Dynamics of mimicry evolution

NOÉL M. A. HOLMGREN* AND MAGNUS ENQUIST

Department of Zoology, Division of Ethology, Stockholm University, SE-106 91 Stockholm, Sweden

We simulated mimicry evolution by allowing three populations to coevolve: two populations of senders and one of receivers. Artificial neural networks were used to model receivers, and it was assumed that recognition was inherited. The senders' signals consisted of nine dimensions. Changes to receivers and senders were caused by random mutations during the course of the simulation. Whereas it paid both types of senders to elicit the same response from the receiver, it benefited the receiver to respond in this way only towards one of the sender types. The receiver was thus in conflict with one of the senders, e.g. as in Batesian mimicry. Monotonically increasing response gradients caused the appearance of the model and the mimic to move in the same direction. Mimicry evolved because the mimic approached the model faster than the model moved away. Even after mimicry was established the model and the mimic were constantly changing in appearance. Our results conform with what is known in comparative psychology and ethology about how animals respond to stimuli. Several of our results are a direct consequence of recognition and have not, to our knowledge, been reported before, showing the importance of considering the recognition mechanism in detail when studying mimicry.

© 1999 The Linnean Society of London

ADDITIONAL KEY WORDS:—artificial neural network – Batesian mimicry – communication – imitation – signalling.

CONTENTS

*Corresponding author. Current address: Department of Natural Sciences, University of Skövde, Box 408, SE-541 28 Skövde, Sweden. Email: noel.holmgren@inv.his.se

INTRODUCTION

Mimicry was suggested as one of the first major pieces of evidence in support of Darwin's theory of evolution through natural selection (Wallace, 1866; Poulton, 1898). As such it has received much attention (e.g. Müller, 1879; Dixey, 1897; Marshall, 1908; Dixey, 1909; Fisher, 1927, 1930; Nur, 1970; Turner, 1977, 1987; Sheppard *et al.*, 1985).

Mimicry typically involves three players: two senders and one receiver. For mimicry to evolve it is essential that both senders benefit from eliciting the same response from the receiver. This provides the incentive for looking alike. Two situations can occur depending on whether it pays the receiver to react with a favoured response towards only one of the senders or to both. Classical examples are Batesian and Müllerian mimicry. Both involve aposematic/warning signalling to a predator, and the favoured response is to not be attacked. In Batesian mimicry a palatable prey species mimics the appearance of another species noxious to predators thus reducing its risk of being attacked (Bates, 1862). In Müllerian mimicry, two aposematic noxious forms conform to the same aposematic signal to their mutual benefit (Müller, 1879) and there are no conflicts with the receiver.

Several hypotheses have been generated to explain how the appearance of the senders (or their signals) changes or moves relative to each other during evolution. When conflicts occur, the terms 'model' and 'mimic' are generally used to identify the two senders; this terminology is also adopted here. The model and the receiver have mutual interests, whereas the mimic is the player with interests that conflict with those of the receiver. These terms suggest that it is the appearance of the mimic that changes and becomes similar to that of the model. In fact, it has been argued that the appearance of the model does not change at all during evolution (e.g Dixey, 1897). This reasoning is based on the assumption that the response gradient reaches its maximum with the average appearance of the model. Thus variations (mutations) in the model will be generally less efficient. Throughout this paper we will refer to response gradients when we describe how the receiver's sensitivity to stimulation varies along different dimensions.

Fisher (1930) argued against Dixey and claimed that "selection will tend to modify the model so as to render it different from the mimic and as conspicuous as possible". However, he retained the idea that the gradient reaches its maximum with the average appearance of the model. His explanation was based on the assumption that the gradient is asymmetrical so that mutations away from the mimic are less disadvantageous than those towards the mimic.

More recently, researchers (Nur, 1970; Turner, 1977, 1987; Sheppard *et al.*, 1985) have argued that the model moves away from the mimic but at a slower rate. The mimic must be able to approach the model faster than the model moves away for mimicry to become established (Fisher, 1930). The explanation of why the model tends not to move remains the one offered by Dixey (1897). Attempts to explain why models should move away from the mimic are however more vague. It seems necessary for the appearance of the model to change during evolution, otherwise Müllerian mimicry becomes problematic. Here both senders can be viewed as models for each other. If the appearances of the two senders are dissimilar and both are at their own response maximum can they ever evolve into one signal? A way out of this dilemma has been to introduce a two-step process (Nicholson, 1927; Fisher, 1930; Ford, 1963). The first step is taken by the mutation of a modifier gene

regulating a super-gene complex, bringing one species closer in appearance to the other. In the second step, minor adjustments occur resulting in a perfect resemblance.

We believe that the issues discussed above can only be resolved if we consider in detail the properties of the receiver's recognition mechanism. Most studies of mimicry have focused on strategic factors (but see Fisher, 1927; Turner, 1981; Dittrich et al., 1993). Those studies taking behaviour mechanism into account have focused on learning and forgetting rules (e.g. Turner & Speed, 1996), not on the abilities of generalization and discrimination of the receiver. An important reason why mechanisms of recognition traditionally have not been included in models of mimicry evolution was the lack of useful models of biological recognition.

In this paper we will make use of recent advances in modelling nervous systems—termed artificial neural networks—and use them as a model of the receiver's recognition mechanism. Artificial neural networks have provided realistic models of stimulus control (Maki & Abunawass, 1991; Pearce, 1994; Enquist & Arak, 1998; Ghirlanda & Enquist, 1998; but see Dawkins & Guilford, 1995; Kamo et al., 1998) and have been used for studying the evolution of signal form (see e.g. Enquist & Arak, 1993). One decisive advantage of artificial neural network models when compared with other models of recognition is their ability to generalize in a realistic fashion (Ghirlanda & Enquist, 1998). This allows us to simulate the evolution of signals in which the form of signals can change freely in many dimensions.

The aim of this paper is to gain insight into the role of recognition in mimicry evolution. In evolutionary simulations we have investigated a case with a conflict between the receiver and one of the senders. We assume that signals as well as recognition are genetically inherited.

THEORY

The simulation involves three populations (species): one population of receivers and two of senders. The receivers are artificial neural networks that are able to respond to the signals of the senders. All populations consist of a number of different types of individuals and the number of each type increases or decreases depending on their fitness. New types originate by a mutation of existing types. The fitness of a receiver type is determined by its responses to the senders and a stimulus that does not evolve. The non-evolving stimulus is a stimulus in the receiver's surroundings that requires a different response from that of the model. In reality, many stimulus-response relationships are coded together in the nervous system. We included this stimulus in order to prevent the network from becoming oversensitive to stimulation. The fitness of a sender type is based on the ability to elicit the favourable response ('identified as model') from the receivers.

The signals are complex in the sense that they are made up of an array of values, i.e. the signals are multidimensional (Shepard, 1987). We will often describe a signal as a point in a signal space, and signal evolution as movement in this space. The distance between two points in the signal space is expressed by the Euclidean distance. The signals may represent any kind of stimuli: visual, olfactory or acoustic. The neural network coding of signals does not deal with sense specific properties of the signals or sense organs, but focuses instead on the coding of the information in the neurones. Hence, the model is equally applicable to all sensory systems.

TABLE 1. Simulation parameters for the receiver

Population size	1000
Max number of types	100
Mutants per generation	1
Fitness, response to model	1.0
Fitness, response to mimic	−1.0
Fitness, response to irrelevant	−1.0
Fitness, no response	0.0
Network, input cells	9+1
Network, hidden cells	5
Network, output cells	1

Details of types

The individuals of the receiver population are modelled as artificial neural networks or more specifically as multi-layered perceptrons (see e.g. Haykin, 1994). The network is organized in three layers: a receptor layer consisting of nine receptor cells plus a bias neurone, an intermediate layer of five cells, and an output layer of one cell. Each cell in a layer is connected with a synapse to all cells in the next layer. The total number of synapses is thus 55. The output of the network ranges between 0 and 1 and is interpreted as the probability of response—hereafter called the response.

The signals and the irrelevant stimulus consist of nine dimensions, with intensity between 0 and 1 in each dimension. The irrelevant stimulus has the value 0.5 in all nine dimensions. The populations of receiver and senders can consist of up to 100 different types. The total number of individuals is fixed, whereas the number of individuals of each type varies during the simulations. For more details see Tables 1 and 2.

Mutations of the types

From each of the three populations, a proportion (0.001) of the individuals is randomly selected for mutation in each generation. If the number of mutants does not turn out as an integer a random number between 0 and 1 is added and then the new number is truncated. Thus, the nearest upper or lower integer is alternatively used so that the average equals the original number. When an individual is selected for mutation, each of its units (i.e. synaptic weights or signal dimensions) is mutated with a probability that on average results in two units being modified. Receivers are mutated changing each synaptic weight with a probability of 2/55 (55 is the number of synapses in the network) by adding a random normal distributed number ($\mu=0$, $\delta=0.2$) to its previous value. Senders are mutated by modifying each dimension value with a probability of 2/9 (nine is the number of dimensions) by adding a random normal distributed number ($\mu=0$, $\delta=0.05$) limited to between 0 and 1. These parameters are selected to be large enough to enhance the speed of simulation, but low enough to have an efficient selection. The number of types is kept within boundaries due to a balance between mutations and the extinction of rare types (i.e. drift-mutation equilibrium).

TABLE 2. Stimulation parameters for senders and other stimuli

	Model	Mimic	Irrelevant stimulus†
Population size	1000	500	1000
Max number of types	100	100	1
Mutants per generation (average)	1	0.5	
Fitness, response	1.0	1.0	--
Fitness, response	0.0	0.0	
Stimulus dimensions	9	9	9

†Technically the irrelevant stimulus is treated as a population.

Next generation propagation

Reproduction is simplified as semelparous and asexual. The individuals of the next generation are copies of their parent type, or if they are selected for mutation, a new type slightly modified from the parent type. Parents are selected randomly weighted by their relative fitness.

The three populations are denoted R, for the receiver population, and M and N for the two sender populations. The types in each population are denoted by small letters so that $r \in R$, $m \in M$, and $n \in N$. We express the fitness contribution to an individual of type i from the interaction of members of population J as the sum of all responses to all individuals member of J:

$$w_j^i = \sum_j \xi_j \rho(i,j). \tag{1}$$

Here, ξ is the number of individuals of type j, and $\rho(i,j)$ is the response probability (the receiver output) from the interaction between types i and j. One of the types must be a sender and the other must be a receiver.

The fitness function for a receiver of type r is then given by the sum of its interactions with the two sender populations and the irrelevant stimulus i

$$W_r = dw_M^i + ew_N^i + hw_i^i. \tag{2}$$

The parameters, d, e and h are interaction parameters determining the payoff for the receiver to interact with the two different senders and the irrelevant stimulus. The fitness parameters were set so that the receiver would benefit when responding to the model ($d = 1$) and punished when responding to the mimic ($e = -1$). Response to an irrelevant stimulus was also punished ($h = -1$). The population size of the model was set to 1000 and the population size of the mimic was set to 500. For the irrelevant stimulus, which occurs in only one type, its 'encounter frequency' ξ_i is set to 1000. For more details see Table 2. Similarly the fitness functions for a sender of type m and n are

$$W_m = fw_R^m \tag{3}$$
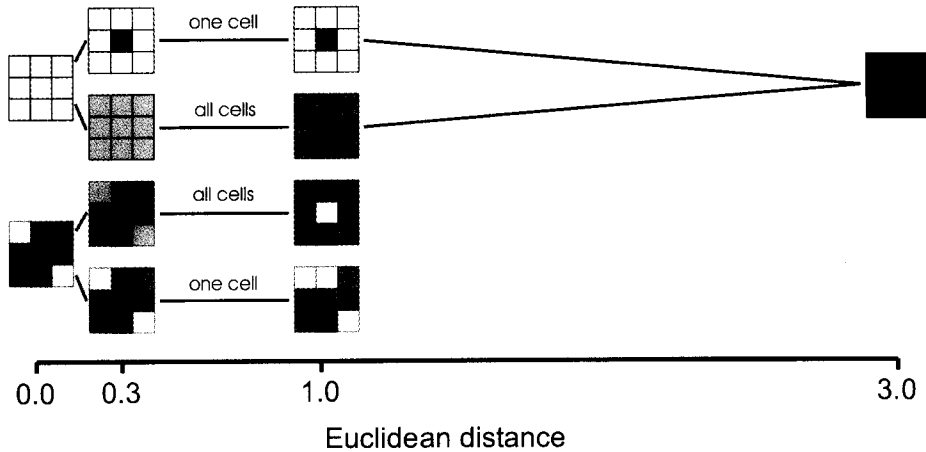
$$W_n = gw_R^n \tag{4}$$

respectively.

Figure 1. Some visualized examples of Euclidean distances in signals of nine grey scale dimensions. Two patterns, one homogeneous and one heterogeneous, are changed in one or several dimensions.

The simulations are initiated with all senders in each population belonging to one type, assigned a random signal. In addition, the receiver population is initiated as one type. The synaptic weights are here set randomly to a value between 0.1 and -0.1. This is a standard procedure to avoid initial stalling in the simulations (Haykin, 1994). Each individual simulation was run for 50 000 generations.

Euclidean distance

In order to compare more easily the signals of the two sender populations we calculated the average signal in each population. The average signal is simply the signal with the average value in each dimension. The distance between two signals is measured by the Euclidean distance:

$$D(\mathbf{a},\mathbf{b}) = \sqrt{\sum_i (a_i - b_i)^2} \tag{5}$$

where i is a dimension of the signal vectors $\mathbf{a}$ and $\mathbf{b}$ emitted by two senders. When using nine dimensions as in our case, the maximum dissimilarity is 3. The expected distance between random signals is 1.2. Figure 1 gives examples of Euclidean distances if the signals are given a visual interpretation.

RESULTS

The simulations were run for 50 000 generations and repeated 100 times. The number of generations simulated was well beyond that at which they reached
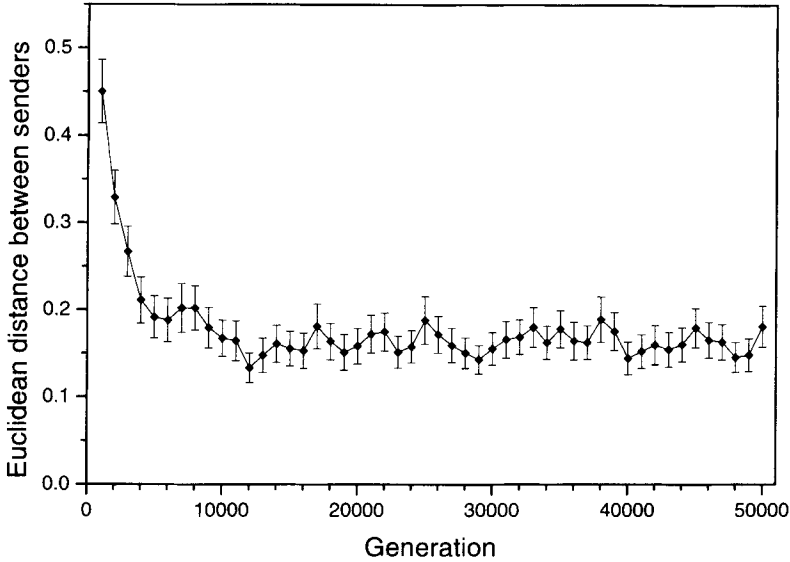
Figure 2. Euclidean distance between the average signals of the two senders during the course of simulations. Simulated data is shown for every 1000th generation. Error bars indicate standard error of estimated mean.

equilibrium in terms of distance between the senders (Fig. 2). The distance between the model and the mimic settled to an average of $D=0.16$ after about 10000 generations. The receiver learned very quickly to have a low response to the non-evolving stimulus, which it kept throughout the simulations (data not shown).

The movements of the model and the mimic depend on their relative distance (Fig. 3). When the model and the mimic are rather similar ($D<0.2$), they both move away from each other. Otherwise, the mimic tends to approach the model, while at the same time the model tends to move away from the mimic. The strength of these movements weakens as the distance between the types decreases. The movement away by the model is, however, for most of the range less than the approaching tendency by the mimic. Calculated over the whole data set the average escaping movement of the model is on average smaller ($\Delta D=4.4 \cdot 10^{-5}$ per generation) than the average approaching movement by the mimic ($\Delta D=5.8 \cdot 10^{-5}$ per generation). The dynamics suggest an equilibrium difference between the model and the mimic in the interval of $D=0.15$–0.2. This fits with the observed equilibrium obtained after 10000 generations of $D=0.16$ (Fig. 2). However, the distribution of Euclidean distances between the average signal of the model and the mimic has its maximum in the interval $D=0$–0.05 (Fig. 3). This probably reflects the truncation of the distribution to the left, i.e. the Euclidean distance can not be negative.

The response gradients typically show a monotonical increase (Fig. 4). This suggests that in order to elicit the highest response, the model and the mimic move in the same direction. The average response to the model and the mimic is 0.97 and 0.95 respectively, which indicates that the response gradients most commonly increase towards the model. Not surprisingly, the receiver is generally better able to separate the two senders (higher response difference) the more dissimilar they are in appearance (Fig. 5). The response to the model and the mimic decreases the
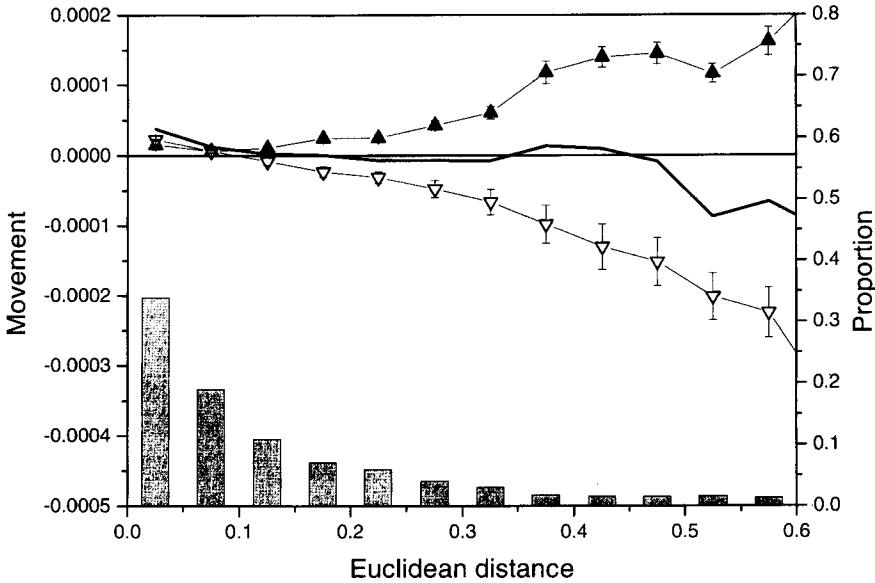
Figure 3. Average relative movement of model and mimic as a function of average Euclidean distance. The relative movement of the model (solid upward triangle) is calculated as $D(\tilde{m}_{t+1}, \tilde{n}_u) - D(\tilde{m}_t, \tilde{n}_u)$, where D is the Euclidean distance, $\tilde{m}$ the average signal vector in the model population, $\tilde{n}$ is the average signal vector in the mimic population, t is the current generation and hence $t+1$ is the next. The movement of the mimic (open downward triangle) is calculated in a corresponding way. The net movement is denoted by the line. Positive movement indicates movement resulting in an increased distance between the two senders (i.e. away), and hence a negative movement indicates a decrease in sender distance (i.e. approach). Error bars are standard errors of estimated mean. Bars at the bottom shows the proportion of all observations ($n = 4\,999\,900$) in each distance range.

more dissimilar they are (Fig. 5). As the distance between the model and the mimic increases, the response gradient becomes steeper (Fig. 5), and consequently the rate of signal change increases (Fig. 3). Consequently, events when the receiver is able to make a clear distinction between the model and the mimic are rare. At these instances, there is a strong selection for the mimic to become more similar to the model. When the two senders are very similar, the generalization gradient in the vicinity is rather flat. In this range, both senders move away from each other, mainly due to random mutations.

The monotonically increasing gradients could be the cause for the model and the mimic signal to evolve towards the corners of the signal space (i.e. all dimensions are at one end of the extreme; Fig. 6). In the second half of the simulations the model is on average a distance of $D = 0.15$ away from the nearest corner in this space. The corresponding figure for the mimic is $D = 0.16$. In most cases, the senders were at one edge in the stimulus space (i.e. the senders were at the extreme in least one dimension).

Figure 3 may give the impression that the model always moves away from the mimic. If we resolve movement data on the basis of the response difference we can actually see that the roles as 'model' and 'mimic' can occasionally be reversed (Fig. 7). In fact, the response to the mimic is higher than to the model in 39.5% of all generations. Note however that in the majority of these cases the model and the mimic are very similar in appearance (Fig. 7).
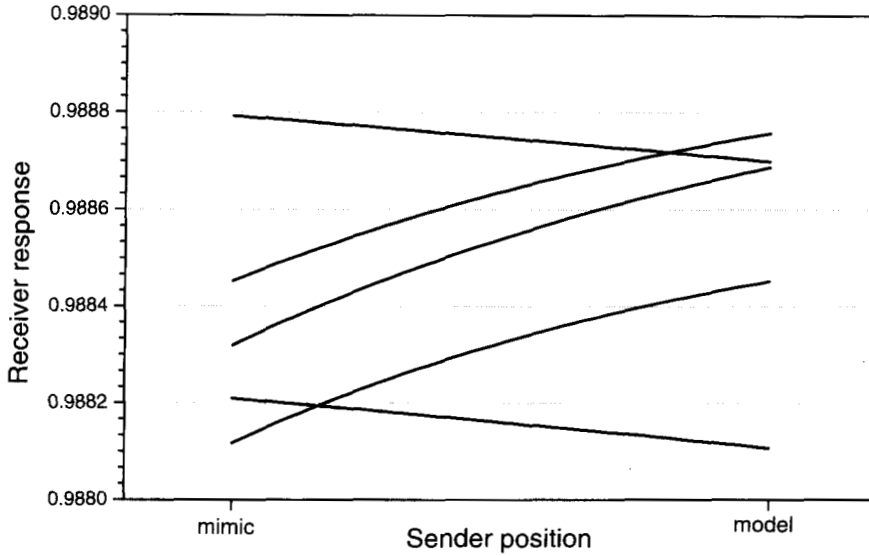
Figure 4. Five examples of response gradients showing their typically monotonical increase. The gradients were obtained by saving the networks of the most common type together with the average signal of the model and the mimic at regular intervals during the simulations. After the simulations, a set of stimuli was generated. These stimuli are points on a vector in the stimulus space which intersects through the average model and the average mimic signal. The vector does not extend beyond the model of the mimic, which indicates that both senders are at the edge of the stimulus space. The response functions shown are thus each the response of one of the saved networks to the average model signal, the average mimic signal and the signals in between.
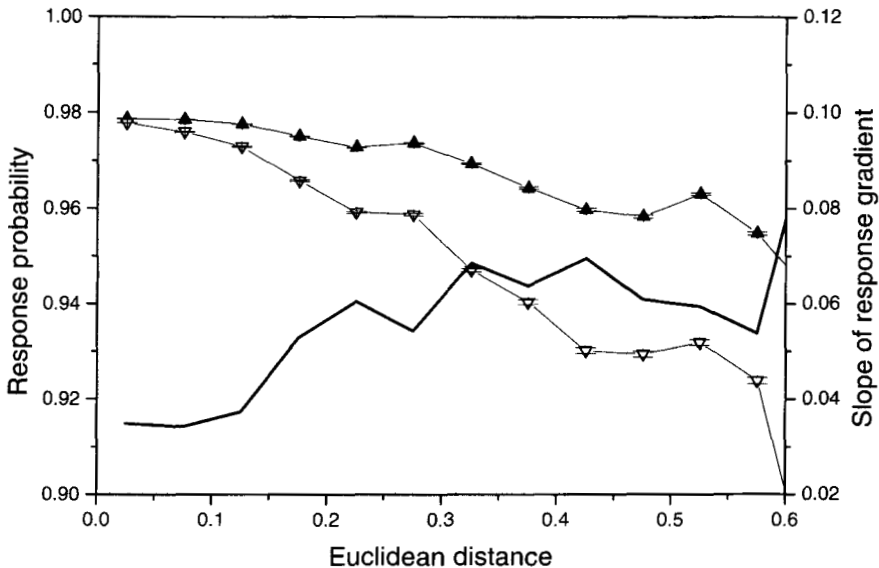

Figure 5. Average receiver response to model (▲) and mimic (▽) as a function of the average Euclidean distance between the two senders. The thick solid line indicates the slope of an approximated linear response gradient along the vector intersecting model and mimic average averages. Error bars are standard errors of estimated mean.
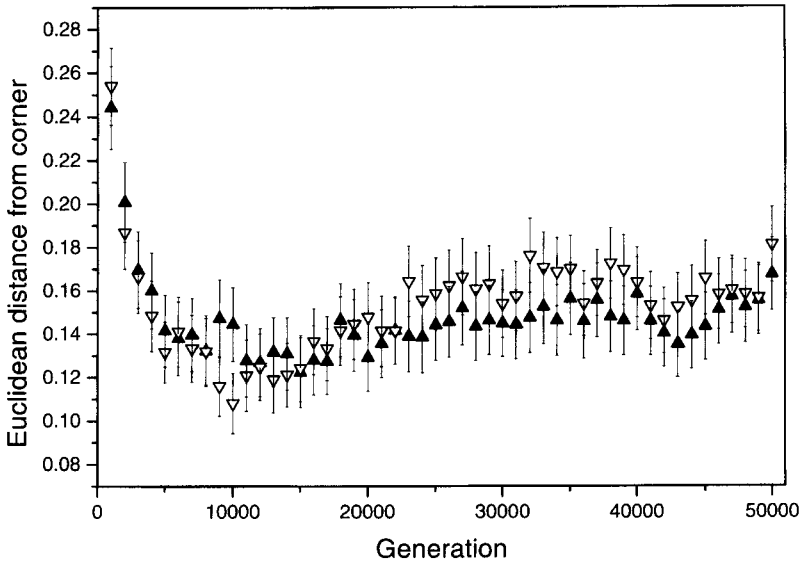
Figure 6. Average Euclidean distance of model (▲) and mimic (▽) from the nearest corner at every 1000th generation of the simulation. Error bars are standard errors of estimated mean.
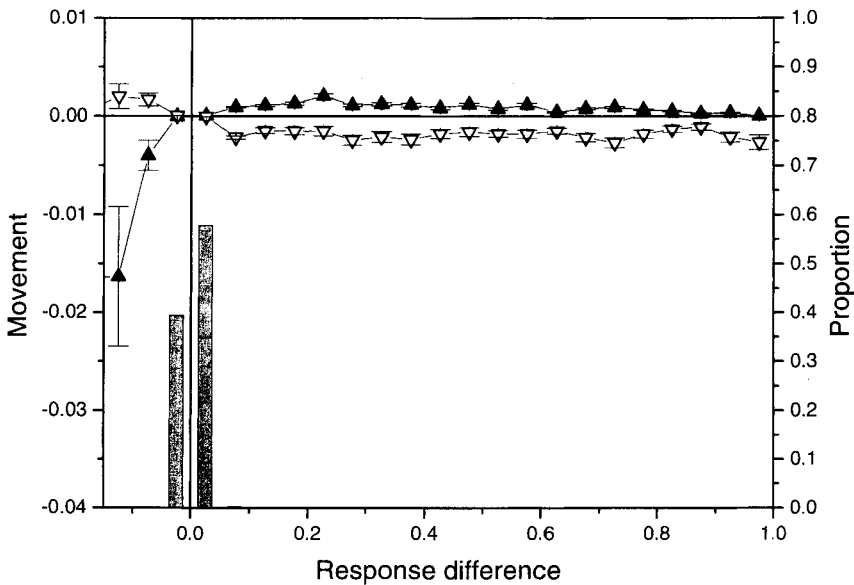

Figure 7. Average change per generation in the Euclidean distance between the model and the mimic as a function of difference in average response. The response difference is measured as the difference in average response of the receiver population to the model minus the average response to the mimic. The movements are calculated as in Fig. 3. Error bars are standard errors of estimated mean. Bars at the bottom shows the proportion of all observations ($n = 4999900$) in each difference range.
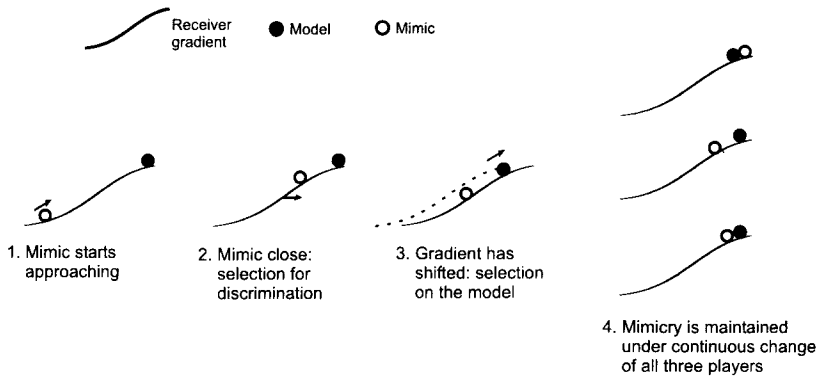
Figure 8. Principle of most common response gradients along one dimension. The network usually develops a response to the model and the mimic along a dimension in which they are distinguishable. The response to the model is then higher than to the mimic (left graph), but the selection gradient (the derivative of the response curve) may be highest for the mimic (Fig. 4). As a consequence, the mimic will move faster to the right than the model, and occasionally appear ahead of the model eliciting a higher response than the former. In these reversed positions, the roles of the two senders are also reversed, the mimic now being a model for the 'model'.

DISCUSSION

Our results show that the modelling approach attempted here is a profitable one for studying mimicry, an evolutionary problem that offers a number of challenges to theoretical work. Of course, since this is the first attempt at modelling mimicry with a new technique the results generate as many questions as they answer. It is clear, though, that in contrast to earlier thinking, mimicry games may be more dynamic and that the players constantly change in appearance.

One of our most interesting results is that while in most generations the mimic approached the model, the model simultaneously tended to move away from the mimic. Occasionally this pattern was reversed. We can understand these results in light of the recognition mechanism by considering response gradients. Typically, neither the model nor the mimic is located where the receiver is most likely to react with the favoured response. Instead both players are evolving towards the response peak with the model usually being somewhat ahead, i.e. closer to the peak than the mimic (See Fig. 8). The reason why the model is not at the response peak is that the receiver constantly tries to discriminate between the model and the mimic. In doing so it may pay the receiver to take the cost of being slightly less sensitive to the model if this is compensated by a larger drop in sensitivity towards the mimic. In practice, this reorganization of the memory causes a displacement of the response peak away from both the model and the mimic. This displacement or peak shift causes the model to depart from the mimic. The existence of peak shift and supernormal stimuli are general results from studies of how stimuli control behaviour (e.g. Mackintosh, 1974; Hinde, 1970; Rilling, 1977, Enquist & Arak, 1998).

For mimicry to evolve it is of course necessary that the mimic approaches the model faster than the model moves away as long as the appearance of the two players is different (e.g. Nur, 1970). This was also clearly seen in our simulations (See Fig. 3). Two factors seem to contribute to this result. One is the tendency for the response gradient to be steeper for the mimic than the model (Fig. 4). The other

factor is the edges and corners of the stimulus space which limit the movement of both players and in particular hinders the model from moving away. For instance, if red hue is selected, this dimension will eventually be exhausted. When this happens, an edge of the stimulus space has been reached and no further evolution is possible along that particular direction. This restricts the mobility of the model; while evolution may occur in other dimensions, it gives the mimic an increased possibility for approaching the model in appearance.

Recent theory on Batesian mimicry evolution views response gradients as bell-shaped and considers the problem along a single dimension (Sheppard *et al.*, 1985; Turner, 1981; Fig. 3). Although this picture is intuitively appealing, it ignores the fact that multi-dimensional signal space separation occurs in many dimensions and many evolutionary trajectories are possible. It has been discussed that both empirical data and theory suggest that the steepest response gradient is a monotonically increasing function and not bell shaped (Ghirlanda & Enquist, 1998, in prep. See also intensity generalization; Mackintosh; 1974). To compare the single dimensional theory by Sheppard *et al.* (1985) and Turner (1981), and our multi-dimensional theory, we have to acknowledge that mimicy evolution has two phases: the initial phase when a signal of a species becomes similar to another species signal, and a second phase in which the mimicry is maintained. The single dimensional theory assumes that the two species are initially separated in one dimension and are protected by species-specific bell-shaped response (i.e. the predators' neural responses) which curves around each species mean appearance. Assuming that the species can only change its appearance along the single dimension, this creates a problem of how the second species could become a mimic of the model since intermediate forms would be less protected.

Monotonically increasing gradients and multidimensionality may explain why a mimic can approach the model when being very dissimilar in appearance. The multi-dimensional theory of mimicry evolution based on our simulations is depicted in Figure 8. Mimicry evolution starts with the mimic beginning to approach the model (Step 1 in Fig. 8). When the mimic comes close in appearance to the model selection starts to act on the receiver to discriminate between the model and the mimic (Step 2 in Fig. 8) leading to a change in the response gradients. One consequence of the latter is that the appearance of the model no longer resembles the appearance that will elicit the highest response rate from the receiver. Thus the model is now under directional selection (Step 3 in Fig. 8), typically away from the appearance of the mimic. Eventually, mimicry is established, although not as a stable situation. Instead the three players constantly change; the degree of mimicry varies within limits and it is not always the model that is most efficient in eliciting the favoured response (Step 4 in Fig. 8). The implication of this is that there is no need for a special mechanism such as the mutations of modifier genes (see e.g. Turner, 1977) to bring the mimic reasonably close to the model.

The concept of dimensions may be abstract, but this is deliberate in order to stress the generality of the mechanisms. In a sender-receiver interaction the dimensions of the signal may differ considerably between its source of generation in the sender and its implementation in the receiver. For example, the dimensions of a signal may be defined on the basis of the genes coding for its expression, but are rather different during its transmission in a physical element—for instance, as sound or light. On the receptor level of the receiver, visual impressions have a lot of spatial dimensions in the form of cones and rods behind the retina. Sound impressions have as many

dimensions as frequency-unique receptor hairs in the cochlea. It is apparent that the dimensions at each receiver's receptor level does not correspond directly to the dimensions of the sender genes that control the signal expression. At higher perceptual levels, the sensory dimensions have been reduced to a smaller number of feature dimensions such as colour or shape, which may correspond more closely to the dimensions on the sender level. Our artificial neural networks represent this higher level of perceptual processing.

The continuous movements of the model and the mimic have several consequences. The speed of evolution will be higher. There are also likely to be consequences for signal form. The presence of the mimic may push the model to a more extreme appearance (e.g. saturated colours, contrast with background, wild morphological structures). This is likely to be particularly important when such exaggerations are costly (See Enquist & Arak, 1998, for a general discussion of this).

Recognition was inherited rather than learned in our simulations. This requires comment since learning is important in many famous mimicry cases such as palatable butterflies mimicking poisonous ones, thereby avoiding attacks from birds (Brower, 1958a–c; Benson, 1972). Mimicry, however, also occurs when recognition is primarily inherited. For instance, mimics have exploited mate recognition in insects many times. One such example is male fireflies of the genus *Photinus* that are attracted to the flesh-responses of predatory females of the related genus *Photuris* which mimic those of *Photinus* females (Lloyd, 1965). Another example is the ability of some orchids to mimic female wasps and thereby attract males, which as a consequence carry pollen between flowers (Dafni, 1987). Obviously, our results are relevant to cases such as those just described in which recognition is mainly inherited. One may argue on theoretical grounds that the results may also have relevance for cases with learned response. If each receiver interacts many times with senders, the receiver will rarely be in the learning phase when it is encountered by a sender. When the selection for sender signals comes mainly from trained receivers, the difference depends less on whether the response is inherited or learned.

REFERENCES

Bates HW. 1862. Contributions to an insect fauna of the Amazon Valley. Lepidoptera: Heliconidae. *Transactions of the Linnean Society of London* **23:** 495–566.
Benson WW. 1972. Natural selection for Müllerian mimicry in *Heliconius erato* in Costa Rica. *Science* **176:** 936–939.
Brower JVZ. 1958a. Experimental studies of mimicry in some North American butterflies. Part I. The monarch *Danaus plexippus*, and the viceroy, *Limenitis archippus archippus. Evolution* **12:** 32–47.
Brower JVZ. 1958b. Experimental studies of mimicry in some North American butterflies. Part II. *Battus philenor* and *Papilio troilus, P. polyxenes* and *P. glaucus. Evolution* **12:** 123–136.
Brower JVZ. 1958c. Experimental studies of mimicry in some North American butterflies. Part III. *Danaus gilippus bernice* and *Limenitis archippus floridensis. Evolution* **12:** 273–285.

Dafni A. 1987. Pollination in Orchis and related genera: Evolution from rewards to deception. In: Arditti J, ed. *Orchid Biology, Reviews and Perspectives* Ithaca: Cornell University Press, 80–104.

Dawkins MS, Guilford, T. 1995. An exaggerated preference for simple neural network models of signal evolution? *Proceedings of the Royal Society of London* **B261:** 357–360.

Dittrich W, Gilbert F, Green P, McGregor P, Grewcock D. 1993. Imperfect mimicry: a pigeon's perspective. *Proceedings of the Royal Society of London* **B251:** 195–200.

Dixey FA. 1897. Mimetic attraction. *Transactions of the Entomological Society of London* **1897:** 317–332.

Dixey FA. 1909. On Müllerian mimicry and diaposematism. *Transactions of the Entomological Society of London* **1909:** 559–583.

Enquist M, Arak A. 1993. Selection of exaggerated male traits by female aesthetic senses. *Nature* **361:** 446–448.

Enquist M, Arak A. 1998. Neural representation and the evolution of signal form In: Dukas R, ed. *Cognitive ecology* Chicago: Chicago University Press.

Fisher RA. 1927. On some objections to mimicry theory; statistical and genetic. *Transactions of the Entomological Society of London* **1927:** 269–278.

Fisher RA. 1930. *The genetical theory of natural selection.* Oxford: Clarendon Press.

Ford EB. 1963. Mimicry. *Proceedings of the XVI International Congress of Zoology,* 184–186.

Ghirlanda S, Enquist M. 1998. Artificial neural networks as models of stimulus control. *Animal Behaviour* (in press).

Haykin S. 1994. *Neural networks, a comprehensive foundation.* New York: MacMillan College Publishing Company.

Hinde RA. 1970. *Animal behaviour: a synthesis of ethology and comperative psychology.* Tokyo: McGraw-Hill Kogukusha.

Kamo M, Kubo T, Iwasa Y. 1998. Neural network for female mate preference, trained by a genetic algorithm. *Philosophical Transactions Royal Society of London* **B353:** 399–406.

Lloyd JE. 1965. Aggressive mimicry in *Photuris:* firefly femmes fatales. *Science* **149:** 653–654.

Mackintosh NJ. 1974. *The psychology of animal learning.* London: Academic Press.

Maki WS, Abunawass AM. 1991. A connectionist approach to conditional discriminations: Learning, short-term memory, and attention. In: Commons ML, Grossberg S, Staddon JER, eds. *Neural network models of conditioning and action.* Hillsdale, NJ: Earlbaum, 241–278.

Marshall GAK. 1908. On diaposematism, with reference to some limitations of the Müllerian hypothesis of mimicry. *Transactions of the Entomological Society of London* **1908:** 93–142.

Müller F. 1879. *Ituna and Thyridis;* a remarkable case of mimicry in butterflies. *Transactions of the Entomological Society of London* **1879:** xx–xxix.

Nicholson AJ. 1927. A new theory of mimicry in insects. *Austral. Zoology* **5:** 10–10.

Nur U. 1970. Evolutionary rates of models and mimics in Batesian mimicry. *American Naturalist* **104:** 477–486.

Pearce JM. 1994. Similarity and discrimiation: a selective review and a connectionist model. *Psychological Review* **101:** 587–607.

Poulton EB. 1898. Natural selection: the cause of mimetic resemblance and common warning colours. *Linnean Society Journal of Zoology* **26:** 558–612.

Rilling M. 1977. Stimulus control and inhibitory process In: Honig WK, Staddon JER, eds. *Handbook of operant behavior.* Englewood Cliffs: Prentice-Hall, 432–480.

Shepard RN. 1987. Toward a universal law of generalization for psychological science. *Science* **237:** 1317–1323.

Sheppard PM, Turner JRG, Brown KS, Benson WW, Singer MC. 1985. Genetics and the evolution of Müllerian mimicry in *Heliconius* butterflies. *Philosophical Transactions Royal Society of London* **308:** 433–613.

Turner JRG. 1977. Butterfly mimicry: the genetical evolution of an adaptation. In: Hecht MK, Steere WC, Wallace B, eds. *Evolutionary Biology* New York: Plenum Press, 163–206.

Turner JRG. 1981. Adaptation and the evolution in *Heliconius:* a defence of neo-Darwinism. *Annual Review of Ecology and Systematics* **12:** 99–121.

Turner JRG. 1987. The evolutionary dynamics of Batesian and Muellerian mimicry: similarities and differences. *Ecol. Entomology* **12:** 81–95.

Turner JRG, Speed MP. 1996. Learning and memory in mimicry. I. Simulations of laboratory experiments. *Philosophical Transactions Royal Society of London* **B351:** 1157–1170.

Wallace AR. 1866. On the phenomena of variation and geographical distribution, as illustrated by the Papilionidae of the Malayan region. *Transactions Linnean Society of London* **25:** 1–72.