

**Warning Signal Evolution in Natural and Virtual Populations**

by  
**Christopher David Beatty**

A thesis submitted to  
the Faculty of Graduate Studies and Research  
in partial fulfillment of  
the requirements for the degree of  
Doctor of Philosophy

Department of Biology

Carleton University

Ottawa, Ontario

June 2006

©2006, Christopher D. Beatty



Library and  
Archives Canada

Bibliothèque et  
Archives Canada

Published Heritage  
Branch

Direction du  
Patrimoine de l'édition

395 Wellington Street  
Ottawa ON K1A 0N4  
Canada

395, rue Wellington  
Ottawa ON K1A 0N4  
Canada

*Your file* *Votre référence*  
*ISBN: 978-0-494-18209-3*  
*Our file* *Notre référence*  
*ISBN: 978-0-494-18209-3*

**NOTICE:**

The author has granted a non-exclusive license allowing Library and Archives Canada to reproduce, publish, archive, preserve, conserve, communicate to the public by telecommunication or on the Internet, loan, distribute and sell theses worldwide, for commercial or non-commercial purposes, in microform, paper, electronic and/or any other formats.

The author retains copyright ownership and moral rights in this thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without the author's permission.

**AVIS:**

L'auteur a accordé une licence non exclusive permettant à la Bibliothèque et Archives Canada de reproduire, publier, archiver, sauvegarder, conserver, transmettre au public par télécommunication ou par l'Internet, prêter, distribuer et vendre des thèses partout dans le monde, à des fins commerciales ou autres, sur support microforme, papier, électronique et/ou autres formats.

L'auteur conserve la propriété du droit d'auteur et des droits moraux qui protègent cette thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

---

In compliance with the Canadian Privacy Act some supporting forms may have been removed from this thesis.

Conformément à la loi canadienne sur la protection de la vie privée, quelques formulaires secondaires ont été enlevés de cette thèse.

While these forms may be included in the document page count, their removal does not represent any loss of content from the thesis.

Bien que ces formulaires aient inclus dans la pagination, il n'y aura aucun contenu manquant.

  
**Canada**

## ABSTRACT

My research program investigated two interrelated phenomena: the evolution of aposematism (conspicuous warning signals) and Müllerian mimicry. Both of these phenomena involve the communication of unprofitability of prey to predators (or to potential mates) and both have been of interest to evolutionary biologists since the beginning of the Darwinian era. Despite this, many fundamental questions about these phenomena still remain.

My thesis research included two specific research efforts: **(1)** the investigation of aposematism and Müllerian mimicry using interactive computer simulations and **(2)** research on the unique case of intraspecific aposematism in the Coenagrionid damselflies. Through computer simulation experiments I asked questions about the conditions under which warning signals could have initially evolved, considering the fate of rare conspicuous mutants, the case in which mimicry by palatable species is possible (Batesian mimicry) and the interaction between the selection forces affecting the evolution of warning signals and prey aggregation. I also investigated the conditions under which selection for Müllerian mimicry was strong. Findings in these experiments included the importance of signal reliability on the evolution of warning signals, as well as the importance of chance effects in the establishment of rare conspicuous mutants. My research also revealed that Müllerian mimicry is most likely to evolve when a large number of prey species are present, such that it may be difficult for predators to remember each individual prey type.

Most species employ warning signals to advertise their unprofitability as prey to potential predators. But might warning signals function within species? It had been previously suggested that Coenagrionid damselflies, which are sexually dimorphic with brightly coloured males, display colour differences due to signaling between males, with bright coloration acting as a badge of maleness and thus warding off the unwanted approaches of other males. Here we tested some of the predictions of this theory, asking whether males showing different colours received increased harassment by conspecific males, and whether this harassment equates to some kind of cost. We have found that males displaying colours similar to females experience significantly higher harassment, though we did not detect a significant difference in survivorship related to this harassment.

## ACKNOWLEDGEMENTS

It is difficult to adequately thank all of the people who have made this thesis possible. I should first thank my supervisor, Dr. Tom Sherratt. I could not have asked for a better advisor than I have found in Tom. His keen scientific mind, his enthusiasm for science and his personable nature have made work in his lab both scientifically productive and a pleasure.

I am also grateful for the guidance of the members of my committee: Dr. Mark Forbes, Dr. Jaroslav Picman and Dr. Rees Kassen. Their advice and support have helped to make my doctoral research a successful endeavour. I thank Dr. Mike Speed for serving as the external examiner for my thesis defence. I also am grateful to my office-mate, Arash Rashed, for his camaraderie during our shared time as PhD students—I have had great fun working with Arash, and I hope I have offered him at least a portion of the support that he has offered me.

A large proportion of the research contained in this thesis has already been published and I am grateful to those colleagues who have assisted in this process through reviewing manuscripts. For Chapter 1 I thank Drs. J.-G. Godin, G. Roberts, M. Speed and D. Wilkinson. For Chapter 2 I thank my co-author Dr. Rod Bain, R. Tittler for statistical advice and Dr. O. Leimar for reviewing the manuscript. For Chapter 3 I wish to thank my co-author Kirsten Beirinckx, as well as F. Batiste, A. Rashed and Drs. G. Ruxton, M. Speed, H. Van Gossum and D. Wilkinson for comments on our manuscript. As the first three chapters of this thesis represent research utilizing simulation experiments of warning signals,

I need to also thank the students of Carleton University for serving as 'predators'. For Chapter 4 I would like to thank F. Phelan and the staff of the Queen's University Biological Station for their assistance in implementing my research, as well as D. Lowe and C. Warmington for assistance in the field. I also thank Dr. D. Paulson for providing the image for Figure 4-1.

Finally, I need to thank my wife, Katie Harding, for her constant support during the past four years. During long days and late nights performing research, drafting manuscripts or preparing lectures she has kept me going with kind words, thorough proofreading and preparation of a great deal more than her share of dinners. I couldn't have finished this thesis without her and am grateful to have her in my life.

## Table of Contents

	Page
Abstract .....	iii
Acknowledgements .....	v
Table of Contents .....	vii
List of Tables.....	ix
List of Figures .....	x
General Introduction .....	1
<b>Chapter 1. The evolution of warning signals as reliable indicators of prey defence</b>	
Preface .....	12
Introduction .....	13
Methods .....	15
Results .....	22
Discussion .....	32
<b>Chapter 2. The evolution of aggregation in profitable and unprofitable prey</b>	
Preface .....	39
Introduction .....	40
Methods .....	43
Results .....	48
Discussion .....	58

## Table of Contents

(continued)

	Page
Chapter 3. The evolution of Müllerian mimicry in multispecies communities	
Preface .....	64
Introduction .....	65
Methods .....	65
Results .....	70
Discussion .....	80
Chapter 4. Conspicuous coloration in males of the damselfly <i>Nehalennia irene</i> (Zygoptera: Coenagrionidae): do males signal their unprofitability to other males?	
Preface .....	83
Introduction .....	84
Methods .....	86
Results .....	93
Discussion .....	101
General Conclusions .....	106
Literature Cited .....	113

## List of Tables

Table		Page
1-1	A summary of the experimental trials conducted and their rationale.....	16
2-1	A summary of experimental conditions exploring the relationship between aggregation, conspicuousness and profitability .....	47
2-2	Results of repeated-measures ANOVAs on arcsine-transformed survivorship in experiment 1 through 3.....	51
3-1	The prey frequencies presented, mean number of prey encountered, mean number of prey attacked and proportion attacked of the first and last unprofitable prey encountered by predators in Experiment 1 .....	72
4-1	Replicated G-test results examining the distribution of painted colour morphs (blue, orange and black) for individuals involved in interactions in experiment 1 .....	95
4-2	Replicated G-test results examining the survivorship of different painted colour morphs (blue, orange and black).....	95

## List of Figures

Figure	Page
<p>1-1 A typical screen shot showing randomly-distributed conspicuous (with 60% red and 60% blue pixels) and cryptic prey (with 15% and 25% green pixels) on a background with 20% green pixels. The complementary pixel color for both the background and the prey was white .....</p>	19
<p>1-2 The mean numbers (<math>\pm 1</math> SE) of the cryptic morphs of undefended and defended prey attacked per replicate in the first generation of trials 2, 9 and 11 and in the first generation of trial 4 where cryptic forms were reversed .....</p>	24
<p>1-3 Results of a single replicate of trial 2. Here cryptic morphs of undefended and defended prey were similar in appearance but their conspicuous morphs were not. (a) The changes in frequencies of the conspicuous forms of undefended (dotted line, circles) and defended (continuous line, diamonds) prey when subject to continued selection by a human volunteer (b) The proportions of particular morphs (circles-cryptic undefended; triangles-conspicuous undefended; squares-cryptic defended; diamonds - conspicuous defended) attacked each generation in the same replicate (only calculated when there were at least 2 such morphs initially present) .....</p>	26

## List of Figures

(continued)

Figure	Page
1-4 Results of a single replicate of trial 11. The graphs show the changes in frequencies of the forms of (a) undefended (circles – cryptic, squares - conspicuous yellow, diamonds – conspicuous blue) and (b) defended prey (circles – cryptic, squares – conspicuous yellow, diamonds – conspicuous red) .....	29
1-5 Results of a single replicate of trial 13. All prey items could occur in an aggregated or a dispersed form .....	31
2-1 Relative mortality index (a) of aposematic, unpalatable cryptic and palatable cryptic prey items when presented solitarily and aggregated in the ‘initial origin’ experiment of Alatalo & Mappes (1996) (b) Relative mortality of aposematic, unprofitable cryptic and profitable cryptic prey items in the computer version of the ‘initial origin’ experiment.....	50
2-2 Mean $\pm$ SE proportions of prey that survived by trial in (a) experiment 2a, (b) experiment 2b, (c) experiment 2c and (d) experiment 2d.....	53
2-3 Mean $\pm$ SE proportions of prey that survived by trial in (a) experiment 3a and (b) experiment 3b.....	56
2-4 Plot of the path taken by predators in attacking prey in (a) experiment 3a, trial III and (b) experiment 3b, trial III.....	57

## List of Figures

(continued)

Figure	Page
3-1 A screen shot of the computer interface used by our human “predators” to forage for prey .....	67
3-2 A summary of experiments 2–5, in which human subjects foraged in a community of profitable and unprofitable prey.....	69
3-3 Results from Experiment 1. (a), The mean number of red and blue unprofitable prey attacked by predators ( $\pm 1$ SE) in relation to the number of these prey items initially available; (b), the implications of this behaviour for their mean per capita ‘mortality’ ( $\pm 1$ SE).....	71
3-4 The mean proportions ( $\pm 1$ SE) of each of the three forms of focal unprofitable prey attacked in experiments 2–4.....	75
3-5 The mean probabilities of predators attacking each of the focal forms of unprofitable prey on encounter, at each stage in the encounter sequence (first encountered prey of this type, second encountered prey etc) .....	76
3-6 The mean total scores ( $\pm 1$ SE) that predators achieved in each of the three systems (prey communities with plain, striped or spotted focal prey) for experiments 2-4 .....	78
4-1 Image of a male <i>Nehalennia irene</i> (courtesy D. Paulson).....	88
4-2 Experimental design for experiments 1 and 2. a) Four-cage design for experiment 1. b) Three-cage design for experiment 2.....	91

## List of Figures

(continued)

Figure	Page
4-3 Results of behavioural observations over the 5 replicates of experiment 1 (total of 20 cages).....	94
4-4 Survivorship data from experiment 1. a) Mean proportion survivorship ( $\pm 1$ SE) for each type of painted male (blue, orange and black) and females per day. b) Total proportion survived ( $\pm 1$ SE) of each type of male (blue, orange and black) and females.....	97
4-5 Results of behavioural observations over the 8 replicates of experiment 2, showing total interactions of painted males and the unpainted males placed in the cage with them.....	98
4-6 Mean proportion survived ( $\pm 1$ SE) of each type of male (blue, black and orange painted, as well as the unaltered males placed in the cage with each type of male) in experiment 2.....	100

## GENERAL INTRODUCTION

Throughout nature, we see spectacular examples of organisms that display bright colouration. While much of the colour that we see in the animal world is attributed to sexual selection (Darwin 1971), since Wallace (1867) coloration has also been interpreted as a warning that these organisms are defended (unprofitable as prey). Defences come in a variety of forms: organisms may be noxious, toxic or simply difficult to eat; they may possess a sting, or they may possess a harmful, sometimes venomous bite. Rather than hide themselves to avoid predation, these individuals advertise their defence to would-be predators, warning of their unprofitability. Warning colouration (termed “aposematism” by the biologist E. B. Poulton (1890)) is abundant throughout the world and is common to many different taxonomic groups. Aposematism and its evolutionary implications have intrigued biologists for over a century (Bates 1862, Darwin 1871, Poulton 1890, Cott 1940, Tinbergen 1984), and yet many questions remain as to its origin.

### *A Few Key Questions:*

#### *The Conspicuous nature of warning signals*

One important question about the evolution of warning signals is the nature of the signal—namely, why are warning signals conspicuous? In their research on domestic chicks, Gittleman & Harvey (1980) found that chicks learned to avoid unpalatable feed (treated with chloroquine) more readily when the feed was

conspicuous (a colour different from that of the background on which it was presented) rather than cryptic. Many authors (Guilford 1988, Alatalo & Mappes 1996, Rowe & Guilford 2000) have explored this 'receiver psychology' theory in the evolution of aposematism: namely, that aposematism capitalizes on innate biases in predators against conspicuous colours. For instance, research has demonstrated that animals are slow to forget relationships between conspicuous colours and unpalatability (Speed 2000), and that conspicuousness reduces recognition errors by predators (Guilford 1985a).

Thus, warning signals may be conspicuous because there is 'something special' about conspicuous colouration (Guilford 1988) and its effect on a predator's ability to learn. This special educational properties explanation for the evolution of warning signals has become broadly accepted, to the point that some researchers express it as an established fact: "Aposematic species are conspicuously colored because predators learn faster to avoid conspicuous patterns..." (Lindström et al. 2001)

Sherratt (2002) has suggested an alternative explanation for the evolution of conspicuous signals of unpalatability: conspicuous warning signals have arisen because signals are selected to be reliable. Reliable signals are generally considered to be honest signals (Maynard-Smith & Harper 2003). While this does not require that bright coloration in all instances be associated with a form of defence (as in the case of Batesian mimics) if bright coloration is consistently correlated in nature with defences then it is a reliable indicator of unprofitability. Warning signals may tend to be conspicuous because conspicuousness reliably

distinguishes defended species from their undefended conspecifics. This idea is not entirely new; Wallace (1867) suggested that distasteful lepidopteran larvae “...required some distinctive mark, something by which they may be contrasted with and separated from the agreeable larvae, in order that they might be freed from the attacks of birds. Brilliant colouration would be such a distinction as was required...” (p. lxxx). However, what is new is the notion that defended prey should capitalize not only on traits that distinguish them from undefended prey, but traits that are also difficult for undefended prey to exploit (Sherratt & Franks 2005). This offers a potential new explanation for why warning signals tend to be conspicuous: conspicuous defended prey are able to survive the increased visibility to predators brought on by their colouration, while conspicuous undefended prey would not survive for very long if easily spotted. In this way defended species enter a “phenotypic space” into which it is difficult for undefended species to follow.

While the influences of receiver psychology and signal reliability on the evolution of warning signals are not mutually exclusive, testing for their separate influences requires subtle differences in experimental approach. Most significantly, testing for signals reliability requires the presence not only of cryptic and conspicuous prey, but also palatable and unpalatable prey in the experimental system. Previous studies on prey conspicuousness have often presented only unpalatable prey to predators (Gittleman & Harvey 1980, Gittleman et al. 1980). Here, selection for conspicuous coloration is not only a function of predator learning about a particular prey type, but how the predator

learns about unpalatable prey while simultaneously learning about palatable prey.

*In the beginning....*

Another question about aposematism concerns its initial evolution. If aposematic forms evolve as conspicuous mutants in populations of defended, cryptic prey, then naïve predators will see these rare mutants more readily than their cryptic conspecifics, but will not have learned to associate their conspicuousness with their defence. If these mutants are sufficiently rare, they would likely be extirpated by predators and not allowed to establish. Thus, how could a population of aposematic organisms become established?

Fisher (1930) suggested that the evolution of distastefulness might have been facilitated by prey aggregation: while naïve predators might extirpate unusual, solitary mutant individuals of a defended species, defended prey in groups would be more likely to establish. This in turn could provide a means for aposematism to establish. While some have asserted that the evidence from phylogenetic studies indicates that aggregation often evolved *after* the evolution of defence and warning colouration (Sillén-Tullberg 1988, Tullberg & Hunter 1996, but see Ruxton & Sherratt 2006), others suggest that aggregation may be an important step in the path to aposematism (Alatalo & Mappes 1996, Riipi et al. 2001). We are therefore left with a further question: what role does aggregation play in the evolution of prey defence and aposematism?

### *The Red Badge of...Defence?*

A related phenomenon to aposematism is Müllerian mimicry. Often in nature, groups of aposematic species will share visual characteristics such as colour and form. These organisms are usually sympatric (often relying on the same host plant for food and for the compounds that are the source of their defence) and while they look very similar may indeed be taxonomically quite different; the “milkweed insects”, those that feed on *Brassicum sp.*, are made up of coleopterans, hemipterans and lepidopterans that have all evolved to look amazingly similar to one another.

An explanation for this convergence of defended species was first offered by Johannes “Fritz” Müller (1879). Müller postulated that when learning to identify unpalatable prey, naïve predators would be required to consume a number  $n$  of each type of unpalatable prey they encountered to learn to associate the conspicuousness of that prey with its defence. If these species looked similar, then a naïve predator would only need to consume  $n$  of the entire group of species to learn to associate the particular signal presented by these species with unpalatability. Müller demonstrated the validity of his theory mathematically—probably the first evolutionary relationship to be presented in this way. Thus **Müllerian mimicry** occurs when a group of species that are all defended look similar to one another. This is generally considered to differ from **Batesian mimicry**, in which palatable mimics resemble unpalatable “model” species, gaining protection from this similarity without possessing the concomitant defence (Bates 1862, Turner 1984, but see Speed 1993).

While Müllerian mimicry is a well-established concept in ecology (Poulton 1890, Joron 2003) many questions remain about the conditions under which it evolves. While the phenomenon of phenotypic similarity of defended species is often observed, uncertainty remains as to whether these “mimicry rings” came about through convergence of two unpalatable species toward one another (as Müller envisioned), or through the advergence of one or more unpalatable species toward a central “model” species (Mallet & Gilbert 1995).

Also, while field studies have found the effects of purifying selection necessary to maintain Müllerian mimicry (Benson 1972, Mallet & Barton 1989, Kapan 2001) and while further studies have found that common forms of unpalatable prey experience a lower per capita attack probability than rare forms (Greenwood et al. 1989, Linström et al. 2001), studies have found little or no evidence to indicate that the common form is at a selective advantage over the rarer form (Greenwood et al. 1981, Rowe et al. 2004). These findings contradict the assumption that there should be selection towards uniformity because one colour pattern is easier to learn than two (Rowe et al. 2004).

For Müller's theory to be correct, if unprofitable species are dissimilar, then predators would need to become familiar with each type of unprofitable prey that they encounter. If there are many potential prey species available to a predator, and if some of these species are sufficiently rare, it seems unlikely that a predator would be able to effectively remember their palatabilities, and may also make recognition errors (MacDougall & Dawkins 1998). Thus it seems unlikely that selection for Müllerian mimicry is generated in this way.

Another possibility is that when the number of defended species becomes too great for predators to learn each type of prey, then predators will generalize based on cues that help indicate the unprofitability of prey. Defended prey species may capitalize on this generalization by evolving to share general traits that predators can easily identify and associate with unprofitability. If predators indeed generalize in this way, then Müllerian mimicry will be more likely to evolve in conditions where predators are faced with a large number of unprofitable prey, such that they cannot learn to identify all individual types. Thus predators in this situation may rely more on discriminative learning (in which prey are categorized into classes) than associative learning (in which the visual attributes of a species are associated with its palatability) (Pearce 1994, Shettleworth 1998).

If generalization is indeed involved in the evolution of Müllerian mimicry, this may also explain how “intermediate” forms of defended species could survive. The question of the fate of intermediates in the evolution of Müllerian mimicry is still under debate (Turner 1984, Joron 2003, Balogh & Leimar 2005). If two distinct aposematic species exist, and if it would be advantageous for these two species to resemble one another, individual mutants that are intermediate in appearance would be at a disadvantage: they may not look enough like either type of aposematic prey to be protected. It is possible that Müllerian mimicry (and Batesian mimicry as well) may evolve through a two-step process: a single large mutation may bring a “mimic” close enough to a “model” that the similarity affords the mimic protection; selection may then improve the

mimic toward greater similarity to the model. If, however, predators use a discriminative rather than an associative approach to prey assessment, then perfect mimicry may not be required to establish Müllerian mimicry. Members of a mimicry ring may simply need to possess the discriminative characteristics that predators use to group them as unpalatable species. While this is an intriguing possibility, it then raises a new question: if discriminative learning is important in the evolution of Müllerian mimicry, then why are Müllerian mimics consistently such “perfect” mimics?

*And now for something completely different...intraspecific warning signals*

The traditional view of warning signals is of a prey organism signaling its unprofitability to a potential predator. But is not possible that warning signals function within a species, signaling unprofitability of another kind? One group in which this is a possibility is the insect order Odonata, commonly referred to as the dragonflies (suborder Anisoptera) and damselflies (suborder Zygoptera) (Corbet 1999). Sexual dimorphism is a widespread phenomenon among the Odonates. This dimorphism is manifested in a number of traits, in differences in overall body size, colouration and behaviour (Lajeunesse & Forbes 2003).

Males are often brightly-coloured, while females tend to be drab. These differences are usually attributed to sexual selection: males are brightly coloured as an indication to females of their fitness as potential mates, or simply as a consequence of female preference. Males perform courtship displays for females, to show off their colours and to entice females to mate with them. If

females use these cues to select mates, those males with the brightest colouration are more likely to reproduce, thus reinforcing this sexual dimorphism in later generations.

While sexual selection is likely the cause of dimorphism in other damselfly families, such as the Calopterygidae, male-female interactions in many species within the family Coenagrionidae (Odonata: Zygoptera) do not fit the standard pattern of female mate choice. Males do not defend oviposition territories, nor do they perform any courting behaviour. Males are highly aggressive toward females, often pursuing and harassing them, sometimes dragging them toward an oviposition site (Gorb 1998). Males are brightly coloured and conspicuous, but much of their colouration is oriented on the dorsal side of the body, and would not be viewed by a female in tandem (copulatory) position with the male.

Females of the Coenagrionids often demonstrate a colour polymorphism; some individuals exhibit female colouration (heteromorphs) and behaviour, while a percentage of the female population look—and sometimes behave—more like males (andromorphs). Research has shown similar female polymorphisms to be present in other groups of insects (Cook et al. 1994). Differences exist in the level of similarity of andromorphic females and males; these differences are observed both between and within species (Miller & Fincke 1999). The percentage of females that are andromorphic can also vary considerably among and within species (McKee et al. 2005).

Sherratt & Forbes (2001) suggested that this complex of colours has developed as a result of signaling between *males*, and if true represents the first

documented case of intraspecific aposematism. In their theory, males are brightly coloured as a means to signal their "maleness" –and their concomitant unprofitability as mates—to other males, and to thus deter other males from attempting to mate with them. Through verbal and quantitative models, Sherratt and Forbes demonstrated that in a situation where males engage in scramble competition for mates, sexual dimorphism is likely to evolve. The dimorphism serves as a signaling system between males, allowing them to avoid harassment by conspecifics attempting to mate with them. These mating errors, if common, would result in energetic costs and lost opportunity costs in mating as well as physical damage.

While previously research has focused on the fate of andromorphic and heteromorphic females in their interactions with males (and whether or not andromorphic females function as male mimics) (Cordero 1992, Fincke 1994, Cordero et al. 1998 Miller & Fincke 1999, Van Gossum et al. 1999, Van Gossum et al. 2001), to investigate whether the intraspecific aposematism theory may explain male conspicuousness (and possibly female polymorphisms) we must consider interactions between *males*. Firstly, we must determine whether males demonstrate effective avoidance of other males based on visual cues. Secondly, we must determine whether excessive harassment of males by other males equates to a fitness cost to the recipient of the harassment.

For my research, I have pursued a number of these questions. In Chapter 1, I explore the conspicuousness and reliability of warning signals, as well as their initial evolution. In Chapter 2, I explore the interaction of aggregation,

signaling and unprofitability. In Chapter 3 I consider the evolution of Müllerian mimicry and the conditions that can generate selection for a multi-species signal. In Chapter 4 I explore the possibility of an intra-specific warning signal through studies of the damselfly *Nehalennia irene*.

In the evolution and maintenance of warning signaling, the evolution of Müllerian mimicry in multi-species systems, and the evolution of aposematic male colouration in damselflies to avoid harassment by other males, signals appear to be selected such that they reliably indicate an individual's unprofitability and simplify the process of predator education. As Fisher observed (1930), "*being recognized as unpalatable is equivalent to avoiding confusion with palatable prey*". This common theme of avoiding confusion through reliable signaling draws together the superficially disparate phenomena that I have investigated in my research program. My results thus far indicate that this is indeed a fundamental factor in the evolution of warning signals.

## Chapter 1:

The evolution of warning signals as reliable indicators of prey defence

### Preface

In the following chapter I ask a number of basic questions about the conditions under which warning signals could evolve. This is performed through a novel experimental approach: utilizing undergraduate students as 'predators' foraging on computer-generated prey. These prey could reproduce, with mutations allowing for the introduction of conspicuous forms. This approach proved to be quite successful, and allowed us to consider experimental conditions that would be intractable to study in the field. This Chapter was published in the *The American Naturalist* in October of 2003 (v. **162(4)**: 377-389). This article was subsequently reviewed in a News & Views Article in *Nature*.

## Introduction

It has long been recognized that prey which possess significant defenses against predators tend to be conspicuous in some way (Wallace 1867; Darwin 1871; Poulton 1890). The contemporary explanation for this phenomenon, termed aposematism (Poulton 1890), is that there is “something special” (Guilford 1990; Speed 2000) about the educational properties of conspicuous traits as a signal of defense (Guilford 1988; Endler 1991; Mallet & Joron 1999). For example, it has been repeatedly shown that predators learn to avoid unpalatable prey more quickly when they are conspicuous than when they are cryptic (Gittleman & Harvey 1980; Gittleman et al. 1980; Roper & Wistow 1986). This theory for the evolution of aposematism is plausible, but there is an important caveat.

Whatever the underlying cause of aposematism, it is likely that predators would evolve an enhanced psychological predisposition to learn to avoid conspicuous prey precisely because such prey tend to be defended (Turner 1984; Sherratt 2002).

While the idea that prey signals evolve to capitalize on the pre-existing psychological biases of predators has become one of the most popular theories for the evolution of aposematism, there are several alternative theories (Guilford 1988; Endler 1991; Mallet & Joron 1999). One important rival explanation is that conspicuous traits simply make defended prey less liable to be confused with undefended prey, which are typically cryptic (Wallace 1867; Fisher 1930; Edmunds 1974; Turner 1975; Chai 1996). The idea was first hinted at by Alfred

R. Wallace (Wallace 1867; Guilford 1990), but articulated most explicitly by R.A. Fisher (1930), who compared the evolution of warning signals to the tendency of unpalatable models to evolve away from palatable mimics (page 148): *“to be recognized as unpalatable is equivalent to avoiding confusion with palatable species”*. The advantages of looking different from undefended prey are clear: predators are more likely to learn to react appropriately to a signal if it is a reliable one (Zahavi 1993), and being distinguishable is essentially creating a reliable signal. However, it is less obvious why selection to avoid confusion might act on conspicuousness and not some other distinguishing characteristic, such as pattern. One possibility is that such traits are amongst the least exploitable by undefended cheats (since non-mimetic conspicuous mutants of undefended prey would be readily detected and attacked), and therefore the most persistent in evolutionary time. In sum, conspicuous traits may have evolved as stable warning signals because these characteristics: (i) help distinguish defended prey from undefended prey and (ii) are difficult to exploit by undefended prey.

To test Fisher’s theory of evolution of warning signals as a means of avoiding confusion, and its corollary (namely, that defended prey evolve traits that are difficult to exploit by undefended prey), we investigated how the morphological and behavioral attributes of “undefended” (profitable) and “defended” (unprofitable) computer-generated prey evolved when subject to selection by foraging humans (see Bond & Kamil 1998, 2002 for an analogous

approach to crypsis using blue jays as predators). Human models have long been used to test and refine ideas relating to predation (e.g. Holling 1959; Dill 1975; Knill & Allen 1995; Glanville & Allen 1997) and here their use was desirable not just because of the flexibility they allowed, but also because participants would have few preconceptions concerning the profitability of novel prey they encountered (a “novel world”, cf Alatalo & Mappes 1996; Lindström et al. 1999; Riipi et al. 2002).

Our computer experiments allowed us to independently control both the reliability of a given trait as a signal of defence, and its conspicuousness. Many experiments have shown that conspicuous defended prey are attacked less frequently than cryptic conspecifics (e.g. Sillén-Tullberg 1985), particularly after a period of learning (e.g. Alatalo & Mappes 1996). However, we know of no experimental study on warning signals that has attempted to formally evaluate the relative importance of signal reliability compared to signal conspicuousness, or one that has simulated changes in morph frequencies of prey from generation to generation as a result of selection.

## **Methods**

To address the contrasting theories we have outlined, we conducted a total of 14 related experiments (“trials”, Table 1-1). Each of these individual trials were replicated a minimum of 5 times using different human volunteers as predators. The volunteers (more than 100 undergraduate students) had no knowledge of

**Table 1-1** A summary of the experimental trials conducted, and their rationale.

Trial no.	Morphs of undefended prey	Morphs of defended prey	Generations of prey	Mutation rate	Replication	Starting conditions *	Question
1	15% green 25% green 60% red 60% blue	None	1	Not applicable	31	5 of each prey type on each of 5 screens	Are the green morphs more cryptic than the non-green morphs, and are the 2 green morphs equally cryptic?
2	15% green 60% blue	25% green 60% red	30	5%	5	All prey cryptic	Do defended prey evolve the morph that reliably signals their defense, or do they remain cryptic?
3	15% green 60% blue	25% green 60% red	30	5%	100	All prey cryptic	Control for (2) with no predation.
4	25% green 60% red	15% green 60% blue	30	5%	5	All prey cryptic	Do we get similar results to (2) when morphs are reversed?
5	15% green 60% blue	25% green 60% red	30	5%	5	All prey cryptic	Do we get similar results to (2) when new generations occur after 30 seconds?
6	25% green 60% red	15% green 60% blue	30	5%	5	All prey conspicuous	Do we get similar results to (2) with different starting conditions?
7	65% red 30% yellow	55% red 30% green	30	5%	5	All prey in red conspicuous form	Do defended prey adopt the cryptic morph if it reliably signals defense?
8	65% red 30% yellow	55% red 30% green	30	5%	100	All prey conspicuous	Control for (7) with no predation
9	15% green 60% red	25% green 60% red	30	5%	5	All prey cryptic	Do defended prey evolve conspicuous morphs if they can be mimicked?
10	Species 1: 15% green 60% red Species 2: 15% green 60% blue	25% green 60% red	30	5%	5	All prey cryptic	Can additional non-mimetic prey allow defended prey and their undefended mimics to evolve conspicuousness?
11	15% green 60% blue 60% yellow	25% green 60% red 60% yellow	50	5%	5	All prey cryptic	Do defended prey adopt the morph that most reliably signals defense?
12	15% green, 60% blue (gen 2), 60% magenta (gen 3), 60% yellow (gen 4)	25% green, 60% red (gen 2), 60% cyan (gen 3), 60% light gray (gen 4)	30	$x = 1$ or 2 mutations per species, generations 2-4 only, no other mutations	10 replicates for each value of $x$ ( $= 1$ or 2) and $s$ ( $= 0$ or 0.5)	All prey cryptic	Do defended prey evolve conspicuous morphs when mutations are extremely rare?
13	50% green, aggregative, 50% green, dispersed	50% green, aggregative, 50% green, dispersed	30	5%	5	All prey aggregated	Do defended prey remain aggregated because it makes undefended prey vulnerable?
14	50% green, aggregative	50% green, aggregative; 50% green, dispersed	30	5% for defended prey only	5	All prey aggregated	Do defended prey become dispersed when undefended prey cannot?

\* All green morphs were considered cryptic, while morphs with 60% non-green colour were considered conspicuous. In all trials except trial 1 there were initially 50 of each species in total, with 10 of each species distributed at random on each of 5 screens

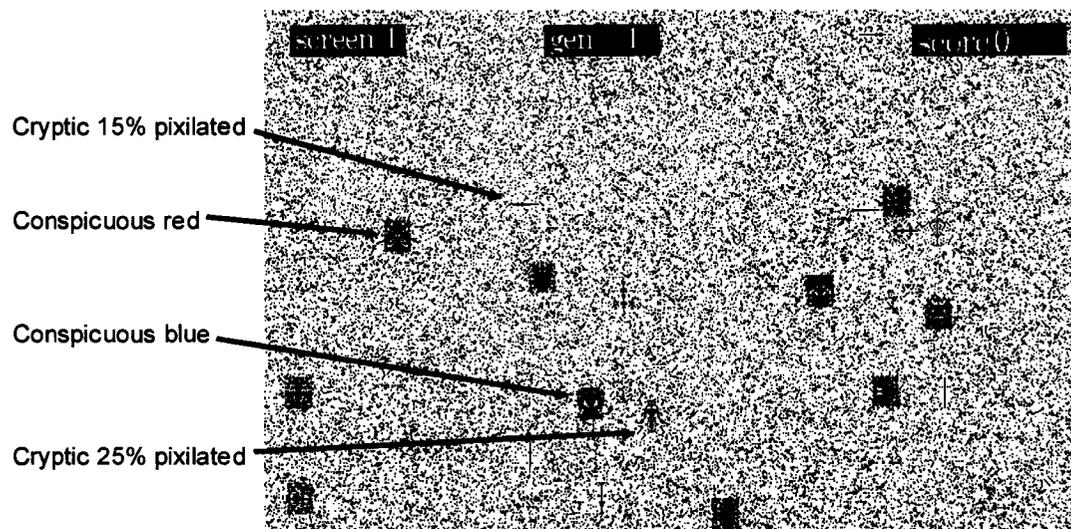
the experimental aims (the majority were non-biologists) and no individual participated in more than one replicate or trial. In each replicate, a single predator typically foraged for prey until 30 prey generations were complete. We therefore assumed that predators live longer than prey, and that predators can potentially remember their experiences across prey generations (cf Servedio 2000; Bond & Kamil 2002).

All experiments except trial 1 and 10 involved a single species of artificial undefended prey, and a single species of defended prey (trial 1 had no defended species; trial 10 had an additional undefended species). On attacking an undefended prey item a human predator added  $b$  ( $= 1$ ) points to its total displayed score, while on attacking a defended prey item a predator lost  $c$  ( $= 1$ ) points from its score. To enhance the stimulus (Rowe 2002) and to allow predators to know what they had attacked without having to view their score, attacks on undefended and defended prey generated distinct sounds (undefended - high pitch, rising scale; defended - low pitch, falling scale). Prey items that were killed simply disappeared from the screen. Undefended prey were always killed on attack, but defended prey survived attack (Järvi et al. 1981; Wiklund & Järvi 1982) with a fixed probability ( $s$ ), in which case they temporarily changed colour (to highlight the fact that a defended prey had been attacked), then returned to their former appearance. We set  $s$  to zero except

where stated, since the evolution of aposematism is least likely to arise under these extreme conditions.

The foraging environment of our volunteers consisted of a series of 5 separate screens, which could be scrolled through by pressing on the keyboard space bar. Each screen consisted of a random mosaic of 20% green and 80% white pixels. At the start of each generation there were 10 individual prey items of each of the two prey species, randomly distributed on each of these screens (except where stated) (Figure 1-1). Presenting predators with mixtures of prey species is the standard experimental practice in studies of this kind (e.g. Alatalo & Mappes 1996; Lindström et al. 1999), and we chose to distribute our artificial prey across several screens because natural predators cannot search in all areas of their environment simultaneously, and because it is quite possible that some individual prey never come into contact with their key predators at all.

All prey items were square (20 x 20 pixels, except in our aggregation trials - see below), and symmetrical in pattern around a central vertical or horizontal line (a “backbone”), which was drawn in its non-white colour. Other aspects of appearance of each individual prey were controlled by two genes. The first gene gave the percentage of pixels in that individual that were non-white (with the exception of its backbone), while the second gene coded for that non-white colour. We appreciate that the genetics underlying animal coloration are much more complex than this, but aposematism is a taxonomically widespread phenomenon which is likely to be controlled by many different genetic systems,



**Figure 1-1** A typical screen shot showing randomly-distributed conspicuous (with 60% red and 60% blue pixels) and cryptic prey (with 15% and 25% green pixels) on a background with 20% green pixels. The complementary pixel colour for both the background and the prey was white.

and we saw no need to invoke anything more sophisticated to test the intuitive ideas we have described.

At the outset, players were shown how to move between screens and attack prey. They were made aware that prey could “mutate”, and asked simply to maximize their personal scores by attempting to attack profitable (undefended) prey, while avoiding unprofitable (defended) prey. Generations came to an end when 30% of the total prey population had been killed or (in trial 5) after a short fixed time (30 seconds) - no species extinctions occurred using these algorithms. Surviving individuals of each prey species at the end of each generation were allowed to reproduce back to their starting density by randomly selecting a parent for each of the new generation from the total pool of available survivors. These offspring were randomly distributed amongst screens (10 per screen). During reproduction there was a 5% chance, analogous to a mutation, that an offspring would have a different appearance to that of its parent. In most cases this meant a radical increase or decrease in conspicuousness, which was brought about by altering the offspring genotype for both the percentage of non-white pixels, and the non-white colour (cryptic forms were typically 15% or 25% green, while conspicuous forms were typically 60% blue, red or yellow).

In trial 1 we formally tested whether morphs which we refer to as cryptic were indeed more difficult to detect than forms which we refer to as conspicuous by running 31 separate replicates with 5 of each of the 4 prey types listed in Table 1, distributed on each of 5 screens. The human subjects were informed

that all the prey types were profitable and asked to forage at their discretion. Each replicate ended after 25 prey items had been attacked. All other trials (2-15) included defended prey, and were replicated a minimum of 5 times for a minimum 30 generations each (trial 11 was conducted for 50 generations while trial 12 was replicated 10 times for each of four parameter combinations).

In trials 2-8, members of defended and undefended species could each occur in one of two discrete morphs: a cryptic or a conspicuous form. In trials 2-6 the cryptic morphs of the two prey species were similar in overall appearance, but the conspicuous forms were not. Conversely in trials 7-8, the conspicuous forms of undefended and defended prey were similar in appearance (65% and 55% red respectively), but the more cryptic forms were readily distinguishable (30% yellow and 30% green respectively). In trials 9-11 we investigated what forms evolved in defended and undefended prey when complete mimicry was possible (trial 9), when there were two undefended prey species and only one could be mimetic (trial 10) and when undefended and defended prey had two conspicuous morphs, only one of which was identical (trial 11).

Modifications to the basic experimental design were necessary to test specific ideas. To quantify just how readily aposematism emerged when mutations were extremely rare, and to simultaneously explore the effect of varying  $s$  ( $= 0$  or  $0.5$ ), we began trial 12 with cryptic undefended and defended prey, and introduced  $x$  ( $= 1$  or  $2$ ) novel conspicuous mutants of each species each generation over generations 2-4 (no other mutations were allowed, thereby

restricting the mutation rate to 3 or 6 individual mutations in 1500 offspring over the course of 30 generations). The rare mutational forms were completely distinct in that their colour not only differed between species, but also between generations.

We also modified our approach to understand the evolution of other distinguishing traits in defended prey besides their conspicuousness. Thus, in trials 13 and 14 we kept defended and undefended prey 50% green, but allowed selection on their degree of aggregation by introducing a “clustering” gene: all individuals on a given screen that had allele “1” for this gene were distributed near a randomly selected cluster-point on that screen for that species, while individuals with allele “0” were placed at a random position within the screen. In this set of trials we reduced prey size to 5 x 5 pixels, but all other default parameter values were the same ( $s = 0$  and 5% probability of mutation of the aggregation allele during reproduction).

Student t-tests examined whether defended and undefended species, or the same species in control and treatment, differed significantly in their final mean compositions. All proportion data were arcsine transformed prior to testing. Means were expressed  $\pm 1$  standard error.

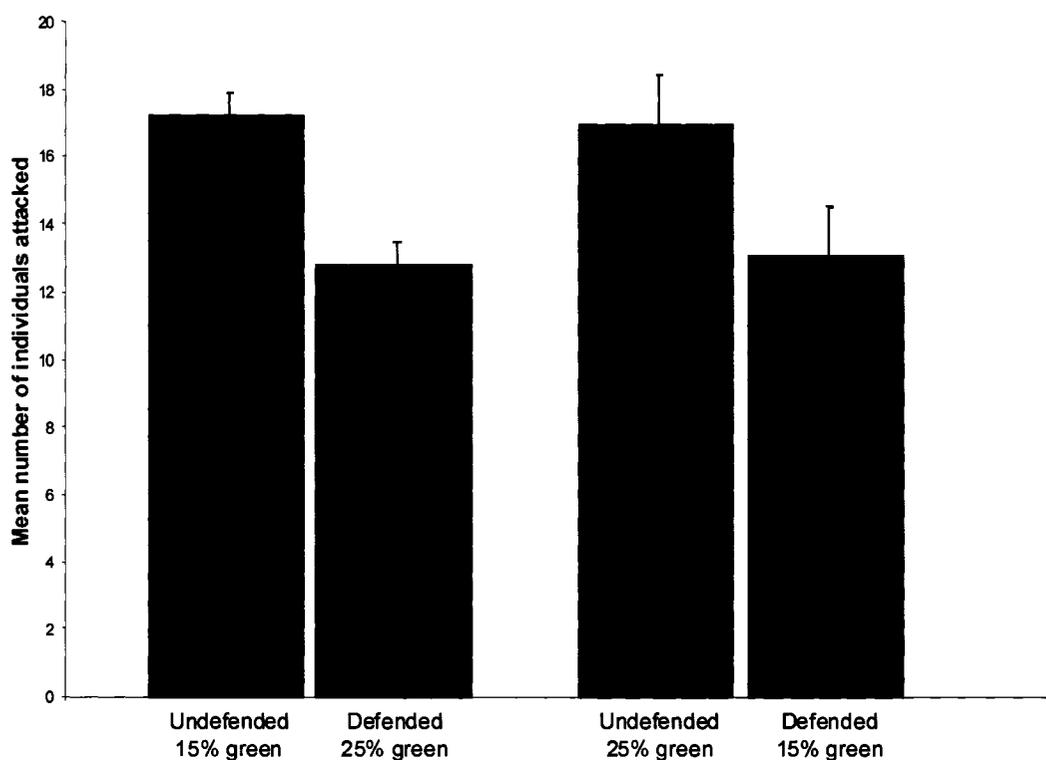
## **Results**

Any prey item with a similar proportion of green pixels as their background was difficult to detect. These subjective impressions were confirmed from our

analysis of trial 1 where the 15% and 25% green morphs comprised 17.4 and 17.0% of the total diet respectively while the red and blue morphs comprised 32.6 and 32.9% (test for homogeneity  $G = 76.25$ ,  $df = 3$ ,  $P < 0.001$ ). The two green morphs were approximately equally cryptic ( $G = 0.034$ ,  $df = 1$ ,  $P > 0.05$ ), as were the two non-green morphs ( $G = 0.008$ ,  $df = 1$ ,  $P > 0.05$ ).

To assess how readily our two types of standard cryptic prey were discriminated, we compared the numbers of undefended (15% green) and defended prey (25% green) attacked in the first generation of trials 2, 9 and 11, before any conspicuous mutants had arisen. If predators could tell the difference between these two prey types, then one would expect more undefended prey to be attacked. As a control, we also compared the attack rates on the two prey types in the first generation of trial 4 where the cryptic morphs were switched for defended and undefended prey (Figure 1-2). In both cases, predators showed clear evidence of being able to discriminate between the two cryptic prey types (one-tailed tests: trials 2,9,11,  $t_s = 3.983$ ,  $df = 28$ ,  $P < 0.001$ ; trial 4,  $t_s = 2.127$ ,  $df = 8$ ,  $P < 0.05$ ), although their discrimination was by no means perfect. Analysis of the attack sequences show that even towards the end of the first generation, mistakes were being made.

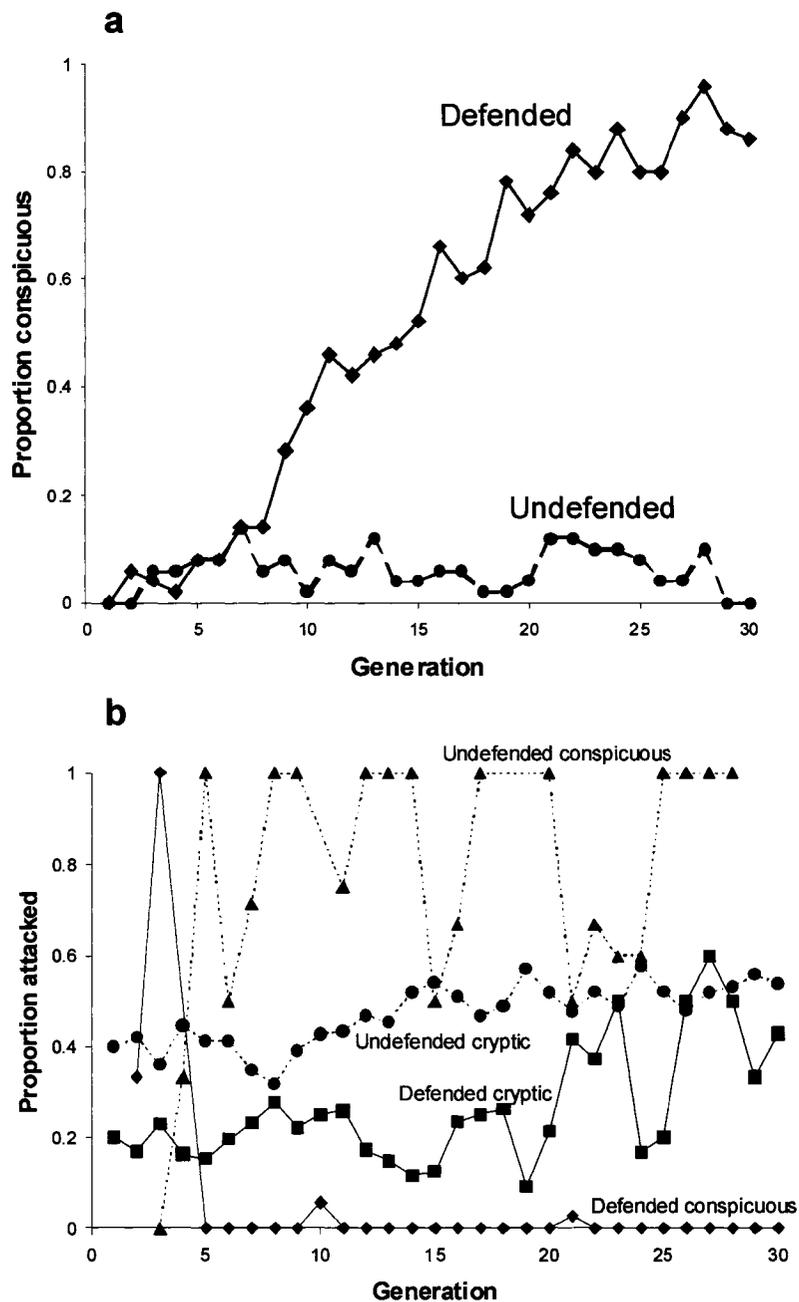
Starting with populations composed entirely of cryptic prey (trial 2), the undefended species tended to remain cryptic over 30 generations of selection, while the defended species invariably became conspicuous (Figure 1-3a shows a typical replicate; the final mean proportions of conspicuous undefended and



**Figure 1-2** The mean numbers ( $\pm 1$  SE) of the cryptic morphs of undefended and defended prey attacked per replicate in the first generation of trials 2, 9 and 11 and in the first generation of trial 4 where cryptic forms were reversed. No other morphs were present. In both cases, significantly more undefended prey were attacked per replicate than defended prey, suggesting that humans could tell these two equally cryptic morphs apart, but not perfectly.

defended prey in all 5 replicates of this trial were 0.068 and 0.864 respectively ( $t_s = 11.816$ ,  $df = 8$ ,  $P < 0.001$ ). This highly significant difference in conspicuousness between defended and undefended species cannot be explained by chance mutation and drift, which would act the same on the two species of prey. Indeed, control simulations with no predation (trial 3, 100 replicates) gave very different results (e.g. final mean proportion of undefended prey that were conspicuous  $0.480 \pm 0.016$ , comparison of trial 2 and 3  $t_s = 7.166$ ,  $df = 103$ ,  $P < 0.001$ ). Most importantly, there was direct evidence of selection in all 5 replicates (and indeed in replicates of related trials): while predators were more likely to attack conspicuous defended prey when they first appeared than their cryptic conspecifics, they eventually avoided such prey types almost entirely (Figure 1-3b). The final mean score of predators in the first generation of prey in this trial was low (1.2) and not significantly different from zero ( $t_s = 0.514$ ,  $df = 4$ ,  $P > 0.05$ ). By contrast, the mean score of predators in the 30<sup>th</sup> prey generation (25.2) was significantly higher than zero ( $t_s = 31.50$ ,  $df = 4$ ,  $P < 0.001$ ).

It is of course possible that the above results arose out of a tendency of the human predators to avoid red, or as some unexpected consequence of the criterion for starting a new generation. To control for these possibilities, we ran further replicated trials in which the two cryptic and two conspicuous mutational forms were switched for defended and undefended prey species (trial 4), and in



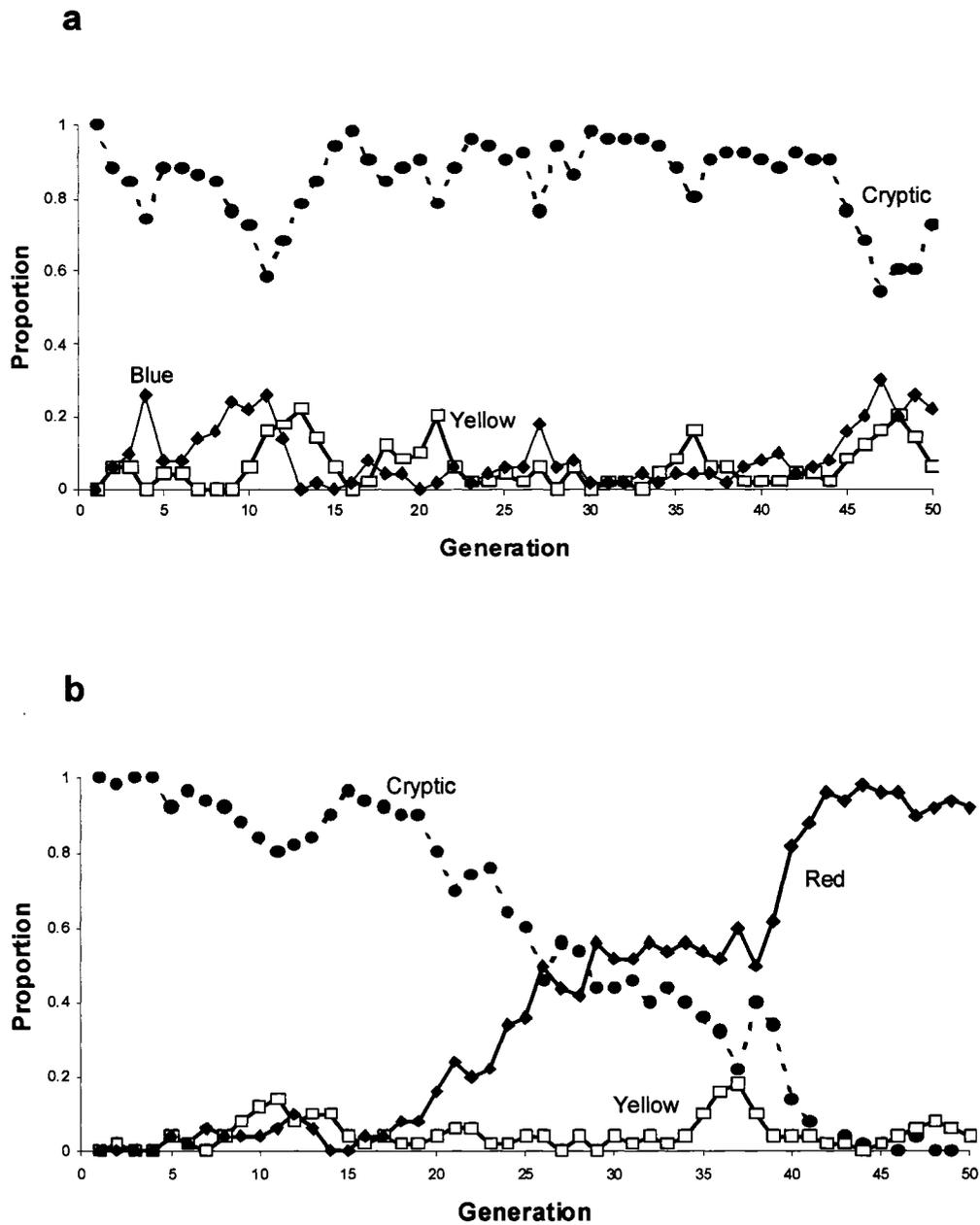
**Figure 1-3** Results of a single replicate of trial 2. Here cryptic morphs of undefended and defended prey were similar in appearance but their conspicuous morphs were not. (a) The changes in frequencies of the conspicuous forms of undefended (dotted line, circles) and defended (continuous line, diamonds) when subject to continued selection by a human volunteer (b) The proportions of particular morphs (circles - cryptic undefended; triangles – conspicuous undefended; squares - cryptic defended; diamonds - conspicuous defended) attacked each generation in the same replicate (only calculated when there were at least 2 such morphs initially present). Note that proportionately more undefended cryptic prey tended to be attacked than defended cryptic prey.

which new generations arose after a fixed time rather than after consuming a fixed proportion of prey (trial 5). In both cases, the final mean proportion of conspicuous defended prey was significantly higher than the final mean proportion of conspicuous undefended prey (when colors were reversed: conspicuous defended prey 0.744, conspicuous undefended prey 0.072,  $t_s = 8.365$ ,  $df = 8$ ,  $P < 0.001$ ; when time based algorithm was employed: conspicuous defended prey 0.728, conspicuous undefended prey 0.252,  $t_s = 3.706$ ,  $df = 8$ ,  $P < 0.01$ ). As a final control we noted that if the green morphs were cryptic, then one would expect undefended prey to evolve their green morph even if they started out in their conspicuous form (trial 6). As anticipated, undefended prey rapidly became cryptic while defended prey remained conspicuous (final mean proportions of conspicuous undefended and defended prey were 0.316 and 0.900 respectively,  $t_s = 13.106$ ,  $df = 8$ ,  $P < 0.001$ ).

In the above trials the cryptic forms of prey were readily confused, but the conspicuous forms were not. In trial 7 we examined how prey evolved when undefended and defended cryptic morphs were distinct in appearance, while their more conspicuous forms were not. Despite the potential for predators to more rapidly learn to associate noxious qualities with conspicuous traits (Guilford 1990; Speed 2000), defended species always evolved their distinct cryptic form (final mean proportion of defended prey that were cryptic  $0.784 \pm 0.06$ , when compared to control trial 8 with no predation (effectively the same as trial 3),  $t_s = 4.474$ ,  $df = 103$ ,  $P < 0.001$ ).

In the preceding experimental trials (2, 4-7) we found that defended prey consistently adopted traits that allowed them to be distinguished from undefended prey. It was therefore of interest to determine what traits tended to evolve in defended prey when many of their characteristics could be mimicked by undefended prey. In trial 9 we ran 5 replicates in which defended and undefended prey had similar but distinguishable cryptic forms and an identical conspicuous form. Aposematism did not evolve under these conditions (final mean proportions of undefended prey and defended conspicuous prey 0.40 and 0.252 respectively,  $t = 0.14$ ,  $df = 8$ ,  $P > 0.05$ ). Interestingly, when we introduced a third, non-mimetic undefended prey species at the same density (trial 10), then predators eventually foraged mainly on this species, and both the defended species and its potential mimic gained selective advantage by being conspicuous (final mean proportions of conspicuous forms in defended and mimetic species  $0.796 \pm 0.135$  and  $0.788 \pm 0.139$ ). Similarly, when we allowed 2 conspicuous colour morphs for each species, one of which was identical between species (trial 11), defended species eventually evolved their unique conspicuous form while undefended species remained cryptic (Figure 1-4a, b). The final mean proportions of undefended and defended prey that were conspicuous following five such replicates were 0.388 (65.6% of which were yellow) and 0.956 (2.9% of which were yellow) respectively ( $t_s = 5.816$ ,  $df = 8$ ,  $P < 0.001$ ).

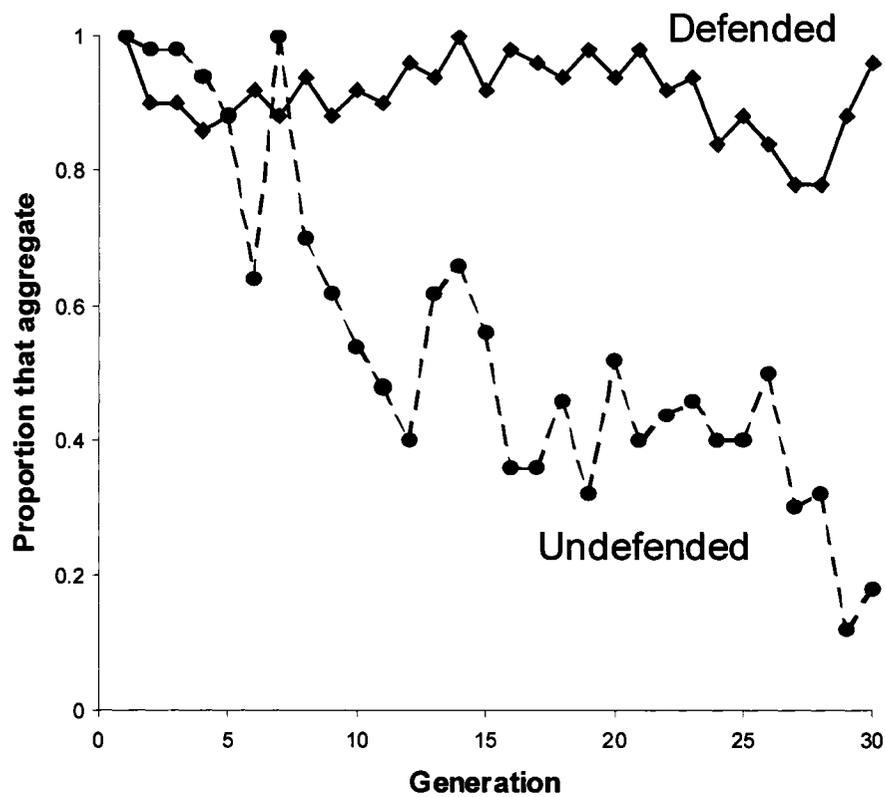
While conspicuous traits readily arose as reliable warning signals in many of the above trials, it is clear that this may have been facilitated by the continued re-occurrence of identical mutational forms in a given species. To quantify just



**Figure 1-4** Results of a single replicate of trial 11. Undefended prey could occur either in a 15% green (cryptic), a 60% blue or a 60% yellow form (both conspicuous). Defended prey could occur either in a 25% green (cryptic), a 60% red or a 60% yellow form (both conspicuous). The graphs show the changes in frequencies of the forms of (a) undefended (circles – cryptic, squares – conspicuous yellow, diamonds – conspicuous blue) and (b) defended prey (circles – cryptic, squares – conspicuous yellow, diamonds – conspicuous red).

how frequently aposematism might arise when mutations were extremely rare, we allowed only  $x$  ( $= 1$  or  $2$ ) distinct conspicuous mutants to appear in each species over generations 2-4 (trial 12). The conspicuous forms of undefended prey never spread. By contrast, the final population of defended prey consisted of more than 50% of the conspicuous form(s) by generation 30 on 3/10 of replicates for  $s = 0, x = 1$ , 1/10 of replicates for  $s = 0.5, x = 1$ , 6/10 of replicates for  $s = 0, x = 2$  and 5/10 of replicates for  $s = 0.5, x = 2$ . Here conspicuous defended mutants managed to spread initially from extreme rarity primarily by drift (the same chance mechanism occasionally reduced small but growing populations of conspicuous defended mutants to extinction), with aposematism more likely to arise when the original mutations first appeared in screens that were unvisited.

In our final trials 13 and 14 we allowed selection on the tendency of prey to aggregate. Starting with aggregated defended and undefended prey, undefended prey tended to lose their aggregation (although not entirely), while defended prey retained theirs (Figure 1-5). The overall final mean proportions of undefended and defended prey that were aggregated following five such replicates were 0.516 and 0.908 respectively ( $t_s = 4.315$ ,  $df = 8$ ,  $P < 0.004$ ). Yet when undefended prey were forced to remain aggregated by preventing their mutation, then defended prey became disaggregated (mean proportion defended prey with cluster gene 0.432, difference in proportions of aggregated defended prey between the trials  $t = 9.746$ ,  $df = 8$ ,  $P < 0.001$ ).



**Figure 1-5** Results of a single replicate of trial 13. All prey items could occur in an aggregated or a dispersed form. The graph shows the changes in the proportion of individuals carrying the aggregation allele in undefended (dotted line, circles) and defended prey (continuous line, diamonds) when subject to selection by a human volunteer.

## Discussion

Although there appears to be a generally accepted hypothesis for the widespread occurrence of aposematism (Lindström et al. 2001, page 9181: “*Aposematic species are conspicuously coloured because predators learn faster to avoid conspicuous patterns...*”) we do not believe that there is enough empirical evidence to accept this theory without reservation, particularly when there are plausible alternatives. Aposematism frequently arose in our study and we believe that its emergence can be understood without the need to invoke any special educational properties of conspicuous signals. Indeed, our study strongly suggests that even cryptic defended prey will be avoided if such prey are distinct and reliably defended (trials 7-8), and this intuitive result is supported by a number of earlier experiments on warning signals (Papageorgis 1975; Gittleman & Harvey 1980; Gittleman et al. 1980). As Mallet & Singer (1987) report, many ithomiine butterflies are highly unpalatable, but to human eyes they appear more distinctive in appearance than conspicuous.

Clearly we have not unequivocally proved that aposematism has arisen as a consequence of selection on defended prey to avoid being confused with undefended prey. However, our results go some way to confirming Fisher’s (1930) intuition that if defended and undefended prey are readily confused, then

there will be strong selection on defended prey to exhibit characteristics that reliably distinguish them from undefended prey. Although novel conspicuous mutants of defended prey were always at an early selective disadvantage compared to cryptic conspecifics (Figure 1-3a), informal student feedback suggested that they rapidly learned to avoid these prey types (with occasional re-evaluations) because they were consistently unprofitable. By contrast, cryptic prey were frequently profitable, giving rise to selection in defended prey to evolve a form which could more reliably reflect their unprofitability.

Fisher's perspective is very different from the contemporary approach which has tended to focus on the evolution of effective educational aids in defended prey alone, rather than on selection to maximize the phenotypic difference between defended and undefended prey. Nevertheless, there is some overlap between Fisher's theory and certain elements of the "receiver psychology" literature as it applies to aposematism. In particular, as Mallet & Joron (1999) noted, Fisher's (1930) view shares many features with the more recent application of the concept of "peak shift" (Hanson 1959; Leimar et al. 1986; Gamberale & Tullberg 1996; Gamberale-Stille & Tullberg 1999), whereby mutants of unpalatable prey are thought to gain selective advantage by being even further apart in appearance from palatable prey than current unpalatable prey. One key distinction however is that the peak shift mechanism is typically understood to work by a gradual increase of the conspicuousness of defended prey (mutations that are only marginally more conspicuous than the current

unpalatable prey are at the greatest selective advantage), whereas Fisher's theory can explain selection for any mutant form that reduces the likelihood of its carrier from being confused with palatable prey.

We have argued that conspicuousness may be particularly effective and stable as a warning trait because it increases prey detectability, a burden which undefended prey items are unlikely to be able to adopt. Indeed in the absence of mimicry, undefended prey consistently evolved their more cryptic morph (trials 2,4-6), because conspicuous mutants of undefended prey were rapidly exterminated. Our experiments have also shown that even a single conspicuous mutant of defended prey, which is easily detected and at first not as readily recognized as defended, may occasionally spread from rarity by chance alone (trial 12). The frequency at which aposematism arose in this particular trial was impressive (up to 50% of replicates) and it is possible to argue that even had aposematism arisen in only 1% of trials, then its evolution would be likely given sufficient time and sufficient unique mutational forms. Although there is an understandable tendency to seek deterministic solutions to evolutionary problems, our results provide some of the first experimental support for the contention that chance alone can allow conspicuous mutants to spread from extreme rarity (Mallet & Singer 1987, see also Mallet & Joron 1999 for further discussion of the application of "shifting balance" theory). Once established in a local area, it is clear that aposematic forms could resist dilution by cryptic immigrant conspecifics since the former would be at a strong selective

advantage. Conversely, the continual emigration of aposematic forms may eventually allow sufficient education of neighbouring predators for aposematic forms to spread in these areas.

Many traits are capable of being exploited by others, and there is often selection to reduce the extent to which this occurs. For example, it is now widely recognized that the occurrence of non-altruists may play an important role in determining the nature of cooperation that evolves between non-relatives (see Dugatkin 1997). Likewise, it has long been appreciated that warning signals become less of a deterrent to predators when there is a high mimetic load (e.g. Sheppard 1959; Brower 1960; Pilecki & O'Donald 1971; Nonacs 1985; Lindström et al. 1997), and that models and mimics may therefore be engaged in an evolutionary "race" (Fisher 1930; Gavrillets & Hastings 1998). Given such observations, it is somewhat surprising that there has been so little discussion of the role of mimicry in shaping the initial evolution of warning signals. The fact that aposematism did not evolve when extensive amounts of mimicry was possible (trial 9) strongly supports the view that the exploitability of a signaling system is an important determinant of its long-term evolutionary success. Note that when not all prey can become mimetic (e.g. they lack the size or shape), then predators may concentrate on prey that are reliably undefended (trial 10), allowing warning signals to evolve. Similarly, we have shown that signaling may still evolve if defended prey can adopt conspicuous phenotypes that for some reason cannot be mimicked (trial 11).

When defended prey and undefended prey could occur either in an aggregated or in a dispersed form, undefended prey rapidly became more dispersed. This arose simply because predators tended to concentrate on the clumps of prey and would pick off all undefended prey in a cluster if the first prey item was found to be undefended. Yet when undefended prey were forced to remain aggregated it was defended prey that became dispersed, because all dispersed prey in this system were reliably defended. Of course there are other costs and benefits to living in groups besides signaling to predators, and groups may also be more conspicuous (Gagliardo & Guilford 1993; Riipi et al. 2002), but our results once again indicate that defended prey will tend to evolve characteristics that undefended prey cannot evolve, and/or would render undefended prey highly vulnerable.

We have primarily explored the possibility that defended prey are conspicuous because such traits set them apart from undefended prey, but it is worth noting that a similar set of arguments might also be applied to understand the evolution of other potential signals of defence, such as locomotory behavior. Slow and predictable movement is common in defended species (Pasteels et al. 1983; Chai & Srygley 1990; Srygely & Chai 1990; Hatle & Faragher 1998; Marden & Chai 1991; Hatle et al. 2002), so much so that it is typically included in the syndrome of traits associated with chemical defense (Whitman et al. 1985). In a recent study Chai (1996) commented on the rarity of Batesian mimics in neotropical butterflies, noting that (p.63) "*the low frequency of cheaters is*

*probably due to the fact that unpalatable butterflies with their associated morphological and behavioural adaptations facilitate detection and capture*". We do not believe that this is coincidental. If defended prey were selected to adopt characteristics that reduced the rate at which they are confused with undefended prey, then one such set of characteristics that might be actively selected for is slow and predictable movement, since it would render any undefended species that possessed these characteristics highly vulnerable. Of course this theory is untested, but it serves to show that the simple idea of honest signaling can explain the evolution of several very different traits regularly associated with defended prey.

It has been argued that warning signals are not handicaps (Zahavi & Zahavi 1997) in the sense of costly extravagances that only defended prey can afford (Guilford & Dawkins 1993). Our data are consistent with this view because once aposematic forms established, then signal reliability was maintained more by the costs of dishonesty in undefended prey, than by the costs of honesty in all prey types (see also Viljugrein 1997; Lachmann et al. 2001). Here all our similar aposematic prey were protected by virtue of sharing a common trait – benefits were transferred among individuals of like phenotype, rather than specifically through relatives (see also Guilford 1985b; Mallet & Joron 1999). We appreciate that once aposematism arises, then it may generate behavioral responses in predators which make its evolution even more likely in other species (Turner 1984; Sherratt 2002; Speed & Ruxton 2002). However, our

experiments suggest that the most fundamental reason why defended prey tend to be conspicuous is because it sets them apart from undefended prey. For this reason, mathematical models of the evolution of aposematism which have concentrated exclusively on the evolution of traits in defended prey (e.g. Harvey et al. 1982; Sillén-Tullberg & Bryant 1983; Yachi & Higashi 1998; Speed 2000), may have missed something important. Indeed, if the reliability of a warning signal is an important determinant of its evolutionary success, and conspicuousness facilitates honesty, then the predisposition of predators to learn more rapidly from conspicuous signals may be more of a symptom, than a primary cause of aposematism.

## Chapter 2:

### The evolution of aggregation in profitable and unprofitable prey

#### Preface

Following on the research performed in Chapter 1, which included some simple experiments studying aggregation and profitability, it seemed appropriate to explore the interplay of aggregation, signalling and profitability more thoroughly. Previous work by Alatalo & Mappes (1996) had established the 'Novel World' methodology, exposing wild-caught birds to novel prey types and studying their ability to learn about palatability and aggregation. This methodology lent itself to reproduction on the computer screen, allowing for comparison between natural predators and human foragers, and also allowing for further experiments based on Alatalo and Mappes' results. This chapter was published in *Animal Behaviour* in 2005 (v. **70**: 199-208)

## Introduction

The conditions under which aposematism—the conspicuous colouration of unpalatable or otherwise defended prey—could evolve have long been a topic of speculation (Wallace 1867; Poulton 1890). A perceived roadblock to the initial establishment of rare, aposematic mutants is the intense predation to which they would be subjected by naïve predators. Aposematic prey, although defended, are much more likely to be seen by predators, and if predators are unaware of their defence (and do not exhibit neophobia) then such prey are more likely to be attacked on encounter (Gittleman & Harvey 1980). This “double-whammy” means that rare conspicuous mutants of defended prey should, on average, be attacked more frequently than their cryptic conspecifics.

A possible solution to this problem, first suggested by Fisher (1930, page 159), is that gregariousness could facilitate the evolution of distastefulness (and hence aposematism). Thus, if prey are warningly coloured and aggregated, then an attack on one individual by a naïve predator could lead to subsequent avoidance of others in the group—often relatives—that share the same trait (this proposal was the initial inspiration for Hamilton’s (1963) theory of kin selection (Hamilton 1996)).

A number of researchers have investigated the relationship between aggregation and aposematism. Some have found evidence that avoidance learning is faster when warningly coloured prey are aggregated (Gittleman & Harvey 1980; Gagliardo & Guilford 1993; Alatalo & Mappes 1996; Riipi et al. 2001), though some of these conclusions have been challenged (Tullberg et al.

2000). Others have argued that aposematic prey in aggregations present a more intense signal to their predators (Gamberale & Tullberg 1996; Gamberale & Tullberg 1998; Sherratt 2002). Furthermore, a dilution effect may improve the survivorship of grouped aposematic prey, especially if the level of detectability of aggregations does not increase linearly with group size or increasing signal conspicuousness (Bertram 1978; Sillén-Tullberg & Leimar 1988; Riipi et al. 2001).

One factor complicating experimental investigation of this phenomenon is the prevalence of aposematic organisms in nature. Current predators are not naïve in the evolutionary sense, and experiments with natural predators are therefore potentially confounded by pre-established responses to defended conspicuous prey. To address this difficulty, Alatalo & Mappes (1996) devised a series of “novel world” experiments, using wild-caught great tits (*Parus major*) as predators. The great tits foraged in an arena where the floor was covered in one of two symbols, either crosses or black squares. Distributed on this background were palatable and unpalatable prey items labelled with one of these two symbols. A prey item was “cryptic” if the symbol it carried was similar in appearance to the background symbol, while unpalatable prey items were classed as “aposematic” if they carried a different (effectively conspicuous) symbol. Through this method, prey items presented signals that bird predators could learn to associate with (un)palatability, but that were unlike signals that the predators would have experienced in nature.

In this study we replicated Alatalo & Mappes' "novel world" experimental conditions in a computer environment, using humans (*Homo sapiens*) as predators (see Tucker & Allen 1993, Glanville & Allen 1997 for parallel approaches using humans to address questions relating to crypsis). We have recently investigated questions about the initial evolution of aposematism with a similar system of computer experiments (Sherratt & Beatty 2003), and wished to ask further questions about aposematism and aggregation by expanding on Alatalo and Mappes' imaginative approach. To compare our results with humans to those found with great tits, we first repeated Alatalo and Mappes' "initial origin" experiments. If their results were truly robust, we might expect the same general outcome with another species, albeit one with a high potential for learning and strategizing.

We then used this experimental format to test several important questions about the relationship between aggregation, conspicuousness and profitability that have so far not been addressed. First, we wished to investigate the effects of profitable prey configuration on the adaptiveness of aggregation in unprofitable prey. We believe that this issue is important to consider because it is quite possible that the form of a signal that is most advantageous to unprofitable prey (e. g. aggregation and/or conspicuousness) is to some degree dependent on the traits exhibited by profitable prey (and vice-versa, Sherratt & Beatty 2003). Therefore, we designed experiments to ask: **(i)** will aggregation evolve in cryptic unprofitable prey if it reliably indicates unprofitability? **(ii)** will aggregation evolve in conspicuous unprofitable prey if it is a reliable, yet

superfluous, indicator of unprofitability? (iii) will aggregation evolve in cryptic unprofitable prey if cryptic profitable prey are also aggregated? Finally, to understand why unprofitability is associated with aggregation, it is clearly helpful to elucidate how readily aggregation evolves in profitable prey. We therefore designed our experiments to help evaluate: (iv) under what conditions will aggregation evolve in profitable prey?

Second, we note that while previous researchers addressed the possible advantages of aggregation in the evolution of aposematism, none of these tests presented aggregated and solitary configurations of unprofitable prey simultaneously (Riipi et al. 2001 used combinations of aggregated and solitary prey in their detectability experiments, but these prey were all palatable).

Perhaps a more appropriate approach would be to compare the survivorship of aggregated forms of unprofitable prey with that of solitary forms of unprofitable prey in the same system, as if a mutation for aggregation had recently arisen. We have therefore addressed this issue with our experiments, presenting solitary and aggregated forms of prey to predators simultaneously, thus allowing for comparison of the survivorship of prey in these two configurations, both between and within experiments.

## **Methods**

The human predators (70 student volunteers) used in these experiments had no knowledge of the experimental aims (the majority were non-biologists). No volunteer participated more than once, performing only a single replicate of one

experiment. Their foraging environment consisted of a 14cm × 14cm white arena on a computer screen. All prey items were 5mm × 5mm, and labelled with either a cross or a black square. Prey items (and background items where appropriate) that were labelled with a cross were enclosed in a square border to aid the user in determining the extent of the prey. Distributed within the arena were background items (either crosses or squares, whichever was the symbol on the “cryptic” form in a particular simulation). Also distributed in the arena were 16 “palatable” (profitable) prey, and 16 “unpalatable” (unprofitable) prey (see below). Prey were distributed on a randomized grid in the arena, with the constraint that no prey item or background item was allowed to overlap with another item.

Prey could be attacked by moving a cursor using a mouse, and clicking when the cursor was directly over the prey (an “attack”). On attacking a profitable (“palatable”) prey item the human predator added 1 point to its total displayed score, while on attacking an unprofitable (“unpalatable”) prey item the predator lost 1 point from its score. To indicate the profitability of an individual prey item (without requiring the predator to continually look at the score) attacks on profitable and unprofitable prey generated distinct sounds. Attacking a prey item “killed” that item; prey that were killed simply disappeared from the screen, while attacks on background items generated no response.

Each participant first took part in a single practice trial lasting 20 seconds, which was administered to familiarize participants with the test conditions and to minimize any effects of neophobia. Here, 16 prey items marked with crosses

and 16 prey items marked with squares along with 16 “cross” and 16 “square” background items were distributed in the arena. All prey were profitable and solitary.

After the practice trial, each volunteer took part in a version of one of 3 experiments, which consisted of three consecutive trials. Each trial had 16 background items (experiment 1) or 9 background items (experiments 2 and 3) (see below for justification), and all started with 16 profitable prey and 16 unprofitable prey distributed within the arena. Each trial had the same experimental conditions and lasted for 20 seconds. After 20 seconds had expired, a message box prompted the participant to proceed to the next trial by clicking on a button, until the third trial was complete. Participants were instructed to maximize their total points by trying to attack profitable prey while avoiding unprofitable prey. As the participants were volunteers they were all eager to take part in the study and we posted top scores on each experiment to enhance their motivation. Since a desire to maximize points, rather than hunger, was the motivation for our human predators, the three trials of each replicate were performed without an intervening time period.

Three experiments were conducted. The structure and rationale behind each of these experiments are summarized in Table 2-1. Each version of each experiment was replicated 7 times. For experiment 1 (the “initial origin” experiment), test conditions were identical to those performed by Alatalo & Mappes (1996). Thus, in experiment 1a and 1b, solitary prey were distributed in the arena, with 16 profitable cryptic prey (similar to the background), 8

unprofitable cryptic prey and 8 unprofitable conspicuous prey (different from the background). Crosses were the background item in experiment 1a, while squares were the background item in experiment 1b. Experiment 1c and 1d were similar to experiment 1a and 1b, except that prey were aggregated in 4 groups of 4, with cryptic profitable, cryptic unprofitable and conspicuous unprofitable items grouped together. Crosses were the background item in experiment 1c, while squares were the background item in experiment 1d.

In our further experiments (2a-d and 3a, b) 16 profitable and 16 unprofitable prey were presented to predators in each of three trials. In these experiments, however, one of the prey types was presented such that half of the prey were in a solitary configuration (8 individuals) and half were in an aggregated configuration (2 clusters of 4). In experiment 2a-d, unprofitable prey were presented in these configurations; in experiment 3a, b profitable prey were presented in this way. In experiment 2a-d, unprofitable prey conspicuousness (cross or square) and profitable prey configuration (solitary or aggregated) were systematically manipulated to study the interactions between these factors and unprofitable prey survivorship. In experiment 3a, b the survivorship of solitary and aggregated profitable prey was compared under different configurations (solitary or aggregated) of unprofitable prey.

**Table 2-1** A summary of experimental conditions

Experiment	Background	Profitable (n=8)*	Profitable (n=8)*	Unprofitable (n=8)*	Unprofitable (n=8)*	Question
1a						Do we get similar results using humans rather than birds?
1b						see above
1c						see above
1d						see above
2a						(i) Will aggregation evolve in cryptic unprofitable prey if it reliably indicates unprofitability?
2b						(ii) Will aggregation evolve in conspicuous unprofitable prey if it is a reliable (yet superfluous) indicator of unprofitability?
2c						(iii) Will cryptic unprofitable prey evolve to be solitary if cryptic profitable prey are aggregated?
2d						(iii) Will conspicuous unprofitable prey evolve to be solitary if cryptic profitable prey are aggregated?
3a						(iv) Will aggregation evolve in profitable prey when conspicuous unprofitable prey are solitary?
3b						(iv) Will aggregation evolve in profitable prey when conspicuous unprofitable prey are aggregated?

\*A total of 16 individuals of each prey type (profitable and unprofitable) were presented at the beginning of each trial. Each prey type has been presented here in two columns (8 prey in each column per trial) to more easily portray the combinations of prey in each experiment. A total of 16 background items were distributed in experiment 1; 9 background items were distributed in experiments 2 and 3. Roman numerals in the Question column refer to a similar question asked in the introduction.

Background items were distributed in a solitary configuration throughout all these experiments. For experiment 2a and 2b, all of the profitable and half of the unprofitable prey were distributed in a solitary, random configuration. For experiment 3a, all of the unprofitable and half of the profitable prey were distributed in the same way. This large number of solitary prey on the screen would sometimes cause the two aggregations of unprofitable prey to appear indistinct. As we wanted aggregations of prey to appear distinctly as aggregations the number of background items distributed on the screen was reduced from 16 (in experiment 1) to 9 (in experiments 2 and 3). This ensured that prey items in an aggregated configuration were easily identified as aggregations.

In preparation for statistical analyses, we tested whether the survivorship of solitary prey was correlated with the survivorship of aggregated prey within each trial of each experiment (for unprofitable prey in experiment 2 and profitable prey in experiment 3). No significant correlations were found, so data on survivorship for within-experiment comparisons were analyzed assuming independence.

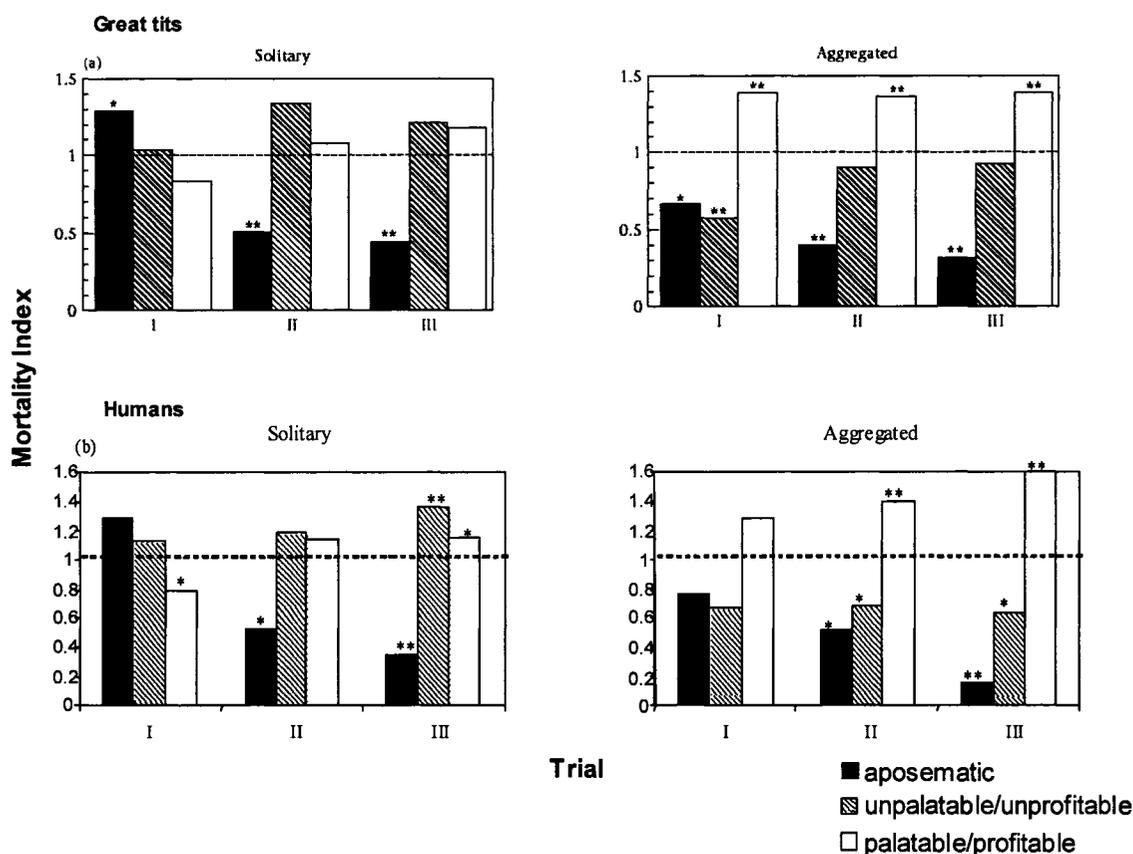
## **Results**

### *Experiment 1—great tits versus humans*

Alatalo & Mappes (1996) reported the relative mortality of each prey type by calculating the proportion of killed prey items in each group (aposematic, unpalatable cryptic and palatable cryptic) divided by the proportions initially

available (0.25 for aposematic and unpalatable cryptic items, and 0.50 for profitable cryptic items)—a “mortality index”. In their calculations, only the first 6 prey attacked in trial 1 and the first 12 prey attacked in each of trials 2 and 3 were included (see Alatalo & Mappes 1996, p. 709). For ease of comparison we have begun by presenting the results of the “initial origin” portion of our work (experiment 1a-d) in precisely the same manner, although we note that this “mortality index”, based on all prey rather than just unpalatable prey, has interpretive limitations (Tullberg et al. 2000; see Alatalo & Mappes 2000 for a re-analysis) and we do not utilize it in the analysis of our new experiments. Our results with human predators foraging on computer-generated prey were surprisingly similar to those found with great tits (Figure 2-1).

To examine how the survivorship of prey varied with experimental factors, we used a repeated-measures ANOVA on arcsine-square-root transformed proportions that survived of each prey type (here the denominator was simply the number of prey items of a given type initially present). Table 2-2 presents the within- and between-subject factors, main effects and significant interactions for all repeated-measures ANOVA analyses conducted. The type of background (cross or square) had no significant effect ( $P > 0.05$ ) in our experiment. The proportion of aposematic prey that survived increased significantly with



**Figure 2-1** (a) Relative mortality index of aposematic (black bar), unpalatable cryptic (striped bar) and palatable cryptic (white bar) prey items when presented solitarily and aggregated in the 'initial origin' experiment of Alatalo & Mappes (1996) (an attacked prey item is counted as 'killed'). Mortalities are standardized by dividing the observed mortality by the randomly expected mortality; horizontal lines show the point at which observed matches expected mortality. Reprinted from Tullberg et al. (2000). (b) Relative mortality of aposematic (black bar), unprofitable cryptic (striped bar) and profitable cryptic (white bar) prey items in the computer version of the 'initial origin' experiment. Mortalities are calculated as in (a). Significant deviations from a randomly expected mortality of 1, as judged from separate *t* tests on arcsine transformed proportions, are denoted by asterisks (\* $P < 0.05$ ; \*\* $P < 0.01$ ).

**Table 2-2** Results of repeated-measures ANOVAs on arcsine-transformed survivorship in experiments 1 through 3

Experiment	Prey type	Main effects*	Significant interactions
1	aposematic	trial: $F_{2,46} = 7.88, P = 0.001$ background: $F_{1,23} = 1.05, P = 0.32$ prey configuration: $F_{1,23} = 0.148, P = 0.70$	
	cryptic unprofitable	trial: $F_{2,46} = 0.10, P = 0.91$ background: $F_{1,23} = 0.06, P = 0.82$ prey configuration: $F_{1,23} = 0.05, P = 0.82$	
	cryptic profitable	trial: $F_{2,46} = 4.38, P = 0.02$ background: $F_{1,23} = 0.02, P = 0.90$ prey configuration: $F_{1,23} = 18.21, P < 0.001$	
2	solitary unprofitable	trial: $F_{2,46} = 3.12, P = 0.05$ profitable prey configuration: $F_{1,23} = 9.21, P = 0.006$ conspicuousness of unprofitable prey: $F_{1,23} = 65.78, P < 0.001$	profitconfig * unprofconspic: $F_{1,23} = 10.10, P = 0.004$ trial * unprofconspic: $F_{2,46} = 5.63, P = 0.006$ trial * profitconfig * unprofconspic: $F_{2,46} = 3.52, P = 0.04$
	aggregated unprofitable	trial: $F_{2,46} = 5.12, P = 0.01$ profitable prey configuration: $F_{1,23} = 3.11, P = 0.09$ conspicuousness of unprofitable prey: $F_{1,23} = 29.90, P < 0.001$	
	profitable	trial: $F_{2,46} = 3.82, P = 0.03$ profitable prey configuration: $F_{1,23} = 42.30, P < 0.001$ conspicuousness of unprofitable prey: $F_{1,23} = 7.90, P = 0.01$	profitconfig * unprofconspic: $F_{1,23} = 8.11, P = 0.009$
3	solitary profitable	trial: $F_{2,10} = 2.96, P = 0.07$ unprofitable prey configuration: $F_{1,5} = 0.29, P = 0.60$	
	aggregated profitable	trial: $F_{2,10} = 3.14, P = 0.06$ unprofitable prey configuration: $F_{1,5} = 0.01, P = 0.91$	
	unprofitable	trial: $F_{2,10} = 19.70, P < 0.001$ unprofitable prey configuration: $F_{1,5} = 1.13, P = 0.31$	
3 (10 seconds)	solitary profitable	trial: $F_{2,10} = 4.15, P = 0.03$ unprofitable prey configuration: $F_{1,5} = 0.001, P = 0.97$	
	aggregated profitable	trial: $F_{2,10} = 4.19, P = 0.03$ unprofitable prey configuration: $F_{1,5} = 0.10, P = 0.76$	
	unprofitable	trial: $F_{2,10} = 13.10, P < 0.001$ unprofitable prey configuration: $F_{1,5} = 0.63, P = 0.44$	

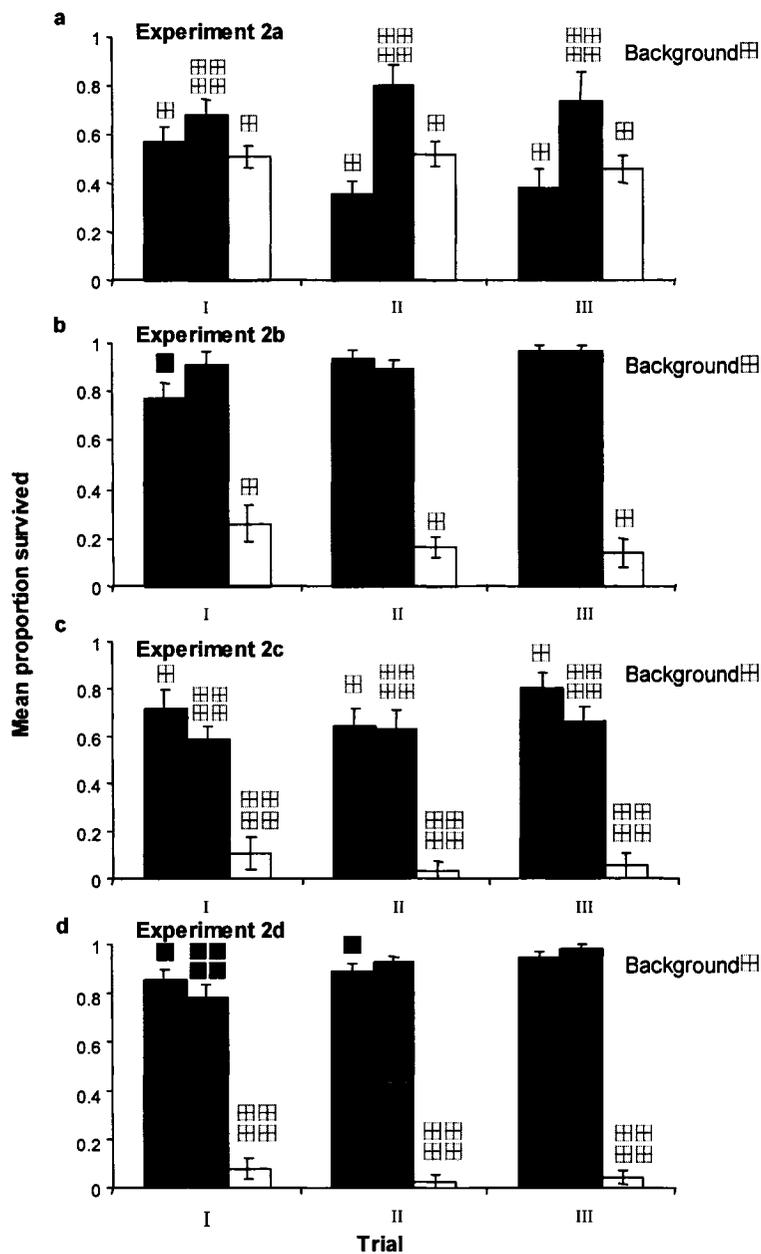
\*The within-subject factor in all experiments was trial. The between-subject factors in experiment 1 were background and prey configuration. The between-subject factors in experiment 2 were profitable prey configuration and unprofitable prey conspicuousness. The between-subject factor in experiment 3 was unprofitable prey configuration. Results for experiment 3 are presented for both the full trial and only the first 10 seconds of each trial.

increasing trial number, but the proportion of these prey that survived was not significantly affected by prey configuration. Neither trial number nor configuration had significant effect on the proportion of unprofitable cryptic prey that survived. The proportion of profitable prey that survived decreased significantly with increasing trial number and was significantly higher when prey were distributed in the solitary configuration. As will be seen, this important result concerning profitable prey is reflected in our later experiments.

*Experiment 2—solitary versus aggregated unprofitable prey*

Figure 2-2 shows the proportion of profitable and unprofitable forms of prey that survived per trial in experiment 2a-d. A repeated-measures ANOVA revealed that the proportion of aggregated unprofitable prey that survived was significantly higher when these prey were conspicuous (and therefore looked different from profitable prey) rather than cryptic (experiment 2b and 2d vs. experiment 2a and 2c) (Table 2-2). The proportion of these prey that survived also increased significantly with increasing trial number. There was no effect of the configuration of profitable prey on aggregated unprofitable prey survivorship, and no interactions were statistically significant.

The proportion of solitary unprofitable prey that survived was significantly higher when these prey were distinct in appearance (experiment 2c and 2d) rather than cryptic (experiment 2a and 2b), and increased significantly with increasing trial number (Table 2-2). In this case the proportion of solitary unprofitable prey that survived was also higher when



**Figure 2-2** Mean  $\pm$  SE proportions of prey that survived by trial in (a) experiment 2a, (b) experiment 2b, (c) experiment 2c and (d) experiment 2d. Dark bars indicate unprofitable prey; open bars indicate profitable prey. Prey conspicuousness (cross or square) relative to the background ( $\boxtimes$ ) and prey configuration (solitary or aggregated) are represented by symbols above or within each bar.

profitable prey were aggregated rather than solitary, and there was a significant interaction effect between the aggregation of profitable prey and the conspicuousness of unprofitable solitary prey. This interaction arose at least in part from the contrast between profitable and unprofitable prey: survivorship was higher for cryptic solitary unprofitable prey when profitable prey were aggregated (experiment 2c vs. experiment 2a). In sum, unprofitable prey in both configurations experienced increasing survivorship as trials advanced, and had higher survivorship when they looked different from profitable prey.

For profitable prey in this experiment (which were always cryptic), a repeated-measures ANOVA (Table 2-2) revealed that the proportion that survived decreased significantly with trial number. Profitable prey had significantly lower survivorship when unprofitable prey were conspicuous (implying that profitable prey were easily distinguished from unprofitable prey); profitable prey also had significantly lower survivorship when they were aggregated rather than solitary. There was also a significant interaction between the aggregation of profitable prey and the conspicuousness of unprofitable prey on the proportion of profitable prey that survived. In effect, profitable prey experienced lower survivorship when unprofitable prey were conspicuous, and much lower survivorship when profitable prey were aggregated, to the point of depletion of aggregated profitable prey (experiment 2c and 2d).

In experiment 2a, an independent-sample *t*-test on arcsine-square-root transformed proportions revealed that aggregated unprofitable prey had significantly higher survivorship than solitary unprofitable prey within trials 2 and

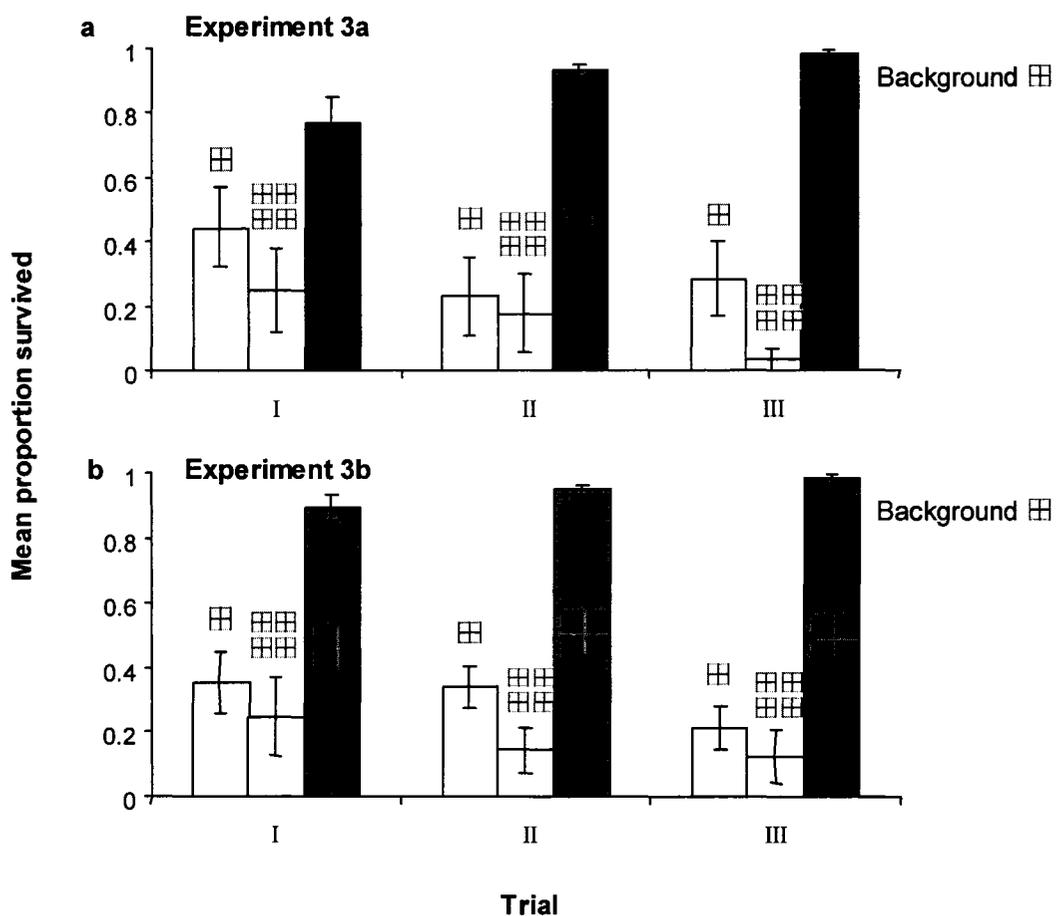
3 ( $t_s = 4.531$ ,  $df = 12$ ,  $P = 0.001$ , and,  $t_s = 2.489$ ,  $df = 12$ ,  $P = 0.03$

respectively). In this case all aggregated prey were reliably unprofitable. In experiment 2b-2d, no significant differences between proportions of solitary unprofitable and aggregated unprofitable prey that survived were found in any of the three trials.

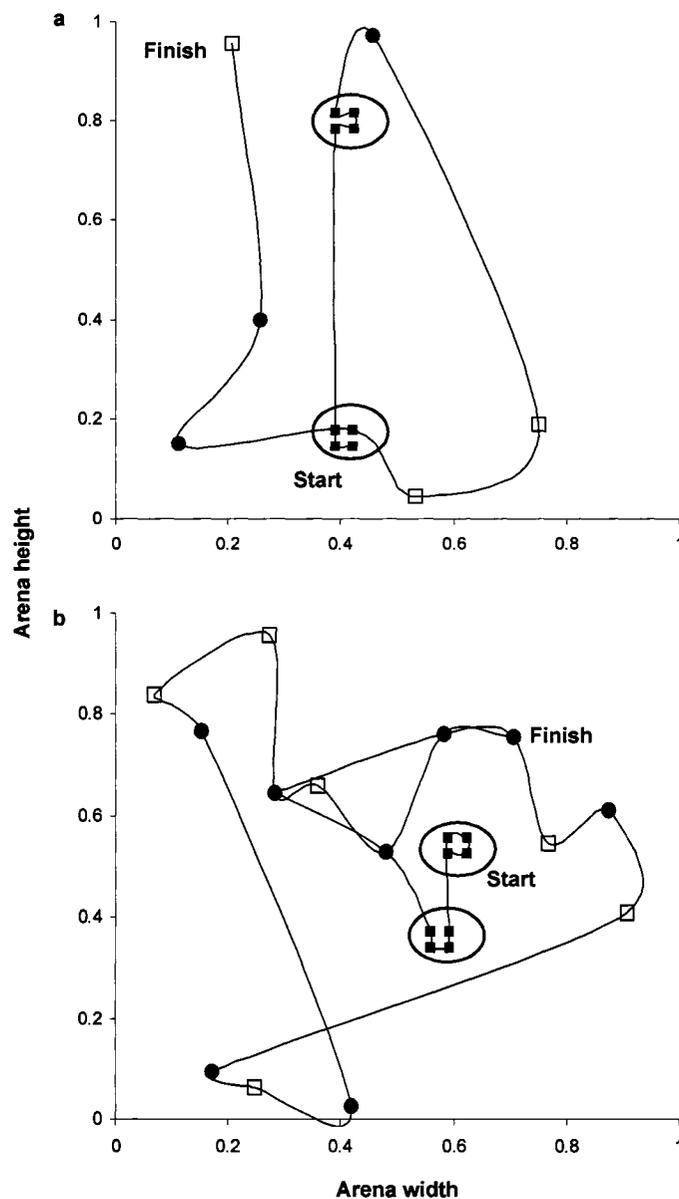
### *Experiment 3—solitary versus aggregated profitable prey*

Figure 2-3 shows the proportion of solitary and aggregated profitable prey that survived in experiment 3, when unprofitable prey were either conspicuous and solitary (experiment 3a) or conspicuous and aggregated (experiment 3b). To examine survivorship of each type of profitable prey (aggregated or solitary), we once again used a repeated-measures ANOVA for arcsine-square-root transformed proportions of prey that survived (Table 2-2). For solitary and aggregated profitable prey, survivorship did not change significantly with increasing trial number and was not affected by the configuration of unprofitable prey. No interactions were significant.

In experiment 3, 20 seconds was often adequate for a predator to deplete all aggregated profitable prey and then move on to solitary profitable prey, such that overall survivorship for solitary and aggregated profitable prey was approximately the same by the end of the trial. To demonstrate the depletion in this experiment, we show the path and attack order of predators (randomly selected from our data set) in trial 3 of experiment 3a and 3b (Figure 2-4) (all



**Figure 2-3** Mean  $\pm$  SE proportions of prey that survived by trial in (a) experiment 3a and (b) experiment 3b. In experiment 3a, unprofitable prey were conspicuous and solitary; in experiment 3b, unprofitable prey were conspicuous and aggregated. Dark bars indicate unprofitable prey; open bars indicate profitable prey. Prey conspicuousness (cross or square) relative to the background ( $\boxplus$ ) and prey configuration (solitary or aggregated) are represented by symbols above or on each bar.



**Figure 2-4** Plot of the path taken by predators in attacking prey in (a) experiment 3a, trial III and (b) experiment 3b, trial III (arena dimensions were 14×14 cm; dimensions are scaled to 1). These replicates were chosen at random from the seven replicates of each experiment (all of the replicates were similar). ■: Aggregated profitable prey that were attacked; □: solitary profitable prey that were attacked; ●: background items that were attacked. Aggregations of profitable prey are indicated with ovals; the start and finish points of the predator path are labelled.

replicates showed the same trend). Note that predators did not rapidly jump to different spatial locations over the screen, but (once aggregated profitable prey were removed) they moved around “zapping” prey in a relatively smooth manner.

Depletion was not an issue in experiments 1 and 2, because profitable prey were not presented in different configurations within an experiment, and so switching from one profitable prey type to another could not occur. To address the issue of depletion of aggregated profitable prey in experiment 3, the survivorship of each prey type was analyzed at 10 seconds into each trial. A repeated-measures ANOVA for arcsine-square-root transformed proportions of each prey item that survived found that the survivorship of aggregated profitable prey decreased significantly with increasing trial number, but was not affected by the configuration of unprofitable prey (always conspicuous in this experiment). Solitary profitable prey survivorship also decreased significantly with increasing trial number and was not affected by the configuration of unprofitable prey. Independent-sample *t*-tests revealed that aggregated profitable prey had significantly lower survivorship than solitary profitable prey within trial 3 of experiment 3a ( $t_s = 3.73$ ,  $df = 12$ ,  $P = 0.003$ ); a similar trend with borderline significance was found within trial 3 of experiment 3b ( $t_s = 2.08$ ,  $df = 12$ ,  $P = 0.06$ ).

## Discussion

Our application of the “novel world” methodology using humans has revealed new insights about the advantages of aggregation to profitable and unprofitable prey. When we replicated the test conditions of Alatalo & Mappes (1996), humans demonstrated a striking similarity to previous data generated with bird predators. Two key reasons for this similarity are that humans, like birds, are capable of responding to features that reliably distinguish profitable from unprofitable prey, and that both react in qualitatively similar ways to clusters of prey. In particular, while both great tits and humans would leave an aggregation of unprofitable prey after attacking the first item in the aggregation, they would systematically deplete aggregations of profitable prey. Tullberg et al. (2000) proposed exactly the same mechanism as an explanation for the results of Alatalo & Mappes (1996).

Our new experiments do not unequivocally support the contention that aggregation is a necessary step in the evolution of aposematism. Aggregated cryptic unprofitable prey did eventually experience a significantly higher survivorship than solitary cryptic unprofitable prey, but only when this aggregation allowed them to be distinguished from profitable prey (experiment 2a). Our results suggest that if prey palatability can be assessed on the basis of aggregation alone, then there may be less of a selective advantage in evolving a separate aposematic signal. Indeed, to take the converse, aggregated and solitary unprofitable prey did not have any marked differences in survivorship when unprofitable prey could be distinguished on the basis of their

conspicuousness (experiment 2b). The trend, albeit non-significant, for unprofitable cryptic prey to survive better when they were solitary when profitable prey were aggregated (experiment 2c), and the significant effect of configuration of profitable prey on attack rate of solitary unprofitable prey (ANOVA results, see Table 2-2), support the contention that unprofitable prey may evolve characteristics that help predators distinguish them from profitable prey (Fisher 1930; Srygley & Chai 1990; Sherratt 2002; Sherratt & Beatty 2003; Sherratt et al. 2004; Srygley 2004).

Some researchers have suggested a potential benefit of gregariousness to aposematic prey (Gagliardo & Guilford 1993; Gamberale & Tullberg 1996; Gamberale & Tullberg 1998; Riipi et al. 2001; Sherratt 2002) in that aposematic prey in aggregations might present a heightened signal, and thereby be more repellent to potential predators. After directly comparing the survivorship of solitary and aggregated aposematic prey (experiment 2b and 2d) within all three trials, we could find no evidence in our particular study that humans were more effectively deterred by aggregated compared to solitary conspicuous signals. In our case, when the profitability of prey was reliably signaled through prey conspicuousness (profitable prey were never conspicuous), then both aggregated and solitary forms were left relatively untouched from an early stage. One reason for this discrepancy may be that our human predators learned quickly to avoid reliably unprofitable prey types, whatever they looked like. It is also possible that our initial instruction trial reduced any potential heightened

response to conspicuous signals. Finally, we note that our conspicuous prey were black squares, and did not exhibit a more natural warning signal.

While most researchers have focused on what traits are advantageous to unprofitable prey, here we have also considered what traits are advantageous to profitable prey. When solitary and aggregated profitable prey were presented to predators simultaneously, aggregated profitable prey were attacked more than solitary profitable prey, regardless of whether unprofitable prey were solitary or aggregated (experiment 3, 10-second trials). Further support for this trend is seen by comparing the fate of aggregated and solitary profitable prey within experiment 2. In this experiment, aggregated profitable prey were attacked at a much higher rate than solitary prey, with aggregated profitable prey consistently being depleted (Figure 2-2). The process by which this occurs is well demonstrated in Figure 2-4, which shows the foraging paths taken by our human predators in experiment 3. Predators rapidly moved to aggregated profitable prey first in these trials, and consequently consumed all members of an aggregation before moving on. In experiment 3b unprofitable prey were also aggregated, but in contrast to profitable prey, predators attacking an unprofitable prey item in an aggregation would quickly leave the aggregation.

With respect to our initial questions (i-iv) the overwhelming pattern that emerges from all of these experiments is a highly intuitive one. Unprofitable prey had higher survivorship when they had traits that helped distinguish them from profitable prey, a case that is supported by both between- and within-experiment comparisons of our results. When unprofitable prey could be distinguished from

profitable prey on the basis of their appearance (e. g. conspicuousness) then there was no significant advantage to them adopting one configuration over another. Leaving aside the possibility of perfect mimicry of unprofitable aggregated prey, profitable prey tend to survive better on a per capita basis when solitary rather than in aggregated form for the simple reason that aggregations of profitable prey represent rich sources of prey that are preferentially exploited. In nature, without the satiating effect of distastefulness, small, edible prey are at a disadvantage in that their aggregations would need to be infinitely large for individual survivorship to be enhanced by predator satiation (Sillén-Tullberg & Leimar 1988). Thus aggregation is severely limited as a means of predation avoidance for small profitable prey.

Phylogenetic studies (Sillén-Tullberg 1988; Tullberg & Hunter 1996; Nilsson & Forsman 2003) have revealed that the majority of repellent defenses and warning colours have evolved in species that continue to be solitary. For instance, in their sample of 578 species of moth, Nilsson & Forsman (2003) found that only 34% of species with conspicuous larvae exhibited an aggregated lifestyle, while less than 2% of species with non-conspicuous larvae were aggregated. Tullberg & Hunter (1996) examined a sample of 800 tree-living macrolepidopteran larvae and argued that their data were consistent with the proposal that the evolution of repellent defenses and warning colouration precede the evolution of gregariousness. Our experiments do not reveal anything directly about the order in which traits such as conspicuousness and gregariousness evolved, but they do show that non-mimetic profitable prey tend

to be vulnerable when in aggregations, while unprofitable prey can sometimes, but not always, gain from being clustered.

Of course, there are more general benefits of gregariousness, regardless of prey defence, such as thermal enhancement, increased developmental rate and protection from desiccation (Clark & Faeth 1997; Clark & Faeth 1998; Klok & Chown 1999; see Krause & Ruxton 2002 for a review). As there are benefits to gregariousness not related to prey defense, could it be that unprofitable prey tend to be aggregated because they can avail themselves of these benefits, while profitable prey simply cannot (Gagliardo & Guilford 1993)? This could explain why some, but not all, aposematic species are aggregated, and why so few profitable species are found in aggregations (Nilsson & Forsman 2003). In other words, the observed association between gregariousness and aposematism might not be a function of prey defence in and of itself, but a case where selection against aggregation has been relaxed for unprofitable prey.

### Chapter 3:

#### The evolution of Müllerian mimicry in multispecies communities

#### Preface

In this chapter I shifted research focus from warning signals to the related phenomenon of Müllerian mimicry. This proved to be a rewarding research effort, and is the component of my thesis that I think holds the greatest potential for future research. The suggestion of a slightly different form of predator learning than that originally invoked by Müller has implications for the nature of selection for mimicry as well as the types of evolutionary responses that could be expected in Müllerian mimics. This chapter was published as a letter in *Nature* in 2004 (v. **431 (7004)**: 63-66).

## **Introduction**

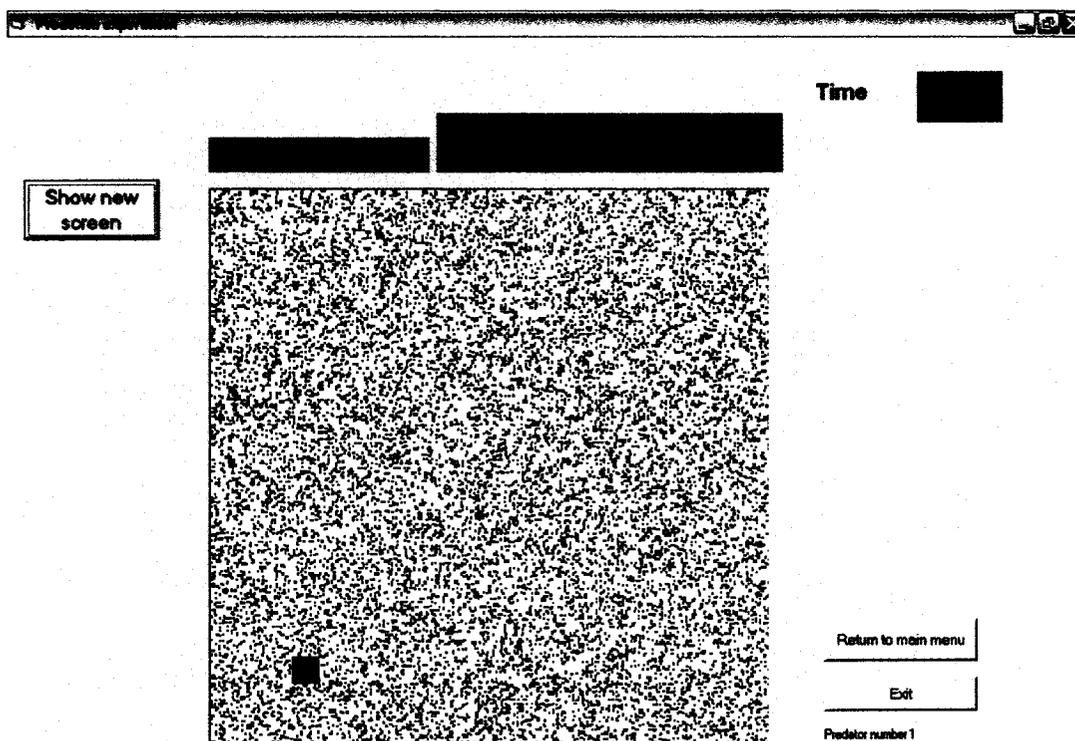
If a predator community needs to attack a fixed number of each distinct form of defended prey (such as those with stings or toxins) before it learns to avoid them, and if this pressure is significant, then there will be selection on unprofitable prey to resemble one another (Müller 1878, Müller 1879, Poulton 1890). Although field experiments have lent support to the idea that common forms of unpalatable prey are at a selective advantage over rare forms of unpalatable prey (Benson 1972, Mallet & Barton 1989, Kapan 2001), the precise mechanisms involved have seldom been investigated. Experiments with avian predators feeding on two distinct forms (occurring at frequencies of 1:9 and 9:1) (Greenwood et al. 1989), and just one form (4%, 12% and 32% of total prey) (Linström et al. 2001) of artificial distasteful prey have found that unpalatable phenotypes have a greater per capita probability of attack when they are rare. However, similar experiments using garden birds (Greenwood et al. 1981), domestic chicks (Greenwood et al. 1981) and captive great tits (Rowe et al. 2004) as predators found little or no evidence to indicate that the common form is at a selective advantage over the rarer form. These findings contradict the traditional assumption that there should be selection towards uniformity because one colour pattern is easier to learn than two (Rowe et al. 2004).

## **Methods**

All human 'predators' were visitors to the Page Break Coffee Bar situated within the MacOdrum Library at Carleton University. A total of 155 human predators (all

volunteers) participated, of which 92% were non-biologists. Predators were not informed about the experimental aims and no subject was allowed to participate more than once.

The computer program was written in MS Visual Basic 6. The artificial foraging environment of each predator consisted of a grid of  $n \times n$  cells. Prey (7×7 mm square and coloured/striped/spotted in a particular way according to type) were distributed within cells of this virtual grid, with no more than one prey item per cell. The predator saw only one randomly selected cell of the grid at a time (the position of the cell in the grid was not displayed) viewed in a square arena (148×148 mm) on the computer screen (see Figure 3-1). The background of the arena was comprised of a mosaic of 10% green and 90% white pixels. Predators changed cells in the search for prey by pressing a command button, causing a new cell to appear in the arena. When predators moved to a new cell that contained a prey item, they could either attack it (by clicking with the mouse cursor on it) or choose to move on without attacking it. On attacking a profitable prey item the predator gained a point and a high-pitched sound was made (which indicated profitability more effectively than using a score alone). On attacking an unprofitable prey item, the predator lost a point and a low-pitched sound was made. Prey items that were attacked disappeared from the system. Predators were asked to forage for prey for five minutes in a way that would maximize their personal scores.



**Figure 3-1** A screen shot of the computer interface used by our human “predators” to forage for prey. One cell of the virtual grid was visible on the screen at any time. At most, one prey item could occur in each cell (in a random position and randomly-selected vertical or horizontal orientation). A striped magenta prey is shown above (lower left corner). Predators attacked prey by clicking on the object using a computer mouse; if a predator chose not to attack a prey item they could move to another cell by clicking the “Show new screen” command button. The time elapsed, number of screens visited and the running score were displayed on the interface.

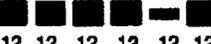
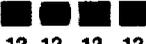
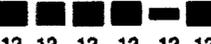
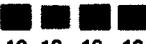
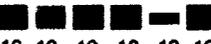
### *Experiment 1*

Profitable prey (mottled green: 70% green, 30% white pixels, 40 items in total) and two forms of unprofitable prey (mottled red and mottled blue: 70% red and blue respectively, 30% white pixels, with a combined initial frequency of 40) were randomly distributed in a 10×10 grid. Overall, nine relative frequencies of red and blue unprofitable forms were presented (0/40, 2/38, 4/36, 8/32, 20/20, 32/8, 36/4, 38/2 and 40/0) and we allowed five different naive human predators to forage at each relative frequency (45 different human subjects in total).

### *Experiment 2–5*

A total of 156 prey items were distributed in a 13×13 grid. Fifteen forms of prey were employed in total, consisting of blue, grey, green, red, yellow and cyan squares with and without a black stripe, and a magenta unprofitable prey type that was either plain, striped or spotted (see Figure 3-2). The attack rate on the focal magenta unprofitable prey species was monitored in each of these experiments, which were all replicated ten times using different human predators (110 naive subjects in total).

As our primary interest was in comparing the survivorship of focal prey under a given set of conditions, we used one-way analyses of variance (ANOVA) to test whether their per capita attack rates (arcsin transformed) differed

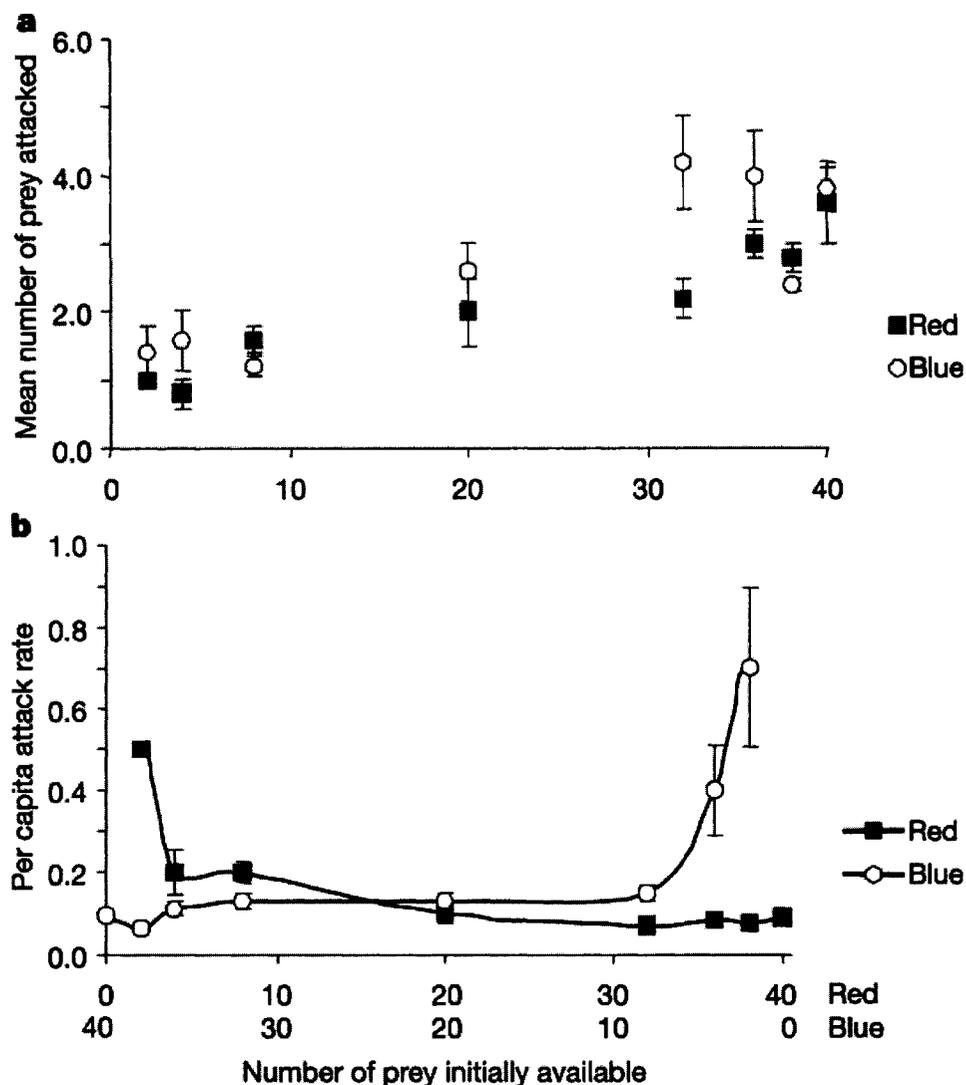
Experiment	Profitable prey available	Non-focal unprofitable prey available	Focal unprofitable prey form(s) available
2	72	72	a  [0.15] 12 ----- b  [0.09] 12 ----- c  [0.14] 12
3	 12 12 12 12 12 12	 12 12 12 12 12 12	a  [0.25] 12 ----- b  [0.05] 12 ----- c  [0.13] 12
4	 12 12 12 12 12 12	 12 12 12 12 12 12	a  [0.28] 12 ----- b  [0.28] 12 ----- c  [0.18] 12
5	 12 12 12 12 12 12	 12 12 12 12 12 12	a  [0.54] 8  [0.00] 4 ----- b  [0.08] 8  [0.58] 4

**Figure 3-2** A summary of experiments 2–5, in which human subjects foraged in a community of profitable and unprofitable prey. The prey forms had colours, with or without a single stripe or spot, as depicted. The numbers underneath each prey type refer to the number of that form available initially. Our analysis centred on the fate of a focal unprofitable prey species that was magenta in colour and either plain, striped or spotted. In experiments 2–4, the focal unprofitable species was monomorphic; in experiment 5 it was dimorphic. The mean proportions attacked of each of the focal types are given in square brackets.

according to their appearance (plain, striped, spotted). To compare the per capita attack rates (arcsin transformed) of the focal prey between experiments, we used a two-way ANOVA with experiment and focal form as fixed factors.

## Results

To evaluate the strength of selection for mimicry in simple systems containing relatively few prey phenotypes, we first conducted several related experiments (experiments 1 and 2a–c). In experiment 1, human ‘predators’ were allowed to search a virtual environment in which they encountered individual prey items selected at random from populations of a profitable form of prey (mottled green, 40 items in total), and two unprofitable forms (mottled red and mottled blue, at nine combinations of frequencies that each totalled 40). As with previous studies (Greenwood et al. 1989, Lindström et al. 2001), predators attacked more of the common form of the unprofitable prey available (Figure 3-3a; number of blue attacked versus frequency of blue available,  $r_{38}=0.42$ ,  $P=0.007$ ; number of red attacked versus frequency of red available  $r_{38}=0.47$ ,  $P=0.002$ ). This relationship has been explained as a consequence of predators not seeing enough of the rarest forms to complete their learning (Greenwood et al. 1989, Mallet 2001). However, in our study we observed how our predators behaved at each and every encounter, and noted that even rare unprofitable prey tended to be rejected by the end of the experiment (Table 3-1). This indicates that avoidance



**Figure 3-3** Results from Experiment 1. a) The mean number of red and blue unprofitable prey attacked by predators ( $\pm 1$  SE) in relation to the number of these prey items initially available; b) the implications of this behaviour for their mean per capita 'mortality' ( $\pm 1$  SE). The proportion of red items attacked was significantly higher than that of blue items only when there were two red and 38 blue items (pairwise  $t$ -test on arcsin-transformed proportions  $t_4=40.77$ ,  $P=0.001$ ). The proportion of blue items attacked was significantly higher than that of red when there were two, four or eight blue items ( $t_4=4.58$ ,  $P=0.010$ ,  $t_4=4.15$ ,  $P=0.014$ ,  $t_4=7.915$ ,  $P=0.001$ , respectively).

**Table 3-1** The prey frequencies presented, mean number of prey encountered, mean number of prey attacked and proportion attacked of the first and last unprofitable prey encountered by predators in Experiment 1 (NA – not applicable). The maximum ratio of mean attacks by predators to mean encounters was 0.5. While the majority of predators attacked the first unprofitable prey encountered of both prey types, the last prey encountered was seldom attacked. Overall, this indicates that predator avoidance learning was largely complete, even when the unprofitable prey frequencies were relatively low.

Frequency		Mean total encounters		Mean total attacks		Proportion attacked: first prey encountered		Proportion attacked: last prey encountered	
<i>red</i>	<i>blue</i>	<i>red</i>	<i>blue</i>	<i>red</i>	<i>blue</i>	<i>red</i>	<i>blue</i>	<i>red</i>	<i>blue</i>
0	40	NA	85.2	NA	3.8	NA	1	NA	0
2	38	2.6	78	1	2.4	0.8	0.8	0.2	0
4	36	7.6	83.6	0.8	4	0.8	1	0	0
8	32	17.2	80.8	1.6	4.2	0.8	1	0.2	0
20	20	47.2	50	2	2.6	0.8	1	0	0
32	8	78	19.2	2.2	1.2	0.8	1	0	0
36	4	88.4	6.6	3	1.6	0.8	1	0	0
38	2	84.2	2.8	2.8	1.4	1	0.8	0	0.4
40	0	102	NA	3.6	NA	1	NA	0	NA

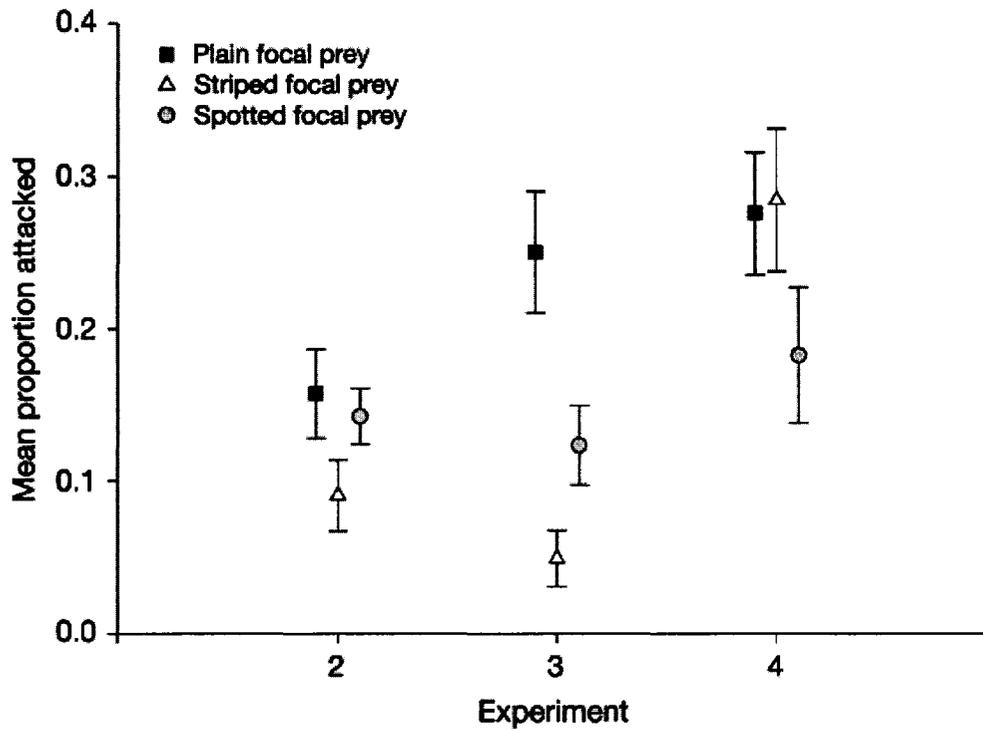
learning was largely complete, even when the unprofitable prey frequencies were relatively low. A more consistent explanation for the phenomenon may be that predators occasionally return to attacking the more common unprofitable type to assure themselves that all of that prey type were unprofitable, because there would be more to lose if some turned out to be profitable.

Despite the increase in number of unprofitable prey attacked with the frequency presented, the per capita 'mortality' of the unprofitable forms declined as their frequency increased (Figure 3-3b). Overall, this observed predatory behaviour would be capable of generating selection for Müllerian mimicry, with particularly rare forms of unprofitable prey selected to resemble common forms of unprofitable prey. Here we have elucidated selection at many more relative frequencies than have so far been examined in a single study, and can confirm, as Müller had anticipated (Müller 1878, Müller 1879) that the difference in survivorship between the rare and common forms was smaller when unprofitable prey types were similar in frequency. Unsurprisingly, significant selection for Müllerian mimicry was not always evident at these intermediate frequency combinations (legend to Figure 3-3b), and whether or not this was a consequence of low statistical power, it remains clear that selection under these intermediate conditions was at best weak.

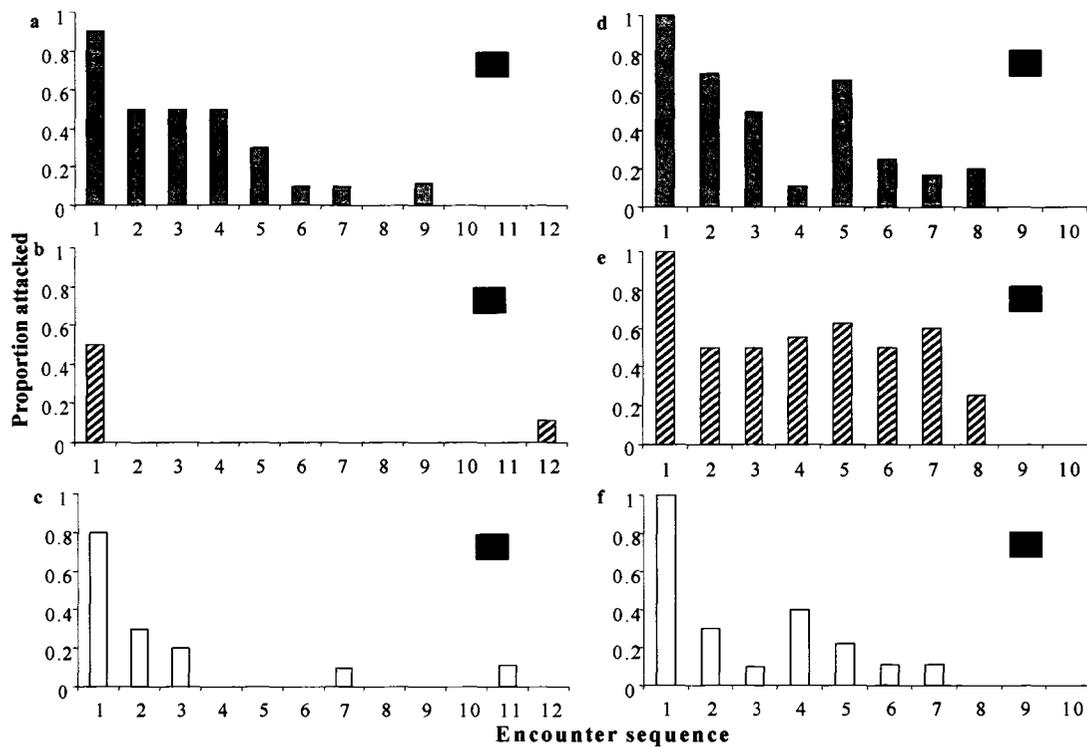
In experiments 2a–c (see Figure 3-2 for outline) we again investigated the strength of selection for Müllerian mimicry in a simple prey community (one

profitable and two unprofitable forms), this time evaluating the relative success of imperfect Müllerian mimics. To do this we compared the survivorship of a distinctly coloured rare unprofitable prey (the 'focal' form) that was: plain, like the profitable form (experiment 2a); striped, like the more common unprofitable form (experiment 2b); or spotted so that it looked like neither the profitable nor the common unprofitable form (experiment 2c). Here, the mean per capita attack rates on each of these three forms in the three separate treatments did not differ significantly (Figure 3-4, ANOVA  $F_{2,27}=2.966$ ,  $P=0.068$ ). Thus, even if selection occurred, it would not strongly favour the rare striped Müllerian mimics in this simple community.

In our next experiments (experiments 3 and 4), we evaluated the strength of selection for Müllerian mimicry in a more complex system with multiple forms of profitable and unprofitable prey present (Figure 3-2). All of our computer-generated prey 'species' were distinct in appearance and were equally common (12 individuals of each). In experiments 3a–c, all of the non-focal unprofitable prey (six species) shared a common pattern element (a stripe) that was not exhibited by any of the profitable prey (six species). By contrast, in experiments 4a–c, a stripe was exhibited by three of the profitable species and three of the non-focal unprofitable species. When a stripe was reliably associated with unprofitability (experiments 3a–c), the focal prey species that carried the stripe had a far greater probability of surviving than the other forms of this species that did not (Figure 3-4, ANOVA  $F_{2,27}=14.52$ ,  $P=0.001$ ; Tukey post hoc comparisons:



**Figure 3-4** The mean proportions ( $\pm 1$  SE) of each of the three forms of focal unprofitable prey attacked in experiments 2–4. An overall two-way analysis of variance on arcsin-transformed proportion of focal prey attacked revealed a highly significant interactive effect of experiment (2, 3 and 4) and focal form (plain, striped, spotted) ( $F_{4,81}=5.12$ ,  $P=0.001$ ) and highly significant main effects (experiment,  $F_{2,81}=11.91$ ,  $P=0.001$ ; focal form  $F_{2,81}=8.77$ ,  $P=0.001$ ).

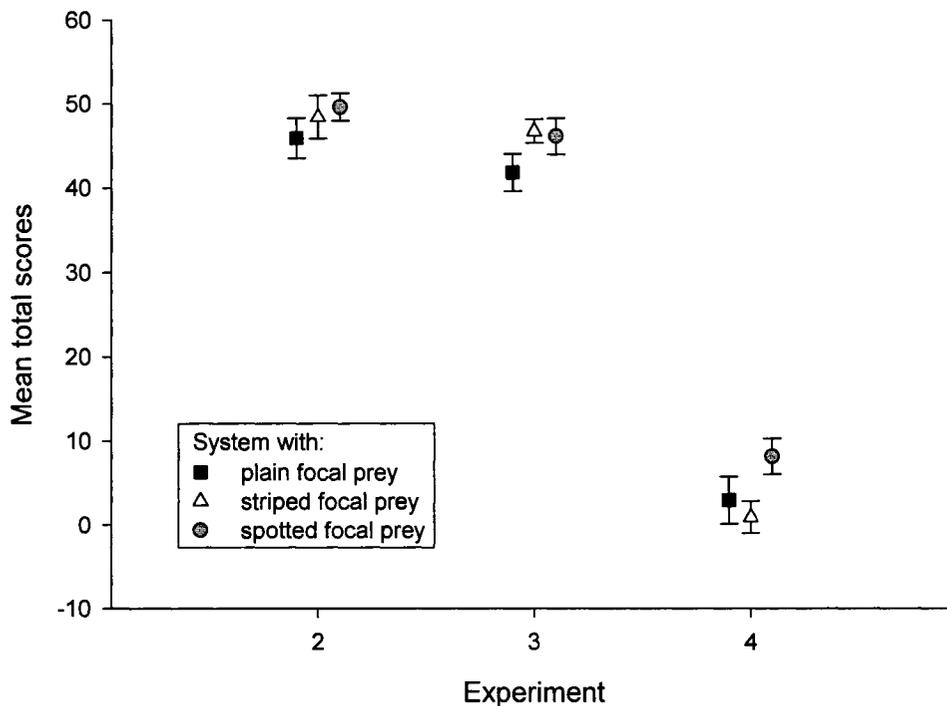


**Figure 3-5** The mean probabilities of predators attacking each of the focal forms of unprofitable prey on encounter, at each stage in the encounter sequence (first encountered prey of this type, second encountered prey etc). (a)-(f) represent the results of experiments 3a, 3b and 3c, and 4a, 4b and 4c, respectively. All predators encountered at least 8 individuals of the unprofitable focal form in experiment 3. Due to the absence of simple cues to profitability, predators foraged more slowly in experiment 4. However all predators encountered at least 3 of the unprofitable focal form, and 80% of predators encountered the focal prey 5 times or more.

striped versus plain,  $P=0.001$ , striped versus spotted,  $P=0.018$ ). In these cases the focal striped prey were not only attacked far less frequently on first encounter, but they were also subsequently avoided more quickly than other forms (see Figure 3-5).

When no trait was reliably associated with unprofitability in these complex systems (experiments 4a–c), there was no significant difference in the proportions of these three different forms attacked in the separate experiments (Figure 4; ANOVA  $F_{2,27}=1.84$ ,  $P=0.179$ ). In this case, predators continued to attack each form of the focal unprofitable prey for some time, generating low mean scores for these foragers (see Figure 3-6) and overall high prey mortality. The observation that striped prey had a greater survivorship in experiment 3b compared to 4b (ANOVA  $F_{1,18}=26.17$ ,  $P=0.001$ ) strongly suggests that predators were not simply learning to avoid all distinct unprofitable prey at equal rates. Instead, predators avoided attacking prey phenotypes at a rate related to the degree to which their traits could be categorized as unprofitable.

It has been argued that a predator's capacity to remember simultaneously the profitabilities of a wide range of prey types may be limited, and this may have an important influence on the nature of selection experienced by potential prey (MacDougall & Dawkins 1998, Ruxton 1998). To test this theory directly, we conducted a two-way ANOVA to compare the proportion of each focal prey type attacked in the simple (experiment 2) and the analogous complex (experiment 3) system. This analysis confirmed that system complexity had a significant



**Figure 3-6** The mean total scores ( $\pm 1$  SE) that predators achieved in each of the three systems (prey communities with plain, striped or spotted focal prey) for experiments 2-4. The scores of our human foragers varied significantly between experiments 2-4, but did not vary significantly with the form of focal unprofitable prey presented (2-way ANOVA: experiment  $F_{2,81} = 382.13$ ,  $P < 0.001$ ; focal prey type  $F_{2,81} = 3.10$ ,  $P > 0.05$ ; interaction  $F_{4,81} = 1.04$ ,  $P > 0.05$ ). The low scores in experiment 4 reflect the difficulty that subjects faced when foraging in an environment with no simple discriminative cues to guide them. The mean scores ( $\pm 1$  SE) in experiments 5a and 5b were  $41.02 \pm 2.22$  and  $40.10 \pm 3.45$  respectively.

influence on the proportion of each focal form attacked (interaction,  $F_{2,54}=3.48$ ,  $P=0.038$ ; experiment,  $F_{1,54}=0.061$ ,  $P=0.807$ ; focal form,  $F_{2,54}=15.69$ ,  $P=0.001$ ), which seems to be driven by a combination of the increased vulnerability of the plain unprofitable form, and the decreased vulnerability of the striped (mimetic) form when the system was complex. Furthermore, given that the mimetic form was attacked at a significantly lower rate than both the plain and spotted forms when the system was complex but not when it was simple, we conclude that selection for mimicry can indeed be more intense when there are multiple species, at least under the specific conditions we have compared.

Taken together, our results suggest that there may be selection for unprofitable species in complex systems to maintain and enhance certain features (such as a colour or stripe in the right place) which happen to be shared by more unprofitable species than profitable species in a given area. To test whether an imperfect mimic could spread from rarity in a small population that contained a more common conspecific which lacked the mimetic trait, we conducted experiments 5a and 5b (Figure 3-2). Despite its rarity, and the fact that it was seen at least twice by predators in all replicates, a rare form that shared a trait (stripe) with the other unprofitable species in the community had a far higher per capita survivorship than more common conspecifics that did not (experiment 5a, paired  $t$ -test on arcsin-transformed proportions  $t_9=9.73$ ,  $P=0.001$ ). Similarly, when the striped focal form was more common than its plain conspecifics, its per capita survivorship was significantly higher (experiment 5b,  $t_9=10.75$ ,  $P=0.001$ ). These results indicate the probable fate of intermediate

phenotypes in the evolution of Müllerian mimicry. If a distinctive mutant form of an unprofitable species arises, we might expect that it would suffer high mortality due to its unique appearance and extreme rarity. However, our results show that if the mutant shares simple signalling traits with more common unprofitable species, then it may well survive at higher rates than conspecifics lacking these traits.

## **Discussion**

Müllerian mimicry is a taxonomically widespread phenomenon (Bates 1862, Müller 1878, Müller 1879, Poulton 1890, Sbordoni et al. 1979, Symula et al. 2001, Turner 2001), so the foraging behaviours that generate it are likely to be exhibited by many different types of predator. Human subjects have long been used to test and refine ideas relating to predation (Götmark & Hohfält 1995, Glanville & Allen 1997, Sherratt & Beatty 2003) and although humans represent visual foragers with a high capacity for learning and strategizing, they share with natural predators a finite capacity to process information (Shepard et al. 1961). Indeed, recent work with human foragers has replicated the qualitative findings of earlier studies using great tits (Alatalo & Mappes 1996, Beatty et al. 2005). The interpretation of our results does not depend on any behavioural traits that are unique to humans, and we hope these findings justify and inspire further work into Müllerian mimicry evolution in complex communities using non-human predators.

One reason for the above results may be that predators are not sufficiently confused to generate selection for mimicry when just two different forms are involved; they may learn rapidly to avoid both types of prey, regardless of their level of resemblance. It has recently been argued that there is likely to be far stronger selection on defended prey to adopt a common form of advertisement when there are multiple forms of prey available (MacDougall & Dawkins 1998). However, despite calls for experiments (Ruxton 1998), no study has investigated selection for mimicry when there are more than two unprofitable prey types. Natural communities often contain a wide variety of prey types that differ in their profitabilities and appearances; Müllerian mimics themselves frequently participate in complex mimicry rings (Mallet & Gilbert 1995). Here we used a system of humans foraging on computer-generated prey to compare the nature and intensity of selection for mimicry when there were few forms of prey, and when there were many different types of prey available.

The relative survivorship of each of our distinct focal forms differed according to the experimental context, which strongly suggests that foragers did not rely entirely on associative learning (as Müller had originally envisioned (Müller 1878, Müller 1879)), but used simple rules to distinguish between profitable and unprofitable prey. Discriminative learning has been widely discussed in the psychological literature (Shettleworth 1998, Shepard et al. 1961, Pearce 1994), but to our knowledge this is the first study to show that this type of behaviour can facilitate the spread of rare mimics with imperfect resemblance only, in a manner that is not apparent in simplified systems with relatively few

prey types. As Fisher observed (Fisher 1930), “being recognized as unpalatable is equivalent to avoiding confusion with palatable prey”. Although Müller did not consider the appearance of profitable prey at all when making his arguments, such considerations may be essential to a full understanding of when and how Müllerian mimicry evolves.

#### **Chapter 4:**

Conspicuous coloration in males of the damselfly *Nehalennia irene* (Zygoptera: Coenagrionidae): do males signal their unprofitability to other males?

#### **Preface**

This final chapter represents the field component of my thesis research, performed at Queens University Biology Station in the summers of 2004-2005. The idea of an intra-specific warning signal is relatively new, and this research represents a contribution to the small but growing body of evidence that warning signals may on occasion function between conspecifics. This chapter is presently in preparation for submission, with an intention to submit to the *Journal of Insect Behaviour*.

## Introduction

Damselflies, like many organisms, demonstrate a sexual colour dimorphism in which males are brightly coloured, whereas females are more characteristically drab (West-Eberhard 1983, Andersson 1994, Corbet 1999). In many groups this conspicuous male coloration is attributed to competition among males to gain access to females (Moore 1990, Gibbons & Pain 1992, Grether 1996a, Hooper et al. 1999, Siva-Jothy 2000), with males displaying their quality as mates, either in their physical attributes or in the quality of the resources that they defend (Emlen & Oring 1977, Ubukata 1987, Conrad & Pritchard 1992, Battin 1993). Among many species in the family Coenagrionidae, however, males do not defend territories or engage in displays to females (Corbet 1999). In these species, mating resembles a “scramble” with males pursuing females and at times coercing them to mate. Also, while males in these species are conspicuously coloured, the colour patterns are displayed primarily on the dorsal surface of the male body, rather than throughout the body or on their wings; as such their coloration is not immediately visible to females as males approach. These conditions beg the question of what selective forces are behind the evolution and maintenance of seemingly bright coloration in males of this group.

Sherratt & Forbes (2001) suggested a possible explanation for the sexual dimorphism in Coenagrionids; namely, that bright coloration has evolved in the males of these damselfly species primarily as a form of sex-related warning coloration. Through verbal and quantitative models they demonstrated that if male-male interactions are costly (through opportunity costs or

energetic/physical costs) and if males utilize coloration to differentiate males from females, then sexual dimorphism with a conspicuously coloured male form is likely to evolve, to avoid confusion and minimize costs to males for 'getting it wrong'. Sherratt and Forbes predict that even under conditions of increased predation of males due to increased conspicuousness to predators, sexual dimorphism would still occur.

While Sherratt and Forbes' model suggests that male conspicuousness to avoid harassment ("antiharassment aposematism" as they termed it) could be an explanation for the sexual dimorphism in Coenagrionids, no study has tested this idea in the field. While it is difficult to demonstrate through experiments whether the conspicuous coloration of males initially evolved as an aposematic signal, predictions can be made about the selective forces that would maintain conspicuousness as a signal of unprofitability.

If coloration serves as an indicator of "maleness" to other males, then we would predict that males showing the common male coloration would be harassed less than males with coloration more similar to that of the female. Further, if increased harassment is costly, we would predict that males displaying colours other than the common male coloration might show lower fitness in some way. While there are many measures of fitness that could be considered (lower foraging time, lowered number of matings) we here consider the most basic of measures of fitness, that of decreased survivorship due to excessive harassment. In this paper we examine whether males of the species *Nehalennia*

*irene* (Odonata: Coenagrionidae) receive significantly different harassment from other males when their conspicuous coloration is altered.

## Methods

### *Study organism*

*Nehalennia irene* is one of the most abundant and widespread damselflies in Ontario (Catling & Brownell 2000). Adults are small in size (approximately 26-28 mm in length). Males are dark in colour with metallic blue-green markings and with a conspicuous area of blue containing a pair of black spots on the dorsal side of abdominal segments 9 and 10 (Figure 4-1). Females are predominantly pale green in colour; females in this region demonstrate a colour polymorphism, with a drab female form and a male-like "andromorph" form (Lam 2004).

In Ontario the adult flight season of *N. irene* lasts from approximately May 30<sup>th</sup> through August 19<sup>th</sup>. *N. irene* is found predominantly in still water habitats, in marshy and boggy ponds, streams and temporary bog pools (Catling & Brownell 2000).

### *Study site/experimental setup*

This research was performed at Queens University Biological Station, approximately 50 km north of Kingston, Ontario. The research site, known as Barb's Marsh (44°31'30"N, 76°22'20") is an approximately 3.7 hectare

permanent beaver pond surrounded by open meadows. A set of 1.8m × 1.8m Bioquip Outdoor Cages were assembled approximately 3m from the wet margins of the pond, in an area where the vegetation had been cut to approximately 10cm in height. In 2004, (Experiment 1) four cages were set up in a 2 × 2 grid, with a one meter gap between the cages. In 2005, (Experiment 2) 3 cages were set up in a straight line along the shore of the pond. A 4L water container was placed in the southwest and northeast corners of each cage, with garden canes placed in the container to serve as perches. Cages were swept in an attempt to remove invertebrate predators (mainly spiders) from the cages to minimize damselfly mortality due to predation. Cages were also regularly seeded with small flying insects collected with sweep nets from the surrounding field to maintain a supply of prey items for the damselflies. Ambient air temperature was monitored both inside and outside the cages. Air temperature and wind-speed readings were taken at the beginning and end of each experiment and during observation periods.



**Figure 4-1** Image of a male *Nehalennia irene* (courtesy D. Paulson).

### *Experimental design*

To test differential harassment and survivorship in males of *N. irene*, we altered male coloration by painting the dorsal side of abdominal segments 9 and 10 with three different colours of non-toxic paint (Ultra Gloss acrylic enamel, DecoArt, Stanford, KY). Abdominal segments 9 and 10 are bright blue in males and andromorphic females of *N. irene* while much of the rest of the body is black in colour; heteromorphic females of *N. irene* lack this blue abdominal coloration. Altering the colour of this relatively small portion of the body makes a large difference in the overall appearance of individuals, at least to human eyes. Previous studies of male response to female coloration (Gorb 1998, Miller & Fincke 1999) have determined that abdominal coloration is the major component used in mate recognition. Altered males in our experiments were painted one of three colours: blue, such that they continued to look like males, orange, an entirely novel colour in the species and black, making males look more like females, at least to human eyes. This does not account for any UV component to coloration, which may exist, and which damselflies could detect.

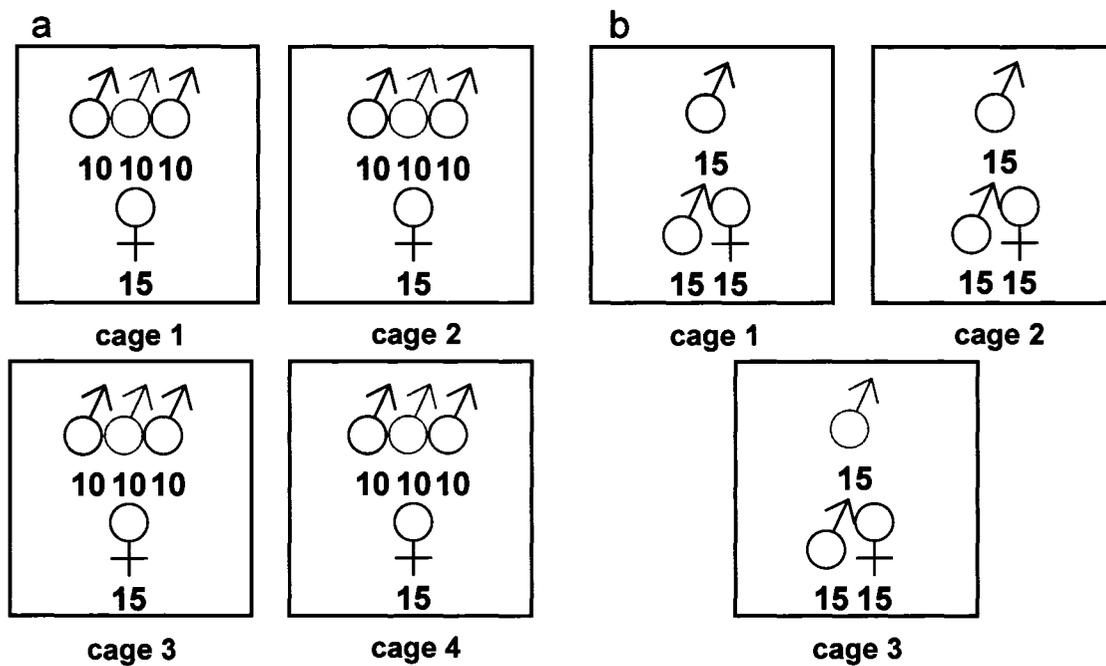
Males and heteromorphic females used for the experiment were collected each morning between 8:45 and 10:30 with sweep nets and placed in 0.6m × 0.6m × 0.6m holding cages, sorted by gender. Males and females were then removed from the holding cages and placed in the experimental cages; males were painted at this time. While all males of each colour were painted together, the order in which paint colours were applied to males was randomized between

cages. The order in which damselflies were placed in cages was also randomized each day.

We conducted two different experiments within our outdoor cages. In experiment 1, 10 males of each colour (blue, orange and black) were placed in each of the four cages along with 15 females (see Figure 4-2a). This experiment allowed us to directly compare the behaviour and survivorship of individuals of each colour type. Individuals remained in the cages for approximately 24 hours, at which time individuals were removed and survivorship of each type of individual (blue male, orange male, black male, female) were recorded. A total of 20 cages (five sets with four cages each) were run between June 23<sup>rd</sup> and July 4<sup>th</sup>, 2004 on sunny/partly cloudy days.

Between 12:30 and 16:00 on the day that each experiment was begun, an observer sat in the corner of each cage for 30 minutes, recording each interaction between individuals. An interaction was recorded if one individual perched atop another individual and attempted to form a tandem. The “harasser” and “harassee” in each interaction were recorded: whether the individual was male or female, and if male what colour it was painted.

In experiment 2, 15 males painted one of the three colours were placed in a cage with 15 unpainted males (all males were handled in the same way; unaltered males were brushed with a dry paintbrush before being introduced to the cage) and 15 females (identical density and sex ratio to the previous experiment). One each of the three male colours was randomly assigned to one



**Figure 4-2** Experimental design for experiments 1 and 2. a) Four-cage design for experiment 1. Each cage contained 10 males of each painted colour and 15 females. b) Three-cage design for experiment 2. Each cage contained 15 painted males, with a different paint colour applied in each of the three cages. Each cage also contained 15 unpainted males and 15 females.

of three cages each day (Figure 4-2b). Behaviour and survivorship were recorded as in experiment 1. This experimental design allowed us not only to compare the behaviour and survivorship of each type of male (blue, orange and black) to each other type, but also allowed us to compare each type of painted male to unpainted males in the same cage, to determine the effects of paint on behaviour and survivorship. Eight replicates (day) of experiment 2 were performed between June 23<sup>rd</sup> and July 5<sup>th</sup>, 2005.

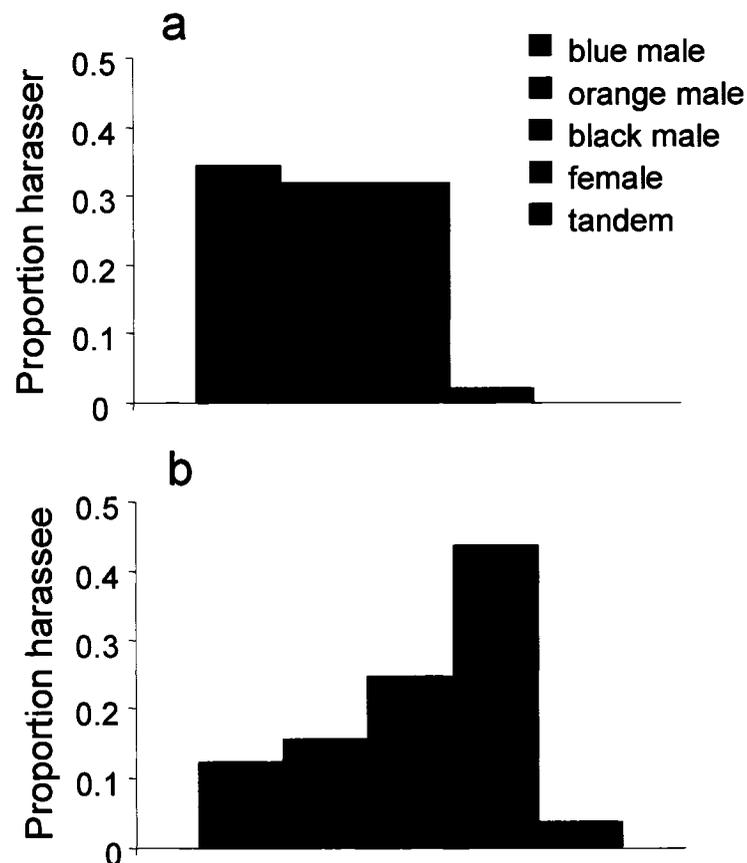
To analyze the behavioural results of experiment 1 we employed goodness-of-fit tests, including a replicated G-test (Sokal & Rohlf 1994, page 720). This approach not only provides an overall (pooled) test of whether different colour morphs were statistically more likely to be 'harassers' or 'harassees' (through  $G_{POOLED}$ ) but also allows us to test whether departures from expectation are in the same direction and magnitude on different days (through  $G_{HETEROGENEITY}$ , equivalent to an interaction term in an analysis of variance). These two G statistics sum to  $G_{TOTAL}$ , a reflection of total divergence from expectation independent of the direction(s) of departure. Survivorship results for experiment 1 were also analyzed with a replicated G-test. Behavioural data from experiment 2 were analyzed with a randomized-block analysis of variance (ANOVA) with day as the block and colour as the explanatory variable. Survivorship data from experiment 2 were also analyzed with a randomized block ANOVA of the same design. To compare survivorship of painted males to unpainted males we employed paired t-tests.

## Results

### *Experiment 1*

#### *Behavioural observations*

A total of 236 observed interactions were recorded during experiment 1. Figure 4-3 shows the proportion of total interactions in experiment 1 in which each type of individual was a harasser (Figure 4-3a) or a harassee (Figure 4-3b). A simple conservative goodness-of-fit test analyzing the single most common harasser (blue male, orange male or black male) and the single most common harassee in each cage (ties were excluded) found no difference between morphs in their frequency of being a harasser ( $G= 1.768$ ,  $df=2$ ,  $P=0.413$ ); also, no difference was found in the likelihood of each type of male to be a harassee ( $G=4.757$ ,  $df=2$ ,  $P=0.093$ ) although there was a trend toward black-painted males receiving a higher level of harassment than orange or blue males. Considering all observations on a per-day basis, however, black-painted males were statistically more likely than blue- or orange-painted males to be the recipients of harassment (replicated G-test—see  $G_{POOLED}$  in Table 4-1). No difference was found between the three male forms in their propensity to be a harasser, though differences were found between days in the overall levels of harassment ( $G_{HETEROGENEITY}$ , Table 4-1), likely due to weather and temperature differences between days.



**Figure 4-3** Results of behavioural observations over the 5 replicates of experiment 1 (total of 20 cages). A total of 236 observations of interactions were made. a) Proportion of total interactions in which each of five possible types (each of the three painted male colours, females and tandems of a male and a female) was the instigator of an interaction, called the 'harasser'. b) Proportion of total interactions in which each of the five possible types was the recipient of harassment, the 'harassee'.

**Table 4-1** Replicated G-test results examining the distribution of colour morphs (blue, orange and black) for individuals involved in interactions. Results for 'harassers' and 'harassees' are analyzed on a per-day basis. Here, differences between different morphs are indicated by the  $G_{POOLED}$  results, while differences between days are reflected in  $G_{HETEROGENEITY}$ .

<b>Interactions</b>	<b>G</b>	<b>df</b>	<b>P</b>	<b>G</b>	<b>df</b>	<b>P</b>
<b>Per-day</b>	<i>'Harasser'</i>			<i>'Harassee'</i>		
Pooled	0.31	2	0.857	10.55	2	0.005
Heterogeneity	23.20	8	0.003	10.48	8	0.233
Total	23.51	10	0.009	21.03	10	0.021

**Table 4-2** Replicated G-test results examining the survivorship of different colour morphs (blue, orange and black). Results for groups by day and by cage found no difference in survivorship between different painted male types (indicated by the  $G_{POOLED}$  results), while differences in survivorship between days or cages are reflected in  $G_{HETEROGENEITY}$ . Overall, no significant differences were found.

<b>Interactions</b>	<b>G</b>	<b>df</b>	<b>P</b>
<b>Per-day</b>			
Pooled	2.98	2	0.225
Heterogeneity	14.40	8	0.072
Total	17.38	10	0.066
<b>Per-cage</b>			
Pooled	1.99	2	0.369
Heterogeneity	4.10	6	0.663
Total	6.10	8	0.636

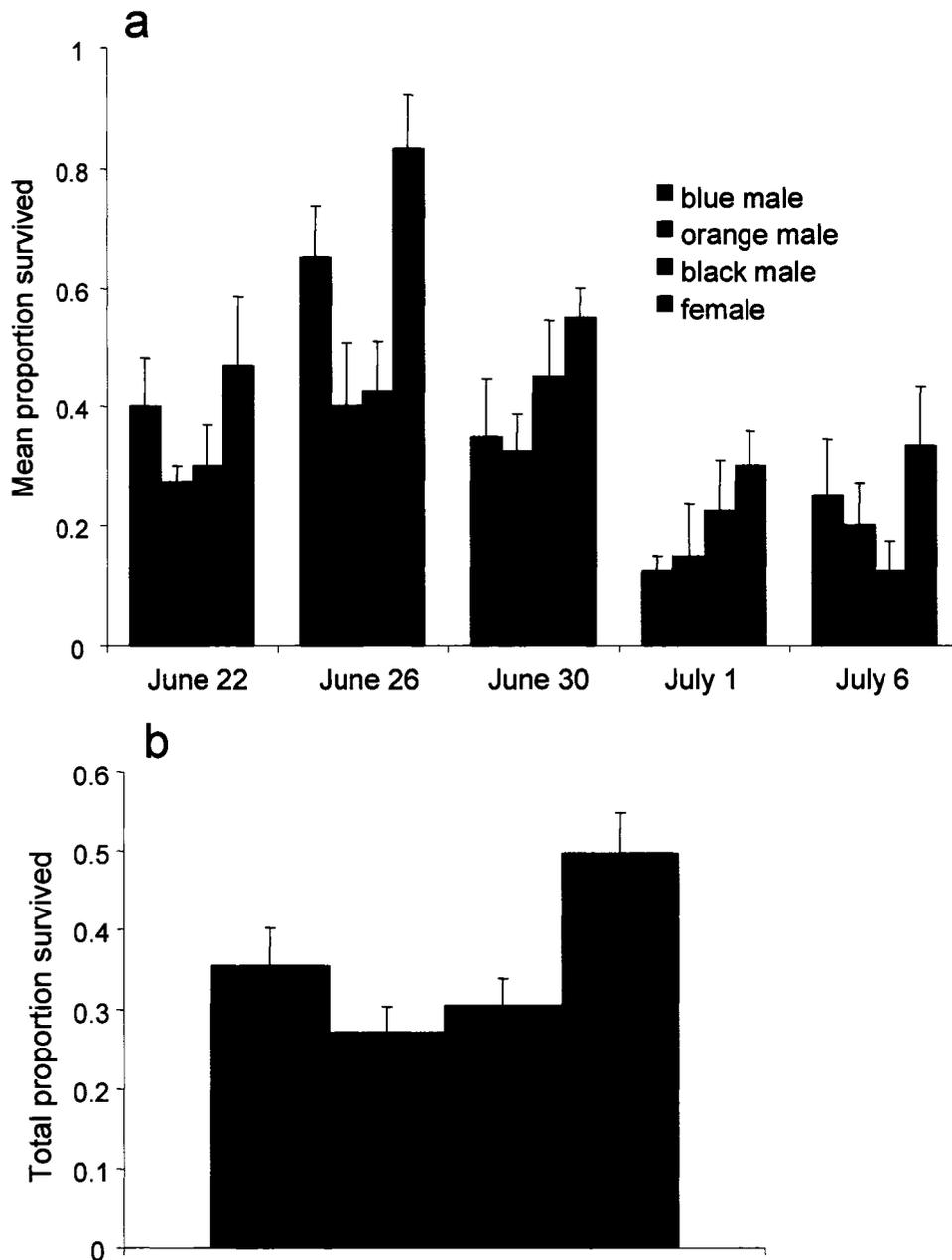
### *Survivorship*

Figure 4-4 presents the mean proportion survived of each type of male and females in experiment 1 (Figure 4-4a), as well as the total proportion survived of each type (Figure 4-4b). A replicated G-test on male survivorship analyzed by day found no significant difference between different colours of painted male, though variation in overall survivorship between days had borderline significance. A replicated G-test on male survivorship analyzed by cage found no significant differences in survivorship between the three painted forms or between cages (see Table 4-2).

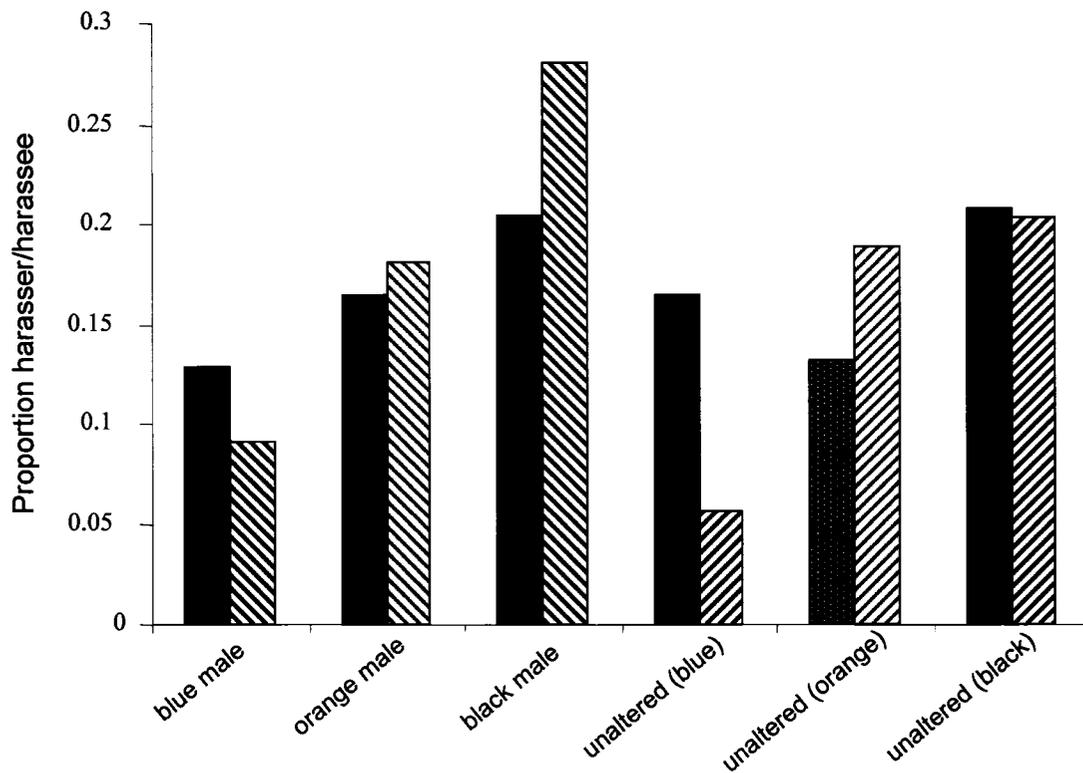
## ***Experiment 2***

### *Behavioural Observations*

A total of 304 behavioural observations were made for experiment 2 (Figure 4-5). In experiment 2, a significant amount of variation was explained by colour (randomized-block ANOVA, colour  $F_{2,24}=4.327$ ,  $P=0.034$ ). Black-painted males again received significantly more harassment than blue- or orange-painted males (Tukey HSD: blue vs. black  $P=0.027$ ; blue vs. orange  $P=0.360$ ; black vs. orange  $P=0.310$ ), while there was no difference in the three forms in their propensity to harass ( $F_{2,24}=0.946$ ,  $P=0.412$ ). In both tests, day (block) had a significant effect on the results, reflecting the fact that overall levels of activity varied between days, likely as a function of weather and temperature conditions (randomized-day



**Figure 4-4** Survivorship data from experiment 1. a) Mean proportion survivorship ( $\pm 1$  SE) for each type of painted male (blue, orange and black) and females per day. b) Total proportion survived ( $\pm 1$  SE) of each type of male (blue, orange and black) and females.

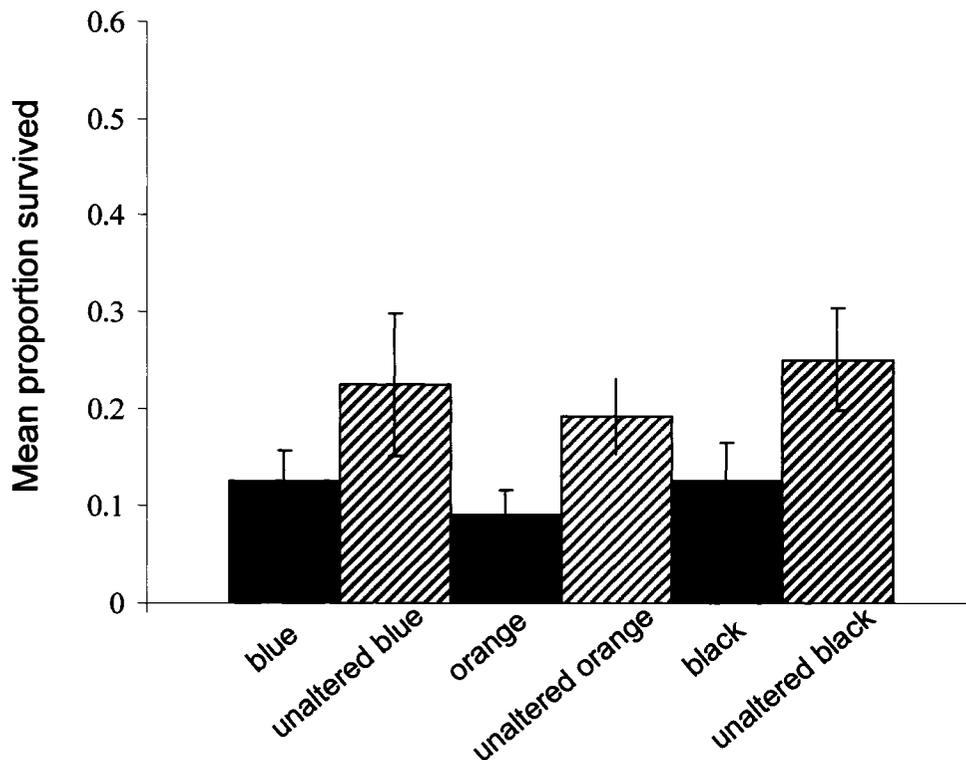


**Figure 4-5** Results of behavioural observations over the 8 replicates of experiment 2, showing total interactions of painted males and the unaltered males placed in the cage with them. A total of 304 observations were made. Solid bars represent painted males that were 'harassers', dotted bars represent unaltered males that were 'harassers' (separated out by the colour of painted male they were caged with). Hatched bars represent males of each type that were 'harassees'—the recipients of harassment.

effect on harassee:  $F_{7,24}=3.68$ ,  $P=0.018$ ; day effect on harasser:  $F_{7,24}=2.768$ ,  $P=0.050$ ). While a consideration of the total interactions indicates a trend toward lower harassment of unpainted males caged with blue painted males (Figure 4-5), this was not statistically significant (randomized-block ANOVA on unpainted males, colour (based on the colour of painted male that males were caged with):  $F_{2,24}=2.72$ ,  $P=0.100$ ).

### *Survivorship*

For experiment 2, a randomized-block ANOVA on arcsine-transformed proportions survived of each of the three types of painted males (blue, orange and black) and the three types of unpainted males (unaltered males each introduced with a different colour painted male) revealed a significant effect of day (block) on survivorship ( $F_{7,48}=3.489$ ,  $P=0.006$ ), but no affect of cage or colour of males ( $F_{5,48}=1.83$ ,  $P=0.132$ ). A direct comparison of arcsine-transformed proportions survived of each painted colour of male with the unaltered males introduced into the cage with them found no difference in survivorship between blue-painted males and unpainted males (paired t-test:  $t_s=1.47$ ,  $df=7$ ,  $P=0.18$ ) or between orange-painted males and unpainted males (paired t-test:  $t_s=2.13$ ,  $df=7$ ,  $P=0.071$ ), but did find a survivorship difference between black-painted males and unpainted males (paired t-test:  $t_s=3.91$ ,  $df=7$ ,  $P=0.006$ ) with black-painted males having lower survivorship than unaltered males introduced into the cages with them (Figure 4-6).



**Figure 4-6** Mean proportion survived ( $\pm 1$  SE) of each type of male (blue, black and orange painted, as well as the unaltered males placed in the cage with each type of male) in experiment 2. While there is a general trend toward higher survivorship of unaltered males, this trend was not significant (randomized-block ANOVA); thus paint did not have a significant effect on the survivorship of painted individuals. Pairwise t-tests did find a significant difference in survivorship between black painted males and the unaltered males placed in the cages with them. Comparisons between blue-painted vs. unaltered males and orange-painted vs. unaltered males found no significant differences.

## Discussion

Previous work on male response to conspecifics has focused predominantly on the coloration and morphology of female models. In a study intended to identify which visual/morphological cues male damselflies use in sex determination, Gorb (1998) presented male and female models with different body components interchanged or painted to look like the opposite sex. Gorb found that males of *Coenagrion puella* tended not to show aggression to one another, and indeed females painted to look like males repelled conspecifics as effectively as male models did. The overall results indicate that abdomen and wing coloration function as the primary cues used by males to identify females.

Miller & Fincke (1999) likewise presented female and male models to males of the species *Enallagma ebrium*, to determine what different responses female morphs (andromorphs and heteromorphs—see Robertson 1985, Fincke 1994, Cordero et al. 1998) elicited from males, to further understand which components of female coloration are used by males in mate recognition. Females modified to look like males in their study received lowered sexual responses from males; likewise (with significant implications for our present study) males modified to look like females—through blackening of conspicuous abdominal coloration—experienced a considerable increase in sexual responses by other males.

Here, we have altered male coloration to observe its effects on the response of males to other males, rather than females. We have also observed these interactions in free-moving study individuals; while our research organisms

were housed in cages, they were allowed to move about freely within cages, thus interactions between individuals include behavioural as well as visual cues. Previous studies predominantly used dead individuals and artificial models (Gorb 1998) or live individuals attached to a perch or tether (Miller & Fincke 1999). Thus we hopefully have studied more realistic male responses to conspecifics. Individuals in our study were quite active (though there was variation in the level of activity between individuals) and quite readily approached other individuals and attempted to mate. Ovipositions were observed in the water features placed in the cages, thus indicating that cage conditions were favourable to completion of the mating sequence.

If conspicuous male coloration in *Nehalennia irene* functions as a warning signal to other males of their unprofitability of mates, then we predicted that males bearing the common coloration would be harassed less than males with coloration similar to that of females. Our results indicate that this is indeed the case. The proportion of approaches to blue-painted males by other males was nearly half that of approaches to black-painted males in experiment 1. The lowered harassment of blue-painted males in our experiments indicates that male coloration may influence other males in the decision to approach them as possible mates. Black-painted males, looking (at least to human eyes) more like females received a much greater proportion of harassment.

The fate of orange painted males in our experiments opens up an interesting question about the exact nature of the signal—the specific information that male coloration provides. Orange-painted males received lower harassment

than black-painted males; their level of harassment was not significantly different from that of blue-painted males. As this is a non-characteristic colour for this species, it could be that these males were simply treated as a different species. It could also be that they were ignored because they did not look like *females*. Thus, our question: does male coloration in this group act as a signal of “maleness”, or is it a signal of “un-femaleness”? Based on our results, the answer could well be the latter, though in fact this would still be in line with the predictions of Sherratt & Forbes’ (2001) model. Males, to distinguish themselves and thereby avoid harassment, display a mark that allows them to be seen as different from females. As with more traditional warning signals, if the coloration of Coenagrionid males functions as a signal of unprofitability, then displaying the signal should help in avoiding confusion with profitable individuals (Fisher 1930, Sherratt 2002, Sherratt & Beatty 2003), which appears to be the case in these experiments.

The fact that we did not find strong significant evidence for differences in mortality between our different types of males is perhaps not surprising. Such differences in survivorship are likely to be subtle and not easy to detect with experiments of this type. For an experiment of such short duration (24 hours) to detect survivorship differences would require the selective pressures against a particular morph to be substantial. Our choice to measure survivorship was based on its ease of measurement and its easy interpretation as a cost associated with carrying an altered signal. Indeed, we do have some suggestion in our survivorship results from experiment 2 that black-painted males have

lower survivorship than unaltered males in the cage with them, though an overall test for differences between all male types did not detect this.

While other studies have found evidence for coloration in damselflies as an intra-sexual signal (Grether 1996b, Grether 1997, Fitzstephens & Getty 2000), these often function as a signal in intraspecific competition between territorial species, with the signal being variable in response to the quality of the individual that carries it. Our system is different in that there is no tendency in *N. irene* (or many other Coenagrionid species) to be territorial, and there appears to be little variability in the signal, in that there is little within-species variability in coloration. Thus while this signal may be maintained by intra-sexual selection, it differs in the message that the signal carries: it is simply an indicator separating males from females.

Others have suggested the possibility of an intra-specific warning signal. Papaj & Newsom (2005) found that the coloration of larvae of the pipevine swallowtail, *Battus philenor*, deterred conspecific females from ovipositing on an already occupied host plant. Papaj & Newsom demonstrated that the colour of the larvae was the significant cue that females used that a plant was occupied and thus avoided it. Poulton (1890) offered a similar example in the case of the *Ichneuemon* wasp, a parasitoid which apparently avoids laying its eggs on host larvae if it detects the bright colours of the eggs of a conspecific, indicating that the larva has already been parasitized. Thus other examples of within-species warning signals do exist.

While we are not able to demonstrate that conspicuous male colouration in Coenagrionid damselflies initially evolved as a warning signal, we have demonstrated that males displaying their common conspicuous coloration due indeed suffer less harassment, a necessary condition if the anti-harassment aposematism hypothesis is correct. Further studies of this system should focus on the actual selective forces that could be behind the maintenance of an intra-specific warning signal. Other costs, such as loss of feeding opportunities, decreased time available for seeking mates or decreased number of copulations could impose selection forces that maintain a within-species warning signal. We plan to experiment with these possibilities in the future, to determine if further evidence exists for the intra-specific aposematism hypothesis.

## GENERAL CONCLUSIONS

In a recent discussion with a colleague I was asked the following question: “Does your focus on research questions that were addressed by biologists a century ago indicate your belief that evolutionary biology is dead?” My response to this provocative salvo was to suggest that addressing unanswered questions, no matter how long they had lingered unanswered, was not a fruitless endeavour. As is oft suggested in the introductory paragraphs of research articles on warning signaling, the interplay between bright coloration and defence fascinated many of the venerable names of 19<sup>th</sup>- and 20<sup>th</sup>-century biology, names such as Darwin, Wallace, Bates, Müller, Poulton, Fisher, Cott and Edmunds. Many of the same questions that they pondered continue to captivate—as the final answers to these questions continue to elude—modern biologists.

My colleague’s position may be attributed to a false sense of resolution in the study of protective coloration: the feeling that the evolution of aposematism, Batesian and Müllerian mimicry (as well as crypsis and disruptive coloration) are all sewn up. This perception could be due in part to the regular use of these colourful (literally) examples in biology textbooks for evolution through Darwinian selection. Just about anyone who has taken a biology course has been told of Kettlewell’s *Biston betularia*, has been explained the parasitic relationship between the Monarch and the Viceroy, has learned that bright colours in nature often cry out “stay away”. But these stories, and many others, are considerably

more complicated than they appear to be, and warning coloration and mimicry are still fascinating fields of study with many unanswered questions.

If there is a common theme running through my thesis it is this: that I have attempted to address some of the fundamental questions concerning the evolution of warning coloration, indeed some of those very questions that have lingered in the minds of biologists since Darwin's time. How did warning colours originally evolve? What role has aggregation played in warning signal evolution? Under what conditions would selection have favoured Müllerian mimicry? Do warning signals function in ways other than communication between prey and predator? Research on aposematism has experienced a resurgence in recent years and I am happy to be a part of this renewed interest in warning signals.

In Chapter 1 I investigated how the attributes of defended (unprofitable) and undefended (profitable) computer-generated prey species tended to evolve when they were subject to selection by foraging humans. When cryptic forms of defended and undefended species were similar in appearance but their conspicuous forms were not, then defended prey became conspicuous while undefended prey remained cryptic. Indeed in all of our experiments, defended prey invariably evolved any trait that enabled them to be distinguished from undefended prey, even if such traits were cryptic. When conspicuous mutants of defended prey were extremely rare they frequently overcame their initial disadvantage by chance—surviving in a part of the foraging area unvisited by a predator, and proliferating in the time prior to their first predator encounter. When Batesian mimicry of defended species was possible, defended prey

evolved unique traits or characteristics that would make undefended prey vulnerable. Overall, this work supports the contention that warning signals are selected for their reliability as indicators of defense, rather than to capitalize on any inherent educational biases of predators.

In Chapter 2 I investigated the response of human predators to profitable and unprofitable prey in a computer version of the “novel world” experiments that addressed the relationship between gregariousness and warning signals in defended species. When all prey types were cryptic and all profitable prey were solitary, then aggregated unprofitable prey had higher per capita survivorship than solitary unprofitable prey. In this case aggregation was the sole reliable distinguishing characteristic of unprofitability. By contrast, when profitable and unprofitable prey could be distinguished in terms of their conspicuousness, then aggregation did not confer a benefit to unprofitable prey. When profitable prey were simultaneously presented in solitary and aggregated configurations, aggregated profitable prey consistently had a lower per capita survivorship than solitary profitable prey, regardless of the configuration of unprofitable prey. I conclude that while aggregation can serve as a signal of unprofitability, it may not be required for the evolution of warning signals. Instead, this research suggested that the common association between aggregation and distastefulness may primarily arise for a “negative” reason: the vulnerability of aggregated palatable prey.

In Chapter 3 I show, using a system of humans hunting for computer-generated prey, that predators do not always generate strong selection for Müllerian mimicry when there are two unprofitable prey types. By contrast, it is demonstrated that when predators are faced with a range of different prey species, selection on unprofitable prey to resemble one another can be intense. Here the primary selective force is not one in which predators evaluate the profitabilities of distinct prey types independently, but one in which predators learn better to avoid unprofitable phenotypes that share traits distinguishing them from profitable prey. This need to simplify decision making readily facilitates the spread of imperfect mimetic forms from rarity, and suggests that Müllerian mimicry is more likely to arise in multispecies communities.

Finally, in Chapter 4, I evaluated some of the underlying assumptions of the intraspecific aposematism theory in Coenagrionid damselflies by comparing the levels of harassment and survivorship of males of the species *Nehalennia irene* that had been painted in such a way as to make them look similar to an unaltered male (painted blue), simply different from the male (painted orange) or more like the heteromorphic female (painted black). I found that in cage experiments with painted males and females that blue-painted males and orange-painted males experience significantly lower harassment than do black-painted males. Generally no significant difference in survivorship in the different male morphs was found, though black-painted males in one experiment had lower survivorship than unpainted males housed with them. While these results do not demonstrate that sexual colour dimorphism initially evolved in this group

as an intra-specific warning signal, they do demonstrate that the level of harassment is influenced by coloration, such that selection could maintain male coloration as a warning signal.

The results of the work presented in this thesis beg a number of new and interesting questions. In Chapter 1 and elsewhere, the reliability of a signal proved to be important in its success as a warning signal. This is in line with predictions from Sherratt's (2002) work, that unprofitable prey evolve signals that profitable prey find difficult to exploit. This also makes sense in the light of Wallace's (1867) original predictions about warning signals allowing unprofitable prey to "be contrasted with and separated from" profitable prey. The two concepts are significantly different, however. Wallace does not require that aposematic prey be more detectable, only that they be distinctive. In their presentation these ideas are distinct, but are they different in their application, and if so, how? Work modelling Wallace's ideas about distinctiveness, in comparison with signal reliability, may bring further understanding as to why warning signals tend to be conspicuous.

In Chapter 3 Müllerian mimicry appeared much more likely to evolve in complex systems of prey, where unprofitable prey share some common, reliable trait. These results beg two significant questions: first, is discriminative learning employed by natural predators in a way similar to its use by humans? Experiments with other predators, such as birds, will be necessary to determine whether this is a common phenomenon. Second, if a simple, common trait

allows Müllerian mimicry to function, why do Müllerian mimics tend to show such perfect similarity? Previous research (Franks & Noble 2004) would indicate that the presence of palatable (Batesian) mimics in a system may not only increase selection for Müllerian mimicry, but may enhance selection for greater similarity between Müllerian mimics. Experiments that include both profitable and unprofitable mimics will be necessary to better understand the selection forces behind Müllerian mimicry.

In Chapter 4, males that looked more female-like (at least to human eyes) experienced increased attention from conspecific males. This increased attention could be harassing, leading to selection for males that look more male-like, in this case showing the conspicuous colours common to males. But what form does this harassment-induced selection take? Results measuring survivorship were equivocal at best, but other endpoints, such as differential physical resources (i.e., as body fat mass), and limits to time spent foraging or searching for mates could also be measured. To further test the concept of intraspecific aposematism, experiments considering these endpoints should be performed.

This has been a fulfilling research programme for me, and I am grateful to have had the opportunity to perform this work and to contribute to the body of knowledge on the evolution of aposematism and Müllerian mimicry. The findings of this thesis beg a number of new questions on signal reliability, selection on

Müllerian mimics and the complexities of evolution. I hope to continue to pursue these questions.

***Lumen Coeleste Sequamur***

**Literature Cited**

- Alatalo, R.V., & Mappes, J. 1996. Tracking the evolution of warning signals. *Nature* **382**:708-710.
- Alatalo, R. V. & Mappes, J. 2000. Initial evolution of warning coloration: comments on the novel world method. *Animal Behaviour* **60**: F1-F2.
- Andersson, M. 1994. *Sexual selection*. Princeton, Princeton University Press.
- Balogh, A. C. & Leimar, O. 2005. Müllerian mimicry: an examination of Fisher's theory of gradual evolutionary change. *Proceedings of the Royal Society of London B* **272**: 2269-2275.
- Bates, H. W. 1862. Contributions to an insect fauna of the Amazon valley. *Lepidoptera: Heliconidae. Transactions of the Linnean Society of London* **23**: 495–566.
- Battin, T. 1993. The odonate mating system, communication, and sexual selection: a review. *Bollettino di Zoologia* **60**: 353-360.
- Beatty, C. D., Bain, R. S. & Sherratt, T. N. 2005. The evolution of aggregation in profitable and unprofitable prey. *Animal Behaviour* **70**: 1195–1202.
- Benson, W. W. 1972. Natural selection for Müllerian mimicry in *Heliconius erato* in Costa Rica. *Science* **176**: 936–939.
- Bertram, B. C. R. 1978. Living in groups: predators and prey. In: J. R. Krebs & N. B. Davies (eds), *Behavioural Ecology: an evolutionary approach*. Sunderland, Mass, Sinauer, pp. 64-96.
- Bond, A.B. & Kamil, A.C. 1998. Apostatic selection by blue jays produces balanced polymorphism in cryptic prey. *Nature* **395**: 594-596.
- Bond, A.B. & Kamil, A.C. 2002. Visual predators select for crypticity and polymorphism in virtual prey. *Nature* **415**: 609-613.

- Brower, J. V. Z. 1960. Experimental studies of mimicry. IV. The reactions of starlings to different proportions of models and mimics. *American Naturalist* **94**: 271-282.
- Catling, P. M. & Brownell, V. R. 2000. *Damselflies and Dragonflies (Odonata) of Ontario: Resource Guide and Annotated List*. ProResources, Ontario Canada.
- Chai, P. 1996. Butterfly visual characteristics and ontogeny of responses to butterflies by a specialized tropical bird. *Biological Journal of the Linnean Society* **59**: 37-67.
- Chai, P. & Srygley, R. B. 1990. Predation and the flight, morphology, and temperature of neotropical rain-forest butterflies. *American Naturalist* **135**: 748-765.
- Clark, B. R. & Faeth, S. H. 1997. The consequences of larval aggregation in the butterfly *Chlosyne lacinia*. *Ecological Entomology* **22**: 408-415.
- Clark, B. R. & Faeth, S. H. 1998. The evolution of egg clustering in butterflies: a test of the egg desiccation hypothesis. *Evolutionary Ecology* **12**: 543-552.
- Conrad, K. F. & Pritchard, G. 1992. An ecological classification of odonate mating systems – the natural influence of natural, intersexual and intrasexual selection on males. *Biological Journal of the Linnean Society* **45**: 255-269.
- Cook, S. E., Vernon, J. G., Bateson, M. & Guilford, T. 1994. Mate choice in the polymorphic African Swallowtail Butterfly, *Papilio dardanus*: male-like females may avoid sexual harassment. *Animal Behaviour* **47**: 389-397.
- Corbet, P. S. 1999. *Dragonflies: Behaviour and Ecology of Odonata*. Ithaca, New York, Comstock Publishing Associates.
- Cordero, A. 1992. Density-dependent mating success and colour polymorphism in females of the damselfly *Ischnura graellsii* (Odonata, Coenagrionidae). *Journal of Animal Ecology* **61**: 769-780.

- Cordero, A., Carbone, S. S., & Utzeri, C. 1998. Mating opportunities and mating costs are reduced in androchrome female damselflies, *Ischnura elegans* (Odonata). *Animal Behaviour* **55**: 185-197.
- Cott, H.B. 1940. *Adaptive Coloration in Animals*. Meuthen
- Darwin, C. 1871. *The Descent of Man and Selection in Relation to Sex*. London, John Murray.
- Dill, L. M. 1975. Calculated risk-taking by predators as a factor in Batesian mimicry. *Canadian Journal of Zoology-Revue Canadienne De Zoologie* **53**:1614-1621.
- Dugatkin, L. A. 1997. *Cooperation among Animals: An Evolutionary Perspective*. New York, Oxford University Press.
- Dukas, R. 2002. Behavioural and ecological consequences of limited attention. *Philosophical Transactions of the Royal Society of London B* **357**: 1539–1548.
- Edmunds, M. 1974. *Defence in Animals*. Essex, Longman.
- Emlen, S. T. & Oring, L. W. 1977. Ecology, sexual selection, and the evolution of mating systems. *Science* **197**: 215-233.
- Endler, J. A. 1991. Interactions between predators and prey. In: J.R. Krebs & N.B. Davies (eds), *Behavioural Ecology, 3<sup>rd</sup> Edition*. Oxford, Blackwell Scientific, pp. 169-196.
- Fincke, O. M. 1994. Female dimorphism in damselflies: failure to reject the null hypothesis. *Animal Behaviour* **47**: 1249-1266.
- Fisher, R. A. 1930. *The Genetical Theory of Natural Selection*, Oxford University Press.
- Fitzstephens, D. M. & Getty T. 2000. Colour, fat and social status in male damselflies *Calopteryx maculata*. *Animal Behaviour* **60**: 851-855.

- Franks, D. W. & Noble, J. 2004. Batesian mimics influence mimicry ring evolution. *Proceedings of the Royal Society of London B* **271**: 191-196.
- Gagliardo, A. & Guilford, T. 1993. Why do warning-colored prey live gregariously? *Proceedings of the Royal Society of London Series B* **251**: 69-74.
- Gamberale, G. & Tullberg, B. S. 1996. Evidence for a more effective signal in aggregated aposematic prey. *Animal Behaviour* **52**: 597-601.
- Gamberale, G. & Tullberg, B. S. 1996. Evidence for peak-shift in predator generalization among aposematic prey. *Proceedings of the Royal Society of London B* **263**: 1329-1334.
- Gamberale, G. & Tullberg, B. S. 1998. Aposematism and gregariousness: the combined effect of group size and coloration on signal repellence. *Proceedings of the Royal Society of London B* **265**: 889-894.
- Gamberale-Stille, G. & Tullberg, B. S. 1999. Experienced chicks show biased avoidance of stronger signals: an experiment with natural colour variation in live aposematic prey. *Evolutionary Ecology* **13**: 579-589.
- Gavrilets, S. & Hastings, A. 1998. Coevolutionary chase in two-species systems with applications to mimicry. *Journal of Theoretical Biology* **191**: 415-427.
- Gibbons, D.W. & Pain, D. 1992. The influence of river flow-rate on the breeding behavior of *Calopteryx* damselflies. *Journal of Animal Ecology* **61**: 283-289.
- Gittleman, J. L., & Harvey, P. H. 1980. Why are distasteful prey not cryptic? *Nature* **286**: 149-150.
- Gittleman, J. L., Harvey, P. H. & Greenwood, P. J. 1980. The evolution of conspicuous coloration - some experiments in bad taste. *Animal Behaviour* **28**: 897-899.
- Glanville, P. W. & Allen, J. A. 1997. Protective polymorphism in populations of computer-simulated moth-like prey. *Oikos* **80**: 565-571.

- Gorb, S. N. 1998. Visual cues in mate recognition by males of the damselfly *Coenagrion puella* (L.) (Odonata: Coenagrionidae). *Journal of Insect Behaviour* **11**: 73-92.
- Götmark, F. & Hohfält, A. 1995. Bright male plumage and predation risk in passerine birds: are males easier to detect than females? *Oikos* **74**: 475–484.
- Greenwood, J. J. D., Cotton, P. A. & Wilson, D. M. 1989. Frequency-dependent selection on aposematic prey - some experiments. *Biological Journal of the Linnean Society* **36**: 213–226.
- Greenwood, J. J. D., Wood, E. M. & Batchelor, S. Apostatic selection of distasteful prey. *Heredity* **47**, 27–34 (1981).
- Grether, G. F. 1996a. Intrasexual competition alone favors a sexually dimorphic ornament in the rubyspot damselfly *Hetaerina americana*. *Evolution* **50**: 1949–1957.
- Grether, G. F. 1996b. Sexual selection and survival selection on wing coloration and body size in the rubyspot damselfly *Haeterina americana*. *Evolution* **50**: 1939-1948.
- Grether, G. F. 1997. Survival cost of an intrasexually selected ornament in a damselfly. *Proceedings of the Royal Society of London Series B* **264**: 207-210.
- Guilford, T. 1985a. How do "warning colours"work? Conspicuousness may reduce recognition errors in experienced predators. *Animal Behaviour* **34**: 286-288.
- Guilford, T. 1985b. Is kin selection involved in the evolution of warning coloration. *Oikos* **45**: 31-36.
- Guilford, T. 1988. The evolution of conspicuous coloration. *American Naturalist* **131** Supplement: S7-S21.

- Guilford, T. 1990. The evolution of aposematism. Pages 23- 61 In: D. L. Evans, and J. O. Schmidt, eds. *Insect Defenses: Adaptive Mechanisms and Strategies of Prey and Predators*. Albany, State University of New York Press.
- Guilford, T., & Dawkins, M. S. 1993. Are warning colors handicaps? *Evolution* **47**: 400-416.
- Hamilton, W. D. 1963. Evolution of Altruistic Behavior. *American Naturalist* **97**: 354-356.
- Hamilton, W. D. 1996. *The Narrow Roads of Gene Land*. Oxford; New York: W.H. Freeman/Spektrum.
- Hanson, H. M. 1959. Effects of discrimination training on stimulus generalization. *Journal of Experimental Psychology* **58**: 321-334.
- Harvey, P. H., Bull, J. J., Pemberton, M. , & Paxton, R. J. 1982. The evolution of aposematic coloration in distasteful prey - a family model. *American Naturalist* **119**: 710-719.
- Hatle, J. D., Salazar, B. A. & Whitman, D. W. 2002. Survival advantage of sluggish individuals in aggregations of aposematic prey, during encounters with ambush predators. *Evolutionary Ecology* **16**: 415-431.
- Hatle, J.D. & Faragher, S.G. 1998. Slow movement increases the survivorship of a chemically defended grasshopper in predatory encounters. *Oecologia* **115**: 260-267.
- Holling, C. S. 1959. Some characteristics of simple types of predation and parasitism. *Canadian Entomologist* **91**: 385-398.
- Hooper, R. E., Tsubaki, Y. & Siva-Jothy, M.T. 1999. Expression of a costly, plastic secondary sexual trait is correlated with age and condition in a damselfly with two male morphs. *Physiological Entomology* **24**: 364-369.

- Järvi, T., Sillén-Tullberg, B. & Wiklund, C. 1981. The cost of being aposematic - an experimental study of predation on larvae of *Papilio machaon* by the Great Tit *Parus major*. *Oikos* **36**: 267-272.
- Joron, M. 2003. Mimicry. In: R. Cardé & V. H. Resh (eds.), *Encyclopedia of Insects*. New York, Academic Press, pp. 714-726.
- Kapan, D. D. 2001. Three-butterfly system provides a field test of Müllerian mimicry. *Nature* **409**: 338-340.
- Klok, C. J. & Chown, S. L. 1999. Assessing the benefits of aggregation: thermal biology and water relations of anomalous Emperor Moth caterpillars. *Functional Ecology* **13**: 417-427.
- Knill, R., & Allen, J. A. 1995. Does polymorphism protect - an experiment with human predators. *Ethology* **99**:127-138.
- Krause, J. & Ruxton, G. D. 2002. *Living In Groups*. Oxford: Oxford University Press.
- Lachmann, M., Szamado, S. & Bergstrom, C. T. 2001. Cost and conflict in animal signals and human language. *Proceedings of the National Academy of Sciences* **98**:13189-13194.
- Lajeunesse, M. J. & Forbes, M. R. 2003. A comparison of structural size and condition in two female morphs of the damselfly *Nehalennia irene* (Hagen) (Zygoptera: Coenagrionidae). *Odonatologica* **3**: 295-301.
- Lam, E. 2004. *Damselflies of the northeast: A guide to the species of eastern Canada and the northeastern United States*. New York, Biodiversity Books.
- Leimar, O., Enquist, M. & Sillén-Tullberg, B. 1986. Evolutionary stability of aposematic coloration and prey unprofitability: a theoretical analysis. *American Naturalist* **128**: 469-490.

- Lindström, L., Alatalo, R. V., Lyytinen, A. & Mappes, J. 2001. Strong antiapostatic selection against novel rare aposematic prey. *Proceedings of the National Academy of Sciences* **98**: 9181-9184.
- Lindström, L., Alatalo, R. V. & Mappes, J. 1997. Imperfect Batesian mimicry - the effects of the frequency and the distastefulness of the model. *Proceedings of the Royal Society of London B* **264**: 149-153.
- Lindström, L., Alatalo, R. V. & Mappes, J. 1999. Reactions of hand-reared and wild-caught predators toward warningly colored, gregarious, and conspicuous prey. *Behavioral Ecology* **10**: 317-322.
- Lindström, L., Alatalo, R. V., Riipi, M. & Vertainen, L. 1999. Can aposematic signal evolve by gradual change? *Nature* **397**: 249-251.
- MacDougall, A. & Dawkins, M. S. 1998. Predator discrimination error and the benefits of Müllerian mimicry. *Animal Behaviour* **55**: 1281-1288.
- Mallet, J. 2001. Mimicry: An interface between psychology and evolution. *Proceedings of the National Academy of Sciences* **98**: 8928-8930.
- Mallet, J. & Barton, N. H. 1989. Strong natural selection in a warning colour hybrid zone. *Evolution* **43**: 421-431.
- Mallet, J. & Gilbert, L. E. 1995. Why are there so many mimicry rings— correlations between habitat, behavior and mimicry in *Heliconius* butterflies. *Biological Journal of the Linnean Society* **55**: 159-180.
- Mallet, J., & Joron, M. 1999. Evolution of diversity in warning color and mimicry: polymorphisms, shifting balance, and speciation. *Annual Review of Ecology and Systematics* **30**: 201-233.
- Mallet, J., & Singer, M. C. 1987. Individual selection, kin selection, and the shifting balance in the evolution of warning colors - the evidence from butterflies. *Biological Journal of the Linnean Society* **32**: 337-350.

- Marden, J. H., & Chai, P. 1991. Aerial predation and butterfly design - how palatability, mimicry, and the need for evasive flight constrain mass allocation. *American Naturalist* **138**: 15-36.
- Maynard-Smith, J. M. & Harper, D. 2003. *Animal Signals* Oxford: Oxford University Press.
- McKee, D., Harvey, I. F., Thompson, D. J. & Sherratt, T. N. 2005. Frequency of female colour morphs in populations of four coenagrionid damselflies (Zygoptera : Coenagrionidae) *Odonatologica* **34**: 37-49.
- Miller, G. A. 1956. The magical number seven, plus or minus two: some limits on our capacity for processing information. *Psychological Review* **63**: 81–97.
- Miller, M. N. & Fincke, O. M. 1999. Cues for mate recognition and the effect of prior experience on mate recognition in *Enallagma* damselflies. *Journal of Insect Behaviour* **12**: 801-814.
- Müller, F. 1878. Über die Vortheile der Mimicry bei Schmetterlingen. *Zool. Anz.* **1**: 54–55.
- Müller, F. 1879. Ituna and Thyridia: a remarkable case of mimicry in butterflies. *Proceedings of the Entomological Society of London* **1879**: xx-xxix.
- Nilsson, M. & Forsman, A. 2003. Evolution of conspicuous colouration, body size and gregariousness: a comparative analysis of lepidopteran larvae. *Evolutionary Ecology* **17**: 51-66.
- Nonacs, P. 1985. Foraging in a dynamic mimicry complex. *American Naturalist* **126**: 165-180.
- Papageorgis, C. 1975. Mimicry in neotropical butterflies. *American Scientist* **63**: 522-532.
- Pasteels, J. M., Gregoire, J. C. & Rowellrahier, M. 1983. The chemical ecology of defense in arthropods. *Annual Review of Entomology* **28**: 263-289.

- Pearce, J. M. 1994. pp. 110-134 In: *Animal Learning and Cognition* (Mackintosh, N. J. ed.) Academic, San Diego.
- Pilecki, C., & O'Donald, P. 1971. The effects of predation on artificial mimetic polymorphisms with perfect and imperfect mimics at varying frequencies. *Evolution* **25**: 365-370.
- Poulton, E. B. 1890. *The Colours of Animals: their Meaning and Use especially considered in the case of Insects*. London, Keegan Paul, Trench, Trübner.
- Riipi, M., Alatalo, R. V., Lindström, L. & Mappes, J. 2001. Multiple benefits of gregariousness cover detectability costs in aposematic aggregations. *Nature* **413**: 512-514.
- Robertson, H. M. 1985. Female dimorphism and mating behaviour in a damselfly, *Ischnura ramburi*: females mimicking males. *Animal Behaviour* **33**: 805–809.
- Roper, T. J. & Wistow, R. 1986. Aposematic coloration and avoidance-learning in chicks. *Quarterly Journal of Experimental Psychology Section B-Comparative and Physiological Psychology* **38**: 141-149.
- Rowe, C. 2002. Sound improves visual discrimination learning in avian predators. *Proceedings of the Royal Society of London B* **269**: 1353-1357.
- Rowe, C. & Guilford, T. 2000. Aposematism: to be red or dead. *Trends in Ecology & Evolution* **15**: 261-262.
- Rowe, C., Lindström, L. & Lyytinen, A. 2004. The importance of pattern similarity between Müllerian mimics in predator avoidance learning. *Proceedings of the Royal Society of London B* **271**: 407– 413.
- Ruxton, G. D. 1998. Sheep in wolves' clothing. *Nature* **394**: 833–834.

- Ruxton, G. D. & Sherratt, T. N. 2006. Aggregation, defence and warning signals: the evolutionary relationship. *Proceedings of the Royal Society of London B* DOI:10.1098/rspb.2006.3570.
- Sbordoni, V., Bullini, L., Scarpelli, G., Forestiero, S. & Rampini, M. 1979. Mimicry in the burnet moth *Zygaena ephialtes*: population studies and evidence of a Batesian-Müllerian situation. *Ecological Entomology* **4**: 83–93.
- Servedio, M. R. 2000. The effects of predator learning, forgetting and recognition errors on the evolution of warning coloration. *Evolution* **54**: 751–763.
- Shepard, R. N., Hovland C. I. & Jenkins, H. M. 1961. Learning and memorization of classifications. *Psychological Monographs* **75**: 1–42.
- Sheppard, P. M. 1959. The evolution of mimicry: a problem in ecology and genetics. *Cold Spring Harbor Symposium in Quantitative Biology* **24**:131–140.
- Sherratt, T. N. 2002. The coevolution of warning signals. *Proceedings of the Royal Society of London B* **269**: 741–746.
- Sherratt, T. N. & Beatty, C.D. 2003. The evolution of warning signals as reliable indicators of prey defense. *American Naturalist* **162**: 377–389.
- Sherratt, T. N. & Forbes, M. R. 2001. Sexual difference in coloration of Coenagrionid damselflies (Odonata): a case of intraspecific aposematism? *Animal Behaviour* **62**: 653–659.
- Sherratt, T. N. & Franks, F. W. 2005. Do unprofitable prey evolve traits that profitable prey find difficult to exploit? *Proceedings of the Royal Society of London B* **272**: 2441–2447.
- Sherratt, T. N., Rashed, A. & Beatty, C.D. 2004. The evolution of locomotory behaviour in profitable and unprofitable simulated prey. *Oecologia* **138**: 143–150.

- Shettleworth, S. J. 1998. *Cognition, Evolution, and Behavior* Oxford University Press, New York.
- Sillén-Tullberg, B. 1985. Higher survival of an aposematic than of a cryptic form of a distasteful bug. *Oecologia* **67**: 411-415.
- Sillén-Tullberg, B. S. 1988. Evolution of gregariousness in aposematic butterfly larvae: a phylogenetic analysis. *Evolution* **42**: 293-305.
- Sillén-Tullberg, B., & Bryant, E. H. 1983. The evolution of aposematic coloration in distasteful prey: An individual selection model. *Evolution* **37**: 993-1000.
- Sillén-Tullberg, B. S. & Leimar, O. 1988. The evolution of gregariousness in distasteful insects as a defense against predators. *American Naturalist* **132**: 723-734.
- Siva-Jothy, M. T. 2000. A mechanistic link between parasite resistance and expression of a sexually selected trait in a damselfly. *Proceedings of the Royal Society of London B* **267**: 2523-2527.
- Speed, M. P. 1993. Müllerian mimicry and the psychology of predation. *Animal Behaviour* **45**: 571-580.
- Speed, M. P. 2000. Warning signals, receiver psychology and predator memory. *Animal Behaviour* **60**: 269-278.
- Speed, M. P. 2001. Can receiver psychology explain the evolution of aposematism? *Animal Behaviour* **61**: 205-216.
- Speed, M., & Ruxton, G. D. 2002. Evolution of suicidal signals. *Nature* **416**: 375-375.
- Srygley, R.B. 2004. The aerodynamic costs of warning signals in palatable mimetic butterflies and their distasteful models. *Proceedings of the Royal Society of London B* **271**: 589-594.
- Srygley, R. B., & Chai, P. 1990. Flight morphology of neotropical butterflies - palatability and distribution of mass to the thorax and abdomen. *Oecologia* **84**: 491-499.

- Srygley, R. B. & Chai, P. 1990. Predation and the elevation of thoracic temperature in brightly colored neotropical butterflies. *American Naturalist* **135**: 766–787.
- Symula, R., Schulte, R. & Summers, K. 2001. Molecular phylogenetic evidence for a mimetic radiation in Peruvian poison frogs supports a Müllerian mimicry hypothesis. *Proceedings of the Royal Society of London B* **268**: 2415–2421.
- Tucker, G. M. & Allen, J. A. 1993. The Behavioral basis of apostatic selection by humans searching for computer-generated cryptic prey. *Animal Behaviour* **46**: 713-719.
- Tinbergen, N. 1984. *Curious Naturalists*. Boston, University of Massachusetts Press.
- Tullberg, B. S. & Hunter, A. F. 1996. Evolution of larval gregariousness in relation to repellent defences and warning coloration in tree-feeding Macrolepidoptera: a phylogenetic analysis based on independent contrasts. *Biological Journal of the Linnean Society* **57**: 253-276.
- Tullberg, B. S., Leimar, O. & Gamberale-Stille, G. 2000. Did aggregation favour the initial evolution of warning coloration? A novel world revisited. *Journal of Animal Behaviour* **59**: 281-287.
- Turner, J. R. G. 1975. A tale of two butterflies. *Natural History* **84**: 28-37.
- Turner, J. R. G. 1984. Mimicry: the palatability spectrum and its consequences. In: R. I. Vane-Wright, & P. R. Ackery, (eds), *The Biology of Butterflies*. Symposium of the Royal Entomological Society of London, Academic Press, pp. 141-161.
- Ubukata, H. 1987. Mating system of the dragonfly *Cordulia aenea amurensis* Selys and a model of mate searching and territorial behaviour in Odonata. In: J. L. Brown, Y. Ito & J. Kiawa (eds), *Animal Societies: theories and facts*. Tokyo, Japan Science Society Press, pp. 213-228.

- Van Gossum, H., Stoks, R. & De Bruyn, L. 2001. Reversible frequency-dependent switches in male mate choice. *Proceedings of the Royal Society of London B* **268**: 83-85.
- Van Gossum, H., Stoks, R., Matthysen, E. Valck, F. & De Bruyn, L. 1999. Male choice for female colour morphs in *Ischnura elegans* (Odonata, Coenagrionidae): testing the hypotheses. *Animal Behaviour* **57**: 1229-1232.
- Viljugrein, H. 1997. The cost of dishonesty. *Proceedings of the Royal Society of London B* **264**: 815-821.
- Wallace, A. R. 1867. *Proceedings of the Entomological Society of London, London, March 4th, lxxx-lxxxi.*
- West-Eberhard, M. J. 1983. Sexual selection, social competition, and speciation. *The Quarterly Review of Biology* **58**: 155-183.
- Whitman, D. W., M. S. Blum, & C. G. Jones. 1985. Chemical defense in *Taeniopoda eques* (Orthoptera: Acrididae): role of the metathroacic secretion. *Annals of the Entomological Society of America* **78**: 451-455.
- Wiklund, C., & T. Järvi. 1982. Survival of distasteful insects after being attacked by naive birds - a reappraisal of the theory of aposematic coloration evolving through individual selection. *Evolution* **36**: 998-1002.
- Yachi, S. & M. Higashi. 1998. The evolution of warning signals. *Nature* **394**: 882-884.
- Zahavi, A. 1993. The fallacy of conventional signalling. *Philosophical Transactions of the Royal Society of London B* **340**: 227-230.
- Zahavi, A., & A. Zahavi. 1997. *The Handicap Principle: a Missing Piece of Darwin's Puzzle*. Oxford, Oxford University Press.