the primary driving force behind evolution. Mutation can be substantially effective in spaces percolated by neutral networks: pathways of level fitness through the fitness landscape. In this case genotypes can vary while still producing similar phenotypes and behaviours. When phenotypes of higher fitness are found the population converges onto them. This incremental approach enables species of animats to discover and converge upon an easily accessible solution. However, if there is no neutral or incremental path between the corresponding basic behaviour and fitter ones, the population will struggle to move away from these sub-optimal behaviours. Figure 4.1 depicts a mock example - in this mock example there is an incremental pathway via standard evolutionary operators (crossover and mutation) to a sub-optimal solution, but no incremental or neutral pathway to the optimal solution. Once the population has achieved the incrementally accessible, but sub-optimal, solution, mutation and crossover are unable to drive the population towards new behaviours as fitness based selection will bias reproduction in favour of those agents still on the sub-optimal peak. In order to bridge the gap between the sub-optimal and optimal peaks, the population must be allowed to explore new solutions.

One approach to solving the problem of sub-optimal convergence is to increase the rate at which mutation is applied, potentially allowing the population to explore more of the solution space and so discover new fitness peaks. However, there are problems with this approach: as mutation rates increase, the evolutionary search strategy begins to resemble random search, with larger mutation rates making it increasingly difficult

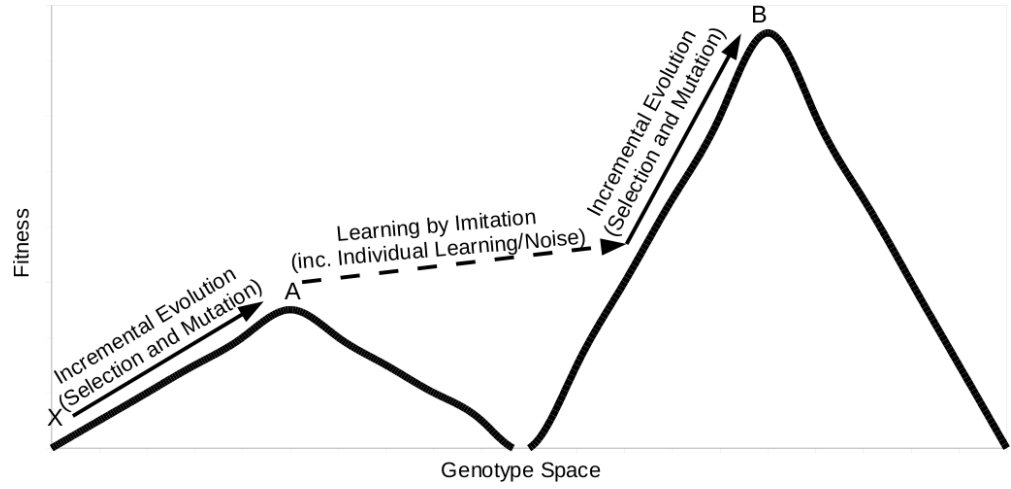


Figure 4.1: Mock fitness landscape with an inaccessible fitness peak. A species starting from point X on the above mock fitness landscape would achieve peak A by way of the hill climbing strategy adopted by incremental genetic evolution (driven primarily by mutation and selection). Gradient-based learning amongst such a species would ordinarily also be restricted to peak A. The inclusion of both noise in the genotype to phenotype map and learning by imitation can enable the species to jump across areas of lower fitness to higher peaks (inaccessible to hill climbing alone), where incremental genetic evolution and learning can resume hill climbing.

for the population to maintain solutions. The point at which mutation becomes so large that favourable structures discovered by evolution are lost more frequently than they are found is known as the error threshold. Ochoa et al. [118] and others have demonstrated a link between error thresholds and optimal mutation rates in evolutionary algorithms.

### 4.1.2 Discovering and Maintaining Inaccessible Solutions: Transcription Errors and Imitation

To solve the issue of sub-optimal population convergence without crossing the error threshold, noise is often added to the fitness landscape via the genotype to fitness map, either by using non-elitist selection or some noisy determination of agent performance. However, where such noise is in the phenotype to fitness section of that map with a phenotype instead of the genotype being used to determine agent behaviour, its ability to aid in the transition between peaks (or more accurately between neutral networks) is limited. By instead incorporating noise into the genotype to phenotype map, as with transcription errors, behaviours inaccessible to incremental genetic evolution may be exhibited reliably by individuals while leaving the genotype untouched. It can be useful to view such noise as a type of unguided individual learning.

In order to maintain successful behaviours in the population, some form of extra-genetic learning needs to take place. The model employed in this work makes use of imitation through interactions between teachers and pupils to facilitate the transmission of learnt behaviours [39, 5, 3, 45]. As in Curran and O’Riordan [45] pupils follow teachers in a mock evaluation on a set of environments. As both teacher and pupil receive the same environmental input the teacher’s output may be used as a target pattern for error back-propagation, reducing the pupil’s output error compared to that of the teacher. By learning in this way pupils are able to imitate the behaviours exhibited by teachers, thus maintaining behaviours in the population that would have been lost in incremental genetic evolution.

### 4.1.3 Neuroevolution of Deliberative Behaviours

This work uses populations of neural networks embodied in animats. The neural network architecture used here is a hybrid of two different networks: the first controlling the high level deliberative behaviours of the animat, and the second controlling the animat’s reactive capabilities [128]. By making use of both reactive and deliberative mechanisms, neural architectures of this sort are able to seek long term goals while also reacting to unforeseen events ultimately enabling the evolution of complex problem solving abilities. To demonstrate these problem solving abilities Robinson et al. [128] developed a complex problem called the ‘river crossing’ or RC task. The RC task required animats to find a single reward-giving *Resource* in a 2D grid-world environment containing a number of obstacles. Alongside Resource objects animats could encounter *Water*, *Grass*, *Traps* and *Stones*. Grass objects made up the majority of the environment and were seen as neutral space for the animats to move across; Trap objects were immediately lethal, as were Water objects, which were placed in such a way to resemble an unbroken river cutting the animat’s path to the Resource. In order to cross the river animats were required to pick up Stone objects, which could be carried at no cost to the animat, and place them in the same cells as Water thus negating their lethality. Once a continuous bridge of Stones over the river had been built animats could access the Resource. To succeed at the RC task animats were required to evolve with no *a priori* knowledge of the world; each new environment was unique and animats had no concept of co-ordinates, making solutions such as ‘move five steps to the right’ impossible, instead animats evolved goals and sub-goals such as ‘go to resource’, ‘avoid traps’ or ‘head to nearest stone’ which then allowed the network to navigate the animat towards these goals. Despite the RC task being reasonably complex, [128] demonstrated that it could be solved by initially converged populations

of animats using only incremental genetic evolution. To test our hypothesis a more complex version of the RC task has been developed: the RC+ task.

#### 4.1.4 The RC+ Task

An important aspect of the RC task was that individuals were evaluated on increasingly difficult environments. In Robinson et al. [128], animats were first shown a map with no river blocking their path; then a river with a width of one cell was introduced, followed by a final environment containing a river with a width of two cells. Stone and Trap objects were of a consistent number throughout all tests giving animats equal exposure in each environment. The RC+ task makes the task harder in regard to both river width and exposure to Stone objects. The number of environments an animat is evaluated on is increased from three to five, with environments becoming increasingly difficult to solve due to river width increasing from zero cells to four cells. To add to the difficulty further, the number of Stone objects gradually decreases from twenty in the first environment to zero in the final environment, making each environment more challenging to the point where the final environment cannot be completed by building a bridge. In order to make the final environment solvable two extra objects, *Object A* and *Object B*, are introduced into the environment. Object A and Object B are rare objects, with only one instance of each found in each environment. Like Stones, Object A and Object B may be carried at no cost to the animat and placed upon any square or object. If an animat happens to place both Object A and Object B on a square containing Water (notionally forming a floating raft that carries the animat to the resource), a reward equal to that of the Resource is received and the animat is considered to have successfully solved the environment. In short, an alternate Resource may be constructed out of the three other objects (Object A, Object B and Water),

removing the need to build bridges but still requiring agents to be driven towards the Resource when Water is not present. The RC+ task is impossible to solve with incremental genetic evolution alone. To solve it, animats are required to engage with Water, Object A and Object B while still avoiding Traps and uncovered Water, and to also be able to reach the Resource in the absence of Water (the simplest sub-solution to evolve). The rarity of both Object A and Object B adds to the difficulty of the RC+ task as animats must now evolve to be driven to towards Object A and Object B despite potentially very little exposure during their time in the environment.

## 4.2 The Model

Animat movement is controlled by a hybrid neural network embodying both reactive and deliberative qualities. This hybrid network may be broken down into two network models: a *shunting network* and a *decision network*, with the decision network passing information on to the shunting network which in turn controls the animat’s movement. The shunting network is not directly exposed to any evolution or learning. The deliberative network on the other hand is exposed to both evolution and learning, enabling the evolution and inheritance of animat behaviour.

### 4.2.1 The Shunting Network

Shunting networks are a specialised form of neural network making use of what is known as the shunting model [167, 168]. The inspiration for the shunting model came from Yang and Meng’s [167, 168] desire to develop motion planning systems capable of reacting quickly in real-time environments, thus allowing robotic agents to exhibit robust and collision-free motion planning behaviours. Instead of directly specifying

behaviours, the shunting model maps network outputs onto environmental outputs (within an internal map of the environment) which are propagated across the environment to form an activity landscape. This activity landscape is used by the agent to control movement through the environment, by dynamic gradient ascent of the landscape. In their model, Yang and Meng [167, 168] demonstrated a neural network composed of an  $n$ -dimensional lattice of neurons, with each neuron representing a possible state in the system. By using neurons to represent states in this way it is possible to represent any system which is capable of being fully described by a set of discrete states.

The environment used for the RC and RC+ tasks is a simple 2D grid-world consisting of  $20 \times 20$  cells, with each cell representing a position in co-ordinate space. Each position in the grid-world may be occupied by any number of objects found in the RC+ environment (Resource, Water, Trap, Grass, Object A and Object B), allowing the system to be fully described by a set of discrete states, thus enabling the use of the shunting model to direct animat movement across the RC+ environment and ensuring a simple one-to-one relationship between neurons and geographical locations.

In Yang and Meng [167, 168], two versions of a transition function for specifying inter-neuron dynamics were developed: one which controlled activity saturation in the network and one which did not. Consistent with the findings of Robinson et al. [128], we found activity saturation not to be a problem exhibited by networks in the RC+ task, enabling the use of the simpler transition function in equation 4.1.

$$\frac{dx_i}{dt} = -Ax_i + I_i + \sum_{j=1}^k w_{ij} [x_j]^+ \quad (4.1)$$

Alpha ( $A$ ) represents the passive decay rate, which determines the degree to which each neuron's activity diminishes towards an idle state. The functions  $[x]^+$  is  $\max(0, x)$ .



The connection weight (or synapse strength)  $w_{i,j}$  between neurons  $i$  and  $j$  is the Euclidean distance between cells  $i$  and  $j$  within the receptive field.  $k$  is the receptive field size and here is set to 4, corresponding to the four cells orthogonally surrounding cell  $i$ . Iota ( $I$ ) is equal to  $E$  in the case of the target, and  $-E$  for an obstacle, where  $E$  is a large integer.

In the case of the RC and RC+ tasks Iota values are limited to 15, -15 and 0, representing the target resource, an obstacle and neutral space respectively. The result of using a transition function with these values are 2D environments with large peaks at the sites of target states, large troughs in cells occupied by obstacles, and large amounts of neutral space through which neuron activity from targets may spread. Using the shunting model to control animat movement allows for goals such as ‘head for resource while avoiding traps’ or ‘place carried stones on water’ to be easily achieved.

#### 4.2.2 The Decision Network

The role of the decision network is to set the Iota values for object states found in the RC and RC+ task. Using the decision network animats can set the desirability of object states in relation to their current environmental inputs, allowing them to manipulate the shunting network’s activity landscape and so combine multiple actions such as ‘pick up the closest stone’ and ‘place stone on water’ to create complex behaviours.

As in Robinson et al. [128], the decision network is simply a feed-forward multi-layer perceptron with one hidden layer comprising of four hidden units. The input layer is capable of representing the animat’s current state in the environment including whether or not the animat is currently carrying a movable object (Stone, Object A, Object B), with each movable object having a dedicated carrying input. Inputs taken by the input layer are single values of 1 or 0, representing the presence of the object

in the same cell as the animat. These input values are fed through to the hidden layer neurons via weighted connections in the range  $[-1, 1]$ . At each hidden unit the weighted sum of inputs is passed through a hyperbolic tangent activation function to produce hidden layer outputs. In the RC+ task the output layer is made up of sixty-seven neurons representing the Iota values of all sixty-four possible environmental states (excluding Grass objects whose Iota values are always set to 0 and therefore do not need be represented in the decision network) and a pick-up/put-down output for each non-static object (Stone, Object A, Object B). At each output neuron the sum of all weighted connections is passed through a hyperbolic tangent activation function with fixed thresholds: neurons outputting within the range  $[-0.3 : 0.3]$  are set to output 0, while all outputs over 0.3 resolve to 1 and all outputs below -0.3 resolve to -1. These fixed threshold values are consistent with those used in Robinson et al. [128].

For outputs representing the pick-up/put-down actions output values of -1 cause the animat to put down the specified object they are carrying, values of +1 causing animats to pick up the movable objects they are currently sharing a cell with providing the animat is not already carrying an object of that type. For all other outputs, resolved output values set the Iota values to be used in the shunting network. So if an output neuron has a negative output, all objects of that class found in the environment at that point in time will have their activations set to -15; for positive outputs to +15. Any object resulting in an Iota value of 0 will remain neutral, causing their activation values in the shunting network to be solely based on the propagated activations of other objects. The resulting environment will contain a number of peaks of high activity and troughs of low activity, gradually propagating activity through neighbouring neutral cells.

Figure 4.2 shows two of the five potential environments an animat may observe in

the RC+ task, and the corresponding activity landscapes given certain outputs from the decision network. The first environment represents the initial challenge an animat must complete, where only traps stand in the way of a resource. As can be seen by this environment’s activity landscape, the Iota value associated with the resource has been set to be positive resulting activity propagating from the resource over the surrounding neutral space. The second environment represents the second challenge, to cross a river before having access to the resource. In this environment, activation propagation from the resource has been impeded by the decision network outputting negative Iota value for Water objects. Negative activity repels animats from objects with negative Iota values; however positive activation can be seen coming from the Object B object, providing a hill-climbing route for the animat to take in activity space.

### 4.2.3 Evolution of the Decision Network

To evolve the decision network a steady-state genetic algorithm was used. At each iteration two animats were selected from the surviving population to be evaluated in tournament selection, with the worst performing animat being replaced by the progeny of the better performer. The competing animats are evaluated in five increasingly difficult environments. If during evaluation an animat fails to complete an environment, the evaluation is terminated. Fitness is set to be the number of environments successfully completed by an animat during evaluation.

An animat’s genotype consists of a set of floating point values each in the range  $[-1, 1]$ , which are transcribed into the connection weights in the animat’s decision network. The genotype and the decision network are stored separately, so any learning that may take place during an animat’s lifetime will only affect the decision network: no changes are made to its genotype after an animat is initially created. New animats

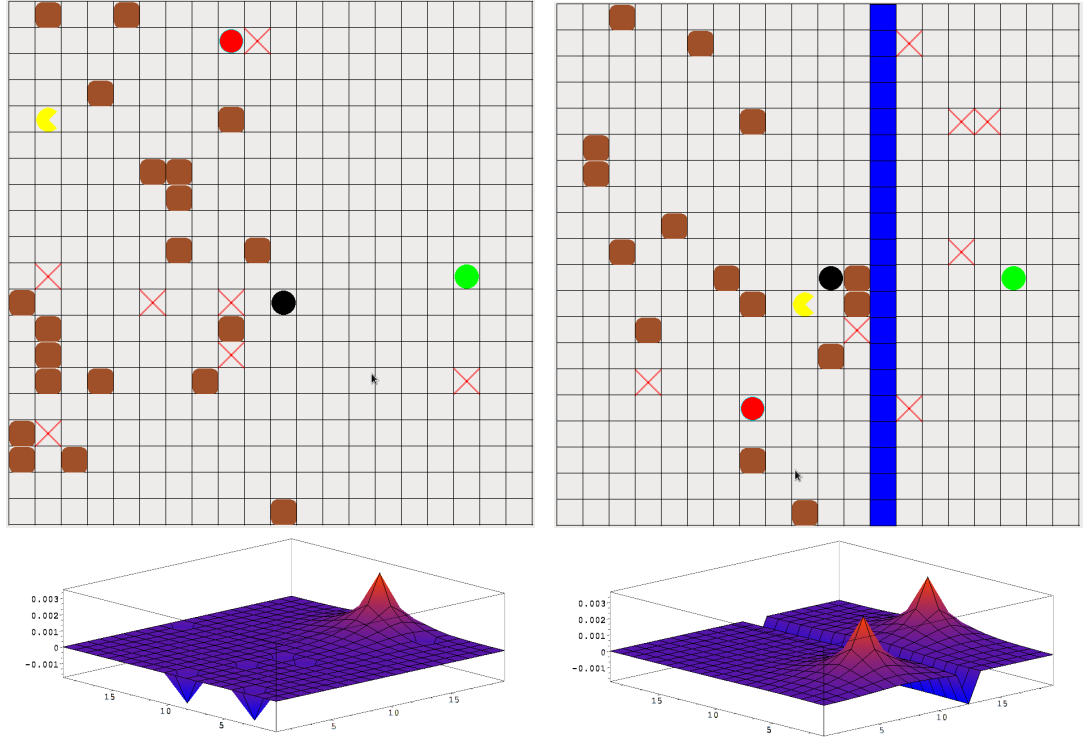


Figure 4.2: RC+ task and activity landscapes. Two environments with their activity landscapes (given certain outputs from the decision network - see main text). Animat=yellow, Stones=brown square, Resource=green circle, Object A=black circle, Object B=red circle, Traps=crosses, Water=blue square.

are the offspring of two other animats from the current population: one tournament winning animat and one randomly selected animat. The child’s genotype is created first through recombination of the parents’ genotypes; for this operation single-point crossover is used with the point of crossover being a randomly selected point in either parent’s genotype. Each loci in an animat’s genotype represents exactly the same connection weight as in any other animat’s genotype, with all genotypes being of length  $L = 308$ . Mutation follows recombination; each point has a probability  $P_{\text{mut}} = 1/L$  of having a random value from  $N(0, 0.4)$  added to it, with the resulting values being bounded within the range  $[-1, 1]$ . Once the genotype has been constructed it is written to the new animat’s decision network; this process is referred to as transcription. During transcription two randomly selected connection weights are overwritten with a new random value selected from a discrete uniform distribution  $U(-1, 1)$ . The weights now present in the decision network dictate the animat’s future behaviours within each environment.

#### 4.2.4 Learning in the Decision Network

Following reproduction new animats are afforded the opportunity to learn from a teacher via error-backpropagation. This method of teacher-pupil back-propagation has been previously employed by Curran and O’Riordan [45]. However, the teacher-pupil scenario used in this work differs in a number of ways. In the learning model used by Curran and O’Riordan [45], teachers were selected from the population based upon their fitness and then assigned  $n$  pupils to teach. We contend that in nature absolute fitness is very difficult to assess. To resolve this issue, the current tournament-winning parent is assigned the role of teacher, with the parent’s most recent progeny assigned the role of pupil.

There are also differences in the way error-backpropagation is used to teach pupils in this model compared to that of Curran and O’Riordan [45]. As with our model, Curran and O’Riordan [45] allowed pupils to hitch-hike on the back of the teacher during a mock evaluation, with inputs shared between teacher and pupil and using the teacher’s output pattern as a target pattern for the pupil to learn. The learning method employed by Curran and O’Riordan [45] permitted pupils to learn from the target pattern until the error between child and parent outputs were minimised to a satisfactory level. In our model pupils are only presented with the current teacher’s output once every simulation time step (immediately after the teacher’s decision network’s inputs, activations and outputs are updated). If a teacher happens to move through the environment in such a way that both inputs and outputs remain the same, the child will be presented with many opportunities to learn a given target input-output pattern. However, if the teacher moves around the environment via many different input combinations, the student will have the opportunity of potentially witnessing many different target outputs but at the cost of having very little time to minimise error. Imitating in this manner enables the population to retain favourable behaviours not coded for genetically, whilst not undermining the incremental genetic evolutionary process.

### 4.3 Experimentation

At each iteration of the model two individuals are taken from the population to be evaluated on a series of five environments/maps. All maps have seven Trap objects placed randomly on the map, one reward-giving Resource, one Object A, one Object B, and  $20 - (5 \times \text{riverwidth})$  Stone objects. River width varies from an initial width of zero, increasing by one cell per map. During evaluation individuals must successfully reach the Resource or place Object A and Object B onto a cell containing Water; any

animat failing to do so within 100 steps or dying by means of a Trap or uncovered Water is not permitted to attempt the next environment.

Fitness in the model is determined to be the number of maps successfully completed in the current tournament iteration, with individual fitness being set to zero before each evaluation. The individual achieving the highest fitness is allowed to reproduce, with the weaker individual being replaced by the progeny of the tournament winner and a randomly selected animat. This steady-state approach maintains the population at a size of 100 individuals.

After reproduction the child is allowed to learn via error-backpropagation from its tournament winning parent. The child follows its parent in a mock evaluation, with the child's inputs being set to those of the parent. Learning takes place for as long as the parent is being evaluated. Once the parent either fails to complete a map or completes all five environments, learning is terminated. At each step through the evaluation the child attempts, via error-backpropagation with a learning rate of  $\delta = 1$ , to learn to imitate the parent's output for the current inputs.

Three strategies are used in this model: two without learning and one with learning. Populations of animats with no access to learning fall into two categories. The first, known as Non-Learners(1), having a mutation rate and transcription error equal to that use by learning populations. As populations of Non-Learners(1) have no way of assimilating transcription errors back into the genotype it may be seen as giving learning populations, known as Learners, an unfair advantage. With this in mind a second of category of non learners, known as *Non-Learners(2)*, are also evaluated. Non-Learners(2) do not have transcription errors, and instead have a mutation rate equal to that of the original mutation rate plus two transcription errors:  $P_{mut2} = 3/L$ .

To test the ability of each strategy to exhibit the behaviour necessary to complete

the most difficult map, fifteen populations of each learning strategy were simulated. Each simulation lasted a maximum of 5,000,000 tournaments. In each simulation the best individual's fitness and the mean population fitness were recorded at intervals of 500 tournaments. The maximum fitness an individual could achieve was five, which directly relates to the successful completion of all five evaluation environments, the fifth environment being impossible to complete by bridge building and so requiring the combination of Object A and Object B on Water. For a population to be considered as adequately completing the fifth map, a fitness of five must have been recorded by the fittest individual at ten recorded tournaments with at least five of these tournaments being unbroken by a sub-optimal result. This ensures that the complex behaviour tested for is not only found but also maintained by the population.

## 4.4 Results

Table 4.1 shows results from the fifteen populations of animats using the Non-Learners(1) strategy: the mean, best and worst number of tournaments required to solve each map, across the fifteen populations (runs), and the proportion of populations that were successful in solving each map. Of the Non-Learners(1) populations over 90% were able to complete maps 1 to 4 but no population was able to demonstrate a successful solution to map 5. Populations of animats using the Non-Learners(2) strategy also demonstrated a high level of proficiency when completing maps where the bridge building solution is effective, though with a lower proportion of populations able to complete map 4 (see table 4.2). This may be due to the higher mutations rate used in the Non-Learners(2) strategy causing the destruction of potentially beneficial behaviours before they can proliferate through the population. To complete map 4 animats had to be stricter (more consistent) in their use of Stone objects. Despite this behaviour being reachable



using incremental genetic evolution it is within a small area of weight-space, causing it to be potentially lost with higher mutation rates. Neither non-learning strategy was able to discover the precise behaviour necessary to complete map 5, so failures recorded in tables 4.1 and 4.2 were not due to a sufficient behaviour being discovered but not maintained: the map 5 solution was simply never found, empirically demonstrating the inaccessibility of map 5 to incremental genetic evolution alone. Observations of runs demonstrated that non-learning agents either ignored, or were actively avoiding Object A and/or Object B, whereas agents in social learning populations who did manage to find a solution to map 5 would find and pick up Object A and Object B before determining that Water tiles were to be approached.

Table 4.3 shows results from animats using the Learners’ strategy. Unlike non-learning strategies, Learners are able to complete map 5 and thus exhibit the complex behaviour tested for in this work a third of the time, proving the hypothesis that learning by imitation is capable of enabling populations of animats to discover behaviours found to be inaccessible to incremental genetic evolution alone. However, Learners are less likely to discover and maintain solutions to maps 3 and 4 than non-learning animats.

Figure 4.3 charts the mean fitness of the best performing population from each learning strategy. From this graph it can be observed that Learners bypassed the sub-optimal bridge building solution once the population had (for some time) been evaluated on maps with rivers. The incremental nature of the evolution in this model causes the majority of the population to rapidly converge on the optimal solution once it has been discovered. Without learning, this optimal behaviour cannot be found. In this model incremental genetic evolution leads to convergence on sub-optimal solutions in non-learning populations, making it impossible for the discovery of the optimal be-

Map	Mean	Best	Worst	Stdev	Success
1	1200	500	3500	996	100%
2	502571	11000	2152500	738090	100%
3	1568000	34000	4429500	1501336	93%
4	1613786	58000	4432500	1506065	93%
5	N/A	N/A	N/A	N/A	0%

Table 4.1: Non-Learners(1) strategy RC+ performance. Mean, best, worst number of tournaments required to solve each map for the Non-Learners (1) strategy.

Map	Mean	Best	Worst	Stdev	Success
1	1400	500	3000	784	100%
2	81692	4500	252500	96805	100%
3	1801286	12500	4987000	1502754	93%
4	2193385	41500	4466500	1497156	87%
5	N/A	N/A	N/A	N/A	0%

Table 4.2: Non-Learners(2) strategy RC+ performance. Mean, best, worst number of tournaments required to solve each map for the Non-Learners (2) strategy.

Map	Mean	Best	Worst	Stdev	Success
1	1533	500	5000	1302	100%
2	512333	9500	2026000	616376	100%
3	2484455	5600	4340500	1395760	73%
4	2458800	88500	4211500	1861794	33%
5	1843200	83500	3851000	1631808	33%

Table 4.3: Learners strategy RC+ performance. Mean, best, worst number of tournaments required to solve each map for Learners strategy.

haviour. By combining learning by imitation and incremental genetic evolution in a model such as the one presented here, it is possible to not only discover complex behaviours inaccessible to incremental evolution alone, but also to have rapid convergence to a population exhibiting and maintaining that behaviour, thus creating a behavioural tradition or culture [164]. The results found here are broadly consistent with those of Acerbi et al. [3], who found that the combination of individual and social learning in artificial embodied agents not only allowed for the development of difficult and costly behaviours, but also provided an adaptive advantage over individual learning alone and lead to cumulative cultural evolution.

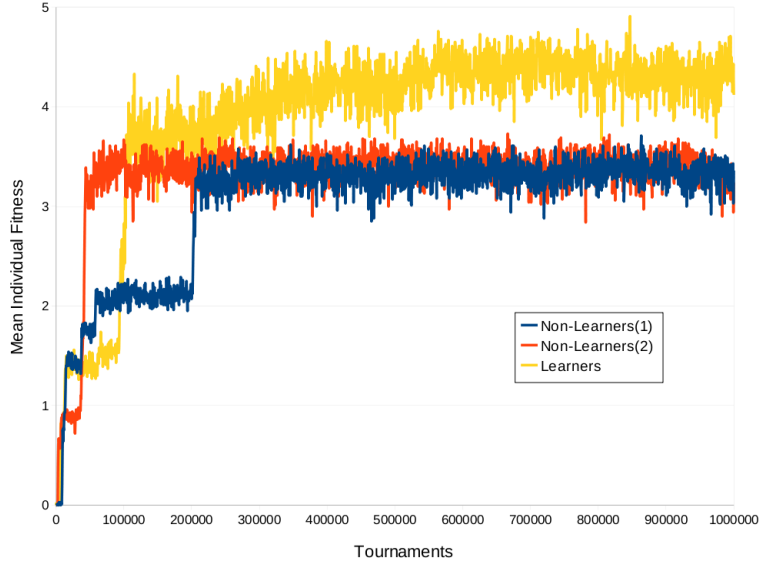


Figure 4.3: RC+ task mean fitness in the best performing populations. Graph showing the mean fitness in the best performing populations for each learning strategy. Populations learning by imitation demonstrated the ability to converge on more complex behaviours, thus achieving a higher fitness. Neither non-learning strategy is capable of producing the more complex behaviour.

## 4.5 Conclusions and Future Work

If a learnt behaviour is exhibited and maintained throughout a population for a number of generations it may tentatively be called a tradition or even a culture. According to

Whiten and Van Schaik [164] traditions are “consistent habits” that make use of social information transfer. In the model demonstrated here, learning by imitation enables social information transfer with behaviours being maintained by converged populations or species giving rise to traditions. The limited set of behaviours observed in this population do not however constitute the category of culture, which is reserved for the maintenance of multiple behaviours by a species. The incremental nature of the model causes sub-optimal behaviours to be phased out of the population. Were greater environmental diversity to be used, it may be possible to evolve a culture rather than a tradition.

The hypothesis presented here was that the introduction of both transcription errors and cultural transmission in the form of learning by imitation are sufficient to discover and maintain the most complex behaviour possible in the model, while incremental genetic evolution alone is not. The results support our hypothesis by demonstrating that without learning by imitation the solution to the final environment is never found but with imitative learning all behaviours can be discovered, exhibited and maintained.

While the RC+ task required considerably complex behaviours to be exhibited, there are a number of issues for the evolution of complex social behaviours. In nature, individuals are not taken away to be tested for fitness, individuals instead live and die in one environment which they constantly interact with. In allowing many individuals to interact with, and potentially terraform, an environment a certain level of dynamism becomes apparent. It is dynamism that forces individuals to evolve increasingly complex survival and strategies, potentially leading to the evolution of cultural adaptation and neural complexity [134, 40, 74]. Another minor drawback to the model used in this work is the limited set of behaviours available to animats. By using a larger environment with a greater variety of potential states available to the animats and evolving

the size and structure of the decision network, it may be possible to demonstrate the evolution of multiple behaviours leading to the emergence of a culture. To investigate more complex behavioural development and the role of imitative learning in the evolution of traditions and cultures, it would be beneficial to implement larger and more dynamic environments and allow for greater evolution in the decision network. A final drawback was the simple vertical social transmission mechanism used. The inclusion of intra-generational or oblique cultural transmission has been shown to be both sufficient [41] and beneficial [5] for the evolution of complex and robust cultural behaviours. Further investigation and application of oblique transmission within models such as that presented here would further benefit our understanding of and ability to achieve the evolution and maintenance of complex cultural traits.

Many of these limitations are considered and improved upon in Chapters 7 and 8 by allowing for agents to interact with one another whilst interacting with, and affecting, their environment. Chapter 6 also builds on the work presented in this Chapter by exploring whether the results presented here are robust across a variety of differing social learning strategies.

## Chapter 5

# The Adoption of Social Learning in Increasingly Variable Environments

### 5.1 Introduction

It is now widely accepted that the species *Homo sapiens*, to which all modern humans belong, evolved in Africa before leaving to populate the rest of world [141]. To successfully populate new and challenging environments hominins must have developed versatile and robust behaviours and survival strategies, with the most apparent adaptation for versatility being the adoption of extra-genetic learning strategies such as social learning [147]. This leads us to ask what was it about the environments in which hominins evolved that enabled them to adapt to be so versatile and ultimately so successful when moving into new and unfamiliar environments. In response to this question numerous authors have suggested a variety of theories and hypotheses regarding the relationship between hominin evolution and the environment [120]. In this work we instead test one of the most prominent theories of hominin evolution and versatility, the Variability Selection Hypothesis [119, 120, 121], using an artificial life simulation.

### 5.1.1 The Variability Selection Hypothesis

The variability selection hypothesis, as proposed by Richard Potts [119, 120, 121], predicts the adoption of versatile behaviours and survival strategies, in response to increasingly variable environments. Over the past seven million years there have been a number of what Potts describes as “large disparities” in environmental conditions, such as fluctuations in temperature and precipitation, and a trend toward increasing climatic variation in and around known early hominin locations in eastern and southern Africa, such as the Turkana and Olduvai basins [120]. Evidence for inter-generational and intra-generational changes, such as changes in forest coverage and the availability of water sources such as lakes and streams, have been found in a variety of geological and climatic indicators including marine oxygen isotope levels [120, 121], providing insight into temperature changes, and ocean dust records [120], providing evidence for dust plumes arising from strong seasonal rainfalls and prevailing wind patterns. Both of these indicators demonstrate an upward trend in environmental variability during the last seven million years in Africa, and around the world in general. Evidence from these, and other climatic indicators, shows that major shifts in the African climate correlate well with important early technological milestones and speciation events in hominin evolutionary history [70]. Key hominin and hominid adaptations such as early bipedality and complex social behaviour emerged during these periods of more pronounced environmental variability [121]. Though the climatic evidence for the variability selection hypothesis is impressive, the hypothesis has had very little theoretical work applied to it. Following the call from Potts [121] for a mathematical framework to explore the variability selection hypothesis, and the work of Grove [70] to that end, we here test the claim that increasing environmental variability is a sufficient selection pressure to elicit the adoption of social learning, in an artificial life simulation.



### 5.1.2 Social Learning

Social learning is not restricted to humans and their ancestors: it is a widely observed natural phenomenon, with many species using a variety of social learning mechanisms such as imitation, emulation, teaching and the use of public information to produce adaptive behaviours in dynamic and challenging environments [93, 123, 164]. It has been suggested that social learning enables animals to better track their environment by assimilating extra-genetic information from others during their lifetimes while avoiding potentially costly individual learning [27].

The effects and benefits of learning have been studied widely in simulation. According to Nolfi and Floreano [116] learning may be seen as having several adaptive functions from an evolutionary perspective. These include allowing individuals to adapt to environmental change, enabling evolution to use information extracted from the environment, and guiding evolution. Hinton and Nowlan [82] demonstrated that by using individual learning, populations are able to solve “needle in a haystack” problems due to learning guiding evolutionary search. Best [16] extended the work of Hinton and Nowlan [82] by demonstrating that, given the same “needle in a haystack” problem, social learning outperforms individual learning. Further work using simulated robots [4], animats [22], autonomous robots [3], ungrounded neural networks [45], and binary strings [89, 88] has contributed further to our understanding of the evolutionary advantages provided by social learning.

### 5.1.3 Social Learning in Increasingly Variable Environments

Numerous models and simulations have demonstrated the adaptive advantages, and highlighted potential failings, of learning strategies in environments exhibiting some level of adversity [10, 11] or variation [22, 25, 27, 70, 89, 160, 88]. In this work we test

the hypothesis that increasing, rather than simply consistent, environmental variability is sufficient to elicit the adoption of social learning. To test this hypothesis populations of individuals, constituted of binary string genotypes and phenotypes, are evaluated on their ability to match a target binary string, nominally known as the environment, with success measured by the Hamming distance between the phenotype and environment. Three classes of environment are used.

1. Static environments in which an environment's target string remains unchanged.
2. Consistently variable environments in which each locus of an environment's target string switches on or off at regular, frequent, intervals.
3. Increasingly variable environments in which the frequency of change increases over the period of evolution.

For each class of environment, populations exhibiting combinations of genetic evolution, individual learning and social learning are evaluated, with the learning rates of both individual and social learning allowed to evolve. Mean population fitness is recorded for each combination of environment and learning strategy, with data also collected on the evolved rates of social and individual learning and the reproductive fitness of individuals exhibiting different learning rates when both extra-genetic learning strategies are combined.

Our expectations were as follows.

1. Social and individual learning strategies, both separately and in combination, will outperform genetic evolution on all environments.
2. When evolved simultaneously social learning will be favoured over individual learning, with individuals exhibiting higher levels of social learning having a higher reproductive fitness, thus showing that social learning is adopted over individual learning in increasing and consistently variable environments.

## 5.2 The Model

The model used is a genetic algorithm with steady state selection, in which individuals, constituted of binary string genotypes and phenotypes of length  $L$ , are assessed on their ability to match a binary target string or, as we shall refer to it here, an environment denoted as  $E$  (also of length  $L$ ). A phenotype is assessed by measuring the Hamming distance between it and the environment. A phenotype is initially a copy of the genotype but can acquire information through learning, which is discussed in more detail later. This may be achieved by one of four strategies.

1. Genetic Evolution - at reproduction random mutations occur with probability  $p_{mut}$  at each locus.
2. Individual Learning - at each epoch (iteration of the steady state genetic algorithm) every individual flips each of the bits in its phenotype with probability  $p_{ind}$ . If a change due to individual learning doesn't lead to a improved fitness then it is reversed.
3. Social Learning - at each epoch every individual copies each locus from a random other individual's phenotype with probability  $p_{soc}$ .
4. Individual and Social Learning (Combined) - at each epoch every individual engages in either individual learning or social learning, with equal probability, at each locus in the phenotype.

The learning rate (per locus probability of flipping or copying) is allowed to evolve independently for each individual. That is to say that a population wide learning rate is not set. Both  $p_{ind}$  and  $p_{soc}$  are floating point values bounded within the range  $[0, 1]$ .

### 5.2.1 Variable Environments

Populations are tested on one of the three environmental set-ups introduced earlier, two of which exhibit some level of variability. Variability is dictated by a sine wave, resulting in cyclic changes. At initialization each locus  $l$  in the environment is assigned a random value  $f$ , which is used to determine the binary value of the environmental locus at each epoch (5.1).

$$E^l = \sin((f^l \times epoch) \times (\pi/180)) \begin{cases} \leq 0 \rightarrow 0 \\ > 0 \rightarrow 1 \end{cases} \quad (5.1)$$

The range of values  $f$  may be initially set to is determined by which environment the population is being tested on:

1. No Variability (static):  $f = 0$
2. Consistent Variability:  $f \in N(1.8, \frac{1.8^2}{2})$
3. Increasing Variability:  $f \in N(0.018, \frac{0.018^2}{2})$

$$f^{epoch} = f^0 + (f^{max} - f^0) \times (\frac{epoch}{epoch^{max}}) \quad (5.2)$$

Values of  $f \approx 1.8$  equate to approximately one change per 100 epochs, with 100 epochs being considered to be one generation of the algorithm (where  $L = 100$ ). A value of  $f \approx 0.018$  equates to approximately one change per 10000 epochs, or one hundred generations. One change per generation is referred to as *high* frequency variability, one change per ten generations as *medium* frequency, and one change per one hundred generations as *low* frequency. As each environmental locus has a unique initial value of  $f$ , the sine wave dictating the value at each locus will be different, thus avoiding uniform environmental change and resulting in changes one each locus that are out of phase and at slightly different rates.

For increasing variability tests the  $f$  values increase over time. The  $f$  value for any environmental locus ( $E^l$ ) during increasing tests is determined by the initial  $f$  value at that locus ( $f^0$ ), the maximum  $f$  value ( $f^{max} = 1.8$ ), the current epoch and the number of epochs the evaluation is permitted to run for (5.2).

### 5.2.2 Evolution and Learning

Each test is populated by  $N$  individuals, each constituted of the following:

- $g \in \{0, 1\}^L$ - *genotype*, an  $L$ -bit string
- $h \in \{0, 1\}^L$ - *phenotype*, an  $L$ -bit string initially equal to  $g$  but subject to learning. The individual's fitness is  $L$  minus the Hamming distance between  $h$  and  $E$ .
- $p_{ind} \in [0, 1]$  - *individual learning rate*, set initially to 0. In populations allowed to learn in this manner  $p_{ind}$  may evolve via mutation.
- $p_{soc} \in [0, 1]$  - *social learning rate*, set initially to 0. In populations allowed to learn in this manner  $p_{soc}$  may evolve via mutation.

These properties are broadly consistent with the properties used by Jones and Blackwell [89]. However, unlike Jones and Blackwell [89] the learning rates are not normalized to sum to unity, instead each rate may evolve to a maximum value of 1.

At each epoch two individuals are selected at random from the population for tournament selection. Reproduction then takes place between the tournament winning individual (the one with the higher fitness) and a random individual from the population, the progeny of this reproduction replacing the tournament loser. Reproduction consists of both recombination and mutation. Recombination is by way of single point crossover, where a random position  $l \in [0, L - 1]$  is selected. Bits 0 to  $l$  being taken

from one of the parents and bits  $l + 1$  to  $L - 1$  from the other, with the order of the parents determined at random at each reproduction. Mutation occurs at each locus in the child’s genotype, with probability  $p_{mut} = 1/L$  of the bit at that locus being flipped. Following reproduction  $g$  is copied without error to  $h$  which from this point in the child individual’s lifetime is used for fitness evaluation and learning. In learning populations parental values of  $p_{ind}$  and  $p_{soc}$  are also inherited (depending on the learning strategy implemented for the population). The child inherits one of its parents’ learning rates at random, with the learning rate then being mutated by the addition of Gaussian random noise (mean 0, standard deviation 0.01).

Learning comes in two distinct strategies: individual and social. At each epoch all individuals from a learning population are afforded the opportunity to learn. Individual learning takes the same form as mutation at reproduction, with each locus in  $h$  bit-flipping with probability  $p_{ind}$ . Social learning on the other hand is a little more involved: for each locus in  $h$  there is a probability  $p_{soc}$  of copying the tournament winning individual’s equivalent locus. Copying the tournament winning individual in social learning strategies may be seen as akin to the “*copy-successful-individuals*” strategy outlined by Laland [93] and implemented (though in a slightly different manner) by Jones and Blackwell [89]. Social learning is not subject to any noise, with socially learnt information being copied exactly in to an individual’s phenotype. In those populations exhibiting both individual and social learning in combination, which of the two learning strategies to use is chosen at random (50:50) for each locus of each individual, and applied with the appropriate learning rate. Individuals are also afforded the opportunity to unlearn any learned information. Each individual maintains a copy of their phenotype from before learning; if after learning their fitness is less than it was during the previous epoch, their previous phenotype is restored. However, the indi-

vidual’s previous phenotype is not assessed in the current epoch, so whilst unlearning information stops individuals obtaining new maladaptive information, an individual’s fitness may still fall due to their previous phenotype performing worse in the current epoch than the previous epoch (should the environment have changed).

### 5.3 Experimentation and Results

Experimentation was initially conducted on the static, consistently variable and increasingly variable environments. Forty initially random populations of size  $N = 100$  were tested for each environmental set-up: ten populations per learning strategy. Each environment, of size  $L = 100$ , was initially identical in its binary composition, as was the random number seed from which the initial  $f$  values were derived. Each population was run for 100000 epochs (1000 generations), with the population being sampled every 100th epoch (once per generation). The data presented here takes the mean performance of each of the ten populations per learning strategy at every generation.

A set of further tests were also conducted to assess in which conditions of environmental variability populations were likely to collapse. These tests were conducted in two differing set-ups. In both set-ups  $N$  was maintained at 100 but before standard tournament selection took place all individuals with a fitness less than  $L/2$  were killed, these individuals being deemed to be unfit. If at this point the new population size  $N' \leq N \times 0.1$  the population is considered to have collapsed and evolution is terminated. If the population does not collapse, tournament selection takes place to replace one surviving individual, and the population is then re-populated to  $N = 100$  by the progeny of randomly selected other surviving individuals. The first test set-up was conducted for a maximum of 100000 epochs, with populations reaching this epoch being considered as surviving populations.

The second population collapse test set-up differs from the first in three distinct ways: tests were simulated for 200000 epochs; only populations exhibiting the individual and social learning strategies combined were tested; and social learning was prohibited from being used or evolving for the first half of each experiment.

### 5.3.1 Static Environments

As can be seen from Figure 5.1(a), under static conditions both social learning and individual and social learning combined to perform much better than genetic evolution and individual learning. These results are broadly consistent with those of Jones and Blackwell [89] who also found social explorations to be advantageous and individual learning to be sub-optimal in static environments. However, unlike Jones and Blackwell [89], in these tests individual learning does not outperform no-learning (genetic evolution alone) over the entire simulation. This result is a little surprising given Hinton and Nowlan [82], which demonstrates that individual learning should be able to better guide evolution than random mutation alone. Individual learning is not highly expressed when used in isolation. Figure 5.4 shows that under unchanging environmental conditions individual learning does not achieve a maximum  $p_{ind}$  of above 0.2, this value being lower than in all other environmental conditions and significantly lower than  $p_{soc}$ , which in static environments achieves a value in excess of 0.7. Individual learning is also marginalized when expressed in combination with social learning. Figure 5.1(b) shows that when evolved together social learning outstrips individual learning by some distance, with individual learning becoming almost unused after an initial spike before 1000 epochs. The fact that mean individual learning is maintained at a value above 0 is likely an artefact of the evolutionary model rather than individual learning being maintained at a low, but significant value; as learning rates cannot evolve to be below



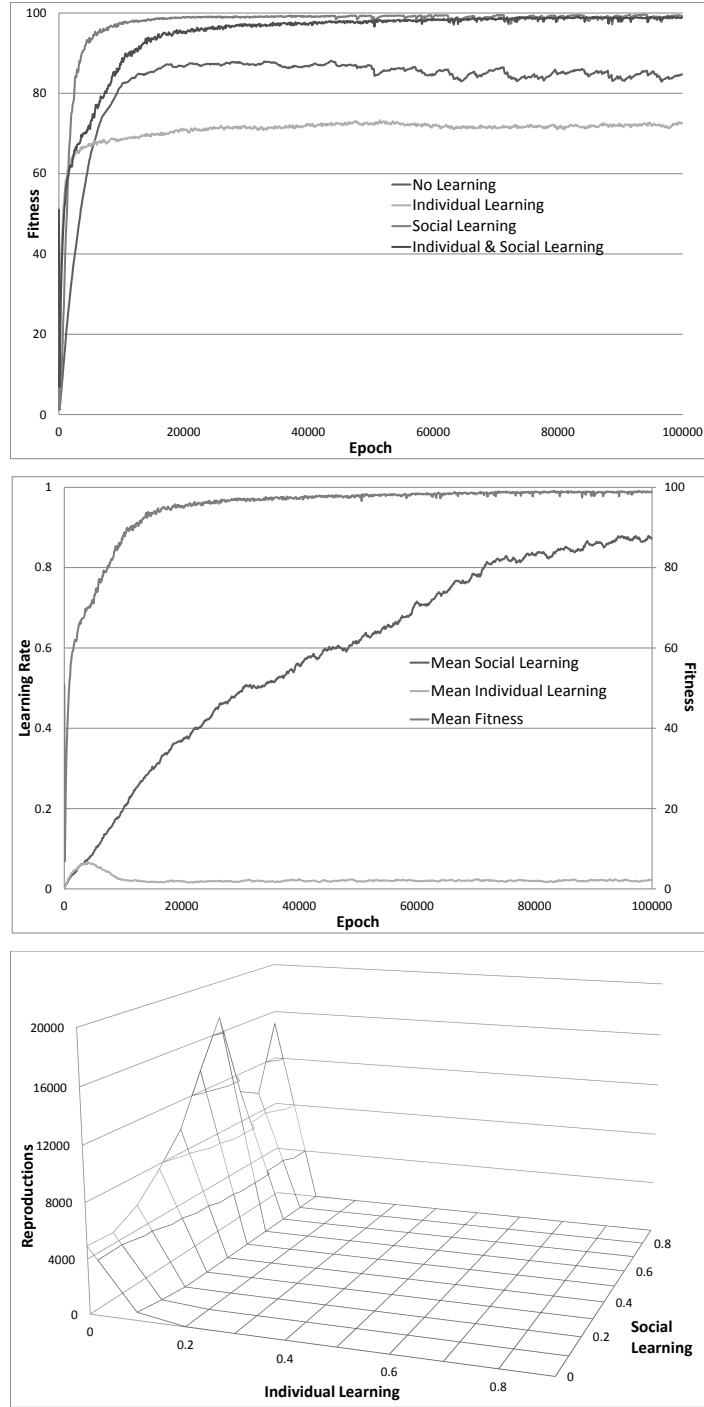


Figure 5.1: Fitness and reproduction rates in static environments. (a) Mean fitness of each learning strategy, (b) Mean fitness of individual and social learning with the evolved learning rates, (c) Reproductive fitness of combined learning rates.

0, any mutation to a learning rate of 0 will have a 50% chance creating a new learning rate of above 0, ultimately resulting in an individual learning rate which bounces off 0 rather than being maintained at 0. Interestingly, for static environments the maximum value of  $p_{soc}$  achieved is larger when individual and social learning are found together, than when social learning is evolved in isolation, implying that social learning requires individual learning to be fully expressed. As hypothesized social learning is adopted over individual learning, this adoption also being reflected by the reproductive fitness of individuals exhibiting the combined learning strategy as shown in Figure 5.1(c). Individuals exhibiting intermediate values for  $p_{soc}$  and low values (below 0.1) of  $p_{ind}$  are shown to be more reproductively fit by contributing to a larger number of reproductions over the evaluation period.

### 5.3.2 Consistently Variable Environments

As shown in Figure 5.2(a), under consistently variable conditions, where  $f$  is maintained at 1.8, the extra-genetic learning strategies all outperform no-learning (genetic evolution alone). In high variability environments non-learners find it difficult to track changes in the environment using mutation and recombination alone, causing populations of non-learners to average out at a fitness of  $L/2$ : no better than random. Of the extra-genetic learning strategies the combined strategy far outperforms individual and social learning alone. Individual learning when exhibited in isolation tends to find a stable value very quickly, but is unable to improve upon it. Social learning on the other hand rapidly (though also rather noisily) finds highly optimal solutions. However, the ever increasing reliance on social learning, as demonstrated by a maximum learning rate of above 0.9 (see Figure 5.4), causes social learners' fitness to decrease to a value equal to that of individual learners, demonstrating that overly conformist learning strategies

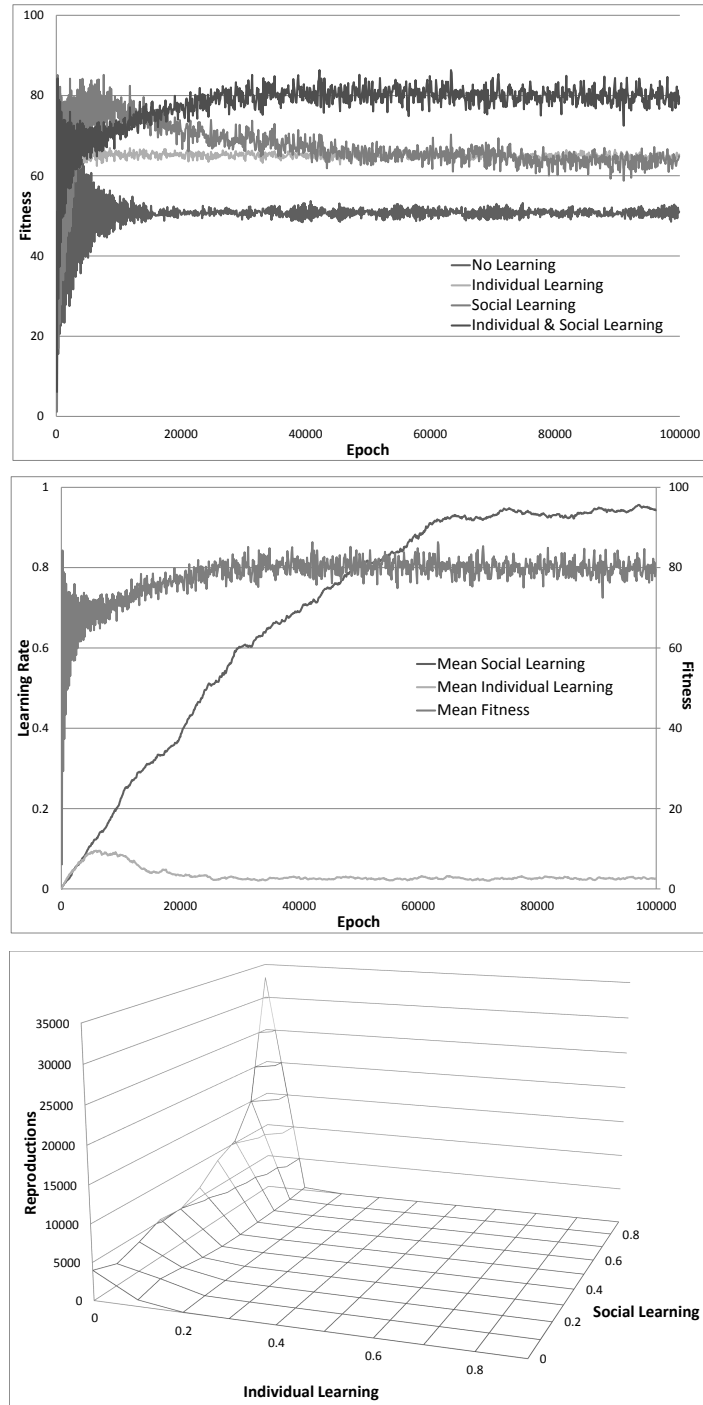


Figure 5.2: Fitness and reproduction rates in consistently variable environments. (a) Mean fitness of each learning strategy, (b) Mean fitness of individual and social learning with the evolved learning rates, (c) Reproductive fitness of combined learning rates.

are no better than trial-and-error personal innovations at tracking high levels of environmental change. By combining individual and social learning the negative aspects of both strategies in isolation vanish: fitness does not stabilize at a sub-optimal value early on and fitness does not decrease over time. This demonstrates that the conformist bias imposed by social learning is tempered by non-social innovation. However, as we can see in Figure 5.2(b and c) social learning is largely adopted over individual learning, with  $p_{ind}$  being sidelined to values well below 0.1 and highly reproductive individuals exhibiting high levels of social learning and low levels of individual learning. The initial spike in individual learning seen early in the combined strategy, while  $p_{soc}$  is also low, may indicate that the vast majority of innovation is introduced into the population before it becomes overly conformist. It is also interesting to note that the spike in  $p_{ind}$  correlates well with the noisiest fitness period. Once enough innovation is introduced into the population innovation appears to be sidelined, although maintained at a low level, and individuals become increasingly reliant on social learning.

### 5.3.3 Environments of Increasing Variability

Unlike in consistently noisy environments, all populations exhibiting extra-genetic learning strategies find it difficult to maintain high levels of fitness when confronted with increasing levels of variability (see Figure 5.3(a)). As the environment becomes more noisy individual learning rates begin to increase, possibly to reintroduce an element of personal innovation to the population, which has become stagnant due to the high levels of conformist learning imposed by large quantities of social learning during times of minimal variability. The reproductive fitness of individuals, as seen in Figure 5.3(c), is also interesting, as reproductively successful individuals tend to exhibit high levels of social learning and increased levels of individual learning, when compared

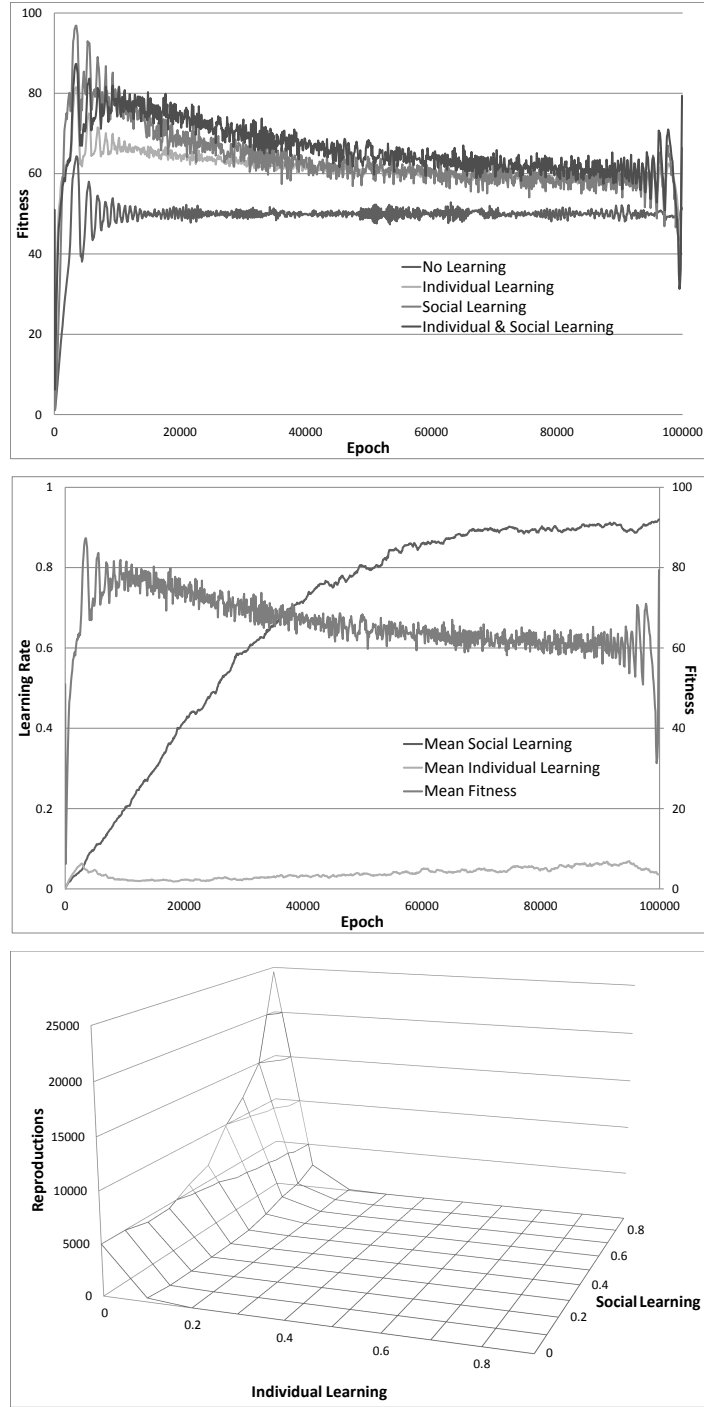


Figure 5.3: Fitness and reproduction rates in increasingly variable environments. (a) Mean fitness of each learning strategy, (b) Mean fitness of individual and social learning with the evolved learning rates, (c) Reproductive fitness of combined learning rates.

to the reproductive fitnesses of individuals in consistently variable or static environments. The comparisons between maximum learning rates for social and individual learning on increasingly variable environments (see Figure 5.4) is worthy of some note: despite individual learning being a necessary component of the combined strategy, it is not exhibited to as high a degree as when found alone; conversely social learning is always exhibited at higher levels when accompanied by individual learning. This again provides evidence that, while social learning is adopted over individual learning, individual learning is necessary for social learning to be used to greatest effect [4, 3]. Evidence from all stages of environmental variability tell a similar story, though to different degrees: social learning is widely adopted over individual learning when found together, with all extra-genetic learning strategies performing better than random on all tests. Extra-genetic learning strategies are also exhibited at higher levels in noisy environments than in static environments. The evidence presented does suggest that increasing variability is sufficient to cause the adoption of versatile survival strategies such as learning, with social learning being the learning strategy of choice.

#### **5.3.4 Population Collapse in Variable Environments (Consistent and Increasing)**

One of the pitfalls of the kind of genetic algorithm used so far is that even when populations are poor at the task, they still survive; of course this is not the case in nature. To explore whether or not the learning strategies implemented in this model are really robust we have also implemented a set of tests where populations may become extinct. The first tests follow the test set-ups above, with populations exhibiting different learning strategies being tested on environments with consistent and increasing variability. Populations falling below  $N \times 0.1$  individuals are considered

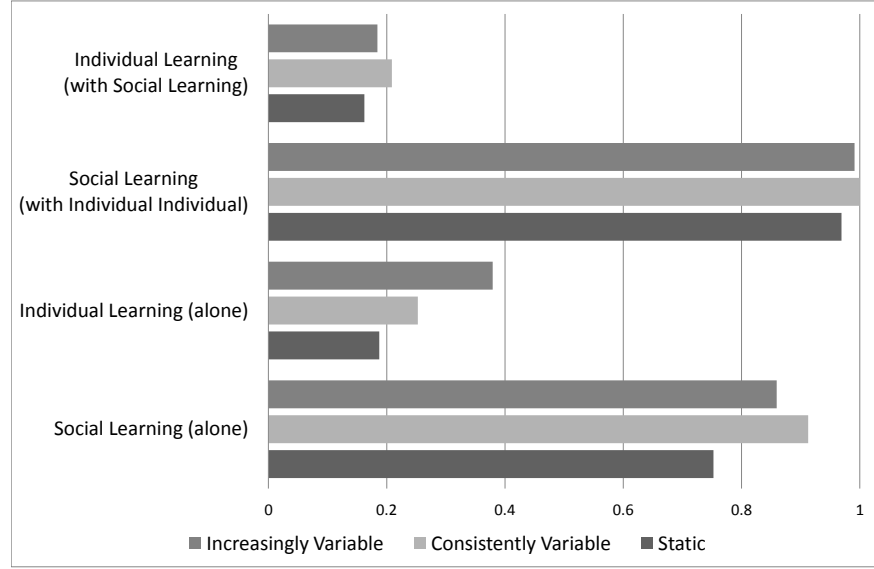


Figure 5.4: Maximum learning rates exhibited over all environmental test cases for all learning strategies.

as being collapsed.

Consistently variable environments were produced with four levels of variability;

1. No variability (static):  $f = 0$
2. Low variability:  $f \in N(0.018, \frac{0.018^2}{2})$
3. Medium variability:  $f \in N(0.18, \frac{0.18^2}{2})$
4. High variability:  $f \in N(1.8, \frac{1.8^2}{2})$

The percentages of populations surviving until the end of evaluation are reported in table 5.1. As may be expected, populations are unable to survive highly variable environments as the increased chance of death makes it all but impossible to re-adapt to new environments. However, individual learning is more robust than all other strategies, achieving a 50% survival rate on high frequency environments; higher rates of

Learning Strategy	Static	Low	Medium	High
Genetic	100%	100%	100%	0%
Individual	100%	100%	100%	50%
Social	100%	100%	90%	0%
Individual & Social	100%	100%	100%	0%

Table 5.1: Population survival in consistently variable environments. % of populations surviving until the end of the simulation for each learning strategy in consistently variable environments.

individual learning, though risky, are better able to deal with sudden environmental shifts. Social learning on the other hand begins to struggle in environments exhibiting medium amounts of variability. As with our earlier tests conformism spreads through the population, increasing the likelihood of population collapse. Combining individual and social learning alleviates the problem to some extent.

Increasingly variable environments were produced at three initial levels of variability: static, low and medium. In these environments variability increase throughout evolution, to a level of high variability.

Learning Strategy	Static	Low	Medium
Genetic	0%	0%	0%
Individual	100%	100%	100%
Social	0%	0%	0%
Individual & Social	0%	0%	0%

Table 5.2: Population survival in increasingly variable environments. % of populations surviving until the end of the simulation for each learning strategy in increasingly variable environments.



Unlike in consistently variable environments all learning strategies, excluding individual learning alone, result in populations that are unable to survive in any increasingly variable environment (see table 5.2); social learning completely undermines individual learning when combined, owing to over-conformism in times of lower variability stagnating the population’s pool of knowledge to the point that the increase in individual learning, usually seen later in increasingly variable environments (see Figure 5.3(b)) is insufficient to redeem the population’s fortunes.

As indicated by tables 5.1 and 5.2, individual learning is the only learning strategy robust enough deal with increasing and high levels of environmental variability. However, in early tests the combined strategy of both individual and social learning was seen to be adaptive in all environmental settings. To investigate whether individual learning is necessary for the successful introduction of social learning we implemented a final set of tests. In these, individual learning was allowed to evolve in isolation for 100000 epochs before the introduction of social learning alongside it for a further 100000 epochs. These tests provide a greater challenge for populations as they are required to survive for twice the evaluation period previously tested. However, this increase in evaluation time does reduce the rate at which environmental variability increases during increasing-variability tests.

As table 5.3 shows, the evolution of individual learning prior to social learning does provide some benefits in increasingly variable environments, but only when beginning from medium levels of variability ( $f = N(0.18, \frac{0.18^2}{2})$ ). Noisier environments provide a greater selection pressure for high levels of innovation, which in turn introduces a larger pool of knowledge for social learning to access; or that the lower rate of increase in variability is significant. A sensitivity analysis will need to be conducted to analyse the precise learning rates, reproductive fitnesses and death rates exhibited in these

“*goldilocks*” conditions.

Variability	Static	Low	Medium	High
Consistent	100%	100%	100%	0%
Increasing	0%	0%	100%	N/A

Table 5.3: Population survival in variable environments when individual learning is pre-evolved. % of populations utilising the combined Individual and Social strategy surviving until the end of the simulation when individual learning is allowed to evolve before the introduction of social learning.

## 5.4 Conclusions and Future Work

Reader and Laland [124] have demonstrated that personal innovations (individual learning) and social learning co-vary across species. The above results go some way to explaining why social learning was adopted most strongly when combined with individual learning; individual learning is necessary for effective social learning and provides an effective mechanism for avoiding population collapse. Whilst social learning alone can maintain adaptive knowledge in the population, over-reliance on it can just as easily reinforce sub-optimal or incorrect knowledge when the environment is highly stochastic, potentially causing the population to collapse [160]. By maintaining a level of personal innovation alongside social learning, populations can maintain non-conformist local search whilst ensuring that useful innovations are transmitted over generations [4]. However, in environments of lower variability conformist social learning ensures a high level of individual fitness. Individual learning on the other hand may impose unnecessary local search which could cause individuals to lose useful adaptations if high levels of individual learning are maintained. The data presented here demonstrates

that when environments are in minimally variable states individual learning plays a smaller role than it does in more variable environments. It is also found to be the case that mortality is greatly increased in environments of high or increasing variability when social learning is exhibited unless individual innovation is allowed to develop in isolation [3].

Our initial hypothesis (developed in order to test Potts’s variability selection hypothesis), that when individual and social learning rates are evolved simultaneously, both increasing and consistently variable environments are sufficient for the adoption of social learning over individual learning, holds true here, though with two main caveats: individual learning is required for successful social learning, and population collapse may only be avoided when individual learning is allowed to pre-evolve in already noisy environments before the introduction of social learning. Both of these caveats require further investigation in steady state genetic algorithms, neural networks [45] and grounded animat simulations [22].

The way noise is implemented also requires further investigation. Sine waves, though used elsewhere to model environmental variation [70], are not the only pattern of environmental variability found in nature. Further tests could include empirically derived data sets [70] or *red noise*, otherwise known as random walk noise (a kind of signal noise produced by Brownian Motion)[160].

# Chapter 6

## Social Learning Strategies

### 6.1 Introduction

Previous research has shown that with the use of social learning, individuals are able to discover more complex behaviours that are not accessible via incremental genetic evolution alone [22]. In this work, and many other simulation models that explore social learning and culture, social learning itself is often limited. These limitations are often centred around who individuals learn from. Here we expand on this previous work (discussed in Chapter 4) to explore whether behaviours inaccessible to incremental genetic evolution alone are still discovered, and maintained, when individuals are permitted to learn from a variety of different individuals. We go on to discuss why these differing teacher-learner social learning strategies solve the task used here in differing ways.

### 6.1.1 Discovering and Maintaining Behaviours Inaccessible to Incremental Genetic Evolution Alone

To solve the issue of sub-optimal population convergence without crossing the error threshold [118], noise can be added to the fitness landscape via the genotype to fitness map. However, depending on where such noise is in the phenotype to fitness section of that map, its ability to aid in the transition between peaks is limited. By instead incorporating noise into the genotype to phenotype map, behaviours inaccessible to incremental genetic evolution may be exhibited reliably by individuals while leaving the genotype untouched. One method for introducing noise in this way is to introduce transcription errors when writing from the genotype to the phenotype in systems with equivalent genotype and phenotype encodings, such as direct artificial neural network weight encodings [22]. By introducing potentially new behaviours to the phenotype we deny the initial possibility of these behaviours being inherited by new individuals through standard Darwinian evolutionary mechanisms. Therefore in order to maintain successful behaviours in the population, some form of extra-genetic learning needs to take place. The extra-genetic learning employed in this model is a combination of the aforementioned genotype to phenotype noise and social learning through interaction between teachers and learners to facilitate the transmission of learned behaviours [5, 45]. As in Borg et al. [22], learners or pupils follow teachers in a mock evaluation on a set of environments or maps. As both teacher and pupil receive the same environmental input the teacher’s output may be used as a target pattern for error back-propagation, reducing the pupil’s output error compared to that of the teacher. By learning in this way pupils are able to partially imitate the behaviours exhibited by teachers, thus maintaining aspects of new behaviours in the population that would have been lost by a stand alone evolutionary process.

The use of teacher-learner social learning has been shown to be sufficient for discovering and maintaining behaviours inaccessible to incremental genetic evolution alone in a grounded simulation [22]. However, these simulations only allowed one form of social learning, in which offspring would learn from their fittest parent. Though a valid approach that has been used in previous work [45], there are other theoretical and empirical models that can be adapted to this work to evaluate whether or not other social learning strategies are still capable of achieving these complex behaviours.

Social learning is seen widely in nature [123] and in a range of species as diverse as humans and nine-spined stickleback fish [90]. The mechanisms and processes that underpin social learning are themselves broad, ranging from teaching, imitation and emulation to stimulus enhancement and exposure [63], with any of these mechanisms potentially being adaptive [112] thus leading to the formation of traditions and cultures [164, 170]. However, within each social learning category there is some dependence on who information is obtained from, be it a teacher or which agent is unintentionally (or intentionally) exposing an individual to something new. As social learning is necessarily conformist, a poor learning model may result in the discovery and propagation of sub-optimal behaviours. In this work we assess whether who you are learning from, otherwise known as ‘who’ social learning strategies [93] (also discussed in Hoppitt and Laland [84] and Rendell et al. [126]), can hinder social learning’s ability to discover and maintain behaviours inaccessible to incremental genetic evolution alone, thus undermining social learning’s adaptive advantage over incremental genetic evolution in complex environments.

Laland [93] assess both ‘who’ and ‘when’ social learning strategies, alongside the complexity of social learning in animals, providing evidence to show its adaptive advantages. Laland [93] has a particular focus on conformity: a population’s ability to

share popular behaviours amongst each other while minimising exploration for new behaviours; the use of conformist social learning can be beneficial or detrimental depending on the environment or task [93, 23, 110]. It has also been suggested that conformist social learning that is not supplemented with non-social exploration can lead to population collapse in temporally varying environments [23], though recent work suggest that conformist learning may be of benefit in spatially varying environments [110]. The ‘who’ social learning strategies (concerned with who an agent should learn from rather than when learning should take place) inspired by Laland [93] are modelled here as three core social learning strategies: ‘Best Parent’, ‘Oldest’ and ‘Fittest’. The ‘Fittest’ strategy selects the fittest individual from the population to be the teacher. The theoretical basis behind this strategy falls partially into the ‘Learning from majority’ category discussed by Laland [93], but also has a wider basis in nature with many animals being shown to learn from more successful individuals. Learning from older individuals derives from the rationale that older individuals must have exhibited successful behaviours to survive, however this does not have to mean the older individual in question is in fact the fittest individual, due to this the ‘Oldest’ strategy is likely to provide a broader range of behaviours than the ‘Fittest’ strategy. The ‘Best Parent’ strategy (as seen in Borg et al. [22]) sets the teacher to be the parent who has won the right to reproduce in a tournament. This is the least conformist strategy of the three as it allows unfit individuals, relative to the rest of the population, to be parents as tournaments only involve a small number of individuals. Additional to these three core strategies we also introduce social learning strategies for learning from random and young individuals. Though not widely evident in nature, the theoretical benefits of learning from a random individual (sometimes described as unbiased social learning) have been have been discussed in numerous works [123, 110]. The theorised benefits of

unbiased social learning arise in temporally varying environments, where learning from a broader set of individuals enables increased access to new behaviours that may be relevant in the specific environmental state being experienced. A ‘Youngest’ strategy, despite no theoretical basis and younger individuals being shown to be poor social models [1], is being evaluated as a contrast to the ‘Oldest’ strategy.

## 6.2 Experimentation

The model used here is fundamentally the same as introduced by Borg et al. [22]. Each iteration/generation has a tournament event in which two individuals from the population of 100 undertake the RC+ task, with each individual’s fitness being determined by the number of maps successfully completed. Each map gets increasingly more complex, so, if an individual is not able to complete a map, they are prevented from continuing on to further maps. Each map has seven Trap objects and  $20 - (5 \times \text{riverwidth})$  Stone objects, both of which are randomly placed, though never on the same space, one reward-giving Resource on the opposite side of the map to the agent starting position, and one instance each of Object A and Object B. The river width varies from an initial width of zero, increasing by one cell per map. Each individual is evaluated on their ability to reach the resource or place Object A and Object B on to a cell containing Water. Agents fail when they come into contact with an uncovered Water or Trap element. Failing to complete a map within 100 steps is also evaluated as a failed attempt. The two tournament individuals are compared, with the fitter agent reproducing with a randomly selected agent from the population, with the child replacing the weaker of the tournament agents. Each loci in an agent’s genotype directly writes to a locus in the agent’s phenotype, which itself directly encodes a weight in the decision network, with all genotypes and phenotypes being of length  $L = 308$ . To



ensure network structures from parents are maintained during reproduction, a single point recombination mechanism is applied. Mutation follows recombination; each loci has a probability  $P_{mut} = 1/L$  of having a random value from  $N(0, 0.4)$  added to it, with the resulting values being bounded within the range  $[-1, 1]$ . Once the child genotype has been constructed it is written to the child agent’s phenotype; this process is referred to as transcription. During transcription two randomly selected connection weights are overwritten with a new random value selected from a discrete uniform distribution  $U(-1, 1)$ . Directly following reproduction the learning strategy is enforced via back-propagation. A mock evaluation of the RC+ task takes place between the teacher and child (now thought of as the learner), with the learner’s inputs being set to those of the teacher. Learning takes place until the teacher either fails or completes all five maps. At each step through the evaluation the learner attempts, via error-back propagation with a learning rate of  $\delta = 1$ , to imitate the teacher’s output for the current inputs. A learning rate of  $\delta = 1$  was deemed necessary in order to enable any learning to take place as at each mock evaluation time step only one iteration of back-propagation is applied. Running multiple iterations of back-propagation per mock evaluation time step was decided against in order to avoid the learning individual becoming essentially a copy of the teacher.

The model in this work utilises five learning strategies, each with a different way of determining teacher selection. The winner of the reproduction tournament being set as the teacher in the ‘Best Parent’ strategy, the fittest individual in the population for the ‘Fittest’ strategy, the individual who has registered the most tournament wins for the ‘Oldest’ strategy, the last animat to be created before the current reproduction event in the ‘Youngest’ strategy, and a random individual for the ‘Random’ strategy. In any case where more than one individual met the criteria to be assigned the role of

teacher, an individual from the valid sub-set was chosen at random, this situation only every arose when using the ‘Fittest’ or ‘Oldest’ strategies. One hundred populations for each learning strategy were evaluated so the results can be aggregated for an overview of each strategy’s performance. Simulations were run for 2,000,000 tournaments, with each simulation recording the fitness of the fittest individual and the mean fitness of the population at every 500th tournament. The highest fitness is five, which indicates an agent completed map five. To indicate the behaviour has not only been achieved but also maintained the fitness of five has to have been recorded a further ten times, without a suboptimal result. Each learning strategy is comprised of 100 populations of agents.

### 6.3 Results

Table 6.1 (top) shows the proportion of populations that were successful in solving each map. The most notable result was that all strategies were able to complete map five, the map which required exhibiting and maintaining a behaviour that in previous work was not obtainable by incremental genetic evolution alone [22], thus demonstrating that discovering and maintaining behaviours inaccessible to genetic evolution alone is possible using various teacher-learner social learning strategies, even those strategies that are either non-conformist (the ‘Random’ strategy) or contrary to strategies observed in nature (the ‘Youngest’ strategy). It should be noted that to complete one map, all preceding maps must have also been completed, therefore the ability to solve map five indicates that a population also managed to successfully complete maps 1-4. In Table 6.1 (top) we do see many instances of learning strategies failing to complete simpler maps; we also see this in Table 6.1 (bottom), which shows how many populations were successful at completing each map as their maximum achievement, that is to say com-

Map	BP	Fittest	Oldest	Random	Youngest
<b>1</b>	99%	99%	99%	99%	99%
<b>2</b>	71%	68%	54%	74%	63%
<b>3</b>	47%	47%	37%	54%	47%
<b>4</b>	39%	46%	34%	49%	38%
<b>5</b>	8%	15%	5%	10%	7%

Map	BP	Fittest	Oldest	Random	Youngest
<b>None</b>	1%	1%	1%	1%	1%
<b>1</b>	28%	31%	45%	25%	36%
<b>2</b>	24%	21%	17%	20%	16%
<b>3</b>	8%	1%	3%	5%	9%
<b>4</b>	31%	31%	29%	39%	31%
<b>5</b>	8%	15%	5%	10%	7%

Table 6.1: RC+ map completion for each social learning strategy. (top) % of populations completing each map for each social learning strategy. (bottom) % of populations achieving each map as their maximum achievement for each social learning strategy. (BP = Best Parent)

pleted map one or two, ... without going on to complete any later maps. Maps 2-4 were all solvable using either a ‘bridge building’ strategy or the more advanced Object A + Object B strategy; therefore some learning strategies sometimes failed to find the sub-optimal, but more incrementally accessible, ‘bridge building’ strategy. We would also expect to see populations that were able to complete map two also completing map four as the behaviour required is the same, the only difference being a wider river, however Table 6.1 (bottom) shows that all strategies had populations that exhibited flawed behaviours which were not as generally applicable as they should have been. In comparable tests by Borg et al. [22], non-learning populations were shown to achieve above 90% success on maps three and four, with 100% success for maps one and two, the failure of the social learning strategies explored here to achieve this rate of success for maps three and four (as indicated by Table 6.1 (top)) indicates that whilst social learning can enable access to, and maintenance of, behaviours inaccessible to incremental genetic evolution, they are less effective at solving simpler, incrementally accessible, tasks. One explanation for this result is that social learning is necessarily conformist, even when unbiased or random, thus running the risk of sub-optimal behaviours being maintained and dispersed within the population.

The results also offer no definitive best strategy for the solving the RC+ task, as all are able to achieve the final map. However both Table 6.1 (bottom) and Fig 6.1 do allow us to begin seeing the differences between strategies. Performance may be viewed from three differing perspectives: (1) the number of populations achieving map five, (2) the distribution of maps achieved by populations, (3) the speed at which populations were capable of completing maps. Both measure (1) and (2) may be considered using the data from Table 6.1 (bottom): from this data we can see that ‘Fittest’ strategy achieves the highest proportion of populations completing map five, however if we conduct a

Chi squared test to find whether the proportion of populations achieving map five is dependent on the social learning strategy applied or not we come out with a p-value of 0.1316, thus indicating that the proportion of populations achieving map five is in fact independent of the strategy applied, therefore we cannot say with any certainty that the ability of the ‘Fittest’ strategy to achieve map five is significantly better than any other strategy (we do find that a Chi squared test that only considers the ‘Fittest’ and ‘Oldest’ strategies does provide a p-value below a significance level of 0.05, but no other pairings do). If we take Table 6.1 (bottom) to be a contingency table on which a Chi Squared test may be conducted we may be able to derive whether the distribution of maps achieved by populations (measure (2)) is dependent or independent of the social learning strategy used. When such a test is conducted a p-value of 0.04739 is produced, demonstrating that the distribution of maps achieved by populations is dependent on the strategy used. This result requires further investigation of the data for each population, for each strategy, in order to provide a robust overview of the dynamics each strategy employs to solve the task. Measure (3) may be considered using the graphs seen in Fig 6.1. From Fig 6.1 it we see that populations employing the ‘Best Parent’ strategy are able to achieve map five quicker than other strategies, with the ‘Youngest’ strategy struggling to achieve map five in any haste. However if we only consider the average number of generations to complete each map both ‘Best Parent’ and ‘Youngest’ give an average performance, with ‘Random’ and ‘Oldest’ giving the best general performance. It is interesting to note that those populations employing the ‘Oldest’ strategy who are able to complete map five, do so quicker on average than ‘Oldest’ strategy populations that complete map two, three or four this result indicates that when individuals in ‘Oldest’ strategy populations do discover the behaviour required to solve map five, it spreads rapidly through the population. As the ‘Oldest’ strategy acts somewhat

like a ‘Dominance’ strategy, with only the dominant tournament winning agent acting as the teacher, it is unsurprising that behaviours can spread rapidly, however the random nature of tournament selection can somewhat undermine this strategy’s ability to guarantee fit behaviours or a consistent teacher. The best performing populations for the ‘Best Parent’, ‘Oldest’ and ‘Fittest’ strategies (as seen on the left of Fig 6.1) also indicates that once a favourable behaviour is discovered using these strategies it is able to spread reasonably quickly. This is unsurprising as each of these strategies can be highly conformist, with successful individuals potentially having a monopoly on being the teacher for new agents. With the ‘Youngest’ strategy, the high turnover of teachers provides little opportunity for beneficial behaviours to take hold, though these teaching agents are the progeny of tournament winning parents, so can be expected to be reasonably fit. The most surprising result is the general performance of the ‘Random’ strategy, given that unlike the other strategies there is no guarantee of the teacher being either consistent nor particularly fit. One reason for the ‘Random’ strategy performing at least as well as the other strategies is the nature of the RC+ task itself. If a population only discovers the ‘bridge-building’ behaviour needed for maps 2-4, whilst forming a dislike for Object A and/or Object B, any conformist strategy will struggle to discover the behaviour required for map five, as the population will tend to conform to the sub-optimal behaviour. However, the very nature of the ‘Random’ strategy allows for a variety of individuals to fulfil the role of teacher, regardless of fitness, thus enabling newer ideas to potentially establish themselves and sub-optimal behaviours to be lost. However, maintaining these newly found optimal behaviours may be difficult in such a strategy. This does suggest that a hybrid approach may be beneficially, whereby numerous conformist and non-conformists strategies may exist within a population thus enabling both innovation and rapid behavioural convergence

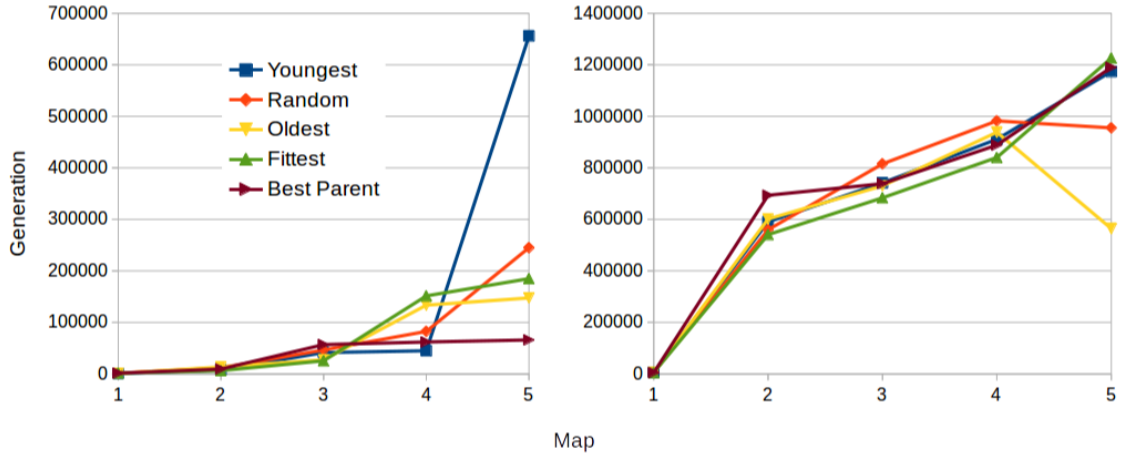


Figure 6.1: Time taken for RC+ map completion for each social learning strategy. (left) Graph showing the first time any population achieved each map. (right) Graph showing the average generation populations achieved each map.

to occur.

## 6.4 Conclusions and Further Work

The aim here was to demonstrate that multiple, varied, social learning strategies would be capable of discovering and maintaining behaviours that are inaccessible to hill-climbing strategies such as incremental genetic evolution. The results presented here echo previous work [22], while extending the research to show that various social learning strategies are capable of both discovering and maintaining inaccessible behaviours. Due to each strategy applied being highly abstracted from behaviours seen in nature, along with the task being highly artificial, this work is unable to draw strong parallels

to observed social behaviours in nature. Achieving a comparable status will require a more complex use of social learning, a sensible progression would be the inclusion of synchronous, distinct learning styles into a single population. A model that allows for multiple social learning strategies to be employed along side genetic evolution has compelling implications for agents, i.e. choosing optimal learning styles for the appropriate task.



## Chapter 7

# Evolutionary Adaptation to Social Information Use without Learning

As discussed earlier in this thesis, social information can broadly be thought of as information derived from the behaviours, actions, cues or signals of other agents [91]. As social information necessarily involves the direct or indirect broadcasting of information in to the public domain, it is sometimes known as (or conflated with) public information [20]. Here we assess whether the use of social information in populations of simulated neuroevolutionary agents is adaptive when decoupled from within-lifetime learning processes. Within-lifetime learning processes confer significant adaptive advantages to agents employing them, be it through the development of a set of robust and flexible behaviours, the rapid adaptation to new environments or circumstances, the quick incorporation of new information, or the guiding of the evolutionary process itself [116]. The adaptive advantages of learning are particularly potent when social information is incorporated alongside innovation and individual learning [22], resulting in social learning and potentially even cultural evolution [164]. However, as beneficial as within-lifetime social learning processes are, it is unclear to what extent social or public

information has an adaptive benefit when decoupled from these learning processes and evolution is left to determine the value of social information. Is the incorporation of social information alone enough to gain an adaptive advantage over non-social agents? Or are learning processes necessary to allow social information to confer any benefits? These individual are the questions that we address in this chapter.

Social learning is seen widely in nature [123] in a range of species as diverse as humans and nine-spined stickleback fish [90]. The mechanisms and processes that underpin social learning are themselves broad, ranging from teaching, imitation and emulation to stimulus enhancement and exposure [63], with any of these mechanisms potentially leading to the formation of traditions and cultures [164, 170]. However, within each social learning category there is some dependence on who information is obtained from, be it a teacher or which agent is unintentionally (or intentionally) exposing an individual to new information. As social learning is necessarily conformist, a poor social information model may result in the discovery and propagation of sub-optimal behaviours [87]. Despite the potential pitfalls of over-conformist social learning, including sub-optimal behaviour development [87] and even population collapse [160, 23], social learning, and therefore social information transfer, can be of great benefit to agents, thus explaining why even simple forms of social information transfer are seen so widely in nature [126, 63, 153] and have been shown to produce complex behaviours that are easily attributed to more complex social learning mechanisms like imitation [111, 113]. At the heart of the problem being addressed here are three core arguments. (1) Information is a fitness enhancing resource [107], even when information suppression is seen to be adaptive [109] or when information is encoded or interpreted incorrectly [107] - any new information about the world enables populations of agents to better adapt to the world they are in, even if this means disregarding or suppressing information.

(2) Incremental evolution is not a process of unguided random variations, but a process that itself can adapt in a way that is analogous to the kind of learning seen in cognitive organisms [157], leading to complex and robust adaptive traits in nature, autonomous robots [57] and simulated agents [42] in the same way that learning can lead to complex and adaptive behaviours (though on a different time scale). (3) Inadvertently expressed public information and simple mechanisms of social information transfer can lead to behaviours that are sufficiently complex to enable cultural evolutionary processes [46, 164]. These three core arguments give us good reason to believe that social information without within-lifetime learning processes should still be adaptive, and therefore lead to evolution adopting the use of social information to the benefit of social agents over non-social agents. Though we must still be mindful that social information may be at odds with personal beliefs [46] or lead to population-level conformism to sub-optimal behaviours [87], thus leading to a trade-off between the accommodation of social information and the evolution of robust evolved behaviours.

This leads us to the hypothesis that agents making use of social information should outperform non-social agents: any additional information, that is not just noise, that provides more information about the environment should lead agents to an improved “performance” in the environment over agents without access to such information. However, social information may only be useful when it is obtained from a trusted or reputable source [127], thus accurately indicating success or indirectly leading to success, and therefore may provide little or no concrete benefit in complicated or less predictable environments - in these more challenging environments learning may be necessary to allow temporarily useful social information to be quickly adopted and then rejected when it is no longer relevant. This hypothesis will be tested by modelling populations of agents who have no social information available to them and popula-

tions of agents with various forms of social information available to them. Each social information strategy will be tested against the non-social strategy, starting initially with the most basic social strategy available: presence, with the null hypothesis in each case being that the social population does not show an improved ability to solve the task at hand compared to non-social agents. The social information strategies used here are: presence, action, health and age. Presence social information simply enables agents to detect the presence of other agents (non-social agents are essentially blind to other agents); action enables agents to see what other agents are currently doing; health enables agents to see the current energy or battery state of others; and age information enables agents to see how long others have lived for.

## 7.1 Environmental Set-up

The task world used here is known as EnVar. EnVar is a bounded (non-toroidal) 2D environment containing a variety of consumable resources known as plants. Plants are recognised by agents simply as an RGB value. Plants are divided into a number of species, each with a base RGB value and a radius in RGB space. Plants are generated within these RGB regions and identified as belonging to the nearest species according to euclidean distance in RGB space. Each plant species is assigned an energy value, which is transferred to agents if the plant of that species is consumed; energy values may be positive or negative. Notionally the EnVar world is broken up in to cells, though here each cell represents a pixel and therefore the world can be considered to be continuous. Plants in the world take up a number of cells, forming a block, with each block only being able to be eaten a certain number of times before being exhausted (here set to be 200 eating events). Once a plant block has been exhausted it is no longer consumable and therefore removed from the world to be replaced by a new block from a random

plant species somewhere else in the world - this maintains a constant number of food blocks in the world at any time. Agents are permitted to share space with a plant resource but cannot overlap with each other, thus removing the possibility of agents piling up on top of one another on valuable food resources. In this work EnVar is set up to create a  $700 \times 700$  pixel sized cell world, containing five hundred  $10 \times 10$  pixel blocks of plants. In order to test our hypothesis we test populations of social and non-social agents in a set of increasingly difficult environments. Environmental difficulty is dictated by the ratio of positive food resources to negative food resources. The simplest world used here has an equal (1 : 1) ratio of positive food species to negative food species. Tests get progressively harder by increasing the number of negative food species, whilst maintaining only one positive food species, resulting in the most difficult world used here having a 1 : 9 ratio of positive food species to negative food species. As each plant species has an equal chance of appearing in the world, and covers approximately the same portion of RGB space, agents in the most difficult environment are nine times more likely to experience a negative plant resources than a positive plant resource. In the results section below, environment 1 relates to a 1 : 1 ratio environment, with environment 9 relating to a 1 : 9 ratio environment. For all tests here negative food species come with an energy value  $E_{neg} = -10.0$ , with positive food species contributing an energy value of  $E_{pos} = 1.0$  when consumed. This provides a strong evolutionary pressure to avoid eating negative food species.

## 7.2 Neuroevolutionary Model

Agents in the EnVar simulation world are grounded 2D simulated agents, controlled by a hybrid neural network architecture known as the *Shunting Model*. The shunting model uses two interacting neural networks to determine agent behaviours, here rep-

resented as a discrete set of agent actions. The two interacting networks are known as the *Decision Network* and the *Shunting Network*. The decision network is simply a feed-forward neural network comprised of an input layer, one hidden layer and an output layer. Outputs from the decision network are used to produce a locally-connected, topologically-organised network of neurons known as the shunting network, which simply places and organises agent preferences for environmental features and states in such a way to allow the agent to hill climb in a shunting space (known as the activity landscape) that directly maps on to their immediate neighbourhood. The shunting network weights are fixed for all agents, whereas the decision network is genetically encoded and is subject to change via evolution.

### 7.2.1 The Shunting Network

The shunting network is a locally-connected, topologically-organised network of neurons that was originally used for collision free motion planning in robots [168, 167] and has been subsequently applied in a number of 2D and 3D artificial life models [128, 22, 139, 87]. Here the shunting network's topology is simply superimposed on to the environment, with each cell in the network topology directly relating to a pixel within an agent's visual field. Using the shunting equation (see equation 7.1) values for each cell (which can be interpreted as representing an environmental feature or state, and are initially set by the Iota output  $I$  obtained from the decision network) are propagated across the cells of the network, producing an activity landscape with peaks and valleys representing desirable and undesirable features in the environment. The result is a landscape which allows the agent to follow a route determined by the higher Iota values while avoiding undesirable valleys. A mock-up example of an activity landscape with a snapshot of the visual field it represents can be seen in Figure 7.1.

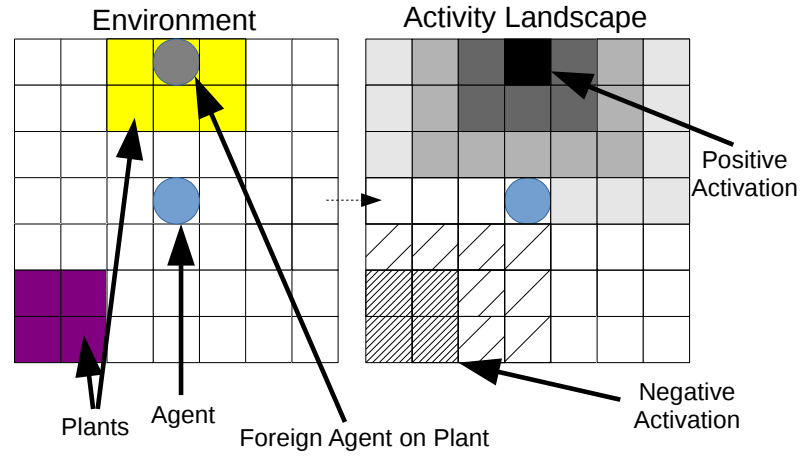


Figure 7.1: Mock-up transition from agent visual field to shunting network activity landscape. The left-hand grid shows the agent's visual field with two plant objects and one other agent occupying the same space as a plant. The right-hand grid shows an example activity landscape for the visual field. The agent determines that an agent on a plant is an interesting feature and therefore assigns it a strong positive Iota value ( $I$ ), whereas the purple plant is seen negatively and is therefore assigned a strong negative Iota value. These Iota values propagate over the activity landscape using equations 7.1 and 7.2. The central agent then chooses to move within its immediate Moore neighbourhood to the cell with highest activity value.

$$\frac{dx_i}{dt} = -Ax_i + \sum_{j \in N_i} w_{ij} [x_j]^+ + I_i \quad (7.1)$$

In equation 7.1 each node in the shunting network corresponds to one pixel within an agent's visual field;  $x_i$  is the activation of neuron  $i$ ;  $A$  is the passive decay rate;  $N_i$  in the receptive field of  $i$ ;  $w_{ij}$  is the connection strength from neuron  $j$  to  $i$ , specified to be set by a monotonically decreasing function of the Euclidean distance between cells  $i$  and  $j$ ; the function  $[x]^+$  is  $\max(0, x)$ ; and  $I_i$  is the external input to neuron  $i$  (known as the Iota value). The shunting network is advantageous as it exhibits computational efficiency by not explicitly searching over all possible paths. In line with the work of Stanton and Channon [139], we use a simplified, stable solution for equation 7.1 as seen in equation 7.2. Here constant  $x_i^{new} = x_i$  for all  $i$ . The maximum Iota value is  $\max I = 15$ , with the resulting value for  $x_i^{new}$  also being capped at a minimum Iota value  $\min I = -15$ . This stops Iota values growing out of control, whilst providing a large enough maximum value (and a small enough minimum value) to ensure activity propagation across the network. In order to allow propagation to occur within a time-step, the shunting equation must be run a number of times, we take this number of iterations to be equal to the diameter of the visual field.

$$x_i^{new} = \min \left( \frac{1}{8} \sum_{j \in N_i} [x_j]^+ + I_i, \max I \right) \quad (7.2)$$

The shunting model implemented here differs in a number of significant ways from previous Artificial Life implementations [128, 22, 139, 87]. In these previous implementations agents see their entire environment, have a set number of discrete environmental features and states to set Iota values for, and are in the environment alone to complete a predetermined task. Here agents have a limited view of the world, have the possibility of needing to set an Iota value for a plant of any given RGB value, and exist as a



population within the environment (leading to possible input states where an agent can be seen on a particular plant). In order to accommodate these differences the shunting model here is run independently for each pixel in an agent’s visual field, which is set here to have a radius of 30 pixels from the centre of the agent, with information about that pixel being included as part of the agent’s decision network input layer. In this way an Iota value is calculated for each unique environmental state within an agent’s visual field (in previous models, each discrete environmental state was included as an output, with only an agent’s internal state or current cell’s state being accommodated in the input layer of the decision network). This change does not change the resulting behaviour of the shunting model or activity landscape, just the way in which information is passed to the shunting network from the decision network. In order to minimise the amount of processing time required to populate and create the activity landscape, Iota values are only collected for unique states experienced by an agent - for a state to be unique it must be a newly experienced set of decision network inputs (discussed below). To further optimise processing time, an agent will only produce an activity landscape if its outputs determine that it should move in the current time step; agents that are not moving do not need an activity landscape.

### **7.2.2 The Decision Network, Neuroevolution and Reproduction**

Evolution in the model is applied only to the decision network. The decision network here is a feed-forward neural network comprised of seven standard input nodes, and an additional social input node in social information tests, eight hidden units, and two output nodes, resulting in 112 - 128 weights. Each network layer is fully connected, with floating point weights in the range  $[-1 : 1]$  being directly encoded from an agent’s

genotype. A standard sigmoid activation function is used at each hidden and output node, though outputs processed for deriving agent actions are then scaled to be within the range  $[0 : 1]$  and the Iota output is scaled linearly to be within the range  $[minI : maxI]$ . As the agent is expected to produce an Iota value to feed in to the shunting network for each unique environmental feature or state within its visual field, inputs into the decision network must accommodate both the internal state of the agent, the state of their current environment, and the state of the environmental feature they are assessing; this leads to there being two sets of input nodes. The first set of input nodes are simply plant RGB inputs - if the agent is viewing empty space these inputs are set to -1, else they are set to be the normalised RGB of the plant being viewed, with RGB values being normalised to be within the range  $[0 : 1]$  by way of linear normalisation. Following these inputs are a series of generic inputs, which are dependent on the agent's internal state and the current environmental state. These inputs are the agent's current battery level in the normalised range  $[0 : 1]$ , a moving average of the agent's battery level over the previous 100 time steps, the agent's current external environmental state and a moving average environmental state, which are both set to be +1 and do not change in the tests presented here (the model is set-up to accommodate external environmental change which is not used here). In social information tests agents have an additional input based on the agent they are viewing.

The genotype, which is essentially an array of weights, is subjected to both mutation and crossover should a reproduction event take place. The crossover mechanism used here is single point crossover, with per locus mutation occurring with probability  $p_{mut} = 1/L$ , where  $L$  is the length of the genotype. Mutation is achieved by way of Gaussian random noise, with a value taken from a normal distribution with  $\mu = 0$ ,  $\sigma = 0.01$  being either subtracted or added to the floating point value at the loci to be

mutated. All weight values are bounded in the range  $[-1 : 1]$ . Reproduction events take place only in response to a death event. Agents can die if they run out of energy, or if they are in the lowest 10% of agents ranked by energy at the end of an epoch. The first method for removing agents from the population ensures that agents cannot remain in the population with no energy, the second method ensures space is made for new agents to be created even if the population as a whole is successful at maintaining above zero energy levels, thus maintaining a selection pressure for task improvement. Both methods of death are not directly related to task ability as it is possible for a good agent to be unlucky and never, or rarely, experience a positive food resource, whereas less able agents may have the fortune to be born near an abundance of food resources or relatively close to the end of an epoch. This method of reproduction maintains a constant population size of 200 agents. The new agent, or child, created to replace the removed agent is the progeny of two agents, one of whom is selected in a tournament, the other of which is selected randomly from the remaining population. The tournament selection mechanism applied here takes two agents from the population, compares their current energy levels, and selects the agent with the higher energy level as a parent. Like in nature, this isn't a perfect measure of fitness as it is possible the agent is young and therefore has not yet had time to lose significant amounts of energy, or the agent could have simply been lucky or unlucky with available food sources. However, in general, agents with more effective behaviours will on average find themselves with better energy levels than agents with less effective behaviours, thus driving evolution toward behaviours that are more suited to the task or environment at hand. The second parent is selected randomly to ensure the population doesn't become dominated by the progeny of a small sub-set of the population, thus maintaining a level of exploration in the genotypic search space. New agents are placed in the world within

the visual field of one of their parents.

### 7.2.3 Agent Actions and Action Energy Costs

The agents in the model have a set of simple, discrete, actions available to them, through the output layer of their decision networks: wait, eat or move. The decision network has two outputs, an Iota output to be fed into the shunting network and an eat/wait output. The agent first considers its current input state at its current position - if the agent produces an Iota value above the threshold  $\theta_a = 0.5$  it indicates the agent is happy with its current state and position and therefore does not move (an activity landscape is therefore not calculated as it not needed). The agent's eat/wait output is then considered; if the output produces a value above the threshold  $\theta_b = 0.5$  the agent attempts to eat whatever may be at its current position; agents are welcome to try and eat at locations where no plant is present, but no benefit for this action is conferred, and the eat action is considered to be an unsuccessful eating attempt rather than a wait action. If an agent decided to eat at a location containing a plant, the plant's energy is transferred to the agent, this does not necessarily lead to the exhaustion of the plant resource, as plants are considered as a mass. The Iota output is in the range  $[-1 : 1]$ , which is then scaled to be within the range  $[minI : maxI]$  for use in the shunting network, whereas the eat/wait output is limited to the range  $[0 : 1]$ . If the eat/wait output gives an output below the expected threshold the agent simply waits at its current location. Waiting and eating both reduce an agents energy by 0.1 energy units (though eating may result in a net energy gain), with moving using up 0.2 energy units per time step. Agents will only move if their Iota output for their current location is below threshold  $\theta_a$ . In this case an activity landscape is created based on the Iota outputs for all visible environmental features. Agents are born with, and are able to

achieve, a maximum energy level of 100 units. As epochs here constitute 1000 time steps, an agent would be able to survive for a maximum of one epoch, or one thousand time steps, by remaining inactive. In order to avoid moving agents moving around in circles, or moving backwards and forwards, in neutral space (where there is no activity gradient from the activity landscape) consecutive neutral move actions maintain the same direction of travel with probability  $p_{dir} = 0.9$ .

Measurements are taken to determine whether an eat event was successful or unsuccessful. Any eat action that does not result in a non-negative energy providing food source being consumed is considered to be unsuccessful, so only eating non-energy reducing plants is a successful eating action. In order to measure a population of agents' success in a given environment, the difference between successful and unsuccessful eating actions is measured. This difference measure is useful as it is possible for agents to spend an equal amount of time eating successfully and unsuccessfully, which would demonstrate a strong performance on a measure of successful eating, but a weak performance on a measure of unsuccessful eating - the difference instead demonstrates a neutral performance, so a population that spends very little time eating, but all of that time eating successfully (so a picky eating strategy) would be a better performing population than a locust-like population that eats everything in sight.

#### 7.2.4 Social Information Strategies

Populations of agents using social information differ only very slightly from non-social populations; social information populations have an additional input unit for social information, thus non-social agents are rendered blind to other agents in the world. The social information strategies explored here, including the no social strategy are discussed below:

NO SOCIAL: No input node is available to the agent to enable social information to be used by the agent’s decision network. Agents proceed with no information about other agents.

PRESENCE: The social information input node receives an input of +1 if another agent is present within the visual field. No other information about the agent being viewed is used. This strategy is not dissimilar to the Inadvertent Information strategy used by agents in the work by Mitri et al. [109], though the agents explored in the work presented here do not have a choice about whether they express social information or not (this is the case for all social information strategies presented here).

ACTION: An input representing the current action state of the agent being viewed. The wait action is input as a value of 0, eat is input as 0.5 and move is represented as 1. Amalgamating these action inputs into one input rather than two or three categorical inputs, whilst not ideal, was implemented in order to ensure the input layer size for all social strategies was equal.

HEALTH: The current energy levels of the agent being viewed are normalised to be within the range [0 : 1] and input to the viewing agent’s decision network.

AGE: The age (in time steps) of agent being viewed is normalised using a hyperbolic tangent function of the logarithm of the age, which is then normalised to be within the range [0:1] (with 1 being an asymptote). Normalising age in this way is necessary as agent’s may live for the entire duration of the simulation, and are not selected against based upon their age. See formula (7.3) where  $a$  represents agent age in time steps.

$$input_a = (\tanh(\log(a)) + 1) / 2 \quad (7.3)$$

## 7.3 Results and Discussion

Forty populations of each social information strategy (including no social) were tested on each environment ( $1 \rightarrow 9$ ). Each population was permitted to evolve in the environment for 100 epochs of 1000 time steps. Reproduction and death events occurred both within and at epoch, meaning all populations were a mix of young and older agents at all stages of evaluation, with agents having no maximum age limit. Population data was accumulated for each epoch, and collected at the end of each epoch. As we are primarily interested here in the final test performance achieved by a population, not the pathway toward this achievement, average metrics were taken for each population, for each environment, for the last 25 epochs of a test, by which point performance had stabilised across measures. The results presented here are the median values of the 40 populations' average last 25 epochs of data - as this data was rarely normally distributed the medians were considered to be of more use than means. In order to derive the statistical significance between population data for each social information strategy a Mann-Whitney U test was used, with p values being derived from the resulting Z-scores. Figure 7.2 presents Z-score values on an inverted secondary y-axis, with p-value being represented by highlighting over Z-score data points. In order to test our hypothesis, that populations of agents making use of social information should outperform non-social agents, we measure the difference between how often agents successfully and unsuccessfully apply their eat actions, thus allowing us to measure the effectiveness of the eating behaviour within populations. Only comparisons for each social information strategy against the no social strategy are undertaken to see if any statistically significant differences arise. We go on to further analyse a wider array of metrics, including successful and unsuccessful eating actions in isolation, agent turnover, and average agent age.

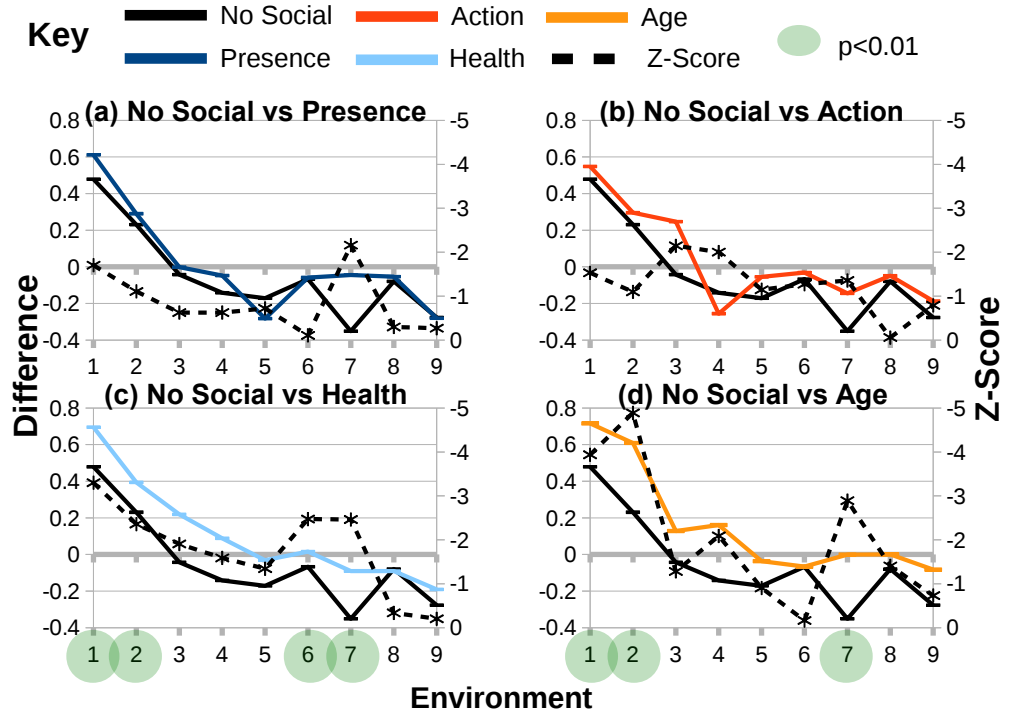


Figure 7.2: The difference in % of actions that qualified as successful eating actions and % of actions that qualified as unsuccessful eating actions in each environment, for each social information strategy compared against no social information. All graphs show the Z-score from an Mann-Whitney U test on the secondary y axis, with highlighting included to indicate statistical significance. Each data point represents the median of the average results for forty populations.

### 7.3.1 Eat Action Performance

In Figure 7.2 we can see the difference between successful and unsuccessful eating actions for each social information strategy compared to results for non-social popula-



tions. Looking first at populations with no social information (black line on all graphs in Figure 7.2) we see that the median difference crosses zero, and therefore indicates the eat action is being applied unsuccessfully more often than successfully, at environment 3 (a 1:3 positive to negative food ratio). All social information strategies manage to maintain the eat action in favour of successful eating until a more difficult environment - this is most notable for both the Health and Age social information strategies where eat actions do not begin to favour unsuccessful eating until environment 5, with the Health strategy re-crossing zero briefly, and the Age strategy maintaining an almost neutral profile for all environments after environment 5. This suggests there is a benefit to social information, in that social information may allow populations to maintain successful behaviours in more challenging and difficult environments. However, if we look more closely at the resulting Z-scores and p-values we see that both the Presence and Action strategies rarely demonstrate a significantly better difference in eat actions over populations of non social agents, and even when significant differences are seen they are with relatively weak and therefore lead us to the conclusion that we cannot say with any certainty that either the Presence or Action social information strategy provides a significant improvement over having no social information at all. Despite the poor performance seen for all strategies in later environments, all strategies were capable of enabling at least one population to achieve a positive eat profile in all environments. It is also worth noting the inconsistent results observed with regard to the No Social strategy in environments 7 and 8. Despite the median result fluctuating in a way that suggest environment 8 was less challenging than environment 7, there was no statistically significant difference between the distribution of results for these environments.

Despite Presence and Action social information being of dubious value, both Health

and Age social information provide a more convincing benefit. We can see in Figure 7.2(c) that populations using social information about the health of other agents demonstrate a significantly better difference in eating actions until environment 7, this performance difference is most noticeable in less difficult environments (environments 1 and 2) where we see a p value  $<0.01$ . Populations using social information about age (as seen in Figure 7.2(d)) also demonstrate a significantly better difference in eating action in less difficult environments, though the statistical significance over environments is less consistent. However, the two most significant Z-scores seen relate to no social information vs. age social information on environments 1 and 2, which demonstrate that social information about age is particularly useful in these less difficult, but still challenging environments. From this data we can begin to see the potential advantages of certain types of social information.

### **7.3.2 Social Information Performance in Less Difficult Environments**

In Figure 7.2 we see that environment 1, where there is a 1:1 ratio of positive to negative plant resources, gives rise to a significant difference in eating performance when social information populations are compared to populations with no access to social information, with this result being extended to environment 2 (a 1:2 ratio) for both Health and Age social information populations. This shows a particular benefit to using social information in less difficult environments. It is worth noting here that whilst environment 1 and 2 are less difficult than later environments used here, they are still themselves reasonably challenging given that we could have tested in environments with positive plant resources in abundance. Having a 1:1 or 1:2 ratio of positive to negative plant resources provides a reasonable challenge, so much so that in environment 3

we see that non-social populations, relying on evolution alone and having no access to social information, now begin to struggle at the task. Figure 7.3 shows a wider range of performance metrics for environment 1, including the breakdown of successful and unsuccessful eating actions in isolation. Here we see that the success of social information populations is as a result of both higher eat success rates and lower eat failure rates, though it is interesting to note that Age, Presence and Action social information populations are capable of demonstrating very low levels of eat success, even when compared to No Social populations, when the full data range is considered. The main driving force behind the success of social populations, especially Health and Age, appears to be consistently low eat fail rates across populations - the upper quartile ranges for both of these strategies not exceeding 0.02 (2% of actions). This suggests that social information is often being used to help agents avoid or not consume negative plant resources. Age and Health information may be particularly useful for this purpose as it would allow agents to avoid or ignore young or unhealthy agents whilst developing a preference for healthy and older agents. Whilst Presence or Action information may also be useful for the purposes of discrimination (move towards areas of high agent presence, or follow moving agents for example), they are both potentially riskier sources of information compared to Health or Age which both provide information about agent success. Figure 7.4, which shows performance metrics for environment 2, also shows that for Age social information this ability to maintain consistently low rates of unsuccessful eating alongside a strong eating success performance is maintained in slightly harder environments. We can also see that for unsuccessful eating actions, the upper quartile range for social information strategies is comparable to the median for non social populations.

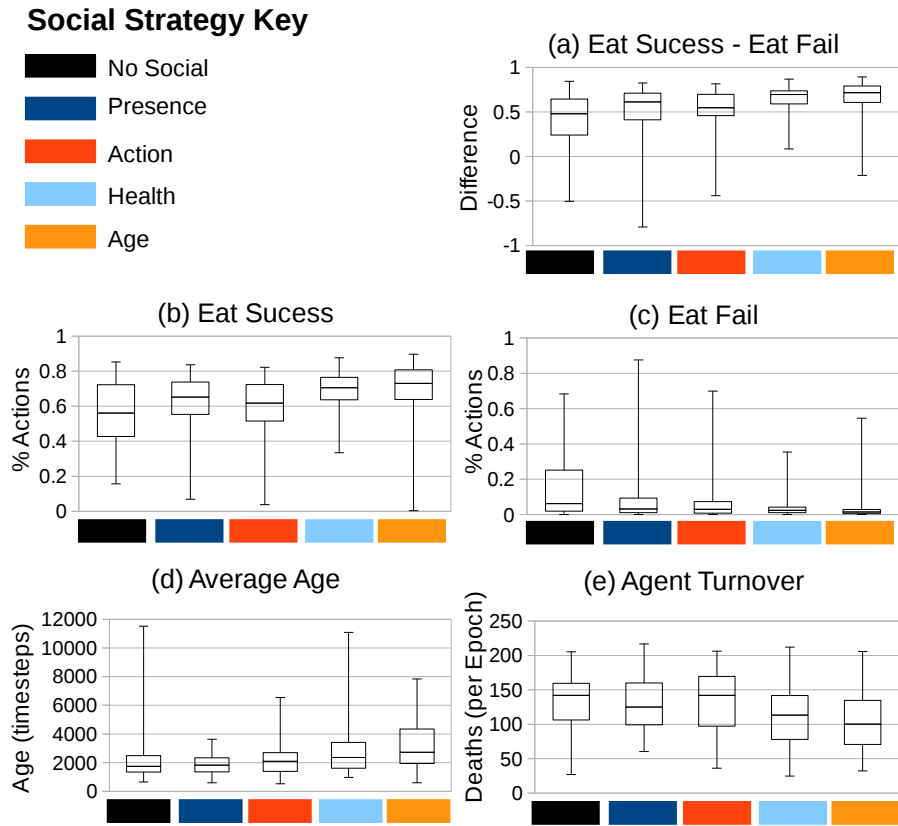


Figure 7.3: Box plots for the eat action and other population metrics, including a breakdown of successful and unsuccessful eating actions, average agent age, and agent turnover, in environment 1.

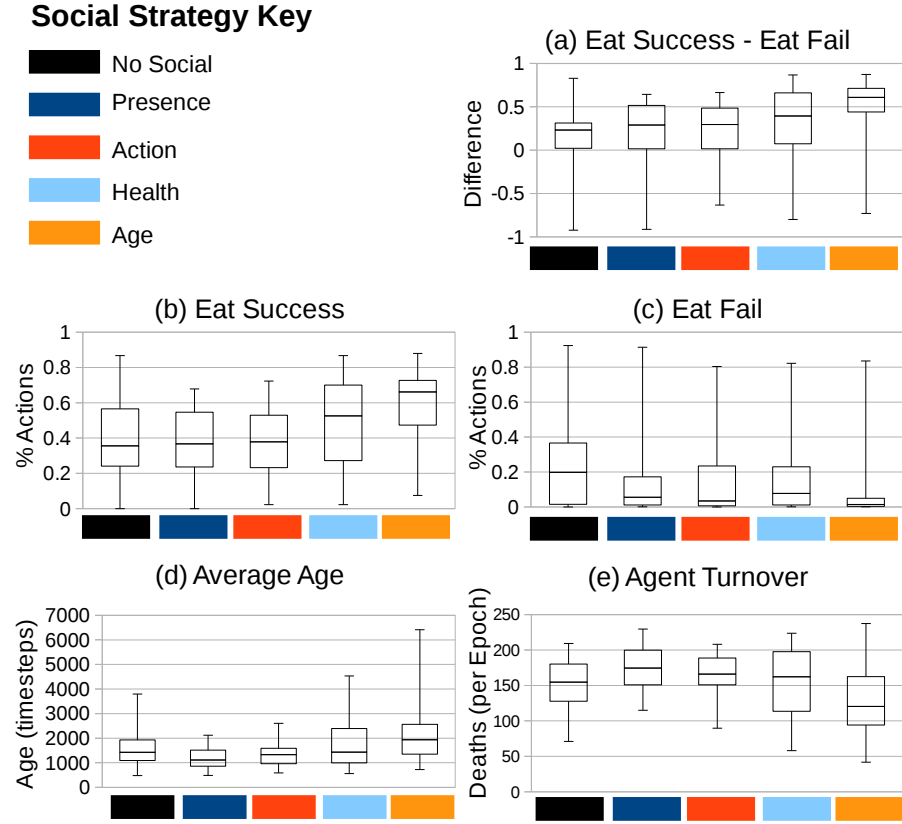


Figure 7.4: Box plots for the eat action and other population metrics, including a breakdown of successful and unsuccessful eating actions, average agent age, and agent turnover, in environment 2.

Alongside information about eating, both Figures 7.3 and 7.4 also give information on average agent age and agent turnover. For both environments 1 and 2 we see both Age and Health social information enabling populations to accomplish a high average agent age with an accompanying reduction in agent turnover (fewer agents dying within an epoch due to running out of energy), though the median agent turnover for Health social information is comparable to the non social tests. Both Presence and Ac-

tion populations fail to distinguish themselves from No Social populations, suggesting the improvements in eating performance seen most notably in environment 1 do not necessarily translate directly to improved survival, this suggests there must be other underlying behaviours that are causing these populations to use more energy, thus resulting in lower average ages and a higher agent turnover when compared to the Age and Health social information populations. The indeterminate quality of both Presence and Action information causes agents using this information to be less discerning about which agents and plant resources they move towards, resulting in less informed movement and therefore less efficient energy expenditure - though further analysis will have to be done to confirm these suspicions. Environment 1 here is not dissimilar to the food foraging task used by Acerbi and Marocco [2] where a reduced mortality rate was also observed in social populations when compared with the mortality rates of non social populations.

## 7.4 Conclusions and Further Work

The work presented here, alongside results from Mitri et al. [109], contribute to the discussion on the adaptive value of social information for evolved simulated agents by demonstrating that social information can provide an adaptive benefit to a neuroevolutionary process when decoupled from a within-lifetime learning process. However, we do see that social information is only of consistent adaptive benefit in less difficult environments, and when the social information itself is informative. This work also demonstrates the potential adaptive benefits of simple social and public information strategies such as social influence, social facilitation, stimulus enhancement, and local enhancement [3, 126, 63, 153], adding further weight to the work by Noble, Todd and Franks [111, 113] in which it was argued that simple social learning mechanisms are

capable of producing complex adaptive behaviours that may easily be confused for the resulting behaviours of more complex social learning mechanisms. The social information strategies implemented here could be argued to be mechanisms of stimulus and local enhancement as the social information inadvertently expressed here by agents could be used by others as an attractor to unfamiliar plant resources or a promoter of eating (or other) behaviours. However, we also see evidence of social information potentially being used to ignore locations or being used to suppress eating (or other) behaviours, which indicates some level of information suppression [109]. In Chapter 8 we undertake a greater analysis of the behaviours being expressed by agents here.

## Chapter 8

# The Effect of Social Information Use without Learning on the Evolution of Behaviour

In Borg and Channon [21] (reported here in Chapter 7) it was shown that social or public information alone, decoupled from any within-lifetime learning process, can result in improved performance on a food foraging task compared to when social information is not available. Here we assess whether access to social information leads to any significant behavioural differences both when this access to social information leads to improved task performance, and when it does not. In short, how strongly is social information used to increase task performance, and do any behaviours resulting from social information use persist even when task performance does not outperform the performance seen when no social information is available?



## 8.1 Social Information and Social Behaviour

The idea that agents may be socially attracted to each other by way of actively seeking each other out in order to benefit from the proximity of others, be it to avoid predators, breed or co-operatively raise their young, or to discover new resources or habitats is a well established one [8, 9]. However, there is still some debate as to why and when social and public information leads to agent aggregation. In reviewing public and social information use, Valone [150] outlines three general hypotheses to explain why individuals might prefer to settle near conspecifics (leading to what may be described as habitat copying via local enhancement):

1. Individual fitness is enhanced via the Allee effect [8, 9, 137]; which is defined by Stephens et al. [140] as “a positive relationship between any component of individual fitness and either numbers or density of conspecifics”. Allee observed that individuals were better able to survive and reproduce when found in groups, concluding that there is a positive correlation between population density or group size and individual fitness (known as the Allee effect). If this effect holds true we would expect there to be selection pressure in favour of aggregation; increased use of public information may therefore be as a result of increased social interaction due to aggregation.
2. Public information based resource discovery results in a reduction in search costs, enabling a more efficient use of energy [137, 69]. As public information may be used to reduce search costs, and increase the chance of experiencing new resources which may have been otherwise overlooked, aggregation may result from a selective pressure to obtain social information rather than increased public information use being a secondary consequence of aggregation itself. The Allee

effect (a positive correlation between individual fitness and aggregation), resulting as a consequence of this selective pressure to access public information.

3. Individuals use the presence of other (established) individuals as an indicator of the high-quality of a habitat without necessarily requiring them to rely on their own (possibly incomplete or poor) evaluation of the habitat [149, 155]. Here public information not only reduces the search costs when discovering resources, but also enables individuals to derive the quality of a unfamiliar resource based on the public information derived about the action, state or presence of others. Again, aggregation and the Allee effect result as a consequence of selective pressures in favour of public information use, rather than public information use resulting as a consequence of a selective pressure in favour of aggregation.

Here we assess three questions regarding agent behaviour in the presence of social information. Firstly, we assess whether the well established notion that social information leads to behaviours that promote agent aggregation is true in simple artificial evolutionary systems such as the one used in Borg and Channon [21]. Secondly, we assess whether agent private information reliability (or environmental predictability) impacts on the agent aggregation and social information use. Finally, we assess whether any observed social behaviours can be seen to persist even when social information use does not lead to an improved task performance.

The question of the persistence of what may be described as non-adaptive social information use, or social learning, was addressed by Higgs [81] in his meme-based simulation study of learning by imitation. One of the many things Higgs [81] concluded was that memes (discrete, replicating, units of “culture” [48, 50, 17, 18]) even when learned blindly (without concern for their adaptive value) provide a selective advantage to imitation. This suggests that behaviour which increases social interactions may

still be adaptive even when task performance is poor. Higgs' result is not necessarily that surprising, as it is more than reasonable to expect to see agents with access to social information of any kind seeking this information out regardless of the contribution this information makes to fitness. Bullinaria [37] rationalises this expectation by stating that "If there exists a set of memes with a range of positive and negative contributions to the overall performance, then not imitating them will leave performance at some baseline, while imitating them will result in a range of performance levels above and below that baseline. Any selection on the basis of performance will then favour those individuals that have imitated the good memes, and hence favour higher imitation rates" - therefore we can see why agents may wish to collect around sources of information; sometimes that information will be useful, so gaining access to it is important. We would therefore expect to see agents attempting to find sources of information even when obtaining that information does not necessarily lead to an improved performance. Agent aggregation and social interaction for the purpose of habitat copying is also found to be adaptive in highly variable environments [152], though with the potential pitfall of population collapse during overly conformist social interaction [160, 23]. It has also been noted by Rendell et al. [125] that strategies that rely heavily on social learning seem to be remarkably successful, even when information obtained from non social sources is no more costly than social information. We would therefore expect behaviours that maximize access to social information to emerge.

In the model set-up used in Borg and Channon [21] which forms the basis for this work there are a large number of possible food resources available to agents, resulting in agents often being uncertain about whether any given food resource will provide a positive or negative amount of energy. As environments in the Borg and Channon [21] model become more difficult, a strategy whereby all food is ignored may evolve,

but this strategy would always be outperformed by a strategy that sought to minimise uncertainty about available food resources in order to discover a positive energy providing resources. Social information, especially about the performance or fitness of an agent, may therefore be sought in order to allow for decisions on whether to consume any given food resource to be influenced by others, thus reducing uncertainty about the safety of a new food resource. This kind of social information seeking behaviour in order to seek out information about new or novel food resources is often seen in Norway rats [64, 114], though it is interesting to note that this social behaviour is only used to develop food preferences and not food aversions; this property of rat social behaviour has been suggested to be as a result of the high levels of lethality associated with poor food choices in rat populations [114], thus resulting in very little social information about negative food resources being available to the population. We may see a similar scenario in the more difficult environments presented here, providing a continued pressure for social behaviour under extreme environmental difficulty.

Turner et al. [148] demonstrates that when models are unreliable children spend more time learning socially, this could also be re-phrased as social learning is more likely to take place when a task is difficult to individually learn. Therefore, it is not unreasonable here to expect agents in populations who have access to social information to seek this information out in order to reduce the unreliability of their own internal models of the world, it is far easier to evolve a rule which states “trust older individuals” than evolve a rule about each possible food resource one may experience, especially when it is likely that any given food resource is new to an agent and therefore is yet to be evaluated. van Bergen et al. [151] also reports that when private information is less reliable, stickleback fish tend to use public rather than private information.

The large amount of evidence to suggest the persistence of social information promoting behaviours in unreliable and challenging environments, and evidence from simulations that social learning mechanisms such as imitation provide a selective advantage even when the information being obtained is not necessarily fitness increasing, along with the well established principle that the desire to obtain social information leads to social aggregation, leads us to postulate the following hypotheses to be assessed to here.

1. Social Information should lead to behaviours that result in increased agent aggregation (i.e. movement to seek to social interactions): We will test this hypothesis by comparing the amount of movement undertaken by agents from social information using populations with non social agents. If we see a significant difference in the amount of movement, we will then assess how often agents from social populations spend around other agents. We require a significantly larger number of movement actions combined with agent aggregation to demonstrate not only socially influenced aggregation, but also behaviours that promote social aggregation. Sergio and Newton [133] provides evidence that in some cases even simple information such as the presence of other individuals (or occupancy) can be a suitable indicator of resource quality and therefore enough to lead to agent aggregation around a food source, therefore we would expect this hypothesis to hold true in all social information strategies presented here (see Chapter 7.2.4); though when the presence of another agent is used as a source of social or public information, some measure of resource quality may still be required, as no information about the success or state of the agent present on the resource is available to act as a proxy for resource quality [150].
2. This aggregation will be somewhat dependent on the reliability of agent private

information: In the model environment used here it could be argued that private information reliability is maximised in the more difficult environments. The most complicated environment tested here has a ratio of one positive food resource to every nine negative food resources, therefore agents have a 90% chance of correctly guessing that a food resource will be dangerous. We may therefore expect agent aggregation (should it be seen) to be at its highest in lower difficulty environments, despite the possibility of non social agents performing well in these environments.

3. Behaviours resulting in increased agent aggregation will persist (though at reduced levels) even when task performance is poor: The adaptive value of social information, even when potentially unreliable, should still be high enough to motivate agents to aggregate. In the more difficult environments tested here we would expect social information to be relatively poor, due to the large quantities of negative food resources populating the environment. However, it would still be beneficial for agents to aggregate in order to provide potential access to any positive behaviour that may emerge in the population. Therefore we would expect behaviours that encourage social aggregation, i.e. movement, to still appear more often in social populations than in non-social ones, in all environments.

We will also go on to assess whether social information leads to any significant difference in the application of the other behaviours available to agents here when compared to non-social populations, and whether task performance has any implications for the application of behaviour - we are especially interested to assess whether a change in task performance from the predominantly successful application of eat actions to the predominately unsuccessful application of eat actions can result in any notable transitions in behaviour.

## 8.2 Experimentation and Results

The experimental set-up matches that used in Borg and Channon [21] (see Chapter 7). Populations of neuroevolutionary agents (making use of the hybrid neural network model known as the shunting model [168, 167, 128, 22, 139, 87]), each population employing a different social information strategy, are tasked with surviving in environments of differing difficulties. The social information strategies used are No Social, Presence, Action, Health and Age (detailed in Chapter 7.2.4), with forty populations of each social information strategy being evaluated. Environmental difficulty is determined by the ratio of positive energy providing food resources to negative energy providing food resources, therefore environment 1 has a 1 : 1 ratio and environment 9 has a 1 : 9 ratio. All data presented here, as in Chapter 7, relates to the final 25 epochs of evolution (of a total of 100 epochs) where population behaviour and fitness had broadly stabilised.

### 8.2.1 Action Profiles

As in Borg and Channon [21] (see Chapter 7) the agents simulated here are capable, via the outputs of their neural networks, of three different actions: eating, waiting and moving, with eating actions being evaluated as either successful (eating a positive energy food resource) or unsuccessful (eating a negative energy food resource or trying to eat when no food resource is present). As outlined in Chapter 7.2.3, each action is accompanied by a small amount of energy loss, in order to ensure agent inactivity is minimised - the move action having a higher residual energy loss than waiting or eating.

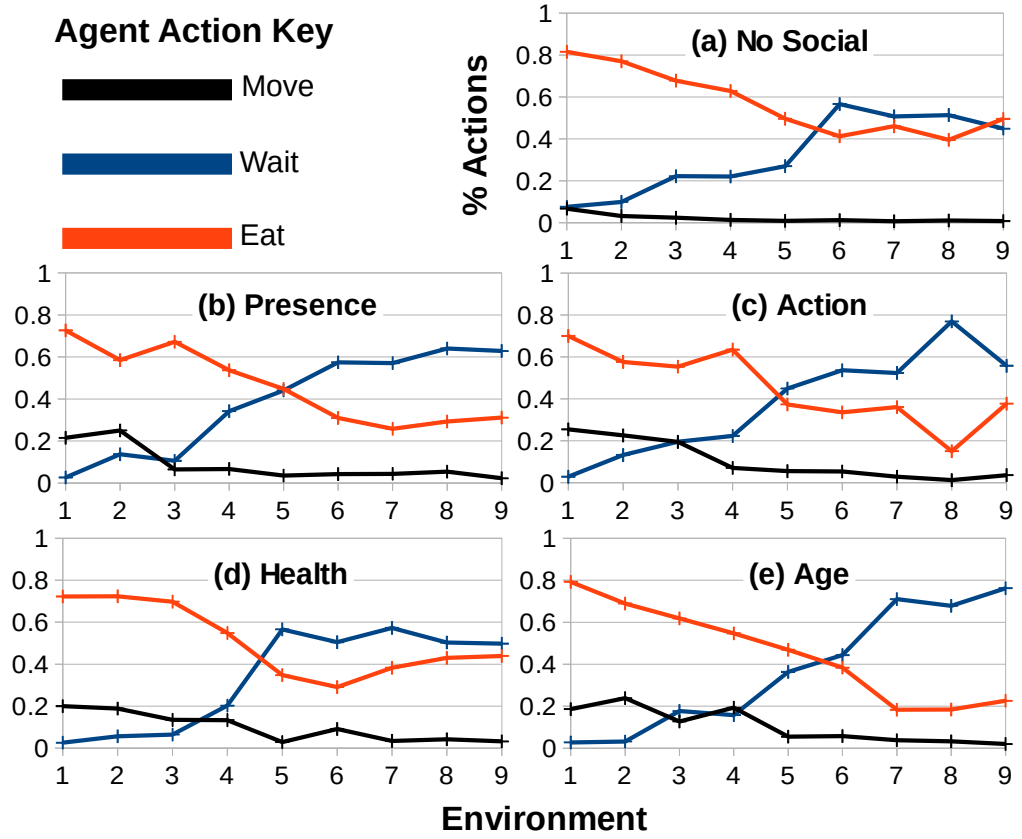


Figure 8.1: Median agent action profiles for each social information strategy over all environments. Data points represent the median % of actions of each particular action type over 40 populations. Data for each population is an average for all agents over the last 25 epochs of the simulation.

Figure 8.1 shows the median action profiles for each social information strategy applied here in each environment, an action profile being the percentage of total actions each individual action contributed. These action profiles can be considered alongside Figure 7.2 from Chapter 7.3.1 to add the context of task performance to the action profiles. The most immediate difference between the social information using popula-



tions and non social populations from Figure 8.1 is the application of the move action. Whilst all populations show a reduction in movement, with an accompanied increase in waiting, non social populations have extremely low levels of movement even in environments of lower difficulty when compared to social information populations. In social populations movement is applied more frequently than waiting in lower difficulty environments. This suggests that the increased performance associated with populations that use social information in simpler environments seen previously [21] (see Chapter 7) is as a consequence of this greater willingness to move, either to find new food resources or to find new sources of social information. As the only difference between social and non social populations is the addition of social inputs to agent neural networks, movement to seek new sources of information is probably closer to the truth; as agents in all populations spend the majority of their time in simpler environments eating, any movement motivated by the desire to be around other agents would lead to a secondary consequence of being around more food resources, enabling agents who are less able to distinguish between positive and negative food resources to defer some of their judgements on the likely pay-off of a food resource, and instead rely on the social information being provided by the agents they now find themselves around to make more informed decisions. However it is not clear from Figure 8.1 whether or not this difference in movement between non social and social populations is significant, and whether this additional movement does lead to more social interactions.

The immediate difference in movement behaviour between non social and social populations seen in Figure 8.1 is demonstrated to be significant by way of Mann-Whitney U tests between the resulting application of move actions for social populations compared to non social populations, this can be seen in Figure 8.3. The continued significance in the difference between social and non social populations regarding movement over

all environments is in contrast to the general lack of significance in task performance difference between social and non social populations in environments past environment 2 (as seen in Figure 7.2, outlined in Chapter 7.3.1); these results indicate that the introduction of social information or public information leads to behavioural differences that persist even when these behaviours do not result in improved task performance.

Regarding the other actions available to agents; eating (see Figure 8.2) and waiting (see Figure 8.4), neither show any particular significant differences (where  $p < 0.01$ ) between social and non social population other than in environment 1 where waiting actions for all social populations are applied significantly less than in non social populations ( $p < 0.01$ ), and eating actions are applied significantly less for social populations using the Presence and Action strategies than in non social populations ( $p < 0.01$ ). This broad lack of any significant differences beyond environment 1, between non social and social populations for eating and waiting, further demonstrates that movement is the primary driving force in the improved task performance seen in earlier environments, especially in environment 2 where only movement is significantly different despite previous work showing a significant difference in task performance (as seen in Figure 7.2, outlined in Chapter 7.3.1); though it should be noted that in environment 1 social information availability also leads to significantly different eating and waiting behaviours, indicating that some adaptive action profile across actions is available to drive improved task performance, rather than just a reliance on movement behaviour. The fact that in environment 1 differences in eat and wait actions result in less eating and waiting taking place in social populations in favour of more movement, also indicates that social agents are willing to risk higher energy expenditure, and are willing to spend less time potentially obtaining energy via eating. This demonstrates that the accommodation of social information leads to a more refined, and ultimately

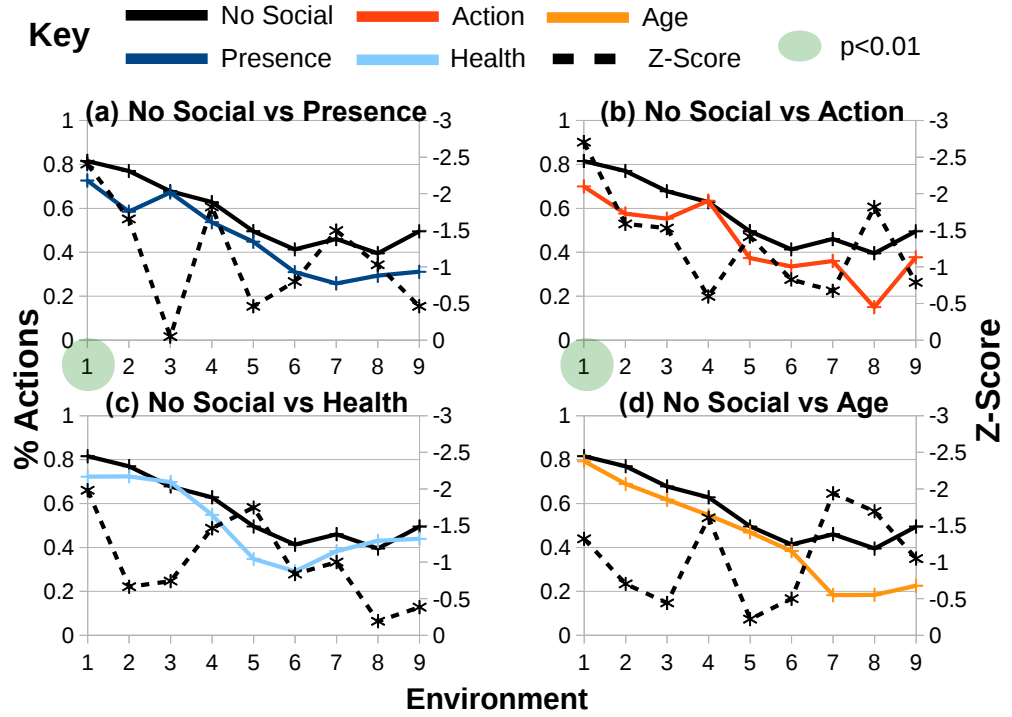


Figure 8.2: Median eat actions for each social information strategy over all environments. Data points on the primary axis represent the median % of the eat action over 40 populations. Data for each population is an average for all agents over the last 25 epochs of the simulation. Data points on the secondary axis represent the Z-score value from a Mann-Whitney U test comparing, for each environment, the median actions for the two social information strategies presented. Z-scores which indicate statistically significant  $p$  values are highlighted.

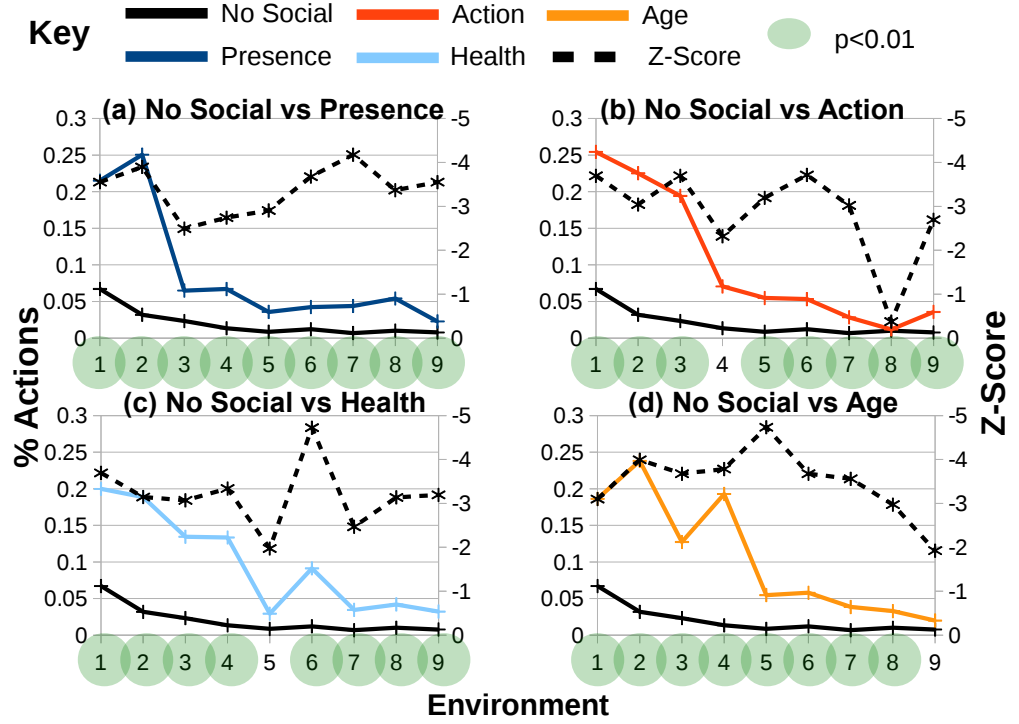


Figure 8.3: Median move actions for each social information strategy over all environments. Data points on the primary axis represent the median % of the move action over 40 populations. Data for each population is an average for all agents over the last 25 epochs of the simulation. Data points on the secondary axis represent the Z-score value from a Mann-Whitney U test comparing, for each environment, the median actions for the two social information strategies presented. Z-scores which indicate statistically significant  $p$  values are highlighted.

more effective, eating strategy as a result of an increased willingness to move. However, as we can see from the action profile box-plots in Figure 8.5, the application of eating and waiting actions is drawn from quite a large range in all populations, though the inter-quartile ranges for all actions do imply some level of consistency in the application of actions in environment 1.

The suggestion here is that the significant improvement in task performance seen in social populations over non social populations in less difficult environments (as discussed in Chapter 7) is as a direct result of the behaviour differences enabled by the accommodation of social information. However, this does lead us to something of a “Chicken and Egg” situation; did social information use follow as a result of good foraging (with good foragers acting as useful sources of social information), or did social information use result in the development of good foraging strategies? As no information about plant resources are communicated by social agents, with only information about the agents themselves being expressed, it would be sensible to assume that the improved task performance seen by social populations in simpler environments is caused by agents developing behaviours that cause greater exposure to other agents (and therefore more sources of social information), which then leads to an improve task performance as a secondary outcome. The fact movement behaviour remains significantly different throughout all tests indicates that some behavioural differences persist despite them providing no improvement in task performance (as seen in Figure 7.2, outlined in Chapter 7.3.1).

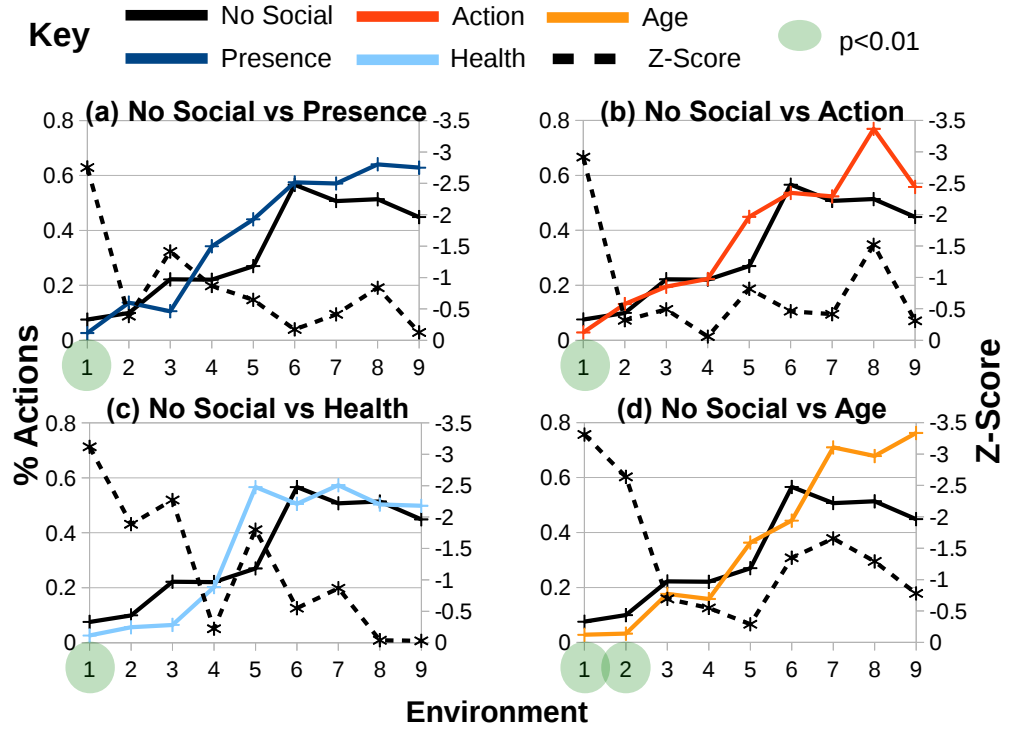


Figure 8.4: Median wait actions for each social information strategy over all environments. Data points on the primary axis represent the median % of the wait action over 40 populations. Data for each population is an average for all agents over the last 25 epochs of the simulation. Data points on the secondary axis represent the Z-score value from a Mann-Whitney U test comparing, for each environment, the median actions for the two social information strategies presented. Z-scores which indicate statistically significant  $p$  values are highlighted.

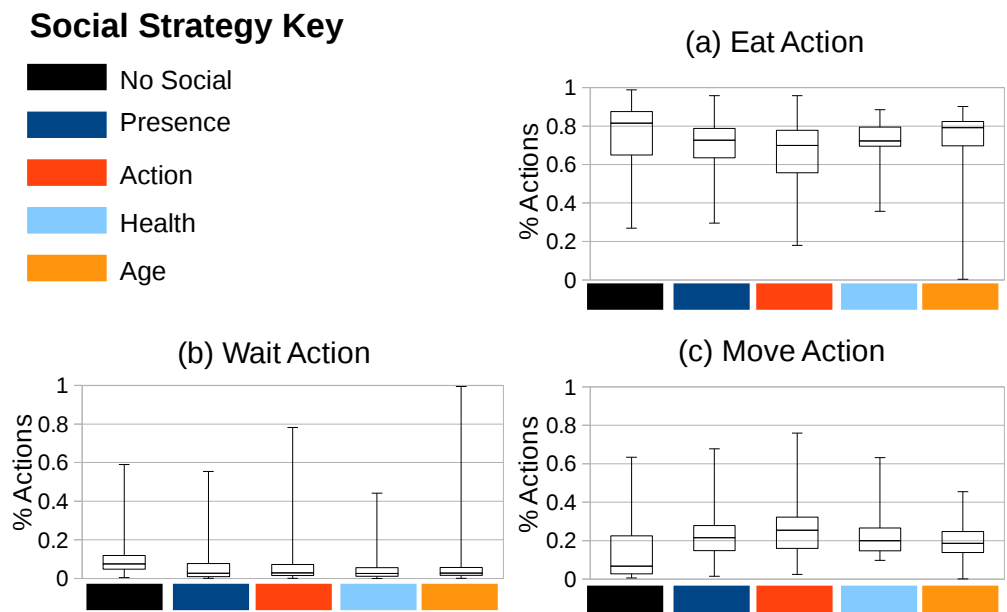


Figure 8.5: Action box-plots for each action, for each social information strategy in environment 1, where there is a 1:1 ratio of positive to negative food resources.

### 8.2.2 Reasons for Moving

It is apparent from Figure 8.3 that movement behaviour for populations permitted to use social information differs significantly from non social populations - this is in contrast to both eating actions (see Figure 8.2) and waiting actions (see Figure 8.4) which only show significant differences between social and non social populations in selected environments. Therefore some analysis on why social agents move is necessary. Alongside action data, data on the number of other agents an agent views over their lifetime is collected. An agent view is counted when an agent registers that another agent is within its visual field. The rationale for collecting this data is that if an agent can register the presence of another agent within its visual field and feed any social information about the agent through its decision network, it is possible for the agent to use this information to either move toward, move away from, or ignore this other agent. Should we see a correlation between the amount of movement undertaken by a population, and the number of agent views over each environment, we may reasonably conclude that this movement serves a purpose of enabling agents to collect together (like a herd, or social group) thus allowing agents greater access to social information. Figure 8.6 plots movement actions (primary y-axis) against agent views (secondary y-axis), showing a high level of similarity between the trajectory of move actions and agent views as environments become more difficult. All social information populations have a Spearman's Product Moment Correlation  $r > 0.982$  when correlating move actions and agent views, this very high level of correlation thus demonstrates that agents in social populations are primarily using movement to bring them closer to other agents, thus enabling greater access to social information; access to social information is the primary motivation for movement, which in simpler environments also results in improved task performance over non social populations (as seen in Figure 7.2, outlined



in Chapter 7.3.1, and reported in [21] ). These may go some way to explaining why social information use is so common in nature [164, 123].

Figure 8.6 shows the number of agent views accumulated between populations using differing social information strategies. We see that the percentage of agent views accumulated by populations using Age information initially exceeds that of populations using Health information, who in turn exceed the agent views accumulated by populations using Action information, with populations using Presence social information accumulating the fewest agent views. This ordering of strategies closely matches the ordering based on task performance seen in Chapter 7. Figure 8.7 shows comparisons of each social information strategy's agent views profile over environments, from these Figures we can see Age populations accumulate statistically more agent views than Presence and Action populations over all environments, with the largest and most statistically significant differences being seen in simpler environments. This coincides with the environments that Age populations also show the highest, and most statistically significant, difference in task performance when compared to non social populations. This indicates that information about age not only leads to improved task performance, but also provides a greater motivation for agents to aggregate. Populations using Health social information also show a significantly larger accumulation of agent views compared to populations simply using information about the Presence of other agents. The results presented in Figure 8.8 are also interesting, in that there are almost no significant differences on any environment when comparing the movement behaviour between social information using populations. Therefore the significant differences seen in the accumulation of agent views for populations using Age information suggests that agents in these populations are using their move actions in response to, and to move toward, other agents more often than in any other social information population.

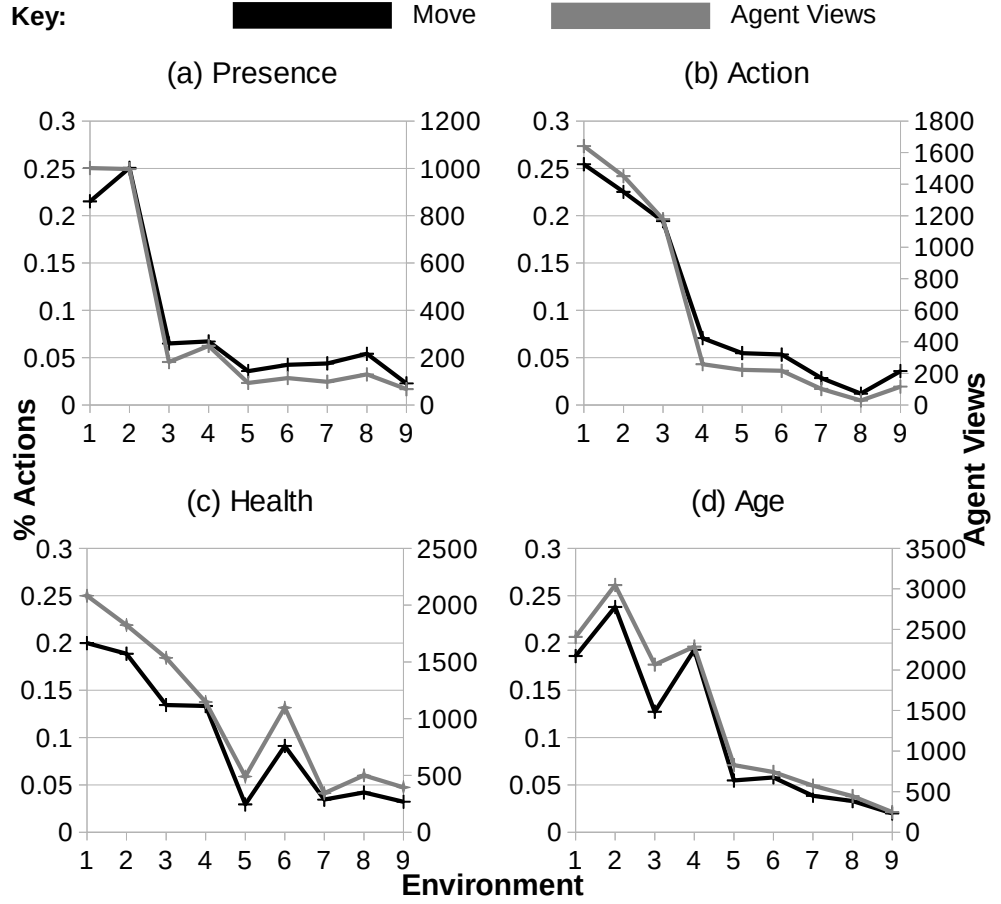


Figure 8.6: Median move actions plotted against median agent views. Median move actions are presented on the primary axis, with the median number agent views per agent presented on the secondary axis. All social information populations have a Spearman's Product Moment Correlation  $r \geq 0.982$  when correlating median move actions against median agent views.

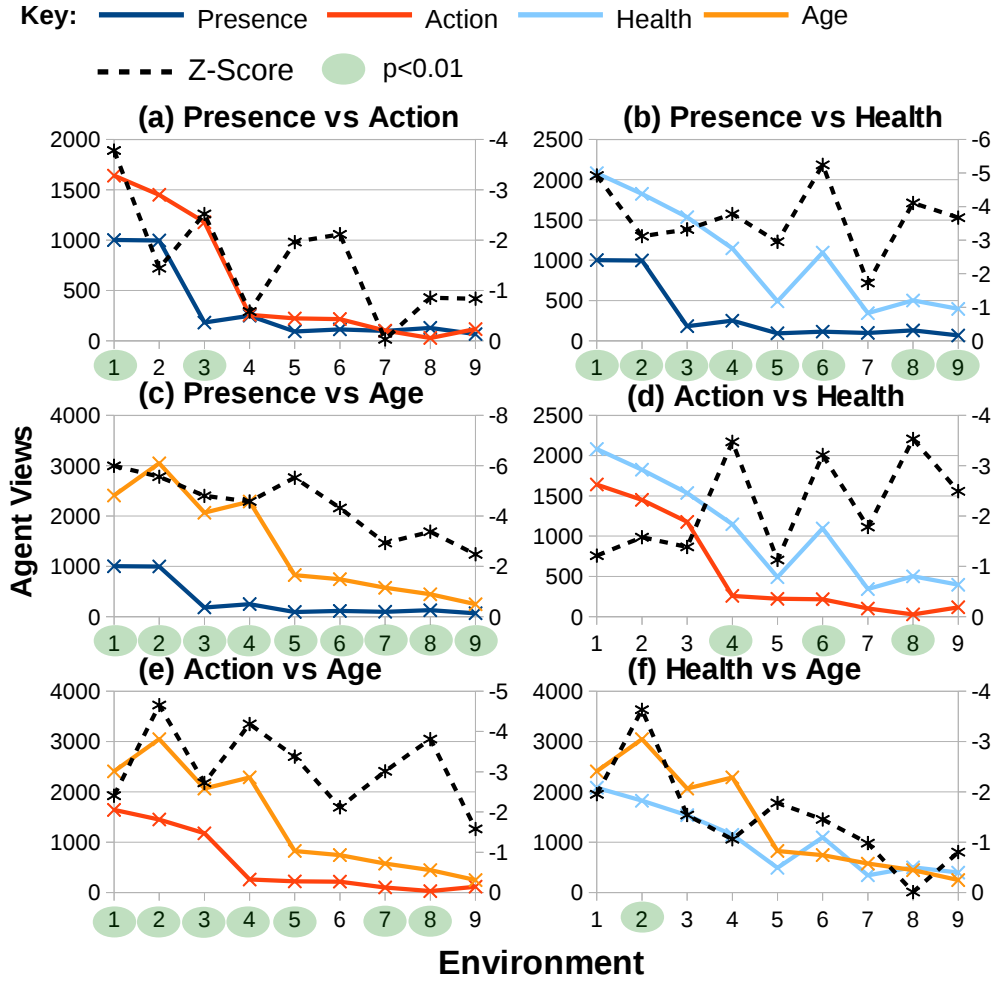


Figure 8.7: Median agent views for each social information strategy over all environments. Data points on the primary axis represent the median of the average number of agent's viewed by each agent over 40 populations. Data for each population is an average for all agents over the last 25 epochs of the simulation. Data points on the secondary axis represent the Z-score value from a Mann-Whitney U test comparing, for each environment, the median agent views for the two social information strategies presented. Z-scores which indicate statistically significant p values are highlighted.

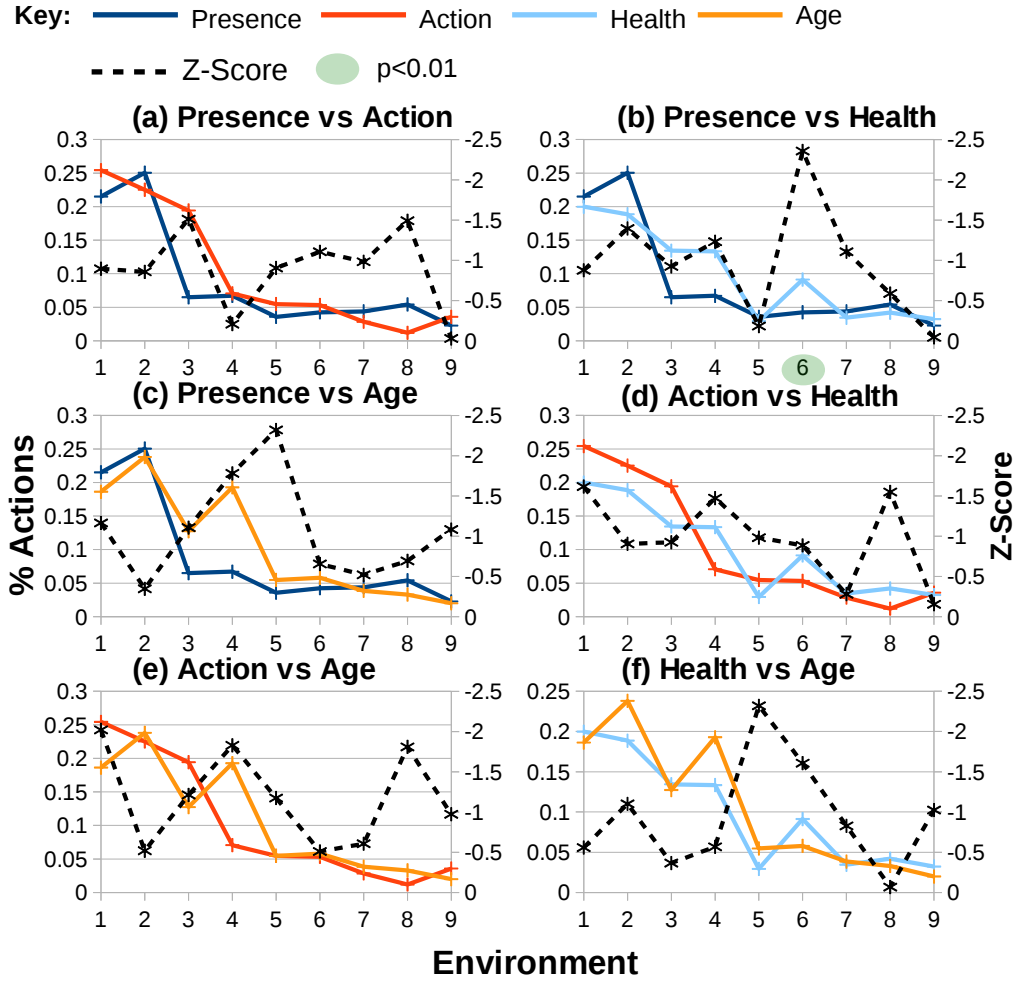


Figure 8.8: Median move actions for each social information strategy (barring non social) over all environments. Data points on the primary axis represent the median % of the move action over 40 populations. Data for each population is an average for all agents over the last 25 epochs of the simulation. Data points on the secondary axis represent the Z-score value from a Mann-Whitney U test comparing, for each environment, the median actions for the two social information strategies presented. Z-scores which indicate statistically significant  $p$  values are highlighted.

### 8.2.3 Behavioural Transitions

From Figure 8.1, and Figures 8.3, 8.2 and 8.4 we can say that agent behaviour changes as environments become more difficult. These behavioural changes lead to a reduction in movement and eating, and an increase in waiting. The primary driving force behind the motivation to eat less, move less and wait more, independent of social information strategy, is that food resources are increasingly likely to be negative in their energy provision, and therefore it makes sense for agents to spend more time conserving their energy waiting for a positive food source to appear near to them or (in the case of social populations) for an agent who's information suggests they can be trusted to move into their visual field. However, in most cases the increase or decrease in actions as environments become more difficult is not necessarily smooth, this being most apparent with move actions (Figure 8.3) which for many social information strategies shows a sudden reduction in action rather than a steady degradation. It is not clear from earlier Figures whether these changes between environments are statistically significant nor what is driving these sudden changes when they occur.

In Chapter 7.3.1 (detailed in Figure 7.2) we showed that task performance (the ability to eat positive food resources more frequently than negative food resources) deteriorates as environments get more difficult - this difficulty being defined by the ratio of positive food resources to negative food resources available in the environment. The point at which task performance changes from successful to unsuccessful (the point at which eating actions result in more negative food resources being consumed than positive food resources) varies depending on the social information strategy being tested, but occurs in all scenarios. For No Social and Presence populations this transition (or zero crossing) occurs between environments 2 and 3, Action populations experience this transition between environments 3 and 4, and both Health and Age populations

experience this transition to primarily negative food resource consumption between environments 4 and 5 (though Health populations do not permanently cross into negative task performance until after environment 6). Here we assess whether any statistically significant changes to behaviour, or behaviour transitions, could be associated with these zero crossing events for food type consumption.

From Figure 8.9 we can see that that non social populations do not exhibit any statistically significant transitions ( $p < 0.01$ ) between environments in regard to movement behaviour. However, statistically significant transitions in movement behaviour between environments can be seen in all social populations. For populations using Presence information (Figure 8.9(b)) we see this statistically significant transition happen between environments 2 and 3; the transition from primarily eating positive food resources to primarily eating negative food resources also occurs between environments 2 and 3. The association between a statistically significant transition in movement behaviour and the transition to primarily consuming negative food resources is also apparent for populations using Action information and populations using Health information (Figures 8.9(c) and (d) respectively) - for Health populations it is also interesting to note that statically significant movement behavioural transitions occur on both occasions when positive food consumption drops below zero. These results demonstrate that movement behaviour in social populations is strongly driven by agent task performance; when agents can no longer successfully solve the task, social populations are less inclined to explore their environment in order to seek out new food resources or new sources of social information. In the case of populations using Age social information, the only significant transition associated with movement behaviour occurs before the transition to non-positive food consumption. The point at which this transition in movement behaviour occurs does correspond with a large drop in task performance

between environments 2 and 3, demonstrating that movement behaviour is still highly sensitive to task performance in Age social information populations.

When considering the total proportion of actions agents dedicate to eating, as seen in Figure 8.10, we do not see any significant changes in eating behaviour that correspond to the point at which task performance transition from predominantly successful application of the eat action to predominantly unsuccessful application of the eat action. Instead, as seen in Figure 8.2, the median total eat action degrades gradually with task performance. It is also worth noting the extremely large data ranges seen with the total application of each action in the box plot data in Figure 8.10. The large inter-quartile ranges especially show that all populations, social and non social, are capable of exhibiting very high and very low levels of eating activity. This is in stark contrast to movement, which we can see from Figure 8.9 has reasonably small inter-quartile ranges for all population types across all environments, and if anything becomes more consistent as environmental difficulty increases, this being in contrast to the general increase in the range of eat action data which generally increases as the environment becomes more difficult. Increasingly large data ranges are also seen when we consider the wait action (as seen in Figure 8.11). Any significant transitions seen in waiting behaviour, in all populations barring Health, do not seem to occur in relation to the transition from positive to negative task performance. These results further indicate that social agents are driven to seek out new sources of social information, but with the caveat that social interactions are likely to result in better task performance; though the fact that social populations move more often than non social populations even when task performance is poor suggests that social populations still persist in a residual amount of socially motivated movement.

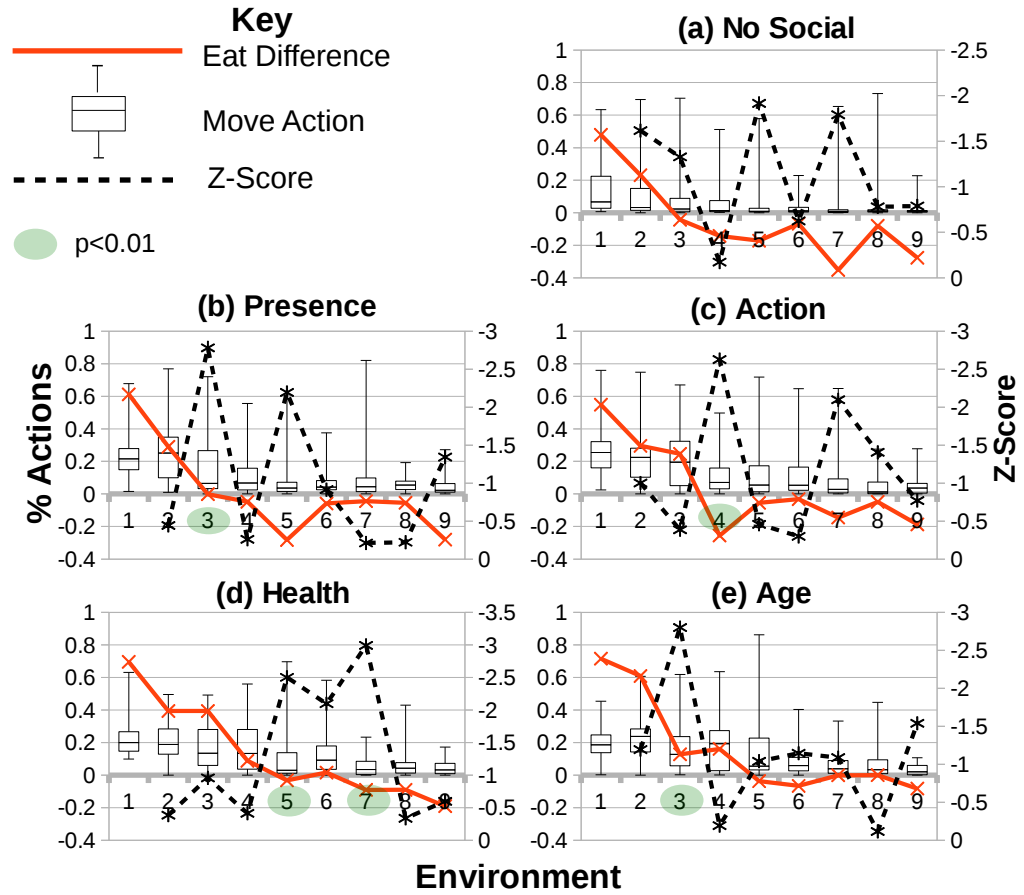


Figure 8.9: The median differences between successful and unsuccessful eat actions (eat difference) is presented on the primary axis along with the box plots for the move action. The Z-score from Mann-Whitney U tests, which compare the action data for the environment on which a data point falls with the previous environment, is presented on the secondary axis. These Z-scores are intended to indicate which transitions in action behaviour between previous environments are significant, thus indicating a significant behavioural transition. Z-scores which indicate statistically significant  $p$  values are highlighted.



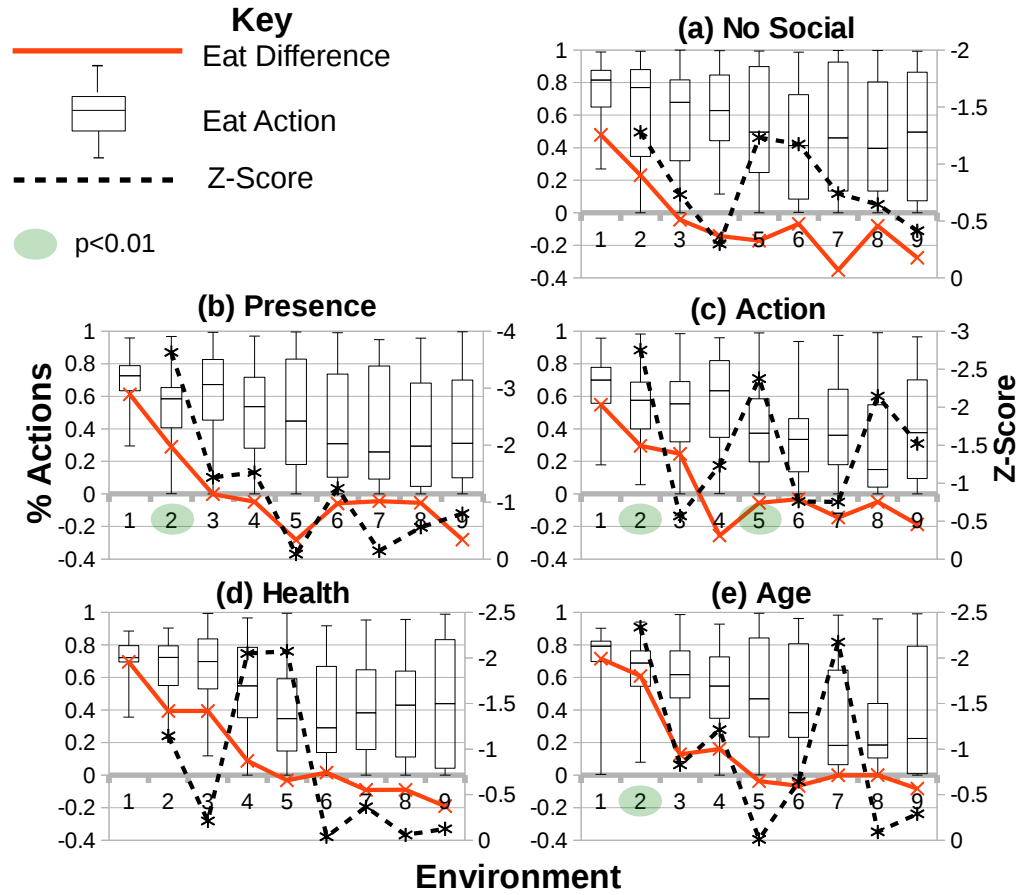


Figure 8.10: The median differences between successful and unsuccessful eat actions (eat difference) is presented on the primary axis along with the box plots for the eat action. The Z-score from Mann-Whitney U tests, which compare the action data for the environment on which a data point falls with the previous environment, is presented on the secondary axis. These Z-scores are intended to indicate which transitions in action behaviour between previous environments are significant, thus indicating a significant behavioural transition. Z-scores which indicate statistically significant  $p$  values are highlighted.

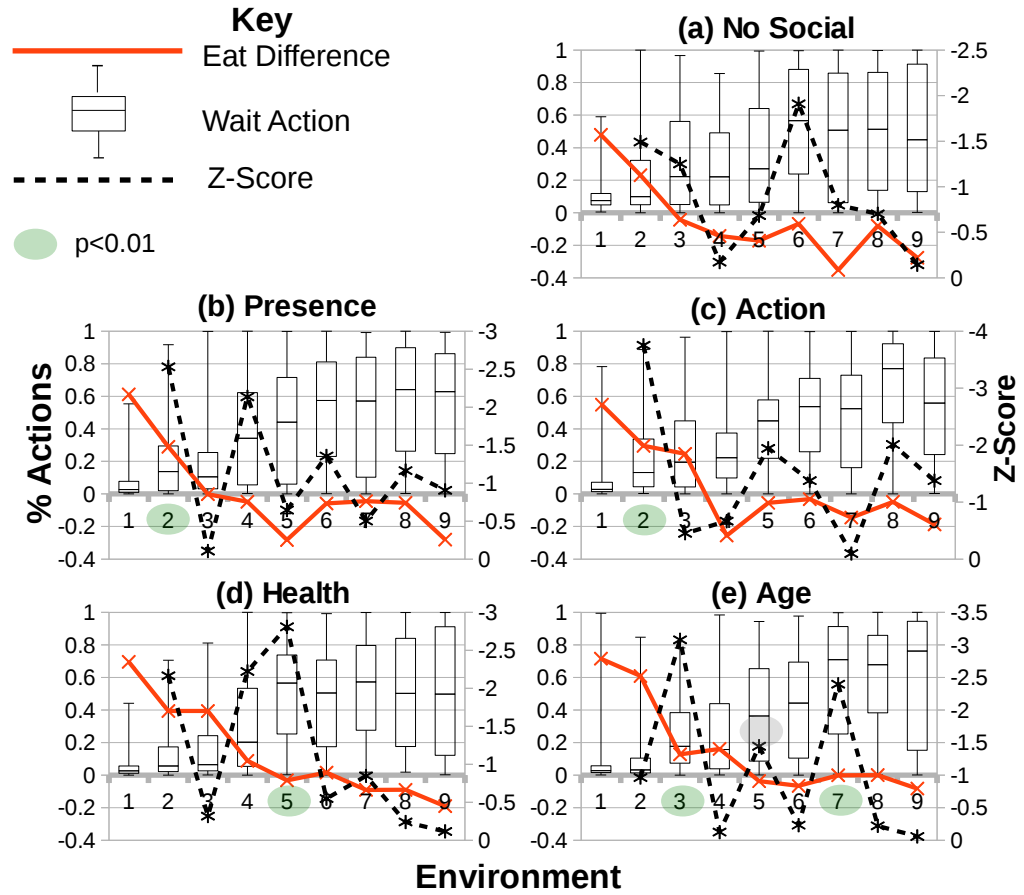


Figure 8.11: The median differences between successful and unsuccessful eat actions (eat difference) is presented on the primary axis along with the box plots for the wait action. The Z-score from Mann-Whitney U tests, which compare the action data for the environment on which a data point falls with the previous environment, is presented on the secondary axis. These Z-scores are intended to indicate which transitions in action behaviour between previous environments are significant, thus indicating a significant behavioural transition. Z-scores which indicate statistically significant  $p$  values are highlighted.

## 8.3 Discussion and Conclusion

In this work we attempted to address three questions. (1) Does social information lead to increased aggregation, as observed in nature [133, 149, 150]? (2) Is agent aggregation, and by extension public information use, dependant on the reliability on agent private information [151, 148]? (3) Do social behaviours persist even when task performance is poor [81]?

Social information transfer is highly prevalent in nature [164], and even the simple presence of other agents have been demonstrated to encourage interesting and novel behaviours in other agents [42], so it not entirely surprising that the results presented in this work provide strong evidence that social information does lead to aggregation promoting behaviours, namely movement for the purpose of increasing the probability of agent interaction, with these increased agent interactions potentially leading to favourable conditions for individual decision making [91]. We also see social behaviours being favoured in the simpler environments tested here. These simpler environments did provide agents with a large variety of food resources that could be either negative or positive with an equal probability, resulting in a task which was reasonably easy to solve (due to an abundance of positive food resources) but also very difficult for individuals to develop a complete set of categorisations for each food resource's edibility. Social behaviours being favoured here are likely to be as a result of public information being reliable but private information being reasonably unreliable. As environments progressed in difficulty, private information about the edibility of any given food resource became more reliable, as it was increasingly likely that any given food resource was energy reducing and therefore not worth consuming - these results are in line with the work of van Bergen et al. [151] and Turner et al. [148] where private and public information reliability was seen as major factors in the expression of social behaviours.

Any social aggregation in later, more difficult, environments would still have yielded some benefits though. In the presence of a food resource in any environment the presence, actions, health or age of other local agents could potentially result in a novel or new food resource being evaluated correctly. Despite private information based on the likelihood of edibility encouraging a conservative policy on eating, this new public information could sometimes yield positive results leading to an adaptive advantage over agents who eschew social aggregation. Here we see a continued preference for movement in social information populations compared to non social populations, even in more difficult environments where task performance between social and non social populations was similar. This continued desire to move for the purpose of aggregation was less apparent in later environments, with waiting actions being preferred due to the risk of unnecessary or un-rewarding energy expenditure in more difficult environments, but still noticeably different from non social cases. These results add additional evidence to the idea that a pressure for evolution to adapt to accommodate social information, be it via social information transfer or imitation, is maintained even when social information is either unreliable or risky [81]. However, it should be noted that all figures presented here show some noisy data trends, therefore more test runs need to be conducted in order to clarify whether this noise is a significant feature worthy of additional exploration, or whether this noise was just a result of needing more tests.

# Chapter 9

## Conclusions

### 9.1 Summary of Conclusions

This thesis sought to investigate the emergence and utility of social behaviour and social learning in artificial evolutionary systems. These investigations were undertaken in reference to a five research questions, which were addressed using a number of grounded and non-grounded Artificial Life simulation models all incorporating artificial evolutionary systems with populations of social agents.

Chapter 4 addressed the question of the utility of social learning over a stand alone incremental genetic evolutionary process. The hypothesis presented in this Chapter was that the introduction of noise in the genotype to phenotype map accompanied by social transmission in the form of parent-child learning by imitation, would be sufficient to discover and maintain complex behaviours which were not accessible to incremental genetic evolution alone, thus demonstrating the adaptive benefits of social learning in the discovery of novel and incrementally inaccessible behaviours. The results support the hypothesis by demonstrating that without social learning the most adaptive behaviours available to agents are never found but with social learning all behaviours

available to agents can be discovered, exhibited and maintained. Chapter 4 therefore demonstrates the adaptability of social learning over incremental genetic evolution alone, thus providing an empirically supported rationale for employing the evolution of social learning in artificial evolutionary systems.

Chapter 5 addresses the question of when social learning will evolve and prove to be adaptive under temporally variable environmental conditions, and goes on to investigate the importance of individual learning to the long term adaptability of social learning. The data presented to address this question suggest that when environments are in minimally variable states, individual learning is required to play a smaller role than it does in more variable environments, leaving social learning to drive population level adaptability. It is also found that the likelihood of population collapse is greatly increased in environments of high or increasing variability when social learning is exhibited unless individual innovation is allowed to first develop in isolation. This is a result of the tendency for social learning to become the dominant form of learning very quickly should both individual and social learning evolve side by side, whereas allowing individual learning to evolve as the sole learning strategy for a time ensures a level of personal innovation remains in the population, thus providing a level of robustness to environmental variability. Individual learning is found to be the key method of information discovery, with social learning acting as a mechanism of rapid information transfer within the population once new and robust solutions have been discovered. The initial hypothesis (developed in order to test Potts's variability selection hypothesis [119, 120, 121]), that when individual and social learning rates are evolved simultaneously, both increasing and consistently variable environments are sufficient for the adoption of social learning over individual learning, is found to hold true, though with two main caveats: individual learning is required for successful so-

cial learning, and population collapse may only be avoided when individual learning is allowed to pre-evolve in already noisy environments before the introduction of social learning, thus providing a sufficient selection pressure for innovation before the benefits of social learning can be realised.

Chapter 6 extends the work presented in Chapter 4 by introducing a number of new social learning strategies. The question to be addressed here was whether behaviours inaccessible to incremental genetic evolution alone are still discovered, and maintained, when agents are permitted to learn from a variety of different individuals, and whether these differing social learning strategies access incrementally inaccessible behaviours in differing ways. The results presented here demonstrated that the results presented in Chapter 4 are robust for a number of different social learning strategies including unbiased social learning and learning from inexperienced individuals. These results demonstrate the general adaptability of social learning when applied in a variety of strategies. It is also noted that the likelihood of social learning discovering and maintaining behaviours inaccessible to incremental genetic evolution alone is dependant on the social learning strategy used, and that social learning populations discover a series of incremental behaviours in a less reliable way than incremental genetic evolution alone, suggesting that the necessarily conformist nature of social learning can lead to social pressures that maintain sub-optimal behaviours within a population in a way that is not seen when incremental genetic evolution alone is utilised.

Chapter 7 addresses the question of whether social information is adaptive in the absence of a within-lifetime learning process. The results presented to address this question demonstrated that social information can provide an adaptive benefit to a neuroevolutionary process in the absence of a within-lifetime learning process. However, social information is only observed to be of any consistently adaptive benefit in

the less difficult environments tested in the Chapter, and when the social information itself is informative. Results also demonstrated the potential adaptive benefits of less complex social information mechanisms such as social influence, social facilitation, stimulus enhancement, and local enhancement, adding further weight to the work of Noble, Todd and Franks [111, 113] in which it was argued that simple social learning mechanisms are capable of producing complex adaptive behaviours equivalent to those exhibited by more complex social learning mechanisms such as imitation and emulation.

Chapter 8 provides further analysis of the artificial evolutionary system tested in Chapter 7 to address the question of to what extent social information affects agent behaviour, especially in regard to agent aggregation for the purposes of increased social interaction. The Chapter goes on to ask whether there are any behaviours exhibited by social populations that may be considered distinct from those exhibited by non-social populations; and whether these behaviours persist when social information use is no longer adaptive? Results presented to address these questions provide strong evidence that social information does lead to aggregation promoting behaviours, namely movement for the purpose of increasing the probability of agent interaction. The behaviours exhibited by social populations were seen to be distinct across populations from non-social populations in regard to aggregation promoting behaviours, with these behaviours being shown to be consistent across environments, even when social information was shown to produce performance results comparable to those of non-social populations.

## 9.2 Contribution of this work

Using artificial evolutionary systems this work:



- Provides the first demonstration of a behaviour inaccessible to incremental genetic evolution alone being evolved through the addition of social learning accompanied by noise in the genotype to phenotype mapping, thus demonstrating the sufficiency of social learning to enable behavioural transitions between sub-optimal and optimal behavioural peaks in the genotypic search space.
- Demonstrates that the sufficiency of social learning to enable behavioural transitions between sub-optimal and optimal behavioural peaks in the genotypic search space may extend to a variety of social learning strategies.
- Provides one of the first definitive answers to the question of whether or not the variability selection hypothesis [119, 120, 121] is sufficient to explain the adoption of social learning in increasingly variable environments. The question was tested empirically using an artificial evolutionary system utilising agents with access to individual learning and social learning and demonstrated the importance of innovation via individual learning to the long term adaptability of social learning under environmental uncertainty.
- Provides a demonstration of the adaptive value of social information when decoupled from a within-lifetime learning process, thus demonstrating that social information alone is sufficient to provide an adaptive advantage over non-social agents.
- Demonstrates the evolution and persistence of social behaviours when social information is decoupled from a within-lifetime learning process, even when social information use is no longer of an adaptive benefit.

## 9.3 Future Work

The work conducted here provides a number of interesting avenues for future work. The work on social learning strategies presented in Chapter 6 could easily be extended to incorporate a wider variety of social learning mechanisms, and to include a variety of social learning mechanisms within the same population, to investigate during which stages of behavioural evolution certain strategies become adaptive. This work could be of some importance to future research in evolutionary robotics, whereby populations of robots could be allowed to evolve social biases and strategies that reflect their current task or circumstance - this work may be of increasing usefulness as automated robotic and artificially intelligent systems become more visible in day to day life, requiring them to not only be environmentally robust but also socially robust, especially should they need to autonomously interact with one another and with humans.

The EnVar model introduced in Chapter 7 and subsequently used in Chapter 8 could easily be extended to answer a number of other questions regarding social behaviour in grounded artificial evolutionary systems. Work to assess the role of discrimination and similarity [55] in adaptive social aggregation is already under way, along with work to combine discrimination and similarity with the social information strategies used in Chapter 7. Investigations into population size, mobility and population density [72], and the effects of mobility costs [2] on the use of social information could also be undertaken. As the EnVar model has only been used thus far to investigate social information without learning it would also be prudent to extend the model to accommodate learning, allowing for investigations into the effect of learning to be undertaken; neuromodulated plasticity [54, 135, 136] is the learning mechanism which seems best placed to be implemented here. Once learning is fully implemented, the EnVar models already implemented functionality for environmental variability can

utilised, the intention being that the work presented in Chapter 5 on temporal environmental variability can be extended to a grounded model to see if the results are robust when populations of agents experience temporal variability collectively.

# Bibliography

- [1] Acerbi, A., Ghirlanda, S., and Enquist, M. (2012). Old and young individuals’ role in cultural change. *Journal of Artificial Societies and Social Simulation*, 15(4):1.
- [2] Acerbi, A. and Marocco, D. (2009). Orienting learning by exploiting sociality: An evolutionary robotics simulation. In *Neural Networks, 2009. IJCNN 2009. International Joint Conference on*, pages 20–27. IEEE Press.
- [3] Acerbi, A., Marocco, D., and Nolfi, S. (2007). Social facilitation on the development of foraging behaviors in a population of autonomous robots. In Almeida e Costa, F., editor, *Advances in Artificial Life: Proceedings of the 9th European Conference on Artificial Life 2007*, pages 625–634. Springer.
- [4] Acerbi, A. and Nolfi, S. (2007). Social learning and cultural evolution in embodied and situated agents. In *Proceedings of the First IEEE Symposium on Artificial Life, ALIFE’07*, pages 333–340. IEEE Press.
- [5] Acerbi, A. and Parisi, D. (2006). Cultural transmission between and within generations. *Journal of Artificial Societies and Social Simulation*, 9(1).
- [6] Acerbi, A., Tennie, C., and Mesoudi, A. (2016). Social learning solves the problem of narrow-peaked search landscapes: experimental evidence in humans. *Royal Society Open Science*, 3(9):160215.

- [7] Acerbi, A., Tennie, C., and Nunn, C. L. (2010). Modelling imitation and emulation in constrained search spaces. *Learning and Behaviour*, 39(2):104–114.
- [8] Allee, W. (1931). *Animal aggregations: A study in general Sociology*. University of Chicago Press Chicago:.
- [9] Allee, W. C. (1939). *The social life of animals*. W. Heinemann.
- [10] Andras, P. (2015). Environment factors and the emergence of cultural - technical innovations. In Andrews, P., Caves, L., Doursat, R., Hickinbotham, S., Polack, F., Stepney, S., Taylor, T., and Timmis, J., editors, *Proceedings of the European Conference on Artificial Life 2015*, pages 130–137. MIT Press.
- [11] Andras, P. (2016). Social learning, environmental adversity and the evolution of cooperation. In Gershenson, C., Froese, T., Siqueiros, J. M., Aguilar, W., Izquierdo, E. J., and Sayama, H., editors, *ALIFE 15: The Fifteenth International Conference on the Synthesis and Simulation of Living Systems*. MIT Press.
- [12] Atton, N. (2010). Social learning in fish. Master’s thesis, University of St Andrews.
- [13] Atton, N. (2014). *Investigations into stickleback social learning*. PhD thesis, University of St Andrews.
- [14] Baggio, J. A. and Hillis, V. (2016). Success biased imitation increases the probability of effectively dealing with ecological disturbances. In *Winter Simulation Conference (WSC), 2016*, pages 1702–1712. IEEE Press.
- [15] Barkow, J. H., Cosmides, L., and Tooby, J. (1992). *The adapted mind: Evolutionary Psychology and the generation of culture*. Oxford University Press.

- [16] Best, M. L. (1999). How culture can guide evolution: An inquiry into gene/meme enhancement and opposition. *Adaptive Behavior*, 7(3–4):289–306.
- [17] Blackmore, S. (1998). Imitation and the definition of a meme. *Journal of Memetics*, 2(11):159–170.
- [18] Blackmore, S. J. (1999). *The meme machine*. Oxford University Press.
- [19] Bond, R. (2005). Group size and conformity. *Group Processes & Intergroup Relations*, 8(4):331–354.
- [20] Bonnie, K. E. and Earley, R. L. (2007). Expanding the scope for social information use. *Animal Behaviour*, 74(2):171–181.
- [21] Borg, J. M. and Channon, A. (2017). Evolutionary adaptation to social information use without learning. In Squillero, G. and Sim, K., editors, *European Conference on the Applications of Evolutionary Computation*, pages 837–852. Springer.
- [22] Borg, J. M., Channon, A., and Day, C. (2011). Discovering and maintaining behaviours inaccessible to incremental genetic evolution through transcription errors and cultural transmission. In Lenaerts, T., Giacobini, M., Bersini, H., Bourguine, P., Dorigo, M., and Doursat, R., editors, *Proceedings of the European Conference on Artificial Life 2011*, pages 102–109. MIT Press.
- [23] Borg, J. M. and Channon, A. D. (2012). Testing the variability selection hypothesis - the adoption of social learning in increasingly variable environments. In Adami, C., Bryson, D. M., Ofria, C., and Pennock, R. T., editors, *ALIFE 13: The Thirteenth International Conference on the Synthesis and Simulation of Living Systems*, pages 317–324. MIT Press.

- [24] Box, H. O. (1984). *Primate Behavior and Social Ecology*. Chapman & Hall, London.
- [25] Boyd, R. and Richerson, P. J. (1983). The cultural transmission of acquired variation: Effects on genetic fitness. *Journal of Theoretical Biology*, 100(4):567–596.
- [26] Boyd, R. and Richerson, P. J. (1985). *Culture and the Evolutionary Process*. University of Chicago Press.
- [27] Boyd, R. and Richerson, P. J. (1995). Why does culture increase human adaptability? *Ethology and Sociobiology*, 16:125–143.
- [28] Boyd, R. and Richerson, P. J. (2002). Group beneficial norms can spread rapidly in a structured population. *Journal of Theoretical Biology*, 215:287–296.
- [29] Braitenberg, V. (1984). *Vehicles: Experiments in synthetic psychology*. MIT Press.
- [30] Bryson, J. J. (2007). Mechanisms of action selection: Introduction to the special issue. *Adaptive Behavior*, 15(1):5–8.
- [31] Bryson, J. J. (2009). Representations underlying social learning and cultural evolution. *Interaction Studies*, 10(1):77–100.
- [32] Bryson, J. J., Ando, Y., and Lehmann, H. (2007). Agent-based modelling as a scientific method: A case study analysing primate social behaviour. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 362:1685–1698.
- [33] Bull, J. J., Meyers, L. A., and Lachmann, M. (2005). Quasispecies made simple. *PLoS Computational Biology*, 1(6):450–460.

- [34] Bullinaria, J. A. (2007). The effect of learning of life history evolution. In Lipson, H., editor, *Proceedings of the 9th Annual Conference on Genetic and Evolutionary Computation*. ACM.
- [35] Bullinaria, J. A. (2009). Lifetime learning as a factor in life history evolution. *Artificial Life*, 15(4):389–409.
- [36] Bullinaria, J. A. (2010). Memes in artificial life simulations of life history evolution. In Fellersmann, H., Dörr, M., Hanczy, M. M., Laursen, L. L., Maurer, S., Merkle, D., Monnard, P.-A., StÅžy, K., and Rasmussen, S., editors, *Proceedings of Artificial Life XII Conference*, pages 823–830. MIT Press.
- [37] Bullinaria, J. A. (2017). Imitative and direct learning as interacting factors in life history evolution. *Artificial Life*, 23.
- [38] Caldwell, C. A. and Millen, A. E. (2009). Social learning mechanisms and cumulative cultural evolution: is imitation necessary? *Psychological Science*, 20(12):1478–1483.
- [39] Cangelosi, A., Hourdakis, E., and Tikhanoff, V. (2006). Language acquisition and symbol grounding transfer with neural networks and cognitive robots. In *Neural Networks, 2006. IJCNN’06. International Joint Conference on*, pages 1576–1582. IEEE.
- [40] Capi, G. and Doya, K. (2005). Evolution of neural architecture fitting dynamics. *Adaptive Behavior*, 13(1):53–66.
- [41] Cavalli-Sforza, L. L. and Feldman, M. W. (1981). *Cultural transmission and evolution: a quantitative approach*. Princeton University Press.



- [42] Channon, A. D. and Damper, R. (1998). The evolutionary emergence of socially intelligent agents. In Edmonds, B. and Dautenhahn, K., editors, *Socially Situated Intelligence: a workshop held at SAB'98, University of Zurich Technical Report*, pages 41–49.
- [43] Cliff, D. T., Harvey, I., and Husbands, P. (1992). Incremental evolution of neural network architectures for adaptive behaviour. In *Proceedings of the European Symposium on Artificial Neural Networks (ESANN'93)*, pages 39–44.
- [44] Confer, J. C., Eastom, J. A., Fleischman, D. S., Goetz, C. D., Lewis, D. M. G., Perilloux, C., and Buss, D. M. (2010). Evolutionary psychology: Controversies, questions, prospects, and limitations. *American Psychologist*, 65(2):110–126.
- [45] Curran, D. and O’Riordan, C. (2007). The effects of cultural learning in populations of neural networks. *Artificial Life*, 13(1):45–67.
- [46] Danchin, É., Giraldeau, L.-A., Valone, T. J., and Wagner, R. H. (2004). Public information: from nosy neighbors to cultural evolution. *Science*, 305(5683):487–491.
- [47] Darwin, C. (1859). *On The Origin of Species*. John Murray.
- [48] Dawkins, R. (1989). *The Selfish Gene*. Oxford University Press, 2nd edition edition.
- [49] Denaro, D. and Parisi, D. (1996). Cultural evolution in a population of neural networks. In Marinaro, M. and Tagliaferri, R., editors, *Proceedings of the 8th Italian Workshop on Neural Nets*, pages 100–111. Springer.
- [50] Dennett, D. C. (1995). *Darwin’s Dangerous Idea: Evolution and the meanings of life*. Simon and Schuster.

- [51] Doncieux, S., Bredeche, N., Mouret, J.-B., and Eiben, A. E. (2015). Evolutionary robotics: what, why, and where to. *Frontiers in R*, 2(4):1–18.
- [52] Dugatkin, L. A. and Godin, J.-G. J. (1993). Female mate copying in the guppy (*poecilia reticulata*): age-dependent effects. *Behavioral Ecology*, 4(4):289–292.
- [53] Dürr, P., Mattiussi, C., and Floreano, D. (2006). Neuroevolution with analog genetic encoding. In Runarsson, T., Beyer, H.-G., Burke, E., Merelo-Guervós, J., Whitley, L., and Yao, X., editors, *Parallel Problem Solving from Nature-PPSN iX*, pages 671–680. Springer.
- [54] Dürr, P., Mattiussi, C., Soltoggio, A., and Floreano, D. (2008). Evolvability of neuromodulated learning for robots. In *Learning and Adaptive Behaviors for Robotic Systems, 2008. LAB-RS’08. ECSIS Symposium on*, pages 41–46. IEEE Press.
- [55] Efferson, C., Lalive, R., Cacault, M. P., and Kistler, D. (2016). The evolution of facultative conformity based on similarity. *PloS One*, 11(12):e0168551.
- [56] Emery, N. J. and Clayton, N. S. (2004). The mentality of crows: convergent evolution of intelligence in corvids and apes. *Science*, 306(5703):1903–1907.
- [57] Floreano, D. and Keller, L. (2010). Evolution of adaptive behaviour in robots by means of darwinian selection. *PLoS Biology*, 8(1):e1000292.
- [58] Gabora, L. (1995). Meme and variations: A computational model of cultural evolution. In Nadel, L. and Stein, D. L., editors, *1993 Lectures in Complex Systems: SFI Studies in the Sciences of Complexity, Lect. Vol. VI*, pages 471–485. Addison-Wesley.
- [59] Gabora, L. (1997). The origin and evolution of culture and creativity. *Journal of Memetics*, 1(1):1–28.

- [60] Galef, B. G. (1988). Imitation in animals: history, definition and interpretation of data from the psychological laboratory. In Zentall, T. R. . and Galef, B. G., editors, *Social Learning :Psychological and Biological Perspectives*, pages 3–28. Erlbaum, Hillsdale, NJ.
- [61] Galef, B. G. (1992). The question of animal culture. *Human Nature*, 3:157–178.
- [62] Galef, B. G. (2009). Strategies for social learning: testing predictions from formal theory. *Advances in the Study of Behavior*, 39:117–151.
- [63] Galef, B. G. (2012). Social learning and traditions in animals: evidence, definitions, and relationship to human culture. *Wiley Interdisciplinary Reviews: Cognitive Science*, 3(6):581–592.
- [64] Galef Jr, B. G. (1996). Social enhancement of food preferences in norway rats: a brief review. In Heyes, C. M. and Galef, B. G., editors, *Social learning in animals: the roots of culture*, pages 49–64. Academic Press San Diego.
- [65] Gintis, H. (2006). A framework for the unification of the behavioural sciences. *Behavioral and Brain Sciences*, 30:1–61.
- [66] Gomez, F. and Miikkuleinen, R. (1997). Incremental evolution of complex general behaviours. *Adaptive Behavior*, 5(3–4):317–342.
- [67] Gonzalez, M., Watson, R., Noble, J., and Bullock, S. (2014). The origin of culture: selective conditions for horizontal information transfer. In Sayama, H., Rieffel, J., Risi, S., Doursat, R., and Lipson, H., editors, *Artificial Life 14: Proceedings of the Fourteenth International Conference on the Synthesis and Simulation of Living Systems*, pages 408–414. MIT Press.

- [68] Goodale, E., Beauchamp, G., Magrath, R. D., Nieh, J. C., and Ruxton, G. D. (2010). Interspecific information transfer influences animal community structure. *Trends in Ecology & Evolution*, 25(6):354–361.
- [69] Greene, C. M. and Stamps, J. A. (2001). Habitat selection at low population densities. *Ecology*, 82(8):2091–2100.
- [70] Grove, M. (2011). Speciation, diversity, and mode 1 technologies: The impact of variability selection. *Journal of Human Evolution*, 61(3):306–319.
- [71] Grove, M. (2014). Evolution and dispersal under climatic instability: a simple evolutionary algorithm. *Adaptive Behavior*, 22(4):235–254.
- [72] Grove, M. (2016). Population density, mobility, and cultural transmission. *Journal of Archaeological Science*, 74:75–84.
- [73] Harvey, I. (2001). Artificial evolution: a continuing saga. In *International Symposium on Evolutionary Robotics*, pages 94–109. Springer.
- [74] Hashimoto, T., Warashina, K., and Yamauchi, H. (2010). New composite evolutionary computation algorithm using interactions among genetic evolution, individual learning and social learning. *Intelligent Data Analysis*, 14:497–514.
- [75] Henrich, J., Boyd, R., and Richerson, P. J. (2008). Five misunderstandings about cultural evolution. *Human Nature*, 19(2):119–137.
- [76] Henrich, J. and Gil-White, F. J. (2001). The evolution of prestige: Freely conferred deference as a mechanism for enhancing the benefits of cultural transmission. *Evolution and human behavior*, 22(3):165–196.

- [77] Henrich, J. and Henrich, N. (2010). The evolution of cultural adaptations: Fijian food taboos protect against dangerous marine toxins. *Proceedings of the Royal Society of London B: Biological Sciences*, 277(1701):3715–3724.
- [78] Henrich, J. and McElreath, R. (2003). The evolution of cultural evolution. *Evolutionary Anthropology*, 12(3):123–135.
- [79] Henrich, J. and McElreath, R. (2007). Dual inheritance theory: The evolution of human cultural capacities and cultural evolution. In Dunbar, R. and Barrett, L., editors, *Oxford Handbook of Evolutionary Psychology*, pages 555–570. Oxford University Press.
- [80] Heyes, C. M. (1994). Social learning in animals: categories and mechanisms. *Biological Reviews*, 69(2):207–231.
- [81] Higgs, P. G. (2000). The mimetic transition: a simulation study of the evolution of learning by imitation. *Proceedings of the Royal Society of London B: Biological Sciences*, 267(1450):1355–1361.
- [82] Hinton, G. E. and Nowlan, S. J. (1987). How learning can guide evolution. *Complex Systems*, 1(3):495–502.
- [83] Holland, J. H. (1975). *Adaptation in Natural and Artificial Systems*. The University of Michigan Press.
- [84] Hoppitt, W. and Laland, K. N. (2008). Social processes influencing learning in animals: a review of the evidence. *Advances in the Study of Behavior*, 38:105–165.
- [85] Isler, K. and van Schaik, C. P. (2006). Metabolic costs of brain size evolution. *Biology Letters*, 2:557–560.

- [86] Isler, K. and van Schaik, C. P. (2009). The expensive brain: A framework for explaining evolutionary changes in brain size. *Journal of Human Evolution*, 57:392–400.
- [87] Jolley, B. P., Borg, J. M., and Channon, A. (2016). Analysis of social learning strategies when discovering and maintaining behaviours inaccessible to incremental genetic evolution. In Tuci, E., Giagkos, A., Wilson, M., and Hallan, J., editors, *From Animals to Animats 14: 14th International Conference on Simulation of Adaptive Behavior*, pages 293–304. Springer.
- [88] Jones, D. (2015). *Evolution and Learning in Heterogeneous Environments*. PhD thesis, Goldsmiths, University of London.
- [89] Jones, D. and Blackwell, T. (2011). Social learning and evolution in a structured environment. In Lenaerts, T., Giacobini, M., Bersini, H., Bourguine, P., Dorigo, M., and Doursat, R., editors, *Proceedings of the European Conference on Artificial Life 2011*, pages 380–387. MIT Press.
- [90] Kendal, J. R., Rendell, L., Pike, T. W., and Laland, K. N. (2009). Nine-spined sticklebacks deploy a hill-climbing social learning strategy. *Behavioral Ecology*, 20(2):238–244.
- [91] King, A. J. and Cowlshaw, G. (2007). When to use social information: the advantage of large group size in individual decision making. *Biology Letters*, 3(2):137–139.
- [92] Laland, K. and Plotkin, H. (1990). Social learning and social transmission of foraging information in norway rats (*rattus norvegicus*). *Animal Learning & Behavior*, 18(3):246–251.

- [93] Laland, K. N. (2004). Social learning strategies. *Animal Learning & Behavior*, 32(1):4–14.
- [94] Langton, C. G. (1989). Artificial life. In Langton, C. G., editor, *Proceedings of an Interdisciplinary Workshop on the Synthesis and Simulation of Living Systems*, pages 94–109. Addison-Wesley Longman Publishing Co.
- [95] Levinson, S. (2006). Introduction: The evolution of culture in a microcosm. In Levinson, S. C. and Jaisson, P., editors, *Evolution and Culture*, pages 1–41. MIT Press.
- [96] Lindblom, J. and Ziemke, T. (2009). Social situatedness of natural and artificial intelligence. *Adaptive Behavior*, 11(2):79–96.
- [97] Lindeyer, C. M. and Reader, S. M. (2010). Social learning of escape routes in zebrafish and the stability of behavioural traditions. *Animal Behaviour*, 79:827–834.
- [98] Lopes, M., Melo, F. S., Kenward, B., and Santos-Victor, J. (2009). A computational model of social-learning mechanisms. *Adaptive Behavior*, 17(6):467–482.
- [99] Lumsden, C. and Wilson, E. O. (1981). *Genes, Mind, and Culture*. Harvard University Press.
- [100] Lycett, S. J., Collard, M., and McGrew, W. C. (2009). Cladistic analyses of behavioural variation in wild pan troglodytes: Exploring the chimpanzee cultural hypothesis. *Journal of Human Evolution*, 57:337–349.
- [101] Marriott, C. and Chebib, J. M. (2014). The effect of social learning on individual learning and evolution. In Sayama, H., Rieffel, J., Risi, S., Doursat, R., and Lipson, H., editors, *ALIFE 14: The Fourteenth International Conference on the Synthesis and Simulation of Living Systems*, pages 736–743. MIT Press.

- [102] Marriott, C. and Chebib, J. M. (2016a). Divergent cumulative cultural evolution. In Gershenson, C., Froese, T., Siqueiros, J. M., Aguilar, W., Izquierdo, E. J., and Sayama, H., editors, *ALIFE 15: The Fifteenth International Conference on the Synthesis and Simulation of Living Systems*, pages 508–516. MIT Press.
- [103] Marriott, C. and Chebib, J. M. (2016b). Modelling the evolution of gene-culture divergence. In Gershenson, C., Froese, T., Siqueiros, J. M., Aguilar, W., Izquierdo, E. J., and Sayama, H., editors, *ALIFE 15: The Fifteenth International Conference on the Synthesis and Simulation of Living Systems*, pages 500–507. MIT Press.
- [104] Marriott, C., Parker, J., and Denzinger, J. (2010). Imitation as a mechanism of cultural transmission. *Artificial Life*, 16(1):21–37.
- [105] Marshall-Pescini, S. and Whiten, A. (2008). Chimpanzees (pan troglodytes) and the question of cumulative culture. *Animal Cognition*, 11:449–456.
- [106] McElreath, R. (2010). The coevolution of genes, innovation and culture in human evolution. In Silk, J. and Kappeler, P., editors, *Mind The Gap: Tracing the Origins of Human Universals*, pages 451–474. Springer.
- [107] McNamara, J. M. and Dall, S. R. (2010). Information is a fitness enhancing resource. *Oikos*, 119(2):231–236.
- [108] Mesoudi, A., Whiten, A., and Laland, K. N. (2006). Towards a unified science of cultural evolution. *Behavioral and Brain Sciences*, 29(4):329–383.
- [109] Mitri, S., Floreano, D., and Keller, L. (2009). The evolution of information suppression in communicating robots with conflicting interests. *Proceedings of the National Academy of Sciences, USA*, 106(37):15786–15790.



- [110] Nakahasi, W., Wakano, J. Y., and Henrich, J. (2013). Apative social learning strategies in temporally and spatially varying environments. *Human Nature*, (23):368–418.
- [111] Noble, J. and Franks, D. W. (2002). Social learning mechanisms compared in a simple environment. In Standish, R., Bedau, M. A., and Abbass, H. A., editors, *Artificial Life VIII: Proceedings of the Eighth International Conference on the Simulation and Synthesis of Living Systems*, pages 379–385. MIT.
- [112] Noble, J. and Franks, D. W. (2004). Social learning in a multi-agent system. *Computing and Informatics*, 22(6):561–574.
- [113] Noble, J. and Todd, P. M. (2002). Imitation or something simpler? modeling simple mechanisms for social information processing. In Nehaniv, C. L. and Dautenhahn, K., editors, *Imitation in Animals and Artifacts*, chapter 16, pages 423–439. MIT Press.
- [114] Noble, J., Todd, P. M., and Tuci, E. (2001). Explaining social learning of food preferences without aversions: an evolutionary simulation model of Norway rats. *Proceedings of the Royal Society of London B: Biological Sciences*, 268(1463):141–149.
- [115] Nocera, J. J., Forbes, G. J., and Giraldeau, L.-A. (2006). Inadvertent social information in breeding site selection of natal dispersing birds. *Proceedings of the Royal Society of London B: Biological Sciences*, 273(1584):349–355.
- [116] Nolfi, S. and Floreano, D. (1999). Learning and evolution. *Autonomous Robots*, 7(1):89–113.

- [117] Nolfi, S. and Parisi, D. (1994). Desired answers do not correspond to good teaching inputs in ecological neural networks. *Neural Processing Letters*, 1(2):1–4.
- [118] Ochoa, G., Harvey, I., and Buxton, H. (1999). Error thresholds and their relation to optimal mutation rates. In Floreano, D., Nicoud, J.-D., and Mondada, F., editors, *Advances in Artificial Life: Proceedings of the European Conference on Artificial Life 1999*, pages 54–63. Springer.
- [119] Potts, R. (1996). Evolution and climate variability. *Science*, 273(5277):922–923.
- [120] Potts, R. (1998a). Environment hypotheses of hominin evolution. *Yearbook of Physical Anthropology*, 41:93–136.
- [121] Potts, R. (1998b). Variability selection in hominid evolution. *Evolutionary Anthropology*, 7(3):81–96.
- [122] Prescott, T. J., Bryson, J. J., and Seth, A. K. (2007). Introduction: Modelling natural action selection. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 362(1485):1521–1529.
- [123] Reader, S. M. and Biro, D. (2010). Experimental identification of social learning in wild animals. *Learning & Behavior*, 38(3):265–283.
- [124] Reader, S. M. and Laland, K. N. (2002). Social intelligence, innovation, and enhanced brain size in primates. *Proceedings of the National Academy of Sciences, USA*, 99(7):4436–4441.
- [125] Rendell, L., Boyd, R., Cownden, D., Enquist, M., Eriksson, K., Feldman, M. W., Fogarty, L., Ghirlanda, S., Lillicrap, T., and Laland, K. N. (2010). Why copy others? insights from the social learning strategies tournament. *Science*, 328(5975):208–213.

- [126] Rendell, L., Fogarty, L., Hoppitt, W. J., Morgan, T. J., Webster, M. M., and Laland, K. N. (2011). Cognitive culture: theoretical and empirical insights into social learning strategies. *Trends in Cognitive Sciences*, 15(2):68–76.
- [127] Rezaei, G. (2011). *Enhanced Social Learning via Trust and Reputation Mechanisms in Multi-agent Systems*. PhD thesis, The University of Melbourne.
- [128] Robinson, E., Ellis, T., and Channon, A. (2007). Neuroevolution of agents capable of reactive and deliberative behaviours in novel and dynamic environments. In Almeida e Costa, F., editor, *Advances in Artificial Life: Proceedings of the 9th European Conference on Artificial Life*, pages 345–354. Springer.
- [129] Rogers, A. R. (1988). Does biology constrain culture? *American Anthropologist*, 90(4):819–831.
- [130] Rose, N. (1998). Controversies in meme theory. *Journal of Memetics-Evolutionary Models of Information Transmission*, 2(1):43.
- [131] Schlag, K. H. (1998). Why imitate, and if so, how? *Journal of Economic Theory*, 78(1):130–156.
- [132] Schut, M. C., Haasdijk, E., and Eiben, A. (2009). What is situated evolution? In *2009 IEEE Congress on Evolutionary Computation (CEC)*, pages 3277–3284. IEEE Press.
- [133] Sergio, F. and Newton, I. (2003). Occupancy as a measure of territory quality. *Journal of Animal Ecology*, 72(5):857–865.
- [134] Seth, A. K. and Edelman, G. M. (2004). Environment and behavior influence the complexity of evolved neural networks. *Adaptive Behavior*, 12(1):5–20.

- [135] Soltoggio, A. (2008). *Evolutionary and Computational Advantages of Neuromodulated Plasticity*. PhD thesis, University of Birmingham, UK.
- [136] Soltoggio, A., Bullinaria, J. A., Mattiussi, C., Dürr, P., and Floreano, D. (2008). Evolutionary advantages of neuromodulated plasticity in dynamic, reward-based scenarios. In Bullock, S., Nobel, J., Wilson, R. A., and Bedau, M. A. A., editors, *Artificial Life XI: The Eleventh International Conference on the Synthesis and Simulation of Living Systems*, pages 569–576. MIT Press.
- [137] Stamps, J. A. (1988). Conspecific attraction and aggregation in territorial species. *The American Naturalist*, 131(3):329–347.
- [138] Stanley, K. O. and Miikkulainen, R. (2002). Evolving neural networks through augmenting topologies. *Evolutionary computation*, 10(2):99–127.
- [139] Stanton, A. and Channon, A. D. (2015). Incremental neuroevolution of reactive and deliberative 3d agents. In *Proceedings of the European Conference on Artificial Life 2015*, pages 341–348.
- [140] Stephens, P. A., Sutherland, W. J., and Freckleton, R. P. (1999). What is the allee effect? *Oikos*, pages 185–190.
- [141] Tattersall, I. (2009). Human origins: Out of africa. *Proceedings of the National Academy of Sciences, USA*, 106(38):16018–16021.
- [142] Templeton, C. N., Philp, K., Guillette, L. M., Laland, K. N., and Benson-Amram, S. (2017). Sex and pairing status impact how zebra finches use social information in foraging. *Behavioural Processes*, 139:38–42.
- [143] Thornton, A. (2008). Social learning about novel foods in young meerkats. *Animal Behaviour*, 76:1411–1421.

- [144] Thornton, A. and Clutton-Brock, T. (2011). Social learning and the development of individual and group behaviour in mammal societies. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 366(1567):978–987.
- [145] Thorpe, W. H. (1956). *Learning and Instinct in Animals*. Harvard University Press.
- [146] Tinbergen, N. (1963). On aims and methods of ethology. *Zeitschrift fur Tierpsychologie*, 20:410–433.
- [147] Tomasello, M. (1999). The human adaptation for culture. *Annual Review of Anthropology*, 28:509–529.
- [148] Turner, C., Giraldeau, L.-A., and Flynn, E. (2016). How does the reliability of a model affect children’s choice to learn socially or individually? *Evolution and Human Behavior*.
- [149] Valone, T. J. (1989). Group foraging, public information, and patch estimation. *Oikos*, pages 357–363.
- [150] Valone, T. J. (2007). From eavesdropping on performance to copying the behavior of others: a review of public information use. *Behavioral Ecology and Sociobiology*, 62(1):1–14.
- [151] van Bergen, Y., Coolen, I., and Laland, K. N. (2004). Nine-spined sticklebacks exploit the most reliable source when public and private information conflict. *Proceedings of the Royal Society of London, Series B: Biological Sciences*, 271(1542):957–962.
- [152] van der Post, D. (2008). *Learning What to Eat: Emerging Cultural Phenomena in Group Foragers*. PhD thesis, Utecht University.

- [153] van der Post, D. J., Franz, M., and Laland, K. N. (2016). Skill learning and the evolution of social learning mechanisms. *BMC Evolutionary Biology*, 16(1):166.
- [154] Vogt, P. and Haasdijk, E. (2010). Modeling social learning of language and skills. *Artificial Life*, 16(4):289–309.
- [155] Wagner, R. H. and Danchin, É. (2003). Conspecific copying: a general mechanism of social aggregation. *Animal Behaviour*, 65:405–408.
- [156] Watson, R. A., Ficiej, S., and Pollack, J. B. (1999). Embodied evolution: Embodying an evolutionary algorithm in a population of robots. In *Evolutionary Computation, 1999. CEC 99. Proceedings of the 1999 Congress on*. IEEE Press.
- [157] Watson, R. A. and Szathmáry, E. (2016). How can evolution learn? *Trends in Ecology & Evolution*, 31(2):147–157.
- [158] Webb, B. (2009). Animals versus animats: Or why not model the real iguana. *Adaptive Behavior*, 17(4):269–286.
- [159] Webster, M. M. and Laland, K. N. (2015). Public information use by foraging ninespine sticklebacks: Social learning or an unlearned social influence on travel direction? *Behaviour*, 152(11):1569–1584.
- [160] Whitehead, H. and Richerson, P. J. (2009). The evolution of conformist social learning can cause population collapse in realistically variable environments. *Evolution and Human Behavior*, 30(4):261–273.
- [161] Whiten, A., Goodall, J., McGrew, W. C., Nishida, T., Reynolds, V., Sugiyama, Y., Tutin, C. E., Wrangham, R. W., and Boesch, C. (1999). Cultures in chimpanzees. *Nature*, 399(6737):682–685.

- [162] Whiten, A. and Ham, R. (1992). On the nature and evolution of imitation in the animal kingdom: Reappraisal of a century of research. *Advances in the Study of Behaviour*, 21:239–283.
- [163] Whiten, A., Hinde, R. A., Laland, K. N., and Stringer, C. B. (2011). Culture evolves. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 366(1567):938–948.
- [164] Whiten, A. and Van Schaik, C. P. (2007). The evolution of animal ‘cultures’ and social intelligence. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, 362(1480):603–620.
- [165] Wilson, E. O. (1975). *Sociobiology: The New Synthesis*. Harvard University Press.
- [166] Yamamoto, S., Humle, T., and Tanaka, M. (2013). Basis for cumulative cultural evolution in chimpanzees: social learning of a more efficient tool-use technique. *PLoS One*, 8(1):e55768.
- [167] Yang, S. M. and Meng, M. (2000a). An efficient neural network method for real-time motion planning with safety consideration. *Robotic and Autonomous Systems*, 32:115–128.
- [168] Yang, S. X. and Meng, M. (2000b). An efficient neural network approach to dynamic robot motion planning. *Neural Networks*, 13(2):143–148.
- [169] Zentall, T. R. (1996). An analysis of imitative learning in animals. In Heyes, C. and Galef, B. J., editors, *Social Learning in Animals: The Roots of Culture*, pages 221–243. Elsevier.

- [170] Zwirner, E. and Thornton, A. (2015). Cognitive requirements of cumulative culture: teaching is useful but not essential. *Scientific Reports*, 5(16781).