

Darwin's principle of antithesis revisited: a role for perceptual biases in the evolution of intraspecific signals

PETER L. HURD, CARL-ADAM WACHTMEISTER
AND MAGNUS ENQUIST

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

SUMMARY

In addition to their tendency towards exaggeration, biological signals also tend to be quite distinct from each other. This is not just true when compared across species and contexts, but also between intraspecific displays of very similar function, such as threat displays. General biases in recognition mechanisms may be responsible for this effect. Previous artificial neural network simulations have demonstrated that perceptual biases in the receiver may cause a coevolving signal to polarize away from another, non-evolving, stimulus. In this paper we extend this work by investigating networks which respond differently to several different stimuli. We show that two relevant signals, ideally producing different responses, also evolve towards converse forms despite being processed through the same network. We found no evidence that attractive signals share common attractive properties. Our results can not be accounted for by existing strategic models of communication.

1. INTRODUCTION

The question of why signals used for species recognition tend to be distinctly different has historically received much attention (see, for example, Dobzhansky 1951; Mayr 1963; Andersson 1994). By comparison there have been relatively few attempts to explain why different signals used by members of the same species often take on very distinct and contrasting forms. Looking beyond the commonly considered property of exaggeration will provide us with more data with which to evaluate hypotheses concerning the evolution of signals.

That different displays within a species tended to be not just different from each other but antithetical was noted by Darwin (1872). As he described it, each component of the threat and submission postures of the dog (head position, tail position, raising of fur, etc.) are all at one exaggerated extreme or the other, and further, they are at opposite extremes in the two displays. A threatening dog raises its ears, tail and hackles, and lowers its head, while a submissive dog lowers its ears, tail and hackles, and raises its head. Similarly the great tit threat displays vary widely in form; different aspects of body erectness, head-up against head-down postures and wing extension are not just extreme, but at reverse extremes in different displays (Blurton Jones 1968). Unlike the dog, the displays of the great tit are all threats, presumably with rather similar functions. To what extent can existing theories explain such observations?

Darwin argued that:

Certain states of mind lead, as we have seen... to certain

habitual movements which are primarily, or may still be, of service; and we shall find that when a directly opposite state of mind is induced, there is a strong and involuntary tendency to the performance of movements which are of a directly opposite nature, though these have never been of service.

Darwin attributed this effect to a psychological mechanism controlling the use of muscles in different states of mind. We are not aware of this hypothesis winning any great support.

The classical explanation has been that exaggeration and distinctiveness of display have both evolved to prevent mistakes (Morris 1957; Cullen 1966). This propagation hypothesis holds that differences between signals are required physically to transmit a signal through a noisy world which tends to degrade it. However, propagation is unable to account for many examples of distinct displays. For instance, an aposematic prey item benefits by deterring predators who have already spotted it, not by gaining the attention of distant predators. The propagation hypothesis is also unable to explain the distinctiveness of short-range displays. Great tits use exaggerated threat postures at distances that are negligible compared with those at which they react to predators, suggesting that signal distinctiveness far exceeds that needed to make them discriminable. Threats displays that carried for long distances might also be disadvantageous (Metz & Weatherhead 1992).

Strategic models of communication generally offer very little in the way of predictions about signal form (Enquist 1985). According to a currently popular strategic model of communication (Zahavi 1975, 1977;

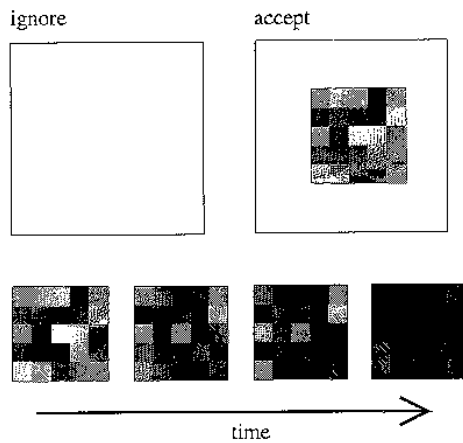


Figure 1. Coevolution of a signal pattern and an artificial neural network receiver during 50 iterations. The receiver's task was to discriminate the signal on a white background from an empty white background. The discriminating receiver drives the coevolving signal towards black, the antithesis of the empty background. Initially each cell in the evolving pattern was assigned a random shade of grey.

Grafen 1990), signallers convey information by using costly signals. The degree of cost is controlled by varying the exaggeration of the signal. Although this offers predictions about the costs associated with different signals and sender states, it provides no prediction about how these costs are to be expressed.

Similarly, the fisherian runaway hypothesis makes little prediction about signal form. It might explain exaggeration of male secondary sexual characteristics, but it is not known if it applies to other intraspecific signalling, and it is not applicable to signalling between species.

Another possibility is that recognition mechanisms select for distinctiveness in signal form. This study has two aims: (i) to explore the role of perceptual bias in intraspecific signalling; and (ii) to develop an artificial neural network technique for modelling situations where several distinct types of response are required. So far neural network studies of biological communication have considered only one type of response (Enquist & Arak 1993, 1994; Arak & Enquist 1994; Johnstone 1994).

Systematic biases in perceptual mechanisms have long been known in psychology (transposition (Köhler 1918); peak shift (Spence 1937; Hanson 1959)) and ethology (supernormal stimuli (Koehler & Zagarus 1937; Tinbergen 1948)). These effects result in a preference for stimuli which are different from the training ideal or the naturally occurring form. Often these biases are for more extreme versions of existing, functional stimuli. Despite the great interest in exaggerated signals with respect to the process of ritualization (Cullen 1966), classical ethologists do not seem to have considered these biases to be important in the evolution of signals.

To our knowledge, the first to suggest the evolutionary importance of perceptual biases was the psychologist Staddon (1975). Recently this idea has been given serious consideration (Leimar *et al.* 1986;

Basolo 1990; Ryan *et al.* 1990; Ryan 1991; Arak & Enquist 1993; Enquist & Arak 1993, 1994).

Biases in perception have been shown to induce polarization of interspecific signals (Enquist & Arak 1993). Interspecific communication occurs between signallers and different recognition mechanisms, whereas intraspecific signalling involves the same recognition mechanisms. In this paper we will investigate the evolution of intraspecific signals, with the expectation that signals will either completely polarize or partly converge since they are filtered through the same receiver mechanism.

2. METHODS

We used artificial neural networks to model the coevolution of signals and recognition mechanisms. We developed the technique used by Enquist & Arak (1993, 1994; Arak & Enquist, 1993) to allow for more than one type of response. The discrimination mechanism consisted of a three-layer perception (Hutchinson 1994) with a 10×10 cell retinal layer, a 12 cell middle layer, and one or two output cells. Images of 5×5 grey scale patterns representing the signal were pasted onto the retina in all possible positions, maintaining the original orientation of the pattern (translations).

The output of a retinal cell is simply the intensity of the projected pattern or background on that cell, a value between 0 and 1. The total activation α of each middle or output layer cell is the sum of all outputs in the previous layer factored by connection weights (Enquist & Arak 1993, 1994; Arak & Enquist, 1993). These cells produce an output of strength s according to the sigmoidal function:

$$s_j = \begin{cases} 1 - 1/[2(1 + a_j)] & a_j \geq 0 \\ 1/[2(1 - a_j)] & a_j < 0. \end{cases} \quad (1)$$

In addition to the effect of the external stimulus, the response is influenced by an internal factor (Arak & Enquist 1993). This factor varies independent of the output according to a normal distribution with mean zero and a standard deviation of 0.025. If the sum of the network output and the internal factor is greater than a threshold of 0.5 then the receiver reacts to the stimulus.

If a network is capable of several different responses (i.e. has more than one output cell), and more than one of these is above threshold, then the cell with the highest output determines the response.

This varying internal factor means that each proximate stimulus, or projection, produces a probabilistic response. Receiver fitness was calculated as the geometric mean of the probability of correctly responding to all presentations of proximate stimuli (reacting to relevant stimuli, and not reacting to irrelevant stimuli). Similarly, signaller fitness was defined as the geometric mean of the probability of eliciting a response across all presentations. The rationale for using the geometric mean rather than the arithmetic mean was partly technical: it tended to avoid evolving to dead-end local optima. The geometric mean has biological justification. It penalizes receivers who react strongly to just a few projections in favour of receivers who react more consistently to a larger set of projections. This situation is consistent with the interpretation of the internal factor as some motivational variable reflecting need. A broad sensitivity is favoured when need is greatest. Signallers and receivers were all scoring at

least 99.99% of maximum fitness when the simulation was ended.

Each iteration of the coevolution between signals and receiver mechanisms produced 100 new mutant networks and 20 new mutant signals. The best of each was retained for use in the following generation. These numbers were chosen so that each parameter had an equal probability of mutation. Network connection weights were mutated with probability 0.01; when a mutation occurred, an increment drawn from a normal distribution of mean 0 and s.d. 0.01 was added to the weight. Component squares of the pattern were mutated with probability 0.05, and an increment from a normal distribution of mean 0 and s.d. 0.3 was added to the intensity of the square (as long as this resulted in a number between 0 and 1).

Obviously the artificial neural networks used here are exceedingly simple in comparison with the biological systems they model. It is not our intention to mimic animal cognition realistically so much as to avoid making *a priori* assumptions about the behaviour of a recognition system.

3. RESULTS

In the most important situation examined above, two different signals coevolve along with a single receiver mechanism. We began with a network coevolving with one signal, then addressed the case of a network capable of making several different responses to different signals. In the first case the task for the network was to discriminate a signal from a blank background. In these simulations the signal evolved quickly towards the opposite of the blank background. Example patterns from one replicate are shown in figure 1*a*.

In the second case we asked what would happen when the signal cannot take this antithetical form. A second irrelevant stimulus was added, so the receiver must discriminate the coevolving pattern from the white background and a pure black pattern. The signals evolve not to a uniform median grey, but to a highly contrasting pattern composed of patches of these extreme colours (figure 1*b*).

To investigate intraspecific communication using more than one signal, a neural network capable of more than one type of response was required. We added a second coevolving stimulus, and required that the receiver respond with different behaviours to different stimuli. Although initially identical, these stimuli polarize, evolving towards converse forms despite being processed by the same network (figure 2*a*).

To investigate the robustness of previous work and its generalizability to this discrimination task, we repeated the last experiment while rotating the patterns (Enquist & Arak 1994). The resulting patterns also evolved towards converse forms, showing a high degree of rotational symmetry (figure 2*b*).

The results of all simulations are summarized in table 1. The general result is that signals diverge to very different forms from all other stimuli, and that all dimensions of the signal space is used.

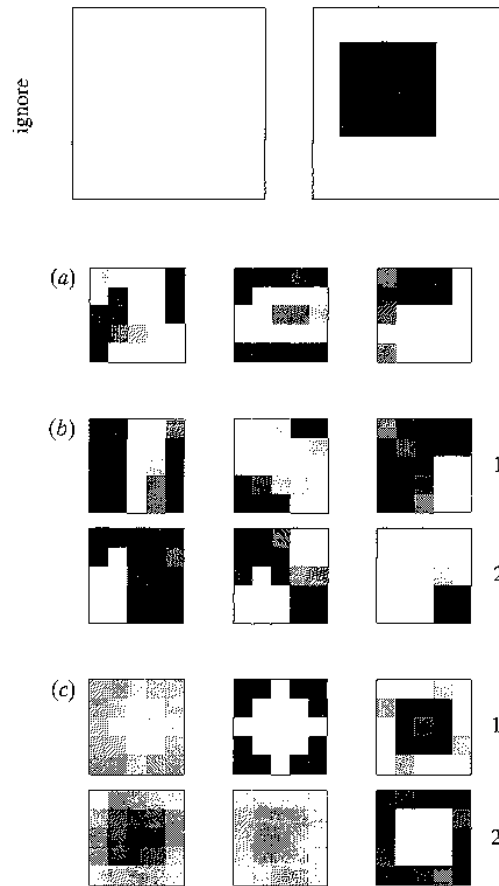


Figure 2. Samples of coevolved patterns from three types of simulations. (a) Patterns that evolved when the network was selected to accept the coevolving signal and ignore both the background and a completely black pattern. (b), (c) Two coevolving stimuli (1 and 2); the network could give two different responses in addition to a non-response. The network was selected to produce different responses to the two coevolving signals and to ignore the background and black pattern. (c) Patterns were not only translated onto different positions on the retina but were also rotated. The figure shows examples of patterns after 500 iterations for each case; (b) and (c) show examples consisting of pairs of patterns.

4. DISCUSSION

We examined the situation in which two different signals coevolve along with a single receiver mechanism. In this case a single receiver must respond appropriately to several different stimuli using the same perceptual mechanism, and cannot specialize on a single target. All intraspecific communication shares this property.

Our simulation results show that coevolution driven by perceptual drive will act to draw evolving patterns away from each other and other stimuli. This occurred despite the fact that they are filtered through the same network.

This mutual selection for exaggeration, and the perceptual pandering to it, occurs without any conflict between players, yet signals do not tend towards

Table 1. *Euclidian distances between signals*

Ten replicates of each simulation are summarized as mean distances (\pm s.e.), measured as the percentage of the longest possible distance, the diagonal between antithetical corners. Each cell in the pattern is treated as a dimension. An R denotes rotation as well as translation of stimuli.

treatment	stimuli			distance		
	coevolving	irrelevant	responses	to white	to black	between signals
1	1	1	1	93.8 \pm 0.7	9.6 \pm 1.0	—
2	1	2	1	57.9 \pm 4.3	76.3 \pm 1.9	—
3	2	2	2	59.5 \pm 4.2	72.1 \pm 3.4	77.0 \pm 2.8
4R	2	2	2	52.1 \pm 5.7	74.7 \pm 3.4	73.5 \pm 2.9

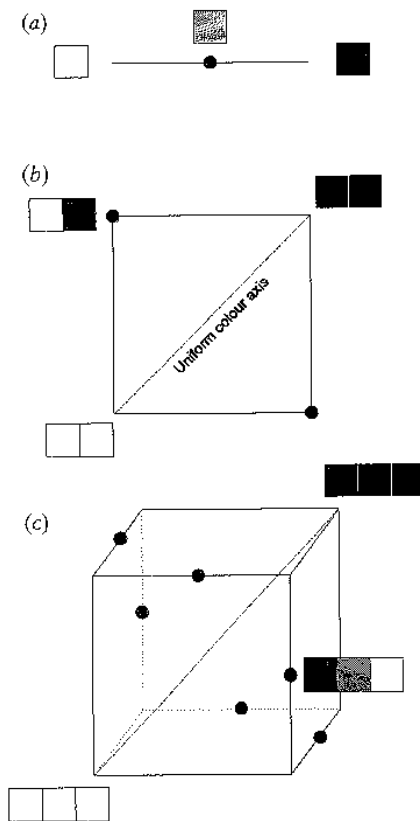


Figure 3. Distance in the signal space. (a) In the one-dimensional case (when the signal consists of one square), a coevolving signal can be most distinct from the antithetical black and white stimuli by being grey. (b) In the case of a two-dimensional signal (two squares), the pattern most different from pure black or white is a combination of these extremes; one black square and one white square. (c) When the signal consists of three dimensions, the most dissimilar pattern contains one white, one grey and one black square. To get these results we have assumed that stimuli 'repel' each other with a force which declines with distance. In the simple cases illustrated here the patterns antithetical to pure black and pure white will be located at equal distances from the black and white corners. These locations are indicated by filled circles. In case 1 the distance is 50% of the distance between black and white (maximum range in the signal space). The corresponding values for cases 2 and 3 are 71% and 65%, respectively.

'conspiratorial whispers' (Dawkins 1982; Krebs & Dawkins 1984). The perceptual imperfections modelled here are very much like those used to explain exaggeration through analogy to advertising (Dawkins 1982).

Figure 3 shows some example signal spaces in which are shown the location of patterns antithetical to pure black and white. This can be compared with Blurton Jones's (1968) depiction of great tit threat displays in a volume space, in which separate axes are used for body position, head position and wing extension. Actual threat display lay at the extreme ends of these axes, in the corners of the display form cube.

It is perhaps surprising that the two stimuli evolved not just to be different but to be each other's opposite. One can imagine that a given nervous system should find as much relevance in the common properties of signals as in that which makes them unique (Ryan 1991). For instance, animals feeding on red berries may have a predisposition to prefer potential mates with some red coloration. This is an important idea and should not be dismissed simply because we found no such effect in our simulation. More complex biological networks may behave differently from our models. However, the limitations of sense organs must shape signals to some degree.

Polarization of displays is seen in many species (Tinbergen 1959; Blurton Jones 1968; Morton 1977; Inglis & Isaacson 1978). However, there are also many examples of signals that appear rather similar. For instance, sex in certain parrots are signalled by very discrete signals despite their otherwise conspicuous coloration (Cinât-Thomson 1926).

What could explain convergence of signals? If signallers always benefit from the same response this should lead to the use of convergent or mimic signals. However, in such a case senders do not communicate by using different signals depending on their state. From the view point of signal form, this is mainly of interest when members of different species try to produce the same signal. Proctor (1992) reports of similarities between male mating signals and copepod prey vibrations. Females respond to these vibrations by closing in on them to feed. The courting male elicits the same response from the female with a similar stimulus.

More relevant to intraspecific communication are

selection pressures common to all signals. For instance, there may be constraints on morphology or coloration. Signals used over longer distances or in noisy environments may be restricted to forms that propagate more efficiently. The cost of signal production and conflicting interests between players will also be important (Enquist & Arak 1993, 1994; Arak & Enquist 1993; A. Arak & M. Enquist, unpublished results). Introducing costs on signal form are likely to decrease polarization along those axes in which it is costly, whereas conflicts are likely to have the opposite effect. Whether or not display costs will overcome the polarization effect in any specific system is an interesting empirical question (A. Arak & M. Enquist, unpublished results).

In conclusion, strategic models are not sufficient to explain all the diversity of signal form. It is necessary to consider the selection pressures emerging from the recognition process itself. Ideally, one would like to consider strategic aspects and the recognition problem simultaneously.

We thank Anthony Arak, Marian Stamp Dawkins and Rufus Johnstone for their comments on the manuscript. This work was supported by an NSERC Canada PGS to P. L. H., grants from Riddarhuset, Enblomska Hjelp-och Stipendiefonden and Kungliga Vetenskapsakademiens fonder to C.-A. W., and by a Swedish Natural Science Research Council grant to M. E.

REFERENCES

- Andersson, M. 1994 *Sexual selection*. Princeton University Press.
- Arak, A. & Enquist, M. 1993 Hidden preferences and the evolution of signals. *Phil. Trans. R. Soc. Lond. B* **340**, 207–213.
- Basolo, A. L. 1990 Female preference predates the evolution of the sword in swordtail fish. *Science, Wash.* **250**, 808–810.
- Blurton Jones, N. G. 1968 Observations and experiments on causation of threat displays of the Great Tit (*Parus major*). *Anim. Behav. Monogr.* **1**, 75–158.
- Cinat-Thomson, H. 1926 Die geschlechtliche Zuchtwahl beim Wellensittich (*Melopsittacus undulatus* Shaw). *Biol. Zbl.* **46**, 545–552.
- Cullen, J. M. 1966 Reduction in ambiguity through ritualization. *Phil. Trans. R. Soc. Lond. B* **251**, 363–374.
- Darwin, C. 1872 *The expression of the emotions in man and animals*. The University of Chicago.
- Dawkins, R. 1982 *The extended phenotype*. London: Oxford University Press.
- Dobzhansky, T. 1951 *Genetics and the origin of species*, 3rd edn. New York: Columbia University Press.
- Enquist, M. 1985 Communication during aggressive interactions with particular reference to variation in choice of behaviour. *Anim. Behav.* **33**, 1152–1161.
- Enquist, M. & Arak, A. 1993 Selection of exaggerated male traits by female aesthetic senses. *Nature, Lond.* **361**, 446–448.
- Enquist, M. & Arak, A. 1994 Symmetry, beauty and evolution. *Nature, Lond.* **372**, 169–172.
- Grafen, A. 1990 Biological signals as handicaps. *J. theor. Biol.* **144**, 517–546.
- Hanson, H. M. 1959 Effects of discrimination training on stimulus generalization. *J. exp. Psychol.* **58**, 321–333.
- Hutchinson, A. 1994 *Algorithmic learning*. Oxford: Clarendon Press.
- Inglis, I. R. & Isaacson, A. J. 1978 The responses of Dark-Bellied Geese to models of geese in various postures. *Anim. Behav.* **26**, 953–958.
- Johnstone, R. 1994 Female preference for symmetrical males as a by-product of selection for male recognition. *Nature, Lond.* **372**, 172–175.
- Koehler, O. & Zagarus, A. 1937 Beiträge zum brutverhalten des halbandregenpfeifers (*Charadrius h. hiaticula* L.). *Beitr. FortpflBiol. Vogel* **13**, 1–9.
- Köhler, W. 1918 Nachweis einfacher strukturfunktionen beim schimpansen und beim haushuhn: über eine neue methode zur untersuchung des bunten farbensystems. *Abh. preuss. Akad. Wis.* **2**, 1–101. (Translated in 1939 as: Simple structural functions in the chimpanzee and in the chicken. In *A source book of Gestalt psychology* (ed. W. D. Ellis), pp. 217–227. New York: Harcourt Brace Jovanovich.)
- Krebs, J. R. & Dawkins, R. 1984 Animal signals: mind reading and manipulation. In *Behavioural ecology: an evolutionary approach*, 2nd edn (ed. J. R. Krebs & N. B. Davies), pp. 380–402. London: Blackwell.
- Leimar, O., Enquist, M. & Sillén-Tullberg, B. 1986 Evolutionary stability of aposematic coloration and prey unprofitability: a theoretical analysis. *Am. Nat.* **128**, 469–490.
- Mayr, E. 1963 *Animal species and evolution*. Cambridge, Massachusetts: Harvard University Press.
- Metz, K. J. & Weatherhead, P. J. 1992 Seeing red: uncovering coverable badges in red-winged blackbirds. *Anim. Behav.* **43**, 223–229.
- Morris, D. J. 1957 “Typical intensity” and its relation to the problem of ritualization. *Behaviour* **11**, 1–12.
- Morton, E. S. 1977 On the occurrence and significance of motivation-structural rule in some birds and mammals. *Am. Nat.* **111**, 855–869.
- Proctor, H. 1992 Sensory exploitation and the evolution of male mating behaviour: a cladistic test using water mites (Acari: Parasitengona). *Anim. Behav.* **44**, 745–752.
- Ryan, M. J. 1991 Sexual selection, sensory systems and sensory exploitation. *Oxf. Surv. Evol. Biol.* **7**, 156–195.
- Ryan, M. J., Fox, J., Wilczynski, W. & Rand, A. S. 1990 Sexual selection, sensory systems and sensory exploitation in the frog *Physalaemus pusulosus*. *Nature, Lond.* **343**, 66–67.
- Spence, K. W. 1937 The differential response in animals to stimuli varying within a single dimension. *Psychol. Rev.* **43**, 427–449.
- Staddon, J. E. R. 1975 A note on the evolutionary significance of “supernormal stimuli”. *Am. Nat.* **109**, 541–545.
- Tinbergen, N. 1948 Social releasers and the experimental method required for their study. *Wilson Bull.* **60**, 6–52.
- Tinbergen, N. 1959 Comparative studies of the behaviour of gulls (Laridae), a progress report. *Behaviour* **15**, 1–70.
- Zahavi, A. 1975 Mate selection – a selection for a handicap. *J. theor. Biol.* **53**, 205–214.
- Zahavi, A. 1977 The cost of honesty. *J. theor. Biol.* **67**, 603–605.

Received 26 October 1994; accepted 17 November 1994