

October 2015

Why are there imperfect Batesian mimics?

Jonathan Warrin

Submitted to the University of Nottingham for the degree of Master of Research in Ecology.

Abstract

Batesian mimicry is when an innocuous species avoids predation by mimicking a defended, harmful species (known as the model). It is expected that natural selection should be towards ever more accurate mimics, but, puzzlingly, mimicry is often imprecise.

One hypothesis that could explain some inaccurate mimicry is that some mimics might be generalist mimics of more than one model species, rather than specialist mimics. Using an online game with humans acting as predators it is shown that in some circumstances intermediate mimics of two models are more protected from predation than specialist mimics of either model.

In order to understand why imperfect mimicry persists, we need to know which aspects of the mimic's appearance are salient to their predators. Humans acting as predators in an online quiz are shown to gain the most information about whether an insect is a defended wasp or an undefended wasp-mimicking hoverfly from the shape of the insect's abdomen. It is also shown that humans appear to pay more attention to the abdomen's colour than its shape.

Acknowledgments

Thank you to my supervisor, Tom Reader, for his support and patience. I'm grateful to Chris Taylor for allowing me to collaborate with him on the work presented in Ch. 3. Chris collected the images, created the quiz and performed experiments with the quiz prior to the one that I was involved with. For the experiment presented in Ch. 3 I created the treatment images and publicised the quiz. Many thanks also to Rohanna Dow for devising and creating the game used in the experiments in Ch. 2. The work presented in Ch.2 is built on experiments she carried out using the game.

Contents

Abstract	2
Acknowledgements	4
1 Introduction: Why are there imperfect Batesian mimics?	
1.1 Predator centred hypotheses	..
1.2 Mimic centred hypotheses	..
1.3 Model centred hypotheses	
1.4 Behavioural studies of mimicry using humans	
2 Testing the multiple models hypothesis using an online game	
2.1 Introduction	
2.2 Methods	
2.3 Results	
2.4 Discussion	
3 Testing predator response to mimetic features	
3.1 Introduction	
3.2 Methods	
3.3 Results	
3.4 Discussion	

1 An introduction to imperfect Batesian mimicry

A Batesian mimic is an unprotected species which gains protection from predators by resembling a defended or dangerous species (Bates, 1862). Batesian mimicry provides an intuitive and compelling example of evolution. Examples of Batesian mimics include hoverflies (Diptera: Syrphidae) which mimic the visual pattern of hymenopterans (Gilbert, 2005) and myrmecomorphic spiders, which mimic the appearance and behaviour of ants (Pekar and Jarab, 2011). Studying mimicry can give excellent insight into the ecological, stochastic and genetic processes that give rise to the phenotypes we see in the natural world (avoiding attack book).

Though mimicry has been widely studied, there remains an unexplained paradox: while some mimics are strikingly accurate, many seem to bear a relatively imprecise resemblance to their supposed models (Cuthill, 2014, Edmunds, 2000). This is puzzling because it would seem that more precise mimicry should always be selected for. How is it that imperfect mimicry persists?

A mimicry complex comprises the mimic species, the aposematic “model” species and the predator species. The hypotheses relating to the persistence of imperfect Batesian mimics can be grouped according to which part of the mimicry complex they focus on (see Table 1). It is worth noting that the hypotheses presented are not mutually exclusive, and it is likely that more than one explanation will play a part in any case of imperfect mimicry.

Table 1 – Descriptions of the main hypotheses proposed to explain imperfect Batesian mimicry. The hypotheses are grouped by the aspect of the mimicry complex upon which they are focussed.

Focus	Hypothesis
Predator	<p>Relaxed selection – Predators do not discern increases in mimetic accuracy above a certain threshold, or they only focus on particular features.</p> <p>Eye of the beholder – mimics appear more accurate to natural predators than they do to us, because of sensory or cognitive differences.</p> <p>Multiple predators – specialist predators of the model species place a selective pressure against similarity to the model.</p>
Mimic	<p>Alternative defences – mimetic accuracy may be traded off against crypsis or speed of escape.</p> <p>Kin selection – increasing accuracy may lead to a higher overall attack rate from the predator on the mimic species, leading to reduced inclusive fitness.</p> <p>Thermoregulation – the colouring/morphology required to mimic accurately may affect the mimics’ capacity to regulate their body temperature.</p> <p>Character displacement – being more similar to the model may increase competition for resources or lead to deleterious reproductive interactions.</p> <p>Evolutionary constraints – there may be genetic or developmental barriers to evolving closer mimicry.</p>
Model	<p>Multiple models – imperfect mimics may be generalist mimics of more than one model species rather than specialist mimics of one.</p>

Predator-centred hypotheses

Signal detection theory (SDT) is often invoked in hypotheses relating to imperfect mimicry. SDT recognises that there is uncertainty in decision making based on received stimuli (Green and Swets, 1966) and, in the case of mimicry, the fact that there is uncertainty means that the likelihood of a predator attacking a prey item will be affected by factors such as the incentive to attack the mimic and the aversiveness of the model. If there are high costs to attacking the model then the predator may be more cautious, and the mimic could be well protected by a not very precise resemblance to the model.

Relaxed selection

Mathematical models, based on SDT, of a predator's incentive to attack at different levels of signal accuracy, predict a non-linear relationship between mimetic accuracy and predator attack rate, and also predict a threshold value above which prey are as protected as if they were perfect mimics (Sherratt, 2002). If a threshold exists then above that level of accuracy there is no selective pressure for accuracy to improve. This can be visualised as a zone of protection in phenotypic space around the phenotype of the model, in which prey are protected against protection.

There are predictions from SDT that we can use to test the validity of the relaxed selection hypothesis. The more aversive and common (relative to the mimic) a model is, the larger the zone of phenotypic protection will be (i.e. predators will avoid less accurate mimics of more aversive models) (Sherratt, 2002). In support of the theory, mimetic fidelity was shown to be under higher selection when the cost of mistakenly eating a model was lower, in an experiment with chickens *Gallus gallus domesticus* sampling artificial prey (Duncan and Sheppard, 1965). Another prediction of SDT is that more profitable mimics will be under greater selection for mimetic accuracy. Penney et al. (2012) found a positive relationship between body size and fidelity, and suggest that this shows reduced selection pressure on smaller, and (it was assumed) less profitable, prey. This, however, does not take into account the fact that search and handling times may mean that the largest prey are not necessarily the most profitable (Pyke et al., 1977). In cases of relaxed selection we would expect there to be high phenotypic variation among mimics because the absence of directional selection would allow the build-up of phenotypic variability (Holloway et al., 2002). Taylor et al. (In Review) found no correlation between accuracy and phenotypic variation in a survey of syrphids, indicating that mimetic inaccuracy in hoverflies is not accounted for by relaxed selection. Holloway et al. (2002) found evidence that this is true for the patterns of some hoverfly species.

Some traits may have higher salience (noticeability) to predators than others. Traits which have lower salience will be under less intense selection to be similar to the model. Variation in feature salience may be due to the sensory limitations (e.g. Cuthill and Bennett 1993) or cognitive biases of the predator (Chittka and Osorio, 2007). It is possible that predators categorise prey based on the presence of certain, because they do not have the cognitive capacity to learn all of the phenotypes of the prey they encounter (Chittka and Osorio, 2007). If this, rather than the signal detection model of continuous variation in mimetic accuracy, were true then it would mean that prey could gain maximum protection just by presenting the feature on which the predator categorises prey, and could be otherwise inaccurate. There is evidence, however, that birds attack mimics of intermediate accuracy at a rate intermediate between the rate they attack perfect mimics and non-mimics, supporting the SDT model (Dittrich et al., 1993).

For further exploration of trait salience and categorisation see Chapter 3.

Eye of the beholder

The Eye of the Beholder hypothesis suggests that predator perception is different enough from our own that the differences we perceive between mimics and their models may be imperceptible to predators (Cuthill and Bennet 1993). A good example of this is the scarlet king snake (*Lampropeltis elapsoides*) which has colour bands arranged in a different order to its model, the coral snake (*Micrurus fulvius*). The order of the colours was shown to not be used by the snake's natural predators to distinguish it from its model (Kikuchi and Pfennig, 2010) though the relative proportions of red and black rings on the snake was selected for (Pfennig et al., 2007, Pfennig et al., 2001).

Other than in *L. elapsoides* there is little evidence of differences in the way humans perceive mimetic signals from natural predators. Dittrich et al. (1993) showed that pigeons (*Columba livia*) rank hoverfly mimics of wasps in an order similar to that of the experimenters, apart from two species, which pigeons classified as accurate whilst people classified them as inaccurate. Penney et al (year) built on this research and asked human volunteers to rank the same images in terms of mimetic accuracy. They also did morphometric analysis of the pictures. The human rankings were correlated with both the morphometric rankings and the bird rankings, suggesting that human judgement of mimetic accuracy is not that dissimilar to the judgement of natural predators.

Multiple predators

Mimics are likely to be under threat of predation from more than one species, and some model species may have specialist predators. If the impact of mimetic accuracy is different on the attack rates of specialist model predators and generalist predators (e.g. if specialist predators only attack highly accurate mimics, but generalists avoid mimics across a wider range of accuracy) then there could be an adaptive peak away from perfect mimicry (Pekar et al., 2011).

(Pekar and Jarab, 2011) studied the quality of mimicry in myrmecomorphic (ant-mimicking) spiders and found mimicry to be less strong in microhabitats dominated by predators which specialise in attacking the model species. Specialist ant predators have been shown to attack ant mimics as often as ants when presented with various static prey (Nelson et al., 2006).

Mimic-centred hypotheses

Most mimic-centred hypotheses involve alternative selective pressures on the mimic species which create trade-offs, leading to an adaptive peak away from perfect mimicry. The selection pressures discussed here are for thermoregulation in insects, alternative defences against predation, kin selection and character displacement (selection against extreme similarity to the model because of the competition for resources and deleterious reproductive interactions that this can cause). The other mimic-based hypothesis discussed here is that there are genetic or developmental constraints preventing the mimic from replicating the aposematic signal more accurately.

Alternative defences

For some mimics, accuracy of mimicry may be traded off with another defence, such as crypsis or speed of escape. Aposematic signals are generally highly conspicuous in order to be effective at deterring predators (Sherratt and Beatty, 2003). There may be a trade-off for the mimic between the benefits of the more accurate mimetic signal and the higher risk of being noticed by a predator, some of which may not avoid the aposematic signal. Changes in body shape or behaviour in order to effect mimicry may lower the mimic's speed of escape (Srygley, 2004). The hoverfly *Syrphus ribesii* has been shown to mimic the slower flight of the common wasp *Vespula vulgaris* (Golding et al.,

2005). A study looking at predation of myrmecomorphic (ant-mimicking) spiders by ant-specific predators found that less accurate mimics had the fastest escape time (Pekar et al 2011). There could also be a trade-off between the allocation of resources to improve mimicry and to improve agility, for example through the creation of molecules required for fast muscle movement. But I am not aware of any studies showing any such physiological differences between accurate and inaccurate mimics.

Kin selection

Johnstone (2002) found through mathematical modelling that there can be inclusive fitness costs of close mimicry. In a situation where models are only moderately costly to attack, or mimics are common relative to models, the predator's incentive to attack is high and signal detection theory predicts that with increasing mimetic accuracy the attack rate of the predators will increase. In a local population that includes close kin there may be an inclusive fitness disadvantage to more accurate mimicry, because of the increased attack level of the predator. In other words the less accurate mimics pay the cost for the survival of their more accurate relatives.

There is a lack of empirical evidence for the kin selection hypothesis. One testable prediction from the hypothesis is that a high abundance of mimics favours imperfect mimicry. There was found to be no evidence for a negative relationship between mimetic fidelity and abundance in an analysis by Penney et al. (2012). Kin selection does not offer a convincing explanation for the existence of imperfect mimicry in hoverflies as their high level of mobility means that they are not likely to be closely related to others in their local environment (Gilbert 2005).

Thermoregulation

The proportion of dark colouration on the bodies of many species is thought to be important for thermoregulation (Clusella Trullas et al., 2007). The effect of pattern colour on thermoregulation in hoverflies has been well studied as they are thought to maintain their body temperature in part by basking in the sun (Morgan and Heinrich, 1987). Taylor et al. (in review) found evidence for a trade-off between the proportion of a hoverfly's abdomen that was dark in colour (darker colouring absorbs solar radiation more effectively (Willmer and Unwin, 1981)) and the mimetic accuracy of the hoverfly. Hoverflies surveyed were almost all found to have a higher proportion of black in their colour pattern than wasps. It is likely that thermoregulation does not constrain the colour patterns of wasp species in the same way, as it is thought they do not rely on basking to maintain their temperature (Heinrich, 1984). The proportion of colour pattern that is dark has been shown to increase in hoverflies in the winter months (Ottenheim and Kuijt, 1998). Dow (2010) showed that there were geographic variations in the amount of dark colouring on the bodies of European hoverfly species, with species found in colder climates having darker colouring. Another species thought to have thermoregulatory colour constraints is the moth *Parasemia plantagenis*: darker individuals were shown to be at greater risk of predation but they were able to warm their bodies more quickly (Hegna et al., 2013).

Character displacement

Organisms compete for resources as well as avoiding predation. Becoming more similar to a model species may lead to greater competition for resources, due to greater convergence of the mimic's ecological niche with the model's and deleterious reproductive interactions (Pfennig and Kikuchi, 2012). Müllerian mimics have been shown to engage in deleterious reproductive interactions with their co-mimics when the aposematic signal is also involved in sexual signalling. For example male individuals of the species *Heliconius erato* and *H. Melpomene*, which are Müllerian co-mimics, have been shown to spend a considerable proportion of their time attempting to court females of the co-

mimic species (Estrada and Jiggins, 2008). But is likely to be less of a problem where Batesian mimics are distantly related to their models (e.g. hoverflies and wasps).

Evolutionary constraints

Imperfect mimics may be unable to evolve greater mimetic accuracy because of developmental or genetic constraints. Studies of the proximate mechanisms used to create a signal can show us whether the imperfection of the signal is because the mimic is creating an analogous physiological or anatomical trait, with associated genetic architecture, rather than a homologous trait, and therefore the trait cannot physically be identical to that of the model. Studies of proximate mechanisms have shown that some species do use homologous methods to create the signal (Kikuchi et al., 2014, Martin et al., 2012). In some instances species have been shown to use entirely different mechanisms to produce the same signal. For example – though it does not involve Batesian mimicry – two species of unrewarding orchid (*Ophrys sphegodes* and *O. exaltata*) produce a chemical compound that mimics a pheromone emitted by female bees in order to attract pollinators (Schluter et al., 2011). It is likely that almost any constraint can be overcome if selection is strong and sustained over a long period of time (Smith et al., 1985).

Model-centred hypothesis

Multiple models

In cases where more than one model species exists in the geographic range of a mimic, a Batesian mimic may mimic more than one model, either through polymorphism (where individuals of the species mimic one model or the other models (Joron and Mallet, 1998)) or through jack-of-all-trades mimicry, where mimics partially resemble more than one model species at once.

Darst and Cummings (2006) found that the Batesian mimetic frog *Allobates* sp's predominant morph in an area depended on the model species (one of two species of the genus *Epipedobates*) present, and that when both species were present the mimic resembled one but not both models. Edmunds and Reader (2014) showed that the occurrence of one colour morph of the polymorphic hoverfly species *Volucella bombylans* was spatially correlated with the distribution of a putative bumblebee model. The authors conclude that the frequency of the different morphs of *V. bombylans* is affected by the composition of the bumblebee community.

The jack-of-all-trades multiple models hypothesis (Edmunds, 2000) is that imperfect mimics may bear a general resemblance to multiple allopatric models, rather than a specific resemblance to one species. This hypothesis is plausible in cases where models are relatively similar to each other. The prevalence of Müllerian mimicry, where species share a very similar aposematic signal, means that there are many opportunities for jack-of-all-trades mimicry to evolve. Sherratt (2002) showed in a theoretical analysis that in a system with two sympatric model species where the phenotypes of the models overlap, mimics intermediate between the two models can have fitness equivalent to that of perfect mimics of one model. But there is little empirical evidence for multiple models mimicry; there have been studies of predator behaviour towards imperfect mimics (Dittrich 1993), but more are needed in order to gain a better understanding of the possible circumstances in which multiple models mimicry will evolve.

See Chapter 2 for further exploration of the multiple models hypothesis.

Using humans in a study of predator behaviour

The experiments presented in Chapters 2 and 3 use humans playing computer-based games in order to test hypotheses about predator perception and the evolution of mimicry of multiple models. Experiments using natural predators to test hypotheses relating to long term evolutionary processes are logistically difficult. Most studies try to understand traits by looking at selection pressures observable in the short term. Mathematical and computational models have also often been used to study predator behaviour (see Ruxton et al. (2004) for a review of attempts to model Batesian mimicry systems). These models, however, must abide by rules (i.e. the structure and parameters of the model) that are likely to oversimplify and contain bias. Humans demonstrate similar behavioural and physiological constraints to other vertebrates, and respond to profitable and unprofitable stimuli similarly to predators foraging for food (need to find reference).

Using people as proxy predators is a common approach; a number of recent studies have used humans foraging in artificial environments to gain insights into defences used by prey. For example, Lichter-Marck et al. (2015) used human predators to study the effectiveness of camouflage and warning signals, and Sherratt et al. (2007) used human predators foraging on computer generated prey in order to investigate the evolution of crypsis.

Behavioural studies of predators can give us an insight into the selective pressures imposed on mimics, and are a means of testing predictions from theoretical models. Chapter 2 is a test of the multiple models hypothesis; an online game simulating the evolution of prey, some of which are defended models, under selection by human predators is used to test whether intermediate mimics are at a selective advantage to specialist mimics. Chapter 3 is a test of the hypothesis that predators pay more attention to some features of a mimic, and therefore less salient features are under relaxed selection. An online quiz is used to test the ability of humans to classify images of hoverflies and wasps as models or mimics, and how their ability to do this changes when presented with aspects of the mimetic signal (such as colour) presented in isolation.

2 Using a game to test the multiple models hypothesis

2.1 Introduction

The multiple models hypothesis is that imperfect mimicry of more than one aposematic model can be more adaptive than close mimicry of one model (Edmunds, 2000, Sherratt, 2002). Edmunds (2000) showed through mathematical modelling that intermediate mimics of multiple models can be as protected as specialist mimics, and have greater geographical ranges. Further modelling based on signal detection theory (SDT) showed that intermediate mimics can be at a selective advantage to specialist mimics if the mimic exists in an environment with multiple sympatric models that bear a resemblance to each other (Sherratt 2002). Species with similar aposematic signals are common due to the prevalence of Müllerian mimicry (where an aposematic signal is generalised over multiple defended species).

SDT, when applied to the situation of Batesian mimicry, is the idea that predators are imperfect at determining whether a prey item is a mimic or a model, and that the decision to attack is therefore based on the perceived probability of the prey being a mimic and the relative costs and benefits of attacking a model or a mimic (see Ch. 1). If the relationship between quality of the mimetic signal and likelihood of attack is non-linear, then mimics may receive significant protection by resembling the model in an imperfect way. If this is the case then mimicking more than one species relatively accurately may confer more protection against predators than mimicking one species perfectly.

It is probable that the phenotype of the egg of the Horsfield's bronze-cuckoo (*Chalcites basalis*) presents an example of intermediate mimicry (Feeney *et al.*, 2014) though clearly this is not an example of Batesian mimicry. The Horsfield's bronze-cuckoo is a brood parasite of multiple species, and the colour and luminance of the cuckoo's eggs were shown to be intermediate between those of their host species. Brood parasites of multiple species are more commonly divided into host-specific races, with eggs that mimic one host specifically. The authors note that the diversity of morphology of the eggs of the different host species of this species of cuckoo was low, and suggest that this is a condition for the evolution of multiple models mimicry. Langmore *et al.* (2011) found that the nestlings of the Horsfield's bronze-cuckoo also appear to be intermediate mimics of the appearance of the host nestlings of multiple species.

A study analysing the colour patterns of hoverflies and hymenopterans, using a tool ("distance transform") that gives an objective comparison of images by comparing them pixel by pixel, found that some hoverfly colour patterns mapped out as intermediate between those of model species such as *Apis mellifera* and *Vespula spp.* (Taylor *et al.*, In Review). If there were a trade-off in the similarity to the different models a species could mimic we would expect to see a negative correlation between similarity to those models. But Taylor *et al.* (In Review) did not find evidence for a trade-off. The mimic species observed could therefore increase their similarity to another model without reducing their similarity to the model to which they are phenotypically closest.

The evidence from Taylor *et al.* (In Review) is in contradiction to previous findings by Penney (*et al.*), which suggested that no hoverfly phenotype mapped out as intermediate between two model species. Penney *et al.* (2012) analysed morphological traits such as antennal length and average RGB colour values of hoverflies and hymenopterans and found that the hoverfly and hymenopteran species were distinct in these characteristics. The authors concluded that no hoverfly had an intermediate phenotype between multiple model species, however it is possible that there are other, more salient features which were overlooked in the study that may be intermediate.

Mimics are often much more widespread in their distributions than their models; this could indicate generalist mimicry, as it is likely that generalist mimics would have a larger potential geographic ranges than specialist mimics (Edmunds, 2000). Edmunds (1978) found that myrmecomorphic (ant-mimicking) spiders which appeared to be generalist mimics of multiple models had a wider geographic range than myrmecomorphs which were specialist mimic of one ant species.

More evidence is needed in order to establish the extent of multiple models mimicry in nature. Studies of the effects of selection pressures in natural environments are logistically difficult to achieve, and modelling approaches adopted typically face constraints. These challenges can be circumvented by using simulated evolution. This has been achieved before, either using computer generated environments (Sherratt and Beatty, 2003, Beatty et al., 2004). Studies have also used human predators and simulated evolution to test the evolution of crypsis (Sherratt et al., 2007) and the predictions of SDT with regard to Batesian mimicry (McGuire et al., 2006).

In this study, the evolution of computer generated prey, some of which were aversive models, selected amongst by human players acting as predators, was studied. The aims were to find whether intermediate mimics evolve, and whether the likelihood that intermediate mimicry evolves is affected by the similarity of the model species to each other.

2.2 Methods

Artificial Prey

The artificial prey were squares (Figure 1) 13mm² in size on a 17" screen. The prey were identical to each other apart from their colour (the feature players used to decide whether they were defended or not). Colour can be defined in red-green-blue (RGB) colour space so it is an easy phenotype to alter, and it is possible to precisely measure the difference between colours as the distance between two points in RGB colour space. Having an abstract object to represent prey (rather than something that looked more like a typical prey item such as an insect) meant that players would have fewer preconceptions and possible biases about the prey (Alatalo and Mappes, 1996, Franks and Noble, 2004).

Phenotypes for the two model species and a random phenotype for each mimic were generated at the start of each game. A mimic's phenotype could fall anywhere in the colour space, as could the phenotype of the first model species. The second model phenotype had to be a certain distance in colour space away from the first. The distance between the model phenotypes was randomly selected at the start of the game to be either 70, 140, 210 or 280 RGB colour units.



Figure 2.2.1. Examples of how prey would appear in the game.

Prey differed in two phenotypic dimensions allowing greater variability in phenotype than is possible in a one-dimensional space (as in Franks and Noble, 2004, McGuire et al., 2006)) Having variation in two dimensions also makes the game more challenging to players, and it enables the results to be examined more easily in relation to mathematical models where prey which differ in two-dimensional phenotypic space are considered (Sherratt, 2002).

The Game

Players were shown the colours of the two model species at the start of the game and told to not attack prey of these colours. Players told to memorise the colours, and were not shown them again once the game had started. In each game there were 8 generations of prey for the player to progress through. A player automatically progressed to the next generation once they had attacked 7 prey. At the start of a new generation, the environment was repopulated with mimics and models. All models from the previous generation reappeared unchanged in a new random position on the screen, as did non-models that had not been attacked. Attacked non-models were given a new phenotype based on that of a randomly selected non-model that hadn't been attacked, but with a small mutation, and appeared in a random position on the screen. This created the conditions required for selection to cause evolution of the phenotypes to occur; mimics bearing a chance likeness to either model species might be mistaken for a model and avoided, potentially creating the proliferation of similar non-model phenotypes.

The phenotype of each model type was constant throughout the game as it is usually assumed that the phenotypes of aposematic species are relatively static, because there is a high fitness cost to deviating from the aposematic phenotype and presenting a signal likely to be less aversive to

predators (Nur, 1970). Players foraging over several generations of prey is representative of the common scenario of vertebrate predators foraging on invertebrate prey.

Generating phenotypes

The phenotype of each prey was set by two values (ϕ and θ) which corresponded to longitude and latitude on a sphere in three dimensional colour space. A spherical colour space was chosen so that there were no edges in phenotypic space that would present boundaries to evolution of the phenotypes. Random points on the surface of the colour sphere were selected using the following equation (Marsaglia, 1972):

$$\theta = 2\pi u \quad \text{Eqn. 1}$$

$$\phi = \cos^{-1}(2v - 1) \quad \text{Eqn. 2}$$

Where theta is the latitude, phi is the longitude, and u and v are uniformly distributed random numbers between 0 and 1.

To set a phenotype for the second model species, the Haversine equation was used to find a point a set distance (d) from the phenotype of the first model (source):

$$\text{hav} \left(\frac{d}{R} \right) = \text{hav}(\phi_2 - \phi_1) + \cos(\phi_1) \cos(\phi_2) \text{hav}(\Delta\lambda), \quad \text{Eqn. 3}$$

$$\text{where: } \text{hav}(\theta) = \sin^2 \left(\frac{\theta}{2} \right) = 1 - \cos \left(\frac{\theta}{2} \right) \quad \text{Eqn. 4}$$

Definitions: d is the distance between two points, R is the radius of the sphere, ϕ_1 is the latitude of point 1, ϕ_2 is the latitude of point 2 and $\Delta\lambda$ is the longitudinal separation between points 1 and 2.

When a mutation occurred, a new phenotype was generated by travelling a small distance, randomly drawn from a normal distribution with mean zero, along a randomly generated bearing between 0 and 2π radians on the surface of the sphere, such that the overall mean mutation size was 0 distance from the original phenotype.

The Game Environment

The environment in which prey appeared was a white area with an approximate height of 10" and width of 7" (the exact size was dependant on the size of the monitor used) surrounded by black background (Figure 2). All elements of the playing environment were greyscale so as not to bias the prey colour choice. Prey distribution in the environment was random, but they could not overlap. The random distribution of the prey made improved the gameplay, as it made it feel more like a search for the prey. Prey were covered by a white square which players clicked to reveal the prey beneath (Figure 2). If the player uncovered a new prey item without eating one they had previously uncovered, then the previous prey item was disabled and covered with a grey square. This meant that prey were viewed sequentially. Players gained points by eating non-models and lost points by eating models. The player's score was displayed on the screen along with a "kill target" telling them how many prey they had left to eat and a level counter telling them which level they were on.

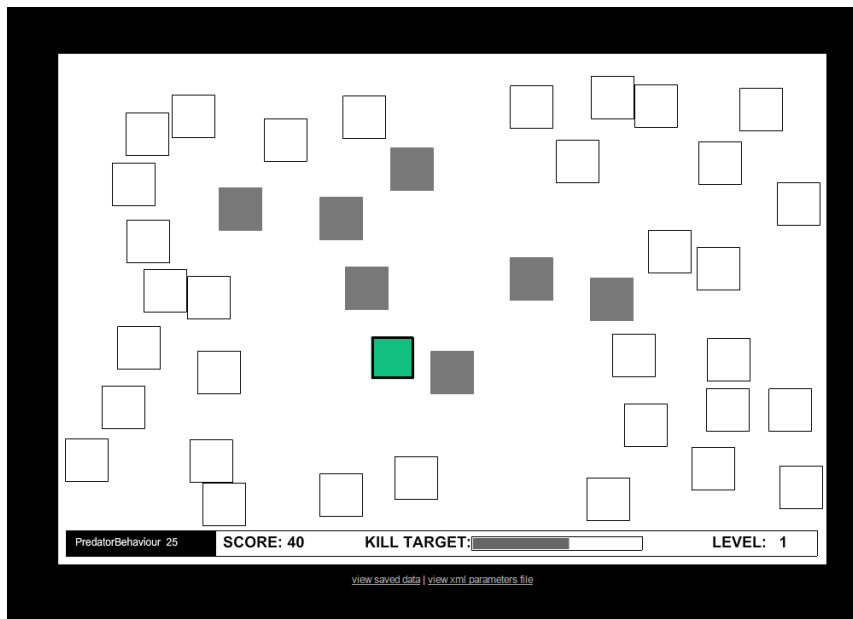


Figure 2.2.2. The game environment. The level number, “kill target” bar and player’s score are displayed at the bottom of the white space. White squares are prey that have not been uncovered, grey squares are disabled prey, and the green square has been uncovered.

Game details

Having the game on the internet meant that a large number of people could participate, but lowered the control of the experimental environment; it was not possible to control the distance at which people sat from their screen, or the resolution of the screens they played the game on. The game could not be played on mobile devices.

Before playing the game participants viewed an instructions screen (Figure 3). Players were told how to play, but were not give any information about the purpose of the research in order to avoid creating any potential bias. Players had the option to view a short demonstration video, and whether they did so or not was recorded. Players had to select: their sex, whether they had played the game before, and whether the sound on their computer was turned on.

At the end of the game, data on the phenotypes of the prey in each level were recorded in a text file. Also recorded was whether each prey was eaten or survived in a level. The IP address of each player was recorded to check whether a person had played the game before. The game was written using Actionscript in Adobe® Flash® CS4 Professional. It was mounted on a website hosted on the University of Nottingham servers and accessed through the url ecology.nottingham.ac.uk/game. Parameters in the game were controlled through a PHP (Hypertext Preprocessor) file.

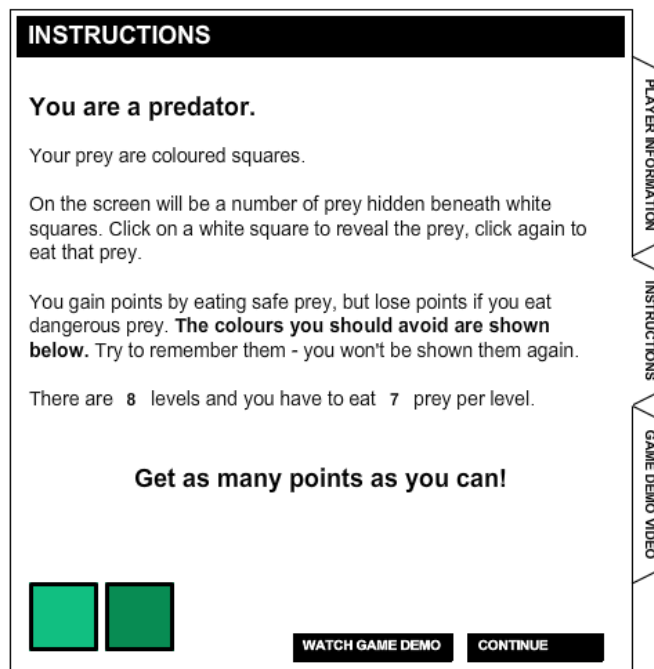


Figure 2.2.3. The instructions screen visible to players before the game commenced.

Players progressed through eight generations of prey and had to eat seven prey items per level. The mutation rate was 20 units, roughly 10% of the average distance between non-model prey in level 1. The number of points gained for eating a non-model was 20, and the number lost for eating a mimic was 5. The number of points was decided upon based on trial games in which the effect of player motivation of different points values was tested. At the start of each level there were 10 of each model species present, and 25 non-models. The width of the colour sphere was 255 units.

Statistical analysis

In order to assess the level of intermediate mimicry that evolved, the density of non-models with phenotypes within phenotypic zones on either side of the two models was calculated (see Figure 2.2.1). The zones were circles 35 RGB units in diameter with their centres in a line that passed through the positions of models A and B. For the treatment with the smallest distance between models (70 RGB units) the two zones on between the models (the “inner” zones) overlapped completely; this was accounted for in calculations of the density of the non-models.

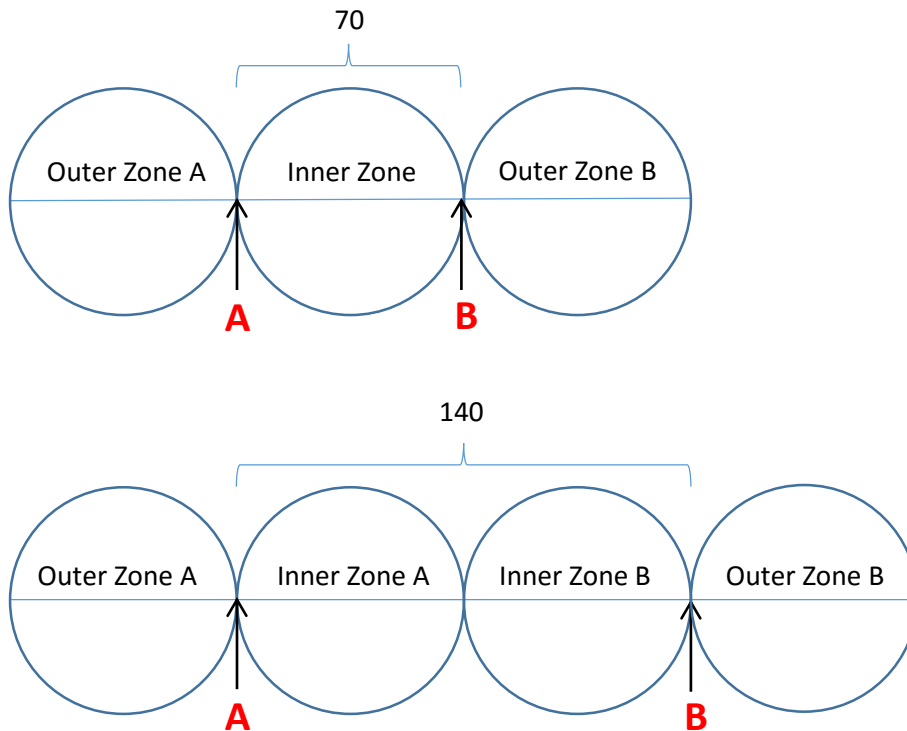


Figure 2.2.4. The phenotypic zones that were used to distinguish non-models in phenotypic space between the models (the inner zones) from non-models that were close to the phenotype of the mimic but not intermediate (in the outer zones). The 70 and 140 units treatments are shown; the positions of models A and B are indicated. The distance between the models is indicated. As the diameter of the zones was 35 RGB colour units, the inner zones for the two models 70 units apart overlap completely.

Data was analysed in R version 3.1.2 (R Core Team, 2014). Responses from participants who had already attempted the quiz were discarded. The response of each individual question was modelled using a binomial generalised linear mixed model. Fixed factors included in the full model were: age of the participant, sex of the participant, species of the specimen in the image, and treatment. Interactions between all factors other than species were also tested; species was not tested because the model was too complex to be analysed with all factors and interactions included. The participant's ID was included as a random factor. Backwards model selection using likelihood ratio tests was used to find the best model (Zur et al 2009).

The data was analysed in R version 3.2.1 (R Core Team, 2014). Game data was not used if the player had previously played the game. Binomial generalised linear mixed models (GLMMs) were used to analyse variables for each individual non-model prey item in each generation. These variables were: whether the prey had been attacked during the generation, the distance of the prey from model A, the distance of the prey from model B, and whether the prey was in the inner or outer zone (see prey outside these zones were excluded from this analysis; figure 2.2.4). Backwards model selection using likelihood ratio tests was used to find the best model (Zur et al 2009). Fixed factors included the full models were generation, treatment and sex. and their two-way and three-way interactions. The phenotypic distance to the nearest model species was also included in the model of whether the prey was attacked. The interactions of sex with the other factors was not included in the attack model because with them the model became too complicated to analyse. Game ID was included as a random factor in all models.

2.3 Results

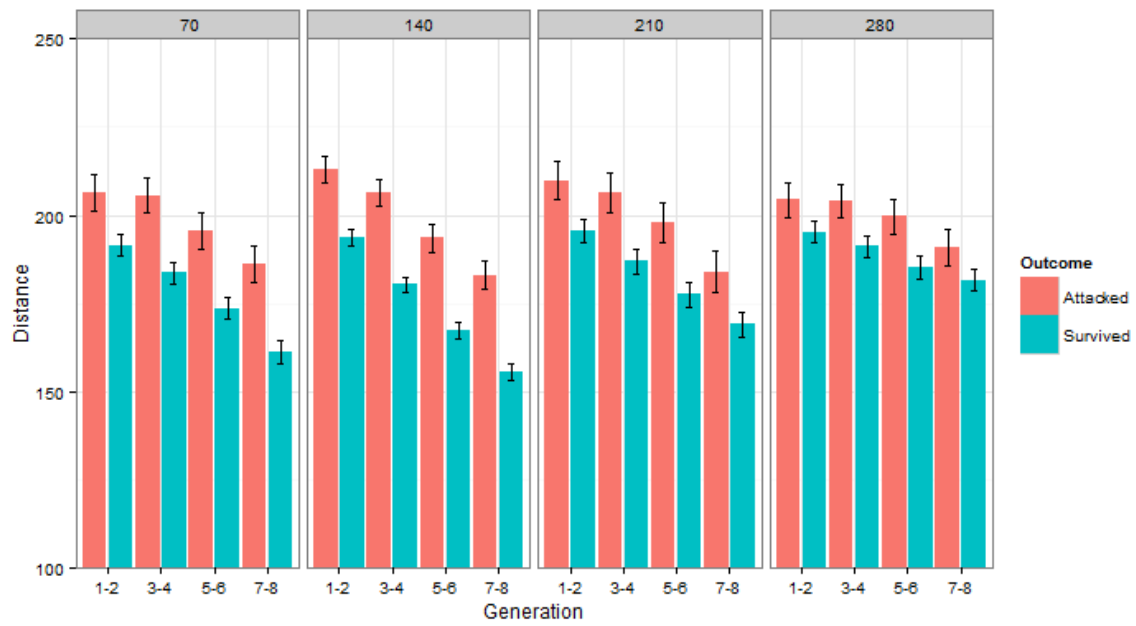
A total of 421 people played the game, with the 70, 140, 210 and 280 treatment distances played by 88, 164, 75 and 94 people respectively. More people played the 140 treatment because of an error in the code which should have randomly selected the treatment type for each game.

Non-models which were more similar to the models were less likely to be attacked by the player, but not for all treatments (table 1). This effect shows that the game is working effectively as players are avoiding non-models that appear similar to models. When the models were less similar to each other the difference between attacked non-models and surviving non-models, in terms of their distance to the nearest model, was smaller (Figure 1). This difference between treatments indicates that there was less strong selection for mimicry when models were less similar. The sex of the player also had an effect on the likelihood that a non-model was attacked (table 1). On average female players attacked more non-model prey items than male players, in every generation (Figure 2).

Table 1. Likelihood ratio tests for GLMMs predicting the likelihood a non-model prey item was attacked. Terms: DModel is the distance of the prey to the nearest model in RGB colour units, Generation is the level (1-8) within the game, Sex is the sex of the player and Treatment is the distance between the two models (70, 140, 210 or 280 RGB colour units).

Terms	Chi	df	p
DModel	620	1	< 0.05
Generation	11.73	1	< 0.05
Sex	8.86	1	< 0.05
Treatment	3.80	3	0.28
DModel:Treatment	83.9	3	< 0.05
DModel:Generation	1.86	1	0.17
Generation:Treatment	1.71	3	0.63
Generation:DModel:Treatment	7.55	3	0.06

1a)



1b)

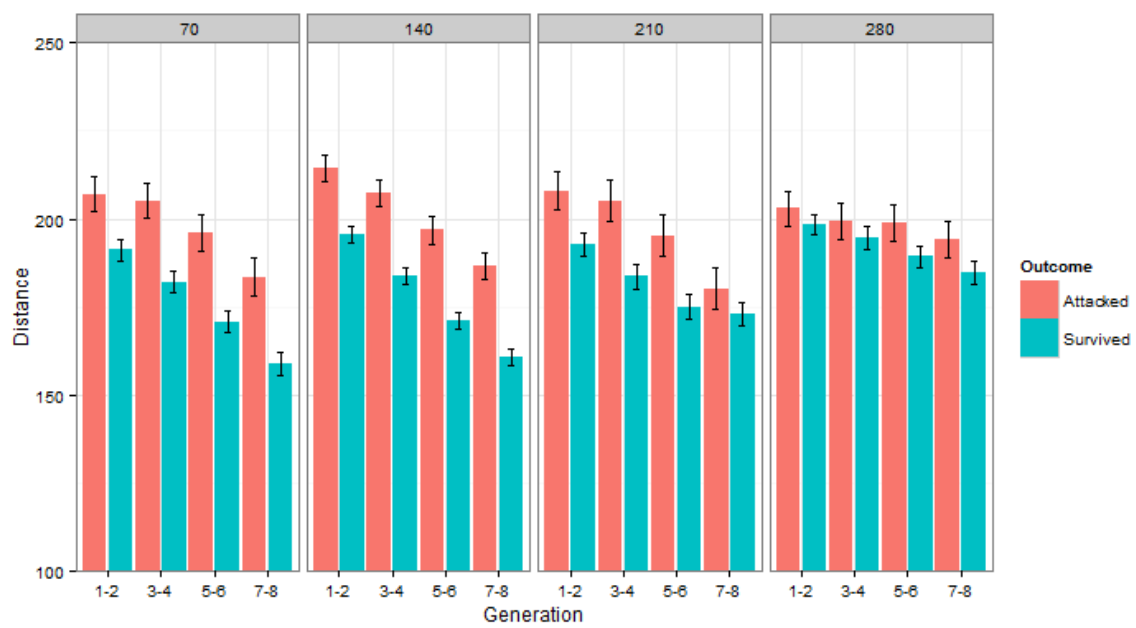


Figure 1. Mean phenotypic distance from model A (a) and model B (b) of attacked and surviving non-models.

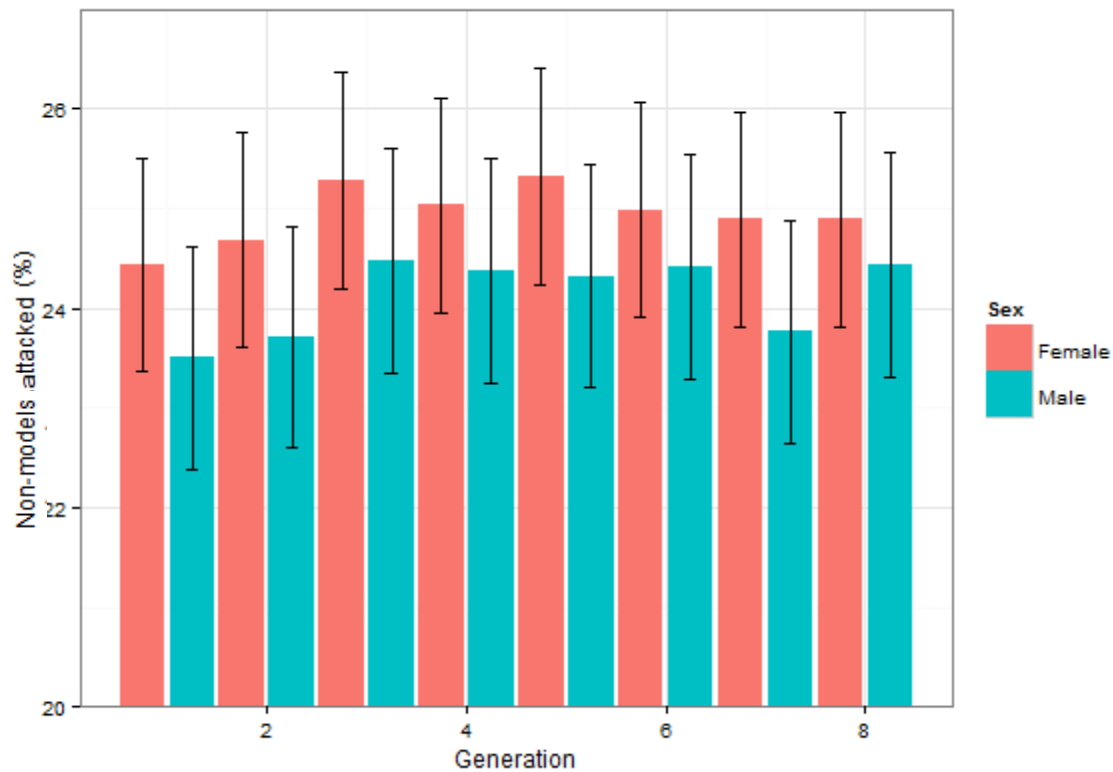


Figure 2. Mean percentage of non-model prey items attacked by female and male players in each generation, with binomial error bars.

The distance in colour space of non-model prey items from models A and B reduced with generation (Table 2). After eight generations they had not reached an equilibrium distance (Figure 3). Treatment type affected the distance of a non-model from model B (Table 2a); the distance reduced more quickly when models were closer together (Figure 3). But treatment type did not affect distance from model A (Table 2b). Importantly, there was an interaction between the effects of generation and treatment on distance of a non-model to both types of model (Table 2): distance to the model changed as the generations progressed for some treatments, but not for all. There was also a three-way interaction between generation, treatment and sex: the observed difference between treatments across generations was not present for male players.

Figure 4 is an illustration of the colours of non-models at the end of the game, and the distance from model A that each of the colours represents. As shown in the figure, in some games nearly all of the non-models evolved close mimicry, but in others mimicry did not evolve.

Table 2a and 2b – Likelihood ratio tests used in backwards model selection for generalised linear mixed models (GLMMs) predicting the phenotypic distance of a non-model prey item to model A (2a) and model B (2b). Definition of terms: Generation is the level of the game to which the player had progressed, Sex is the sex of the player, Treatment is one of four possible distances between models in colour space varying from 70 to 240 RGB units.

a)

Terms	Chi	df	p
Treatment	1274.9	1	< 0.05
Generation	8.49	3	< 0.05
Sex	0.34	1	0.56
Generation:Treatment	123.62	3	< 0.05
Generation:Sex	2.36	1	0.12
Sex:Treatment	2.39	3	0.50
Generation:Sex:Treatment	14.08	3	< 0.05

b)

Terms	Chi	df	p
Generation	1132.3	1	< 0.05
Treatment	8.91	3	0.36
Sex	0.83	1	0.36
Generation:Treatment	116.13	3	< 0.05
Generation:Sex	0.28	1	0.59
Sex:Treatment	4.02	3	0.26
Generation:Sex:Treatment	8.17	3	< 0.05

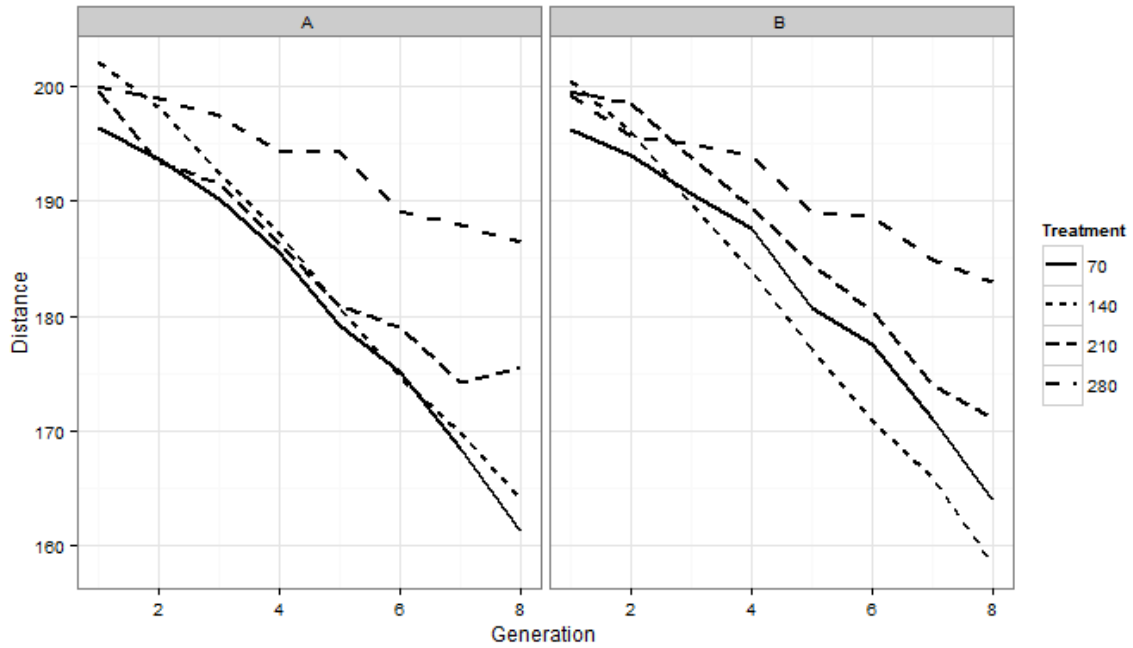


Figure 3 – Mean phenotypic distance, in RGB colour units, of non-models from model A and B in each generation, for each of the four treatments (70, 140, 210 or 280 RGB colour units distance between the models).

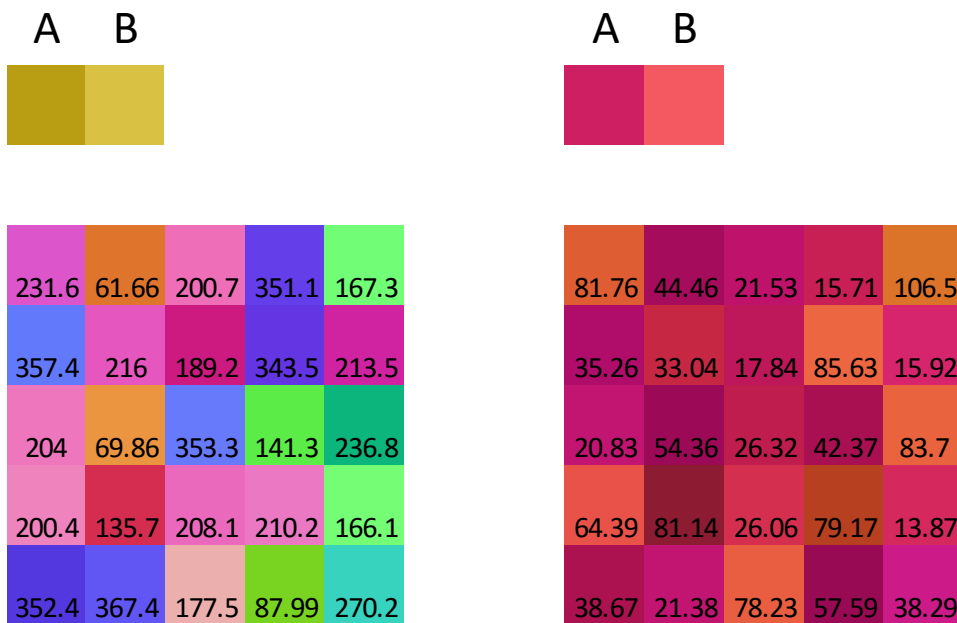


Figure 4. Phenotypes of non-models in the final generation of two games where the distance between the models (labelled A and B) was 70 colour units. Numbers within the squares indicate distance in RGB colour units to model A.

As a method of testing the level of intermediate mimicry, the density of non-models in two areas between the two models (inner zones) was compared to the density of non-models in areas next to the models on the other side (outer zones; see Figure 2.1.1). For all treatments the total number of non-models in the inner and outer zones increased as the generations progressed (Figure 5). When the models were more similar to each other a higher proportion of mimics were found in the inner phenotypic zones (Table 3). A much higher proportion of the mimics were in the inner zones than were in the outer when the models were 70 rgb colour units apart and a similar, less marked difference is present for the 140 units treatment (Figure 5). For the 210 and 280 colour units treatments there is no significant difference in the numbers of prey in the inner or outer zones. Mimicry is evolving for all of the treatments, but intermediate mimicry is only evolving when the models are more similar to each other. There was also a three way interaction between the effects of treatment, sex and generation on the proportion of mimics found in the inner and outer zones; for male players there were fewer mimics in the inner zone for the 70 treatment and more outer mimics for the 280 treatment (figure 6).

Table 3 – Likelihood ratio tests for GLMMs predicting the phenotypic zone (inner or outer) in which mimics were found. The zones were circular areas centred on points either side of each of the models, with the inner zones being the ones between the positions of the models and the outers zone being on the opposite side of the model(see Figure 2.1.1). Terms: Generation is the level the player had progressed to, Sex is the sex of the player and Treatment is the distance between the two models (one of 70, 140, 210 or 280 RGB colour units).

Terms	Chi	df	p
Treatment	16.38	1	< 0.05
Generation	1.13	1	0.29
Sex	2.46	3	0.48
Generation:Treatment	3.37	1	0.07
Generation:Sex	7.26	3	0.06
Sex:Treatment	4.30	3	0.23
Sex: Generation:Treatment	21.98	3	< 0.05

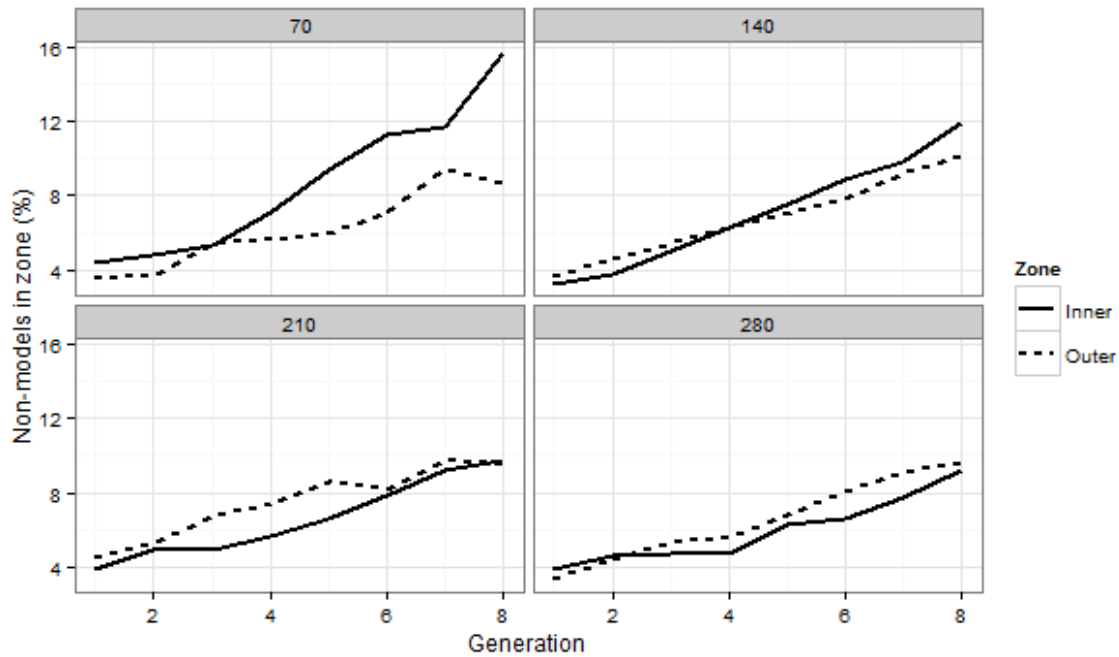


Figure 5. The mean percentage of non-model prey found in inner and outer phenotypic zones (Figure 2.1.1) across the generations for the four treatments.

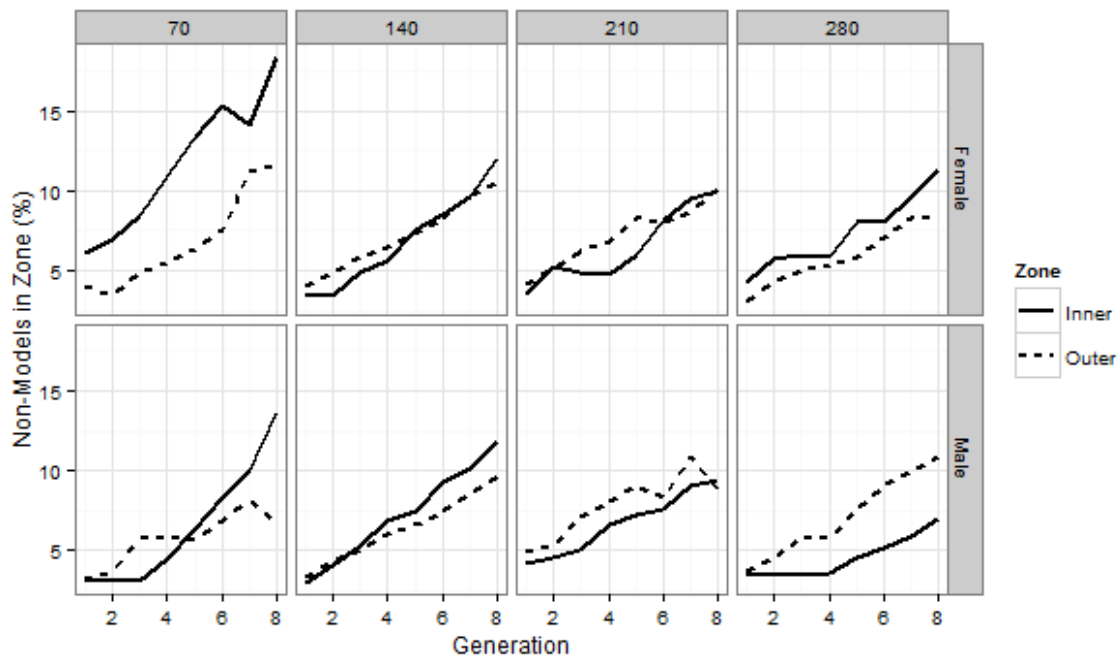


Figure 6 – Mean percentage of mimics in the inner and outer phenotypic zones in each generation, for the four treatment distances, divided by sex.

Creating maps of the phenotypic space with the number of attacks, for all games played, in a given square area indicated by the shading of the square showed that phenotypes close to the models were less likely to be attacked (Figure 7). At the treatment distance of 70 the models create a continuous zone of protection between the, but this is not the case when the models are further apart. Model A also seems to be more strongly protective than model B, for treatment model

distances greater than 70. This is an unexpected result and indicates a bias of the players towards the model A, which was always displayed on the left hand side of the screen when players were showed which colours to avoid. Similar maps showing the positions of the phenotypes of non-models in the last three generations of the games (Figure 8) show that they clustered around the positions of the phenotypes of the models, again with a skew towards model A for treatment distances 140, 210 and 280.

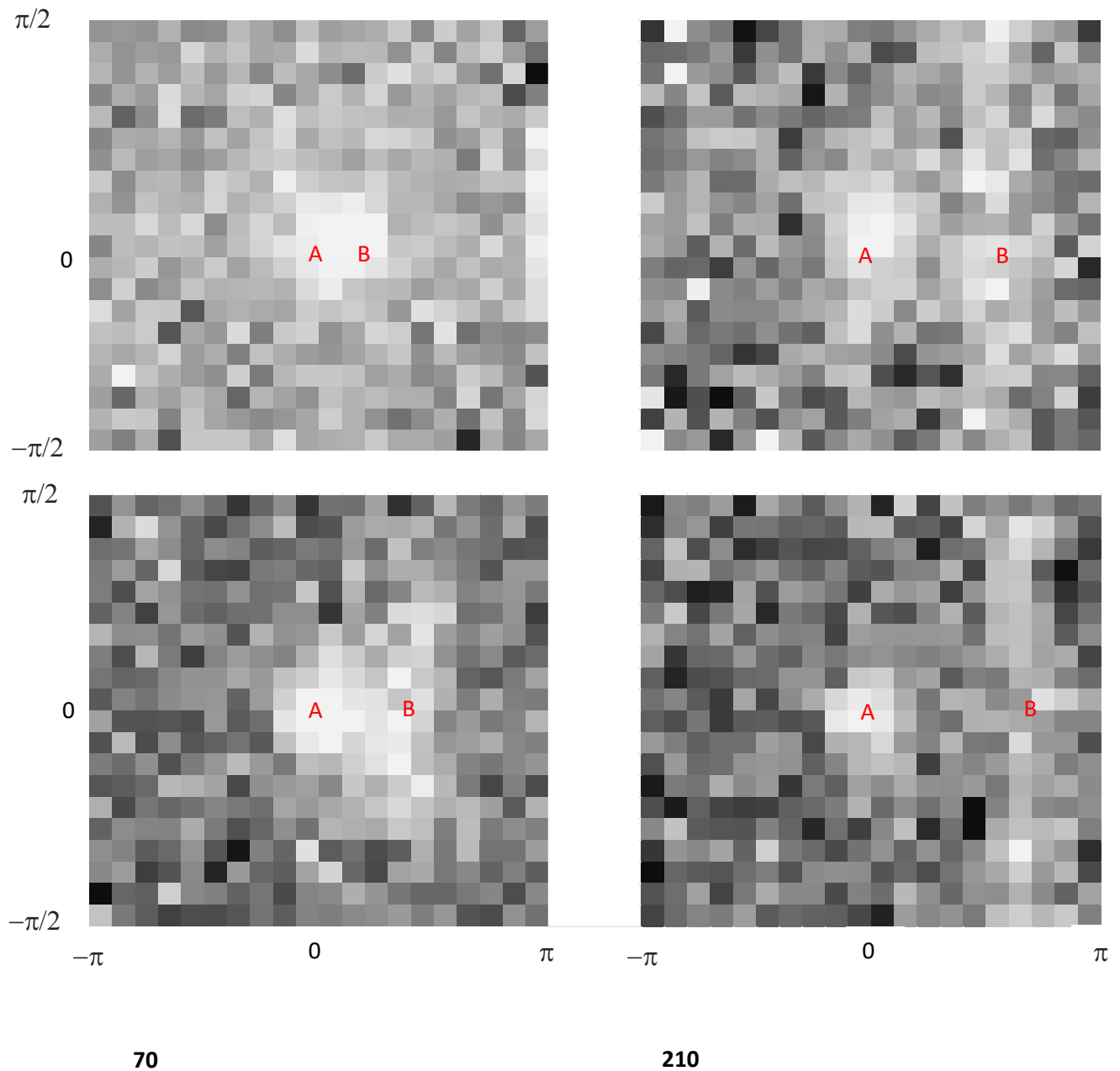


Figure 5 – maps of phenotypic space showing the probability that a non-model prey item was attacked in a given position in phenotypic space, relative to models A and B (positions indicated by letters on chart), for the four treatments. The phenotypic space was spherical: ϕ and λ represent latitude and longitude.

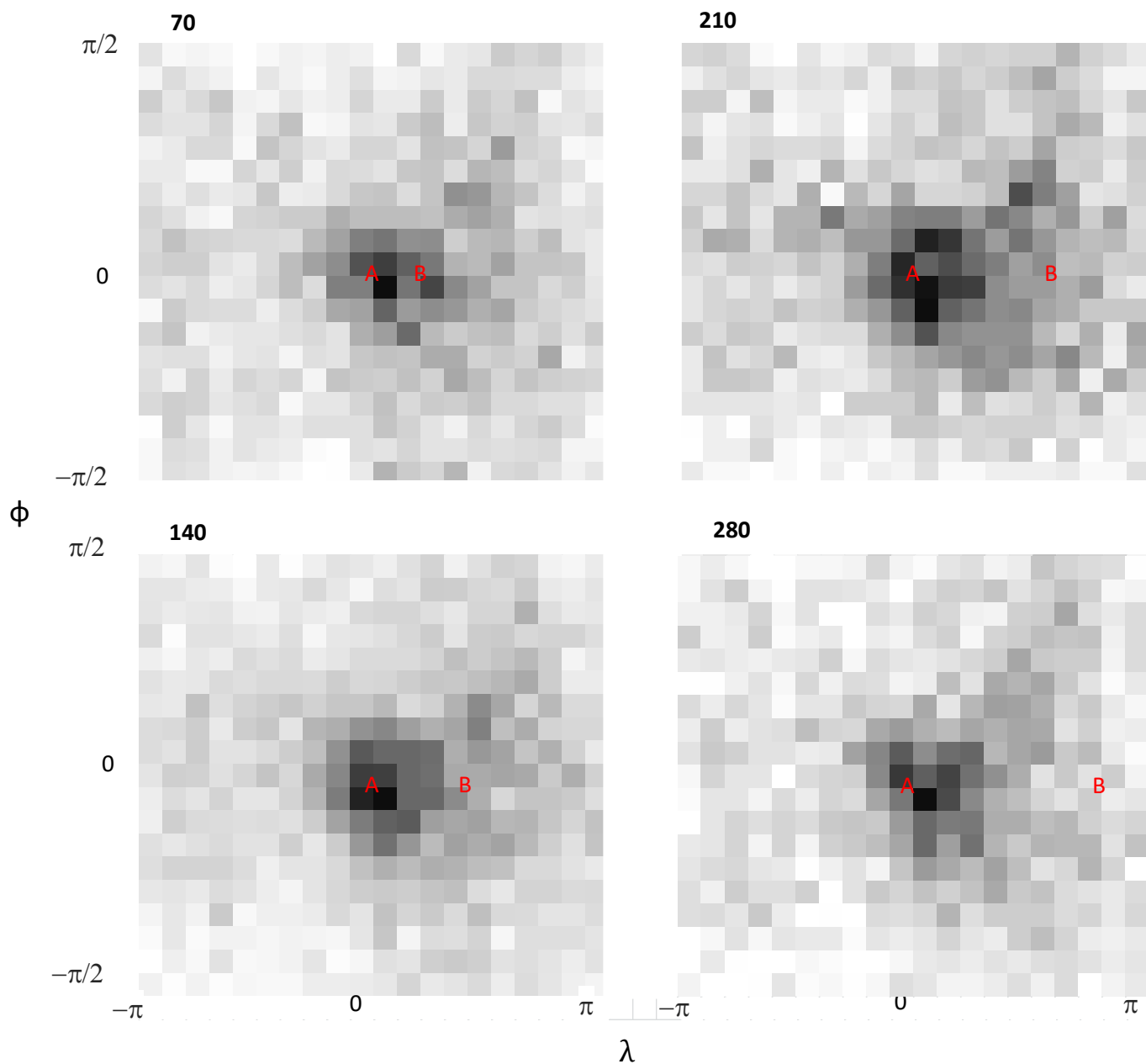


Figure 6 – Maps of phenotypic space showing the density of non-model phenotypes for the four treatments. The positions of models A and B are indicated.

2.4 Game Discussion

The results suggest that intermediate mimics of multiple models can be at a selective advantage to specialist mimics when the models have similar phenotypes to each other. This result supports the multiple models hypothesis of imperfect Batesian mimicry and the predictions of theoretical models (Edmunds 2000, Sherratt 2002).

Mimicry evolved for all treatments. Colours that were similar to those of the models were less likely to be attacked by the players, leading to the proliferation of similar non-model colours. The average phenotype of the non-models also became closer to the phenotypes of the two types of models as the generations progressed. This shows that players were able to learn and remember the colours they were to avoid. This result replicates the success of other studies which attempted to simulate the evolution in computer generated prey (Sherratt et al., 2007, Beatty and Franks, 2012).

It is interesting that mimicry evolved at a faster rate when the models were more similar to each other. This might suggest that players faced with less similar models consciously chose one model type to focus on and consequently the selective pressure for mimicry was less intense. It does not seem to be due to players being less accurate when models were less similar to each other, as there was no difference in the likelihood that a non-model was attacked between treatments. It was harder to remember more dissimilar colours and players were therefore less accurate.

Intermediate mimicry appears to have been adaptive when the two models were similar to each other. A higher proportion of the mimics were intermediate (found in the inner zone) than otherwise for the two treatments where the models were closer together (70 and 140 colour units). This indicates that prey which are similar to both models are more aversive than prey similar to one model only. At the larger distances (210 and 280 RGB units apart) there was no difference between the number of non-models in the inner and outer zones of the phenotypic space, indicating that there was no selective advantage to intermediate mimicry when the models were more dissimilar from each other.

At the end of the eight levels the distance of non-models from models appears to have not reached equilibrium for any of the treatments. It also appears likely that the relative number of non-models in the inner mimicry zone compared to the number in the outer zone would have increased, for the 70 and 140 RGB units treatments.

The maps of attack rate on non-models throughout the game, and percentage of non-models in areas of phenotypic space at the end of the game, show the zone of protection around the models. There was a continuous zone of phenotypic protection created between models when they were 70 RGB units apart, and a less distinct continuous zone when the models were 140 units apart. There is a skew in terms of mimetic protection towards model A: there is a clearer zone of protection and the distribution of non-models in the last generation of the games is more densely clustered around model A than model B for all treatments.

The distribution of mimics in phenotypic space at the end of the games appears to be biased towards model A. This suggests a bias in the players, whether conscious or not, towards the colour presented on the left of the screen when the model colours were presented to the players at the start of the game, as A always was. (evidence of bias in humans).

There was a difference between the sexes in terms of number of attacks on models, the rate at which non-model phenotypes changed, and the number of intermediate mimics that evolved. Males

attacked more models, indicating that they were less able to remember the colours of the models, or that they were less careful in their choices. The lower accuracy of males may be the reason for the observed slower rate of mimicry evolution, as being less accurate would create a lower selective pressure for mimicry. It is not clear why there should be a difference between the sexes in the number of intermediate mimics that evolved.

The colours of non-models in the last generation of most games clearly demonstrate that mimicry evolved, as expected, but there is evidence that some players did not play the game optimally, as there are some games for which very little mimicry evolved. It is likely that players in these games either did not understand the game or were deliberately not playing the game properly. This is a problem with having the game online: there is little control over how the players play the game, or of the environment in which they do so.

Sherratt (2002) showed using mathematical modelling that multiple models mimicry could arise if models were similar to each other in phenotype. The results of the present study demonstrate that selection by human predators can lead to the evolution of intermediate mimicry. Behavioural studies will be invaluable in the future in improving our understanding of the situations in which multiple models mimicry is adaptive.

Humans provide a reasonable substitute for natural predators, and there is a great opportunity for using humans playing online games to explore the evolutionary consequences of selection, but clearly they do not exert selective pressure on mimics in nature and it is necessary to be cautious about interpreting results gained using humans as proxy predators. Future work using birds, which are assumed to be primary predators of many insect mimic species, will be important in improving our knowledge of predator behaviour with regard to Batesian mimics. Future work could also incorporate prey with a greater number of features, allowing more realism in prey appearance.

2.5 Acknowledgements

I'm grateful to Rohanna Dow for the work she did before me on the game used in this chapter. Rohanna created the game with Tom Reader in 2009 for her PhD. The experiments presented here are built on the experiments she carried out.

3 Testing feature saliency using an online quiz

Introduction

Batesian mimics are innocuous species which mimic the aposematic signal of a defended or dangerous species in order to avoid predation. An unsolved puzzle regarding Batesian mimicry is that it often appears to be quite imprecise (Ruxton). In order to study why mimics are imperfect, we need to be able to quantify the accuracy of mimics.

To quantify mimetic accuracy, it is necessary to know which features a predator pays attention to when looking at the mimic. If a predator does not pay attention to a feature then it will not be under selection for mimicry. A feature may be ignored for several reasons: the predator may not have the sensory capacity to notice it, for example if it does not have the necessary types of cone cells in its eyes to distinguish certain colours (Cuthill and Bennett 1993); the predator may only see the mimic for a very short period of time, meaning it cannot resolve certain details of the image; the predator may have a cognitive bias towards certain features (Chittka and Osorio). The visual appearance of insect mimics can be broken down into features such as colour, shape, size and pattern, and studies of predator attention suggest that these features are not equally salient (noticeable). Predators may be able to use *Gestalt* perception (the combination of different features is important; Ikin and Turner 1972).

There is evidence that pattern may be a salient part of the mimetic signal. Dragonflies have been shown to avoid potential prey with black and yellow wasp-like stripes more than they do prey which are plain black or yellow (Kauppinen and Mappes). It is possible that having high contrast between the colours may be more useful than the pattern itself (Aronsson 2013, Zylinski). Contrast between colours in a signal has been shown to have an effect on the speed and persistence of aversion learning (Prudic 2007). Aronsson and Gamberalle 2009, however, found that prey to background colour contrast improved predator avoidance learning, but internal colour pattern was not shown to have the same effect. Bain et al. 2007 analysed traits in images of insects shown to pigeons in a study of where the birds were conditioned to not “attack” images of model species and then their response to mimics recorded (Dittrich et al 1993), and found that the attack rates were predicted well by the colour and also by the pattern morphology of the mimics.

There is evidence that colour is an important trait for avian predators (thought to be major predators for many insect mimics) in distinguishing between prey (Kazemi et al. 2014). For avian predators, it is likely that pattern is used less than colour (Vesely?). Poultry chicks used colour more than pattern in distinguishing between artificial prey (Aronsson and Gamberale-Stille 2012). In an experiment with blue tits (*Parus caeruleus*) feeding on artificial prey with three prominent traits (size, shape and colour) the birds learned the colour trait at a faster rate than the other traits, and perfect colour mimics (which were otherwise imperfect mimics) were as protected as perfect mimics (Kazemi et al. (2014). This indicates that colour had a higher saliency for the birds than the other stimuli. Kraemer and Adams () showed that birds used colour differences to distinguish between two salamander species, one of which was a Batesian mimic of the other. The colours of hoverfly mimics of hymenopterans match their models quite closely ().

The evidence for the importance of shape mimicry is mixed. Experiments by Taylor et al. () using images showing only the outline of Batesian mimetic hoverflies and their models found that people were able to identify prey quite accurately. But Penney et al. (2012) found that hoverfly mimics are usually not similar in their shape to their models.

In the experiments presented in this chapter, an online quiz was used in which players chose whether or not an insect image presented to them was of a wasp or a hoverfly. The experiments follow on from previous experiments using the same method and specimen images (Taylor ...) which showed that people paid attention to the abdomen of the insect when judging mimicry. Some of the

images were altered so that they only conveyed information about one aspect of the abdomen (colour, shape or pattern). The aim was to find whether participants were more accurate for certain treatments than others, and which features people paid most attention to when looking at the full image. Humans have been shown to have similar judgement to bird predators about mimetic accuracy (Dittrich et al. 1993, Penney 2012). Although there are limitations to the conclusions that can be drawn from experiments using humans rather than natural predators, using humans is logistically easier and allows for larger sample sizes.

The study method used has the advantage of using images natural mimics, and allowing experimental manipulation. Using artificial prey (e.g. in Kazemi, Aronsson Gamberalle stille) means that stimuli are not set within natural bounds and this may result in disproportionate effects on predators. It is therefore advantageous to present images of real mimics. The analysis by Bain et al. (2007) on the traits of images used by Dittrich (1993) was retrospective so did not allow experimental manipulation. Presenting the images for a short period of time resembles the natural situation in which predators are likely to have only a short time to view prey (Chittka and Osorio).

2.2 Methods

2.2.1 Images

Specimens for the quiz were collected by Chris Taylor from hoverfly (Diptera: Syrphidae) and stinging Hymenoptera communities around Nottingham, UK and Sobreda, Portugal using a hand net. Hoverfly specimens were collected if they had a similar appearance to Hymenoptera (i.e. a “black” and “yellow” colour pattern, without a large amount of hair). Model specimens collected were the common wasp *Vespula vulgaris* and other stinging Hymenoptera with aposematic colouration.

In order to slow their movement and prevent them from flying, specimens were refrigerated at 5°C for at least an hour after being caught. Photographs were then taken of them in an open petri dish inside a makeshift photo studio (a white, 30 x 18 x 10cm open topped box). An Olympus E420 DSLR camera was used to photograph the specimens. Each specimen was photographed in two ways: in a natural resting position for the “live photo”, and - after being euthanised by freezing - with legs and wings pinned out to the sides for the “abdomen photo”. A 5mm scale bar was placed next to the insect in each photo.

Images of 104 specimens were used in the quiz, from 12 different species. A species was included if there were good live and abdomen photos for at least five individuals of that species. The highest number of specimens for one species was 21. Three species were stinging Hymenoptera (models): two from the genus *Vespula*, and the third, though treated as one species for this investigation, was actually comprised of specimens from several similar species from the genus *Nomada*.

MATLAB (2012) was used for automated processing of the images, and ImageJ (Abràmoff et al., 2004) was used for manual editing. Each image was made into a GIF, in order that the image of the insect flashed for 0.5s then disappeared.

The “live” photos were used to create “Full” images. “Full” images were rotated by eye to a position with the axis of the insect vertical and the head at the top. The contrast of the images was altered using histogram equalisation. Body length was measured from the tip of the head between the antennae to the tip of the abdomen. A scale bar was used for reference. MATLAB was then used to resize the image to a standard scale such that 1mm on the insect was represented by five pixels in the image.

“Abdomen” photos were used to create three types of treatment image: “Shape”, “Colour” and “Pattern” (Figure 1). “Shape” images were created in ImageJ by tracing the outline of the shape of the abdomen in the abdomen images, and then filling in the resulting shape with black while making the background of the image white. “Colour” images were made by using MATLAB to find the average RGB colour values for the “yellow” and “black” sections of the abdomens of the specimens in the abdomen images. These colours were then used to create a standardised image with four horizontal stripes within a 60 x 60 area, alternating between the “yellow” and “black” colours. “Pattern” treatments were created by converting the images to greyscale in MATLAB, and cropping this to an oval shape 60 pixels in height and 40 pixels in width. The image was scaled so that as much of the pattern as possible fitted into the oval cropped area. For insects with narrow abdomens that would not fill the whole of the oval pane, it was necessary to stretch the images horizontally. This did not affect the pattern as these insects tended to have horizontal stripes.

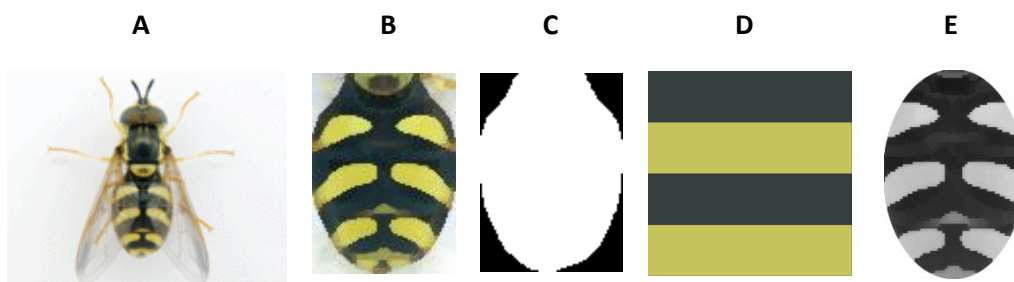


Figure 1 – Example image and treatments. (A) Full image (*Chrysotoxum intermedium*); (B) abdomen image; (C) shape treatment; (D) colour treatment; (E) pattern treatment.

2.2.2 The Quiz

In the quiz participants were sequentially shown 20 images. The first ten images were “Full” images. This section constituted a training period in which participants learned which specimens to avoid. The second ten images were treatment images. For each image the participant had to decide whether to click on a button to “eat” or one to “avoid” (Figure 2). They were then informed whether they had made the correct decision, before progressing to the next image. The aim of the experiment was to test whether there would be a difference in the performance of participants given different treatments.

Quiz - identifying insects

* Required



Q1 of 20: Choose an action *

- Eat
- Avoid

« Back

Continue »

Figure 2 – A question as it would have appeared to a participant. The specimen image flashed for 0.5 seconds before disappearing.

The quiz was publicised to potential participants by sending emails to university mailing lists, Facebook posts on personal accounts and to university societies, and a Twitter account set up for the purpose of publicising the quiz. The quiz was live from January to March 2015, and 134 people participated. Using the internet to publicise the quiz meant that a larger number of people took part than would have been possible if recruiting subjects face to face, but meant sacrificing some control over the experimental environment.

Participants were directed to a webpage which gave them some background to the quiz, but no information about the aim of the experiment, in order to avoid any potential bias this may have caused. They were informed that the data generated would be used for research, and a disclaimer was included stating that the data from the quiz would be anonymous. Participants then clicked through to a second page and entered their details (age, gender and whether they had previously attempted the quiz), before progressing to the quiz.

For each question, the specimen displayed was randomly chosen from all of the specimens. Probabilities of selection were adjusted so that the chance of the chosen specimen being a model or a mimic, of being from a particular species, or of being a particular individual within a species, was equal. An individual specimen could be chosen multiple times in the same quiz.

2.2.3 Statistical Analysis

Data was analysed in R version 3.1.2 (R Core Team, 2014). Responses from participants who had already attempted the quiz were discarded. The response of each individual question was modelled using a binomial generalised linear mixed model. Fixed factors included in the full model were: age of the participant, sex of the participant, species of the specimen in the image, and treatment type. Interactions between all factors other than species were also tested; the interactions of species were

not included because with them the model was too complex to be analysed. The participant's ID was included as a random factor. Backwards model selection using likelihood ratio tests was used to find the best model (Zur et al 2009).

2.3 Results

A total of 133 participants completed the quiz, with 47 presented the colour treatment, 43 the pattern treatment and 43 the shape treatment. 85 of the participants were female, 48 male.

Participants were most accurate when shown the full images in the training phase of the quiz. Response accuracy in the second half of the quiz, when participants were shown the treatment images (showing shape, colour or pattern of the abdomen) was affected by the treatment shown (LR, $\chi^2 = 22.27$, $df = 2$, $p < 0.001$). Participants were more accurate when shown the shape treatment than when shown colour or pattern (Figure 1). This result fits with those of previous experiments carried out by Taylor et al. using the same methodology, which showed that people had a good ability to identify safe prey by looking at treatment images which showed the outline of the full insect.

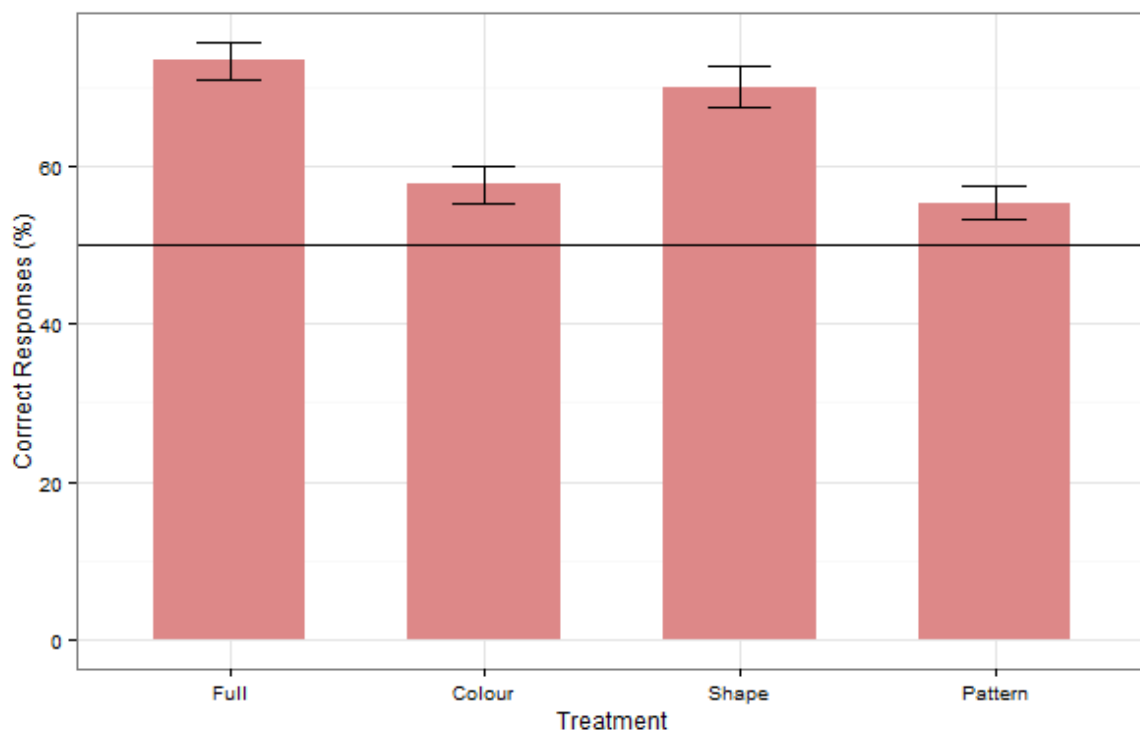


Figure 1. The mean percentage of responses that were correct, for the training phase and the three treatment types, with binomial error bars. The horizontal line indicates the score expected by chance.

The correlation between attack rates on different species in the first half of the quiz, when participants were shown full images of the insects, and the second half, when they were shown the treatment images,....

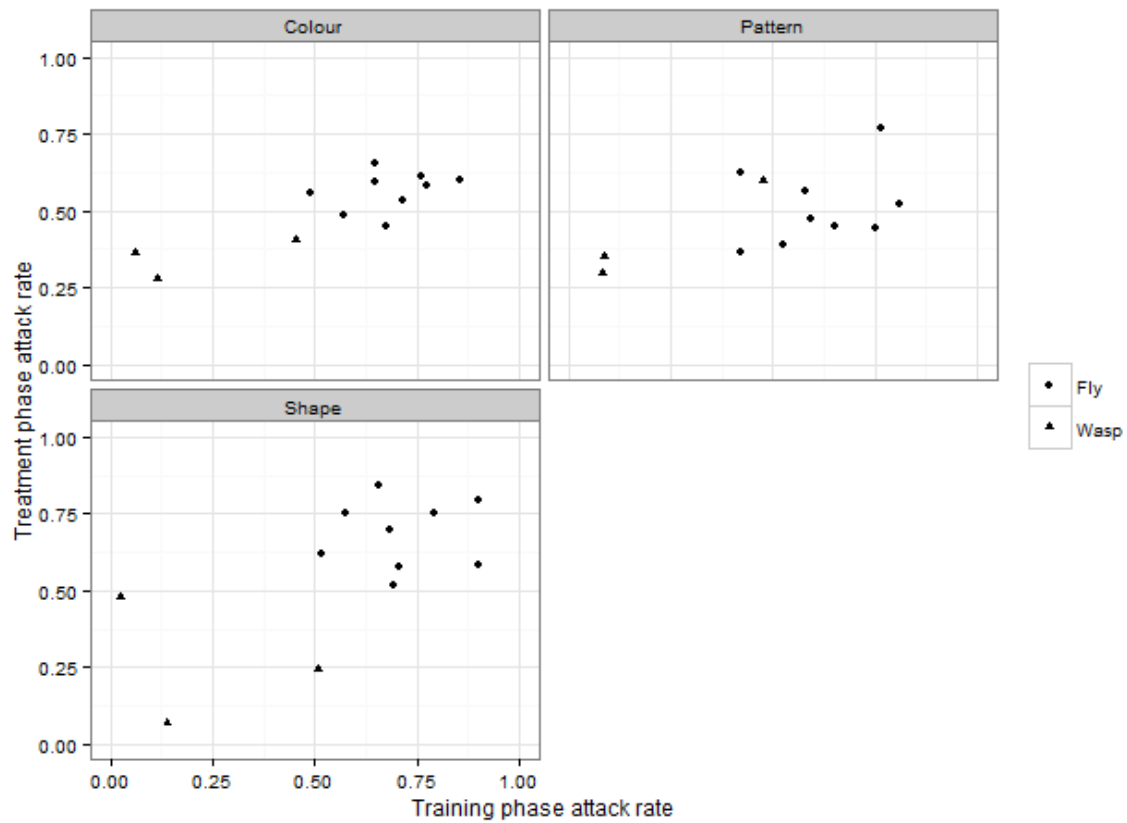


Figure 2. Plots show the attack rates on the different species in the first (training) and second (treatment) halves of the quiz.

The median age of participants was 31. Age did not have a significant effect on response accuracy, but Figure 2 shows that when shown the treatment images accuracy decreases with age from the 26-35 age bracket. No other factor had an effect on the accuracy of the responses to the treatments.

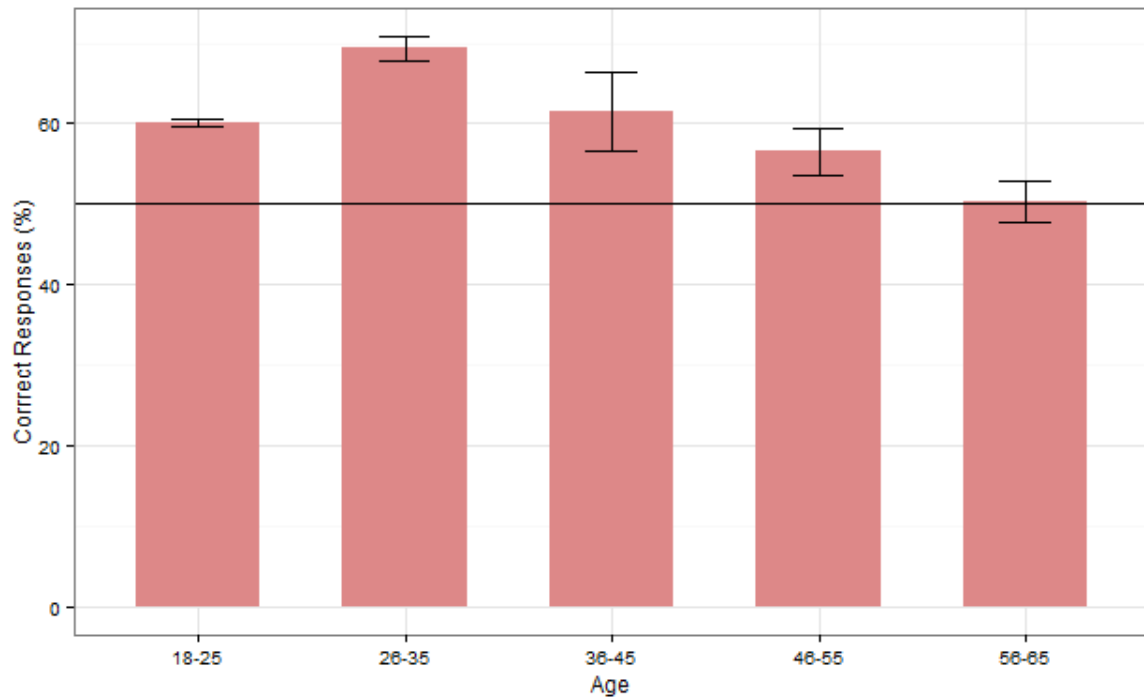


Figure 2 – The mean percentage of correct responses in different age brackets.

Analysis of the results from the training phase (with full images of the specimens) showed that sex had an effect on response accuracy (LR, $X = 14.91$, $df = 1$, $p < 0.001$). Males were more accurate than females (figure ...). Strangely this difference was not present in the responses to treatment images. It is possible that the male participants had better knowledge of the appearance of wasps and hoverflies, but this was not useful to them when shown the treatment images. No other factor (treatment, species or age) had an effect on response accuracy, and there were no significant interactions between factors.

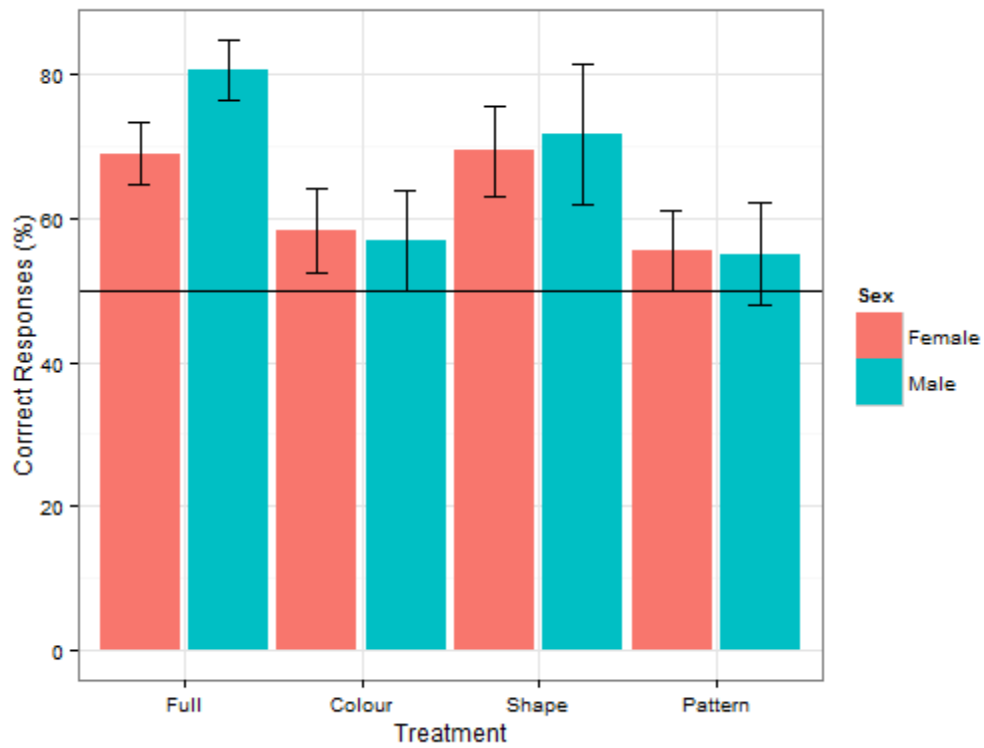


Figure 4 – The mean percentage of correct responses for males and females, with binomial error bars, for the full images shown in the training phase and the three treatment image types.

Discussion

The results give us an insight into how predators perceive visual mimicry. Shape was the most informative feature of the three treatments presented to participants. This result fits with the finding from a previous experiment using the same method which found that the outline of the entire insect was informative to predators (Taylor et al.). In the previous experiment participants seem to have paid less attention to outline than to abdomen's appearance, when viewing the full images. A similar result was found here; though abdomen shape was the most informative of the three features, people appear to pay more attention to colour.

Participants were better at identifying mimics and models by the shape of the abdomen than by the abdomen's colour or pattern. This suggests that shape is an informative feature. This result agrees with the findings of Penney et al (2012) who analysed the shapes of hoverflies and wasps using multivariate analysis and found them to be clearly distinct. A previous experiment, using the same experimental method, found that participants were able to identify mimics and models well when shown the outline of the whole body, rather than just the abdomen (Taylor...). Bain result

Interestingly, species which are good shape mimics (based on attack rate) are not good mimics when the full insect is viewed. This suggests that the information about the shape of the insect is not being used when people view the full image. It may be that people do not pay as much attention to the shape when the full image is shown to them, or that in the full image information about shape is somehow masked by colour and pattern information. Taylor et al. () found a similar result; species that were good mimics when only their outline was shown were not good mimics when they were viewed as full, unaltered images (Taylor...). Taylor ... investigated whether people could be trained to pay more attention to the shape of the abdomen in full images by showing quiz participants the shape treatment first. People were not found to be more averse to good shape mimics during the full image phase; this result does not support the idea that people are simply not paying as much attention to shape when shown full images.

The best mimics in the training phase are also the best mimics in the colour phase. This indicates that colour is a more salient trait than pattern or shape when people view the full specimens. Previous studies have demonstrated the importance of prey colour to predator judgements (Terhune, Kazemi). In experiments using birds trained to avoid artificial models with certain colour, shape and pattern traits, birds learned colour more quickly than the other traits, and colour mimics were more protected than shape or pattern mimics.

People were not able to use abdomen pattern to judge whether the insect was a model or mimic. Previous experiments have suggested that pattern has an important role in the mimetic signal when viewed by birds; Bain () showed that elements of pattern such as number of stripes was important. It is possible that differences in human and bird visual perception account for the lack of importance of pattern in this study.

It is known that in most circumstances humans display global precedence (they process global features of images before local features contained within the image (Navon)). This is thought to not be true, however, for non-mammalian vertebrates (). This difference may play a part in the fact that pattern seems to be less important in studies using humans than those using birds. There is, however, some evidence that birds are capable of *Gestalt* perception, where the combination of features is significant (Ikin and Turner), which would indicate more of a tendency towards global processing.

The sex of the participants was a significant factor in accuracy of identification, but only in the full image phase. The difference between males and females for full images perhaps indicates that males have better prior ability to identify wasps and hoverflies. Whatever difference there is between males and females is not present for the treatment images. There may be sex differences in the way images are processed which mean men are better able to learn the cues to avoid when shown images with multiple features. Alternatively the males who played may have been more experienced at playing online games, or may have better prior knowledge of the appearance of hoverflies and wasps. If the last explanation is correct it would imply that there is a feature other than pattern, colour or shape that males are using to distinguish between the full images, or that viewing the features together gives information that is lost when they are viewed in isolation, which, because of the holistic way that people process images (as discussed above) is not implausible.

The study presented here is unique in that the images of specimens were shown to participants in a realistic way. Low resolution images were shown for a short period of time; this is similar to the natural situation, in which predators are likely to have a short time to view a mimic from a distance (Chittka and Osorio). A previous study which involved humans estimating the similarity of images of a mimic and model used high resolution images and did not set a time limit for people viewing them (Penney). The conclusions drawn from the present study are more reliable because the experiments resembled the natural situation more accurately.

The results presented here lend weight to the argument that to understand imperfect mimicry we need to know which features of the mimic are salient to the predator and therefore under selection for mimicry. The results of previous experiments have shown that some features appear to be more significant to predators than others. In order to establish that mimics are more accurate for a certain feature, such as colour, than they are in their general appearance to the model, it will be necessary to have an objective way of measuring overall mimetic accuracy. Tools such as distance transform, which objectively measure the difference between pattern phenotypes are a step towards being able to do this (Taylor ...). The conclusions that can be drawn from a study using humans are limited, and development of this work in future could use birds as predators. Humans do not exert a selective pressure on mimicry in nature. But there is evidence that the judgements of humans and avian predators on the accuracy of mimics are similar (Dittrich 1993).

Acknowledgements

I'm grateful to Chris Taylor for allowing me to collaborate with him

Bibliography

- ABRÀMOFF, M. D., MAGALHÃES, P. J. & RAM, S. J. 2004. Image processing with ImageJ. *Biophotonics International*, 11, 36-42.
- ALATALO, R. V. & MAPPES, J. 1996. Tracking the evolution of warning signals. *Nature*, 382, 708-710.
- BATES, H. W. 1862. Contributions to an Insect Fauna of the Amazon Valley. Lepidoptera: Heliconidæ. *Transactions of the Linnean Society of London*, 23, 495-566.
- BEATTY, C. D., BEIRINCKX, K. & SHERRATT, T. N. 2004. The evolution of mullerian mimicry in multispecies communities. *Nature*, 431, 63-67.
- BEATTY, C. D. & FRANKS, D. W. 2012. Discriminative predation: Simultaneous and sequential encounter experiments. *Current Zoology*, 58, 649-657.
- CHITTKA, L. & OSORIO, D. 2007. Cognitive dimensions of predator responses to imperfect mimicry? *Plos Biology*, 5, 2754-2758.
- CLUSELLA TRULLAS, S., VAN WYK, J. H. & SPOTILA, J. R. 2007. Thermal melanism in ectotherms. *Journal of Thermal Biology*, 32, 235-245.
- CUTHILL, I. C. 2014. Evolution: The Mystery of Imperfect Mimicry. *Current Biology*, 24, R364-R366.
- DARST, C. R. & CUMMINGS, M. E. 2006. Predator learning favours mimicry of a less-toxic model in poison frogs. *Nature*, 440, 208-211.
- DITTRICH, W., GILBERT, F., GREEN, P., MCGREGOR, P. & GREWCOCK, D. 1993. Imperfect Mimicry - a Pigeons Perspective. *Proceedings of the Royal Society B-Biological Sciences*, 251, 195-+.
- DOW, R. D. 2010. *Why are there imperfect Batesian mimics?*, University of Nottingham.
- DUNCAN, C. J. & SHEPPARD, P. M. 1965. Sensory Discrimination and its Role in the Evolution of Batesian Mimicry. *Behaviour*, 24, 269-282.
- EDMUNDS, M. 1978. On the association between *Myrmarachne* spp. (Salticidae) and ants. *Bulletin of the British Arachnological Society*, 4, 149-160.
- EDMUNDS, M. 2000. Why are there good and poor mimics? *Biological Journal of the Linnean Society*, 70, 459-466.
- EDMUNDS, M. & READER, T. 2014. Evidence for Batesian mimicry in a polymorphic hoverfly. *Evolution*, 68, 827-839.
- ESTRADA, C. & JIGGINS, C. D. 2008. Interspecific sexual attraction because of convergence in warning colouration: is there a conflict between natural and sexual selection in mimetic species? *Journal of Evolutionary Biology*, 21, 749-760.
- FEENEY, W. E., STODDARD, M. C., KILNER, R. M. & LANGMORE, N. E. 2014. "Jack-of-all-trades" egg mimicry in the brood parasitic Horsfield's bronze-cuckoo? *Behavioral Ecology*, 25, 1365-1373.
- FRANKS, D. W. & NOBLE, J. 2004. Warning signals and predator-prey coevolution. *Proceedings of the Royal Society B-Biological Sciences*, 271, 1859-1865.
- GILBERT, F. 2005. The evolution of imperfect mimicry in hoverflies. *Insect Evolutionary Ecology*. CABI.
- GOLDING, Y. C., EDMUNDS, M. & ENNOS, A. R. 2005. Flight behaviour during foraging of the social wasp *Vespula vulgaris* (Hymenoptera : Vespidae) and four mimetic hoverflies (Diptera : Syrphidae) *Sericomyia silentis*, *Myathropa florea*, *Helophilus* sp and *Syrphus* sp. *Journal of Experimental Biology*, 208, 4523-4527.
- GREEN, D. M. & SWETS, J. A. 1966. *Signal Detection Theory and Psychophysics*, New York, Wiley.
- HEGNA, R. H., SAPORITO, R. A. & DONNELLY, M. A. 2013. Not all colors are equal: predation and color polytypism in the aposematic poison frog *Oophaga pumilio*. *Evolutionary Ecology*, 27, 831-845.
- HEINRICH, B. 1984. Strategies of thermoregulation and foraging in 2 vespid wasps, *Dolicovespula Maculata* and *Vespula-vulgaris*. *Journal of Comparative Physiology*, 154, 175-180.
- HOLLOWAY, G., GILBERT, F. & BRANDT, A. 2002. The relationship between mimetic imperfection and phenotypic variation in insect colour patterns. *Proceedings of the Royal Society B-Biological Sciences*, 269, 411-416.
- JOHNSTONE, R. A. 2002. The evolution of inaccurate mimics. *Nature*, 418, 524-526.

- JORON, M. & MALLET, J. L. B. 1998. Diversity in mimicry: paradox or paradigm? *Trends in Ecology & Evolution*, 13, 461-466.
- KIKUCHI, D. W. & PFENNIG, D. W. 2010. High-model abundance may permit the gradual evolution of Batesian mimicry: an experimental test. *Proceedings of the Royal Society B-Biological Sciences*, 277, 1041-1048.
- KIKUCHI, D. W., SEYMOURE, B. M. & PFENNIG, D. W. 2014. Mimicry's palette: widespread use of conserved pigments in the aposematic signals of snakes. *Evolution & Development*, 16, 61-67.
- LANGMORE, N. E., STEVENS, M., MAURER, G., HEINSOHN, R., HALL, M. L., PETERS, A. & KILNER, R. M. 2011. Visual mimicry of host nestlings by cuckoos. *Proceedings of the Royal Society B-Biological Sciences*, 278, 2455-2463.
- LICHTER-MARCK, I. H., WYLDE, M., AARON, E., OLIVER, J. C. & SINGER, M. S. 2015. The struggle for safety: effectiveness of caterpillar defenses against bird predation. *Oikos*, 124, 525-533.
- MARSAGLIA, G. 1972. Choosing a Point from the Surface of a Sphere. *Ann. Math. Statist.*, 43, 645-646.
- MARTIN, A., PAPA, R., NADEAU, N. J., HILL, R. I., COUNTERMAN, B. A., HALDER, G., JIGGINS, C. D., KRONFORST, M. R., LONG, A. D., MCMILLAN, W. O. & REED, R. D. 2012. Diversification of complex butterfly wing patterns by repeated regulatory evolution of a Wnt ligand. *Proceedings of the National Academy of Sciences of the United States of America*, 109, 12632-12637.
- MATLAB 2012. MATLAB. Natick, Massachusetts: The Mathworks.
- MCGUIRE, L., VAN GOSSUM, H., BEIRINCKX, K. & SHERRATT, T. N. 2006. An empirical test of signal detection theory as it applies to Batesian mimicry. *Behavioural Processes*, 73, 299-307.
- MORGAN, K. R. & HEINRICH, B. 1987. Temperature regulation in bee-mimicking and wasp-mimicking syrphid flies. *Journal of Experimental Biology*, 133, 59-71.
- NELSON, X. J., LI, D. Q. & JACKSON, R. R. 2006. Out of the frying pan and into the fire: A novel trade-off for batesian mimics. *Ethology*, 112, 270-277.
- NUR, U. 1970. Evolutionary rates of models and mimics in Batesian mimicry. *American Naturalist*, 104, 477-&.
- OTTENHEIM, M. M. & KUIJT, J. 1998. *Has abdominal colour pattern variation in Eristalis arbustorum any thermoregulatory function?*
- PEKAR, S. & JARAB, M. 2011. Assessment of color and behavioral resemblance to models by inaccurate myrmecomorphic spiders (Araneae). *Invertebrate Biology*, 130, 83-90.
- PEKAR, S., JARAB, M., FROMHAGE, L. & HERBERSTEIN, M. E. 2011. Is the Evolution of Inaccurate Mimicry a Result of Selection by a Suite of Predators? A Case Study Using Myrmecomorphic Spiders. *American Naturalist*, 178, 124-134.
- PENNEY, H. D., HASSALL, C., SKEVINGTON, J. H., ABBOTT, K. R. & SHERRATT, T. N. 2012. A comparative analysis of the evolution of imperfect mimicry. *Nature*, 483, 461-U110.
- PFENNIG, D. W., HARCOMBE, W. R. & PFENNIG, K. S. 2001. Frequency-dependent batesian mimicry - Predators avoid look-alikes of venomous snakes only when the real thing is around. *Nature*, 410, 323-323.
- PFENNIG, D. W., HARPER, G. R., BRUMO, A. F., HARCOMBE, W. R. & PFENNIG, K. S. 2007. Population differences in predation on Batesian mimics in allopatry with their model: selection against mimics is strongest when they are common. *Behavioral Ecology and Sociobiology*, 61, 505-511.
- PFENNIG, D. W. & KIKUCHI, D. W. 2012. Competition and the evolution of imperfect mimicry. *Current Zoology*, 58, 608-619.
- PYKE, G. H., PULLIAM, H. R. & CHARNOV, E. L. 1977. Optimal Foraging - Selective review of theory and tests. *Quarterly Review of Biology*, 52, 137-154.
- R CORE TEAM 2014. R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing.

- RUXTON, G. D., SPEED, M. & SHERRATT, T. N. 2004. Evasive mimicry: when (if ever) could mimicry based on difficulty of capture evolve? *Proceedings of the Royal Society B-Biological Sciences*, 271, 2135-2142.
- SCHLUTER, P. M., XU, S. Q., GAGLIARDINI, V., WHITTLE, E., SHANKLIN, J., GROSSNIKLAUS, U. & SCHIESTL, F. P. 2011. Stearoyl-acyl carrier protein desaturases are associated with floral isolation in sexually deceptive orchids. *Proceedings of the National Academy of Sciences of the United States of America*, 108, 5696-5701.
- SHERRATT, T. N. 2002. The evolution of imperfect mimicry. *Behavioral Ecology*, 13, 821-826.
- 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., POLLITT, D. & WILKINSON, D. M. 2007. The evolution of crypsis in replicating populations of web-based prey. *Oikos*, 116, 449-460.
- SMITH, J. M., BURIAN, R., KAUFFMAN, S., ALBERCH, P., CAMPBELL, J., GOODWIN, B., LANDE, R., RAUP, D. & WOLPERT, L. 1985. Developmental constraints and evolution. *Quarterly Review of Biology*, 60, 265-287.
- SRYGLEY, R. B. 2004. The aerodynamic costs of warning signals in palatable mimetic butterflies and their distasteful models. *Proceedings of the Royal Society B-Biological Sciences*, 271, 589-594.
- STEVENS, M., SEARLE, W. T. L., SEYMOUR, J. E., MARSHALL, K. L. A. & RUXTON, G. D. 2011. Motion dazzle and camouflage as distinct anti-predator defenses. *Bmc Biology*, 9.
- TAYLOR, C. H., READER, T. & GILBERT, F. In Review. Why many Batesian mimics are inaccurate: evidence from hoverfly colour patterns. *Proceedings of the National Academy of Sciences of the United States of America*.
- TEAM, R. C. 2014. R: A language and environment for statistical computing. Vienna, Austria: R foundation for statistical computing.
- WILLMER, P. G. & UNWIN, D. M. 1981. Field analyses of insect heat budgets - reflectance, size and heating rates. *Oecologia*, 50, 250-255.
- ABRÀMOFF, M. D., MAGALHÃES, P. J. & RAM, S. J. 2004. Image processing with ImageJ. *Biophotonics International*, 11, 36-42.**
- ALATALO, R. V. & MAPPE, J. 1996. Tracking the evolution of warning signals. *Nature*, 382, 708-710.**
- BATES, H. W. 1862. Contributions to an Insect Fauna of the Amazon Valley. Lepidoptera: Heliconidæ. . *Transactions of the Linnean Society of London*, 23, 495-566.**
- BEATTY, C. D., BEIRINCKX, K. & SHERRATT, T. N. 2004. The evolution of mullerian mimicry in multispecies communities. *Nature*, 431, 63-67.**
- BEATTY, C. D. & FRANKS, D. W. 2012. Discriminative predation: Simultaneous and sequential encounter experiments. *Current Zoology*, 58, 649-657.**
- CHITTKA, L. & OSORIO, D. 2007. Cognitive dimensions of predator responses to imperfect mimicry? *Plos Biology*, 5, 2754-2758.**
- CLUSELLA TRULLAS, S., VAN WYK, J. H. & SPOTILA, J. R. 2007. Thermal melanism in ectotherms. *Journal of Thermal Biology*, 32, 235-245.**
- CUTHILL, I. C. 2014. Evolution: The Mystery of Imperfect Mimicry. *Current Biology*, 24, R364-R366.**
- DARST, C. R. & CUMMINGS, M. E. 2006. Predator learning favours mimicry of a less-toxic model in poison frogs. *Nature*, 440, 208-211.**
- DITTRICH, W., GILBERT, F., GREEN, P., MCGREGOR, P. & GREWCOCK, D. 1993. Imperfect Mimicry - a Pigeons Perspective. *Proceedings of the Royal Society B-Biological Sciences*, 251, 195-+.**
- DOW, R. D. 2010. *Why are there imperfect Batesian mimics?* , University of Nottingham.**
- DUNCAN, C. J. & SHEPPARD, P. M. 1965. Sensory Discrimination and its Role in the Evolution of Batesian Mimicry. *Behaviour*, 24, 269-282.**
- EDMUNDS, M. 1978. On the association between *Myrmarachne* spp. (Salticidae) and ants. *Bulletin of the British Arachnological Society*, 4, 149-160.**

- EDMUNDS, M. 2000. Why are there good and poor mimics? *Biological Journal of the Linnean Society*, 70, 459-466.
- EDMUNDS, M. & READER, T. 2014. Evidence for Batesian mimicry in a polymorphic hoverfly. *Evolution*, 68, 827-839.
- ESTRADA, C. & JIGGINS, C. D. 2008. Interspecific sexual attraction because of convergence in warning colouration: is there a conflict between natural and sexual selection in mimetic species? *Journal of Evolutionary Biology*, 21, 749-760.
- FEENEY, W. E., STODDARD, M. C., KILNER, R. M. & LANGMORE, N. E. 2014. "Jack-of-all-trades" egg mimicry in the brood parasitic Horsfield's bronze-cuckoo? *Behavioral Ecology*, 25, 1365-1373.
- FRANKS, D. W. & NOBLE, J. 2004. Warning signals and predator-prey coevolution. *Proceedings of the Royal Society B-Biological Sciences*, 271, 1859-1865.
- GILBERT, F. 2005. The evolution of imperfect mimicry in hoverflies. *Insect Evolutionary Ecology*. CABI.
- GOLDING, Y. C., EDMUNDS, M. & ENNOS, A. R. 2005. Flight behaviour during foraging of the social wasp *Vespula vulgaris* (Hymenoptera : Vespidae) and four mimetic hoverflies (Diptera : Syrphidae) *Sericomomyia silentis*, *Myathropa florea*, *Helophilus* sp and *Syrphus* sp. *Journal of Experimental Biology*, 208, 4523-4527.
- GREEN, D. M. & SWETS, J. A. 1966. *Signal Detection Theory and Psychophysics*, New York, Wiley.
- HEGNA, R. H., SAPORITO, R. A. & DONNELLY, M. A. 2013. Not all colors are equal: predation and color polytypism in the aposematic poison frog *Oophaga pumilio*. *Evolutionary Ecology*, 27, 831-845.
- HEINRICH, B. 1984. Strategies of thermoregulation and foraging in 2 vespid wasps, *Dolicovespula-Maculata* and *Vespula-vulgaris*. *Journal of Comparative Physiology*, 154, 175-180.
- HOLLOWAY, G., GILBERT, F. & BRANDT, A. 2002. The relationship between mimetic imperfection and phenotypic variation in insect colour patterns. *Proceedings of the Royal Society B-Biological Sciences*, 269, 411-416.
- JOHNSTONE, R. A. 2002. The evolution of inaccurate mimics. *Nature*, 418, 524-526.
- JORON, M. & MALLETT, J. L. B. 1998. Diversity in mimicry: paradox or paradigm? *Trends in Ecology & Evolution*, 13, 461-466.
- KIKUCHI, D. W. & PFENNIG, D. W. 2010. High-model abundance may permit the gradual evolution of Batesian mimicry: an experimental test. *Proceedings of the Royal Society B-Biological Sciences*, 277, 1041-1048.
- KIKUCHI, D. W., SEYMOURE, B. M. & PFENNIG, D. W. 2014. Mimicry's palette: widespread use of conserved pigments in the aposematic signals of snakes. *Evolution & Development*, 16, 61-67.
- LANGMORE, N. E., STEVENS, M., MAURER, G., HEINSOHN, R., HALL, M. L., PETERS, A. & KILNER, R. M. 2011. Visual mimicry of host nestlings by cuckoos. *Proceedings of the Royal Society B-Biological Sciences*, 278, 2455-2463.
- LICHTER-MARCK, I. H., WYLDE, M., AARON, E., OLIVER, J. C. & SINGER, M. S. 2015. The struggle for safety: effectiveness of caterpillar defenses against bird predation. *Oikos*, 124, 525-533.
- MARSAGLIA, G. 1972. Choosing a Point from the Surface of a Sphere. *Ann. Math. Statist.*, 43, 645-646.
- MARTIN, A., PAPA, R., NADEAU, N. J., HILL, R. I., COUNTERMAN, B. A., HALDER, G., JIGGINS, C. D., KRONFORST, M. R., LONG, A. D., MCMILLAN, W. O. & REED, R. D. 2012. Diversification of complex butterfly wing patterns by repeated regulatory evolution of a Wnt ligand. *Proceedings of the National Academy of Sciences of the United States of America*, 109, 12632-12637.
- MATLAB 2012. MATLAB. Natick, Massachusetts: The Mathworks.
- MCGUIRE, L., VAN GOSSUM, H., BEIRINCKX, K. & SHERRATT, T. N. 2006. An empirical test of signal detection theory as it applies to Batesian mimicry. *Behavioural Processes*, 73, 299-307.

- MORGAN, K. R. & HEINRICH, B. 1987. Temperature regulation in bee-mimicking and wasp-mimicking syrphid flies. *Journal of Experimental Biology*, 133, 59-71.
- NELSON, X. J., LI, D. Q. & JACKSON, R. R. 2006. Out of the frying pan and into the fire: A novel trade-off for batesian mimics. *Ethology*, 112, 270-277.
- NUR, U. 1970. Evolutionary rates of models and mimics in Batesian mimicry. *American Naturalist*, 104, 477-8.
- OTTENHEIM, M. M. & KUIJT, J. 1998. *Has abdominal colour pattern variation in Eristalis arbustorum any thermoregulatory function?*
- PEKAR, S. & JARAB, M. 2011. Assessment of color and behavioral resemblance to models by inaccurate myrmecomorphic spiders (Araneae). *Invertebrate Biology*, 130, 83-90.
- PEKAR, S., JARAB, M., FROMHAGE, L. & HERBERSTEIN, M. E. 2011. Is the Evolution of Inaccurate Mimicry a Result of Selection by a Suite of Predators? A Case Study Using Myrmecomorphic Spiders. *American Naturalist*, 178, 124-134.
- PENNEY, H. D., HASSALL, C., SKEVINGTON, J. H., ABBOTT, K. R. & SHERRATT, T. N. 2012. A comparative analysis of the evolution of imperfect mimicry. *Nature*, 483, 461-U110.
- PFENNIG, D. W., HARCOTBE, W. R. & PFENNIG, K. S. 2001. Frequency-dependent batesian mimicry - Predators avoid look-alikes of venomous snakes only when the real thing is around. *Nature*, 410, 323-323.
- PFENNIG, D. W., HARPER, G. R., BRUMO, A. F., HARCOTBE, W. R. & PFENNIG, K. S. 2007. Population differences in predation on Batesian mimics in allopatry with their model: selection against mimics is strongest when they are common. *Behavioral Ecology and Sociobiology*, 61, 505-511.
- PFENNIG, D. W. & KIKUCHI, D. W. 2012. Competition and the evolution of imperfect mimicry. *Current Zoology*, 58, 608-619.
- PYKE, G. H., PULLIAM, H. R. & CHARNOV, E. L. 1977. Optimal Foraging - Selective review of theory and tests. *Quarterly Review of Biology*, 52, 137-154.
- R CORE TEAM 2014. R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing.
- RUXTON, G. D., SPEED, M. & SHERRATT, T. N. 2004. Evasive mimicry: when (if ever) could mimicry based on difficulty of capture evolve? *Proceedings of the Royal Society B-Biological Sciences*, 271, 2135-2142.
- SCHLUTER, P. M., XU, S. Q., GAGLIARDINI, V., WHITTLE, E., SHANKLIN, J., GROSSNIKLAUS, U. & SCHIESTL, F. P. 2011. Stearoyl-acyl carrier protein desaturases are associated with floral isolation in sexually deceptive orchids. *Proceedings of the National Academy of Sciences of the United States of America*, 108, 5696-5701.
- SHERRATT, T. N. 2002. The evolution of imperfect mimicry. *Behavioral Ecology*, 13, 821-826.
- 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., POLLITT, D. & WILKINSON, D. M. 2007. The evolution of crypsis in replicating populations of web-based prey. *Oikos*, 116, 449-460.
- SMITH, J. M., BURIAN, R., KAUFFMAN, S., ALBERCH, P., CAMPBELL, J., GOODWIN, B., LANDE, R., RAUP, D. & WOLPERT, L. 1985. Developmental constraints and evolution. *Quarterly Review of Biology*, 60, 265-287.
- SRYGLEY, R. B. 2004. The aerodynamic costs of warning signals in palatable mimetic butterflies and their distasteful models. *Proceedings of the Royal Society B-Biological Sciences*, 271, 589-594.
- TAYLOR, C. H., READER, T. & GILBERT, F. In Review. Why many Batesian mimics are inaccurate: evidence from hoverfly colour patterns. *Proceedings of the National Academy of Sciences of the United States of America*.
- TEAM, R. C. 2014. R: A language and environment for statistical computing. Vienna, Austria: R foundation for statistical computing.

WILLMER, P. G. & UNWIN, D. M. 1981. Field analyses of insect heat budgets - reflectance, size and heating rates. *Oecologia*, 50, 250-255.