



## Testing the receiver bias hypothesis empirically with ‘virtual evolution’

LISELOTTE JANSSON & MAGNUS ENQUIST

Department of Zoology, Stockholm University

(Received 2 February 2004; initial acceptance 30 March 2004;  
final acceptance 1 February 2005; published online 29 August 2005; MS. number: 7991R)

Many signals found in nature seem exaggerated, for instance in size or colour. According to the receiver bias hypothesis such signal features evolve as a consequence of nonfunctional response biases in receivers. In this study we tested this hypothesis using chickens, *Gallus gallus domesticus*, in a virtual evolution experiment testing the potentiality of receiver bias to drive the evolution of exaggerated signals. The chickens played the role of receivers that can respond to the preferred stimuli displayed by the sender on a peck-sensitive computer screen. The preferred stimulus was kept and evolved, in the direction specified by the chicken, before being introduced to the next chicken of the successive generation. The chickens were tested on signals changing in three dimensions: length, intensity and area. In all three cases, the signals became considerably exaggerated and beyond what was required for accurate discrimination. Our results support the hypothesis that response biases emerging in discrimination tasks are sufficient to cause the evolution of signal exaggeration.

© 2005 The Association for the Study of Animal Behaviour. Published by Elsevier Ltd. All rights reserved.

The receiver bias hypothesis suggests that biological signals take their form as a result of how recognition mechanisms work when discriminating and responding to different stimuli (e.g. Ryan et al. 1990; Enquist & Arak 1998). The theory is valid for almost all kinds of signals and for interspecific communication. When evolutionary conflicts exist, the receiver bias hypothesis predicts that an evolutionary arms race will occur in which signals and responses to signals change continuously (Dawkins & Krebs 1978; Andersson 1980; Enquist & Arak 1998). In the senders, new signals invariably evolve that elicit more favourable responses from the receivers, because the senders exploit biases in the receivers, while evolution in the receivers reduces sensitivity to new signals. Making signals more efficient at eliciting responses often involves exaggeration in the specific signal dimension. The receiver bias hypothesis may thus explain the prevalent occurrence of costly signals, such as conspicuous and exaggerated signals, the existence of which has been debated since Darwin. Why, for instance, have extravagant signals such as the very long tails of some bird species (e.g. Andersson 1994) and the striking colours of prey species (e.g. Cott 1940) evolved despite their obvious disadvantages.

Correspondence: M. Enquist, Department of Zoology, Stockholm University, 106 91 Stockholm, Sweden (email: [magnus.enquist@intercult.su.se](mailto:magnus.enquist@intercult.su.se)).

Predictions about signal form based on receiver biases come from combining information about the recognition and discrimination problems facing the receivers with an understanding of how recognition mechanisms and biases are generated (Guilford & Dawkins 1991, 1993; Weary et al. 1993; Enquist & Arak 1998). Knowledge about the latter derives from comprehensive studies in ethology and experimental psychology (Mackintosh 1974; Ghirlanda & Enquist 2003). One of the well-studied phenomena central to the issue in this paper is generalization. Generalization means that a novel stimulus generates a similar response as a previously experienced stimulus, provided that they are not too dissimilar. Usually, the strongest response is directed to the familiar stimulus, but in some cases a similar stimulus elicits an even stronger response. The direction, strength and duration of such a bias is determined by several factors, such as experience of stimuli, the presence of negative stimuli, the dimension in which stimuli vary, the degree of variation and also the sensitivity of the perception organs (Ghirlanda & Enquist 2003).

The shift of maximum response from the familiar to the novel stimulus is known in ethology as ‘supernormality’ and in psychology as ‘peak shift’ (Tinbergen 1951; Hanson 1959). The phenomena of peak shifts and supernormality can be illustrated by gradients reflecting the strength of response to the different stimuli along a particular dimension. For instance, in discrimination learning the bias

is located away from a negative stimulus. This means that when the negative stimulus is 'weaker' than the positive, for instance of less intensity, the bias develops for a 'stronger' stimulus, of higher intensity. In the reversed scenario, when the negative stimulus is stronger than the positive stimulus, a bias develops for less intensity (Mackintosh 1974; Ghirlanda & Enquist 2003). The empirical nature of response biases, peak shift or supernormal stimuli depends strongly on the dimension of the stimuli (the strongest biases occur along size and intensity dimensions) and also on the similarity between the discriminated stimuli (Ghirlanda & Enquist 2003). Whether the bias is labelled peak shift or supernormal stimuli does not seem to matter, nor does whether the response is innate or learned (Ghirlanda & Enquist 2003).

Our aim in this study was to test empirically the receiver bias hypothesis, particularly the predictions about costly and exaggerated signals. Testing evolutionary hypotheses is often difficult, especially when they share predictions with alternative hypotheses. Biases, whether adaptive (e.g. Zahavi 1977; Grafen 1990) or by-products (Enquist & Arak 1998), may cause similar selection on signal form, such as increased size or intensity. However, such traits may also be favoured in signals used in long-range communication (Wiley 1983). To be productive, empirical tests of the receiver bias hypothesis must be able to exclude these alternatives. This means that it might be more successful to consider the evolutionary process rather than just the outcome of evolution. One way of doing this is to use comparative data to explore the temporal sequence of events in evolution. This method made it possible to show that receiver biases have existed in receivers prior to the emergence of the signal exploiting the bias (Basolo 1990; Ryan et al. 1990). Phelps & Ryan (1998) have also developed a method of reconstructing the evolution of biases and used this to predict successfully the subsequent direction of signal evolution.

We used another empirical approach that we refer to as 'virtual evolution' that experimentally investigates whether receiver biases can drive the evolution of signal form (Forkman & Enquist 2000; Jansson & Enquist 2003). Such virtual evolution crudely resembles a real evolutionary process. In short, real animals are used as signal receivers and each individual, or group of individuals, represents the receivers in a particular generation. The sender is a computer providing the signals on a computer screen, on which the animal can peck or touch. The receiving animal is adapted to the current signals and other stimuli by training. After training, the animal is tested on a series of signals including the familiar signals and the signal variants assumed to be rare in the current generation. The preferred signal is 'selected' to become the signal of the next generation. The selection process may also include additional fitness factors such as signal cost. By repeating this procedure it may be possible to gain insight into the evolutionary consequences of receiver bias.

In the evolutionary scenario considered in this paper, two senders compete for attention with costly signals. The senders could, for instance, reflect real males displaying for females or flowers displaying to pollinators. The discrimination between the signals by the receiver could

reflect a pollinator's attempt to distinguish between a flower species with or without nectar. The competition between the senders could also be for lack of attention from a receiver, such as a prey displaying unprofitability to a predator. In each of the signal dimensions there is a particular value that is optimal with respect to such consequences as the cost of producing or carrying the signal. Departing in any direction from this value will entail a cost.

We conducted three separate experiments exploring the evolution of signal exaggeration in three common, visual dimensions: length, area and colour intensity. In these dimensions biases are well known to exist (reviewed in Ghirlanda & Enquist 2003) and have also been studied theoretically, for instance in evolutionary simulations using artificial neural networks as models of receivers (length, Enquist & Arak 1993; area and intensity, Hurd et al. 1995; Enquist & Arak 1998). According to the receiver bias hypothesis and what is known about receiver biases in these dimensions (Ghirlanda & Enquist 2003), we predicted that the stimuli would increase despite the associated cost (Arak & Enquist 1995).

## METHODS

### Subjects and Housing

Thirteen female chickens, *Gallus gallus domesticus*, were brought to the laboratory as newly hatched and entered the experiments at an age of 8 months. They were housed in cages (1 × 1 × 1 m), one to three birds per cage, with sawdust bedding, a perch and a water bowl. The cages were placed in a laboratory room, at 20°C, with windows providing daylight in addition to a 12:12 h light:dark artificial light cycle. The birds were given free access to water, except during the experimental sessions, and free access to food (commercial chicken pellets) after their daily sessions, for the rest of the day. Food access was completely free during the session-free weekends. This is a very low restriction schedule, compared to those normally used in operant conditioning experiments, which typically involve keeping the animals at 80–90% of their ad libitum body weight. The current schedule ensured that the birds remained at their ad libitum body weight and were still motivated to work for food (B. Forkman, unpublished data). The study was approved by Stockholms norra försöksdjursetiska nämnd Dnr: N148/98 (the relevant Swedish authority).

### Equipment and Shaping

The birds were trained and tested in an operant chamber placed in front of a touch-sensitive PC-computer monitor (Philips 15 inches, Digital Autoscan Colour Monitor, 105S). The computer registered all pecks made on the screen. The chamber (40 × 50 cm and 44 cm high) was made of a wooden frame coated with chicken wire. A feeder delivered food rewards in a feeding tray (12 × 3 cm) fixed in the lower right corner of the monitor. The displayed stimuli, or the background colour of the

monitor, served as the only light sources in the experimental room.

Prior to the experiment, we trained the birds to peck on the monitor for about 5–10 min twice a day. Initially, all attempts to approach the monitor were rewarded with a piece of mealworm. Next, all attempts to peck on the monitor were rewarded with commercial pellets. Finally, only pecks within an area of 20 mm of a 7-mm grey spot, displayed on different locations, were rewarded. When a bird reached the criterion of pecking 70% of all pecks within the allowed area in three consecutive sessions we considered it to be ready for the experiment.

### Experimental Stimuli

The three experiments were identical except for the dimension along which the stimuli could vary. We used a stimulus in the shape of a line to test biases along a length dimension, a red spot to test biases for colour intensity and a square to test biases for area size (Table 1). During the experiments, the stimuli evolved as a result of the responses from the chickens, which were tested one after the other, each chicken representing a new generation. The line and the red spot could vary along one dimension, whereas the square evolved along both its sides, in two dimensions. The stimuli were displayed randomly in any of eight different locations evenly dispersed on the monitor.

### Discrimination Training

The experiments started with discrimination training. Initially, we trained the chickens to peck on only one stimulus ‘the common signal variant’, which was always rewarding (see Table 2 for the meaning of different labels). Next, we trained them to discriminate between the common signal variant and a nonrewarding stimulus, ‘the negative stimulus’, which was less intense along the measured stimulus dimension than the signal (see stimuli sizes in Table 1). Pecks on the common signal variant




ended the trial, caused the monitor to go white and generated delivery of 0.12 g of commercial pellets. In the case of incorrect pecks, the common signal variant disappeared whereas the negative stimulus remained on the screen until the chicken had ceased to peck for 2 s.

After the delivery of every seventh reward, a ‘false’ selection trial was displayed. It consisted of two common signal variants displayed for 5 s. Irrespective of pecks, this trial never generated any rewards. The purpose was to make the chickens accustomed and motivated to work despite not being rewarded.

### Selection Trials

When the birds had reached a stable performance of at least 70% correct discrimination between the common signal variant and the negative stimulus in three consecutive sessions we replaced the false selection trials with real selection trials. The selection trials were thus interspersed among the familiar training trials. The selection trials consisted of the familiar common signal variant displayed together with one of three novel signal variants. One signal variant was larger than the common variant and one was smaller. The value of the third variant ‘the cost-free signal’ was similar to the value of the common signal variant in generation 0, the start value of the common signal variant for the first chicken of each experiment (Table 1). The selection trials were displayed for 5 s and were never rewarded. Pecks from 150 selection trials were collected from each bird. The chickens were tested one after the other. The value or size of each bird’s common signal variant was decided by the pecks made by the previous chicken. If the previous bird had favoured the larger signal variant, the subsequent chicken started its discrimination training with a common signal variant of larger value and vice versa if the preferred variant was smaller. If the chicken had not expressed any clear preferences the evolution paused and the subsequent chicken was trained on a common signal variant of similar value as that of the chicken representing the previous generation.

Table 1. Appearances of the signals in the three experiments

Test dimension and signal form	Negative stimulus	Value of common signal variant at generation 0	Cost-free signal value $v(0)$	Value of variant signals compared to common signal		Evolutionary step ( $d$ )
				Larger	Smaller	
Length (in pixels) Green line, 2 pixels wide 	7	10	10	+2	-2	2
Intensity of red (in R: colour units of red*) Red spot 8 pixels in diameter 	50	75	75	+20	-20	20
Area, sides of square (in pixels) Blue square 	4	5	5	+1	-1	1

Each colour had a numerical value between 0 and 255. Higher value corresponded to stronger colour intensity.  
\*The colour intensity was made from the Red, Green and Blue (RGB) colour model implemented on computers.

Table 2. Stimuli used and virtual evolution repeated in each successive generation (*t*)

Virtual evolution	Common signal	Negative stimulus	Signal variant		Cost-free signal
			Smaller or less intense	Large or more intense	
<b>Receiver</b> Description Value Selection and procedure	Signal variant dominating in the current generation $v(k) = v(0) + kd$ Receivers selected for responding to the signal but not to the negative stimulus. Discriminating receivers (chickens) were produced by discrimination training	Small stimulus that remains unchanged over generations $< v(0)$	—	—	—
<b>Sender</b> Description Index Value Selection and procedure	Signal variant dominating in the current generation $k_t$ $v(k_t) = v(0) + k_t d$ Signals are selected for eliciting responses from receiver in competition with other signals. Chickens choose between the common signal and variants. The most successful variant, whether common or rare, will be the common signal in the next generation.	—	Mutant signal variant one step smaller than common signal $k_{t-1}$ $v(k_{t-1})$	Mutant signal variant one step larger than common signal $k_{t+1}$ $v(k_{t+1})$	Optimal signal with respect to nonsignalling consequences $0$ $v(0)$ Common signal of the first generation

*v* is the value of the signal along the signal dimension, *k* is the index describing the values of *v* that were used and *d* is the step size at which evolution occurred.

### Participation

Thirteen chickens participated in the study. In each of the three experiments, 10 chickens were tested one after the other. The chickens were randomly assigned a starting order for the first experiment and when a bird had finished one experiment it was transferred to the first available starting order of the next experiment. Owing to differences in how motivated the chickens were to peck, six birds managed to participate in all three experiments, five in two experiments and two birds in only one experiment. The chickens worked for about 20 min, which was the time to become satiated, three to four times per day, 5 days a week.

### Procedure for Virtual Evolution

Here we describe the procedure for virtual evolution used in all three experiments. Table 2 summarizes the labels and use of signals and stimuli. In all experiments, the evolution occurred along a single dimension. Let *v* be the value of the signal along this signal dimension. Only certain values of *v* are used. Let the index *k*, which can take both positive and negative values, describe such values. The index *k* = 0 identifies the cost-free value, that is, the optimal value with respect to the nonsignalling consequences. The relation between signal index and signal value is given by

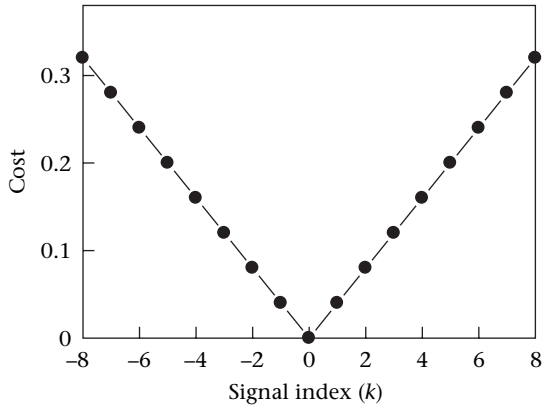
$$v(k) = v(0) + kd$$

where the parameter *d* is the step size at which evolution occurs. In addition to the signals that can evolve, there is another negative and constant stimulus along the same dimension against which the receivers should discriminate. Each dimension has its own *v*(0), negative stimulus and *d* (Table 1) but with respect to the index *k* all three experiments are identical. The benefit of a particular signal depends on how successful the signal is at eliciting pecks from the chicken in competition with another signal present simultaneously on the screen. Formally, the fitness of using signal *k* is

$$\frac{x}{n} - \lambda|k|$$

where the first term *x*/*n* is the relative success of the signal (*x* out of a total of *n* pecks). The second term,  $\lambda|k|$ , is the cost of the signal, which increases with the deviation from the cost-free value (Fig. 1). For  $\lambda$  we used the value 0.04, which imposes considerable cost on signal exaggeration while not completely ruling out the evolution of such exaggeration.

In each generation *t* there is one signal *k<sub>t</sub>* that dominates, referred to as the common signal variant, and two rare variants or mutants *k<sub>t</sub>* + 1 and *k<sub>t</sub>* - 1 representing the current genetic variation in signal value. The signal among these three that has the highest fitness is selected for the next generation. Thus, from one generation to the next, the signal value can remain the same, or increase or decrease by one step. The assumption that only one signal value dominates in each generation simplifies calculation of sender fitness and experimental testing, as we can ignore



**Figure 1.** Signal cost as a function of the signal index  $k$ .  $k = 0$  refers to the cost-free value, that is, the optimal length, size or intensity with respect to such nonsignalling consequences as producing the signal.

interactions with the rare variants. Thus, the fitness of signals  $k_t$ ,  $k_t + 1$  and  $k_t - 1$  depends only on interaction with the current signal  $k_t$  and the fitness of  $k_t$  is 0.5 minus the cost (when competing with itself the probability of being chosen is one half). Let the corresponding fitness values be  $f(k_t, k_t)$ ,  $f(k_t - 1, k_t)$  and  $f_t(k_t + 1, k_t)$ .

The success of the receivers depends on their ability to discriminate between the signal and the negative stimulus. Since only one signal value dominates in each generation, the receiver's fitness depends only on the common signal variant  $k_t$  and the negative stimulus and each chicken needs to be taught this discrimination only when adapting to a new signal.

We repeated the following procedure for 10 generations in each experiment. All experiments started with the signal at the cost-free value, that is,  $k_0 = 0$ .

(1) We trained a chicken to discriminate between signal  $k_t$  as the positive stimulus and the negative stimulus. This adapted the receiver to the current value of the common signal variant.

(2) We then tested the chickens' preferences by letting it choose between the signal  $k_t$  and the variants. The chicken was alternately given choices between the signal  $k_t$  and variant  $k_t - 1$  (selection trial type 1) and between signal  $k_t$  and variant  $k_t + 1$  (selection trial type 2). Each trial was given 50 times. Let  $n_1$  be the total number of pecks during selection trial of type 1 and  $x_1$  the number of pecks on the variant. Let  $n_2$  and  $x_2$  be the corresponding numbers for selection trials of type 2. After the selection trials were completed the chicken was tested in choices between  $k_t$  and the cost-free variant (50 trials).

(3) We decided upon  $k_t + 1$ , the signal of the next generation, by comparing the fitness values  $f_t$  of  $k_t$ ,  $k_t + 1$  and  $k_t - 1$  in the current generation:

$$\begin{aligned}
 f(k_t, k_t) &= 0.5 - \lambda|k_t| \\
 f(k_t - 1, k_t) &= \frac{x_1}{n_1} - \lambda|k_t - 1| \\
 f_t(k_t + 1, k_t) &= \frac{x_2}{n_2} = \lambda|k_t + 1|
 \end{aligned}$$

The signal with the highest fitness was selected for the next generation. If several signals shared the highest

fitness and the smaller variant was included among them, the rule was to approach  $k = 0$  with one step, otherwise not to change the signal.

### Learning Effects

In the first generations, the number of training sessions required to reach the criterion of a stable performance of 70% correct discrimination between the negative stimulus and the common signal variant differed between the dimensions (Fig. 2). The discrimination in the length and area dimensions was initially more difficult than in the intensity dimension and thus required more training sessions. However, once the criterion was reached and the selection trials thus introduced, the degree of correct discrimination between the common signal variant and the negative stimulus was continuously high (Fig. 2).

To analyse whether responding to the signal variants was stable during the selection trials, we divided the data from the selection trials into six parts with respect to time, 25 trials in each, and performed a regression analysis with whole numbers (1,2,...,6) as time units. No effects of time could be found. The slope for the length experiment was  $-0.0074$  ( $r^2 = 0.11$ ,  $P > 0.5$ ) in selection trials including the smaller signal variant and  $0.0063$  ( $r^2 = 0.12$ ,  $P > 0.5$ ) in selection trials including the larger variant. Corresponding values for the intensity experiments were  $0.0042$  ( $r^2 = 0.03$ ,  $P > 0.5$ ) and  $0.0064$  ( $r^2 = 0.04$ ,  $P > 0.5$ ), and in the area experiment  $-0.024$  ( $r^2 = 0.39$ ,  $P = 0.18$ ) and  $0.0033$  ( $r^2 = 0.01$ ,  $P > 0.5$ ).

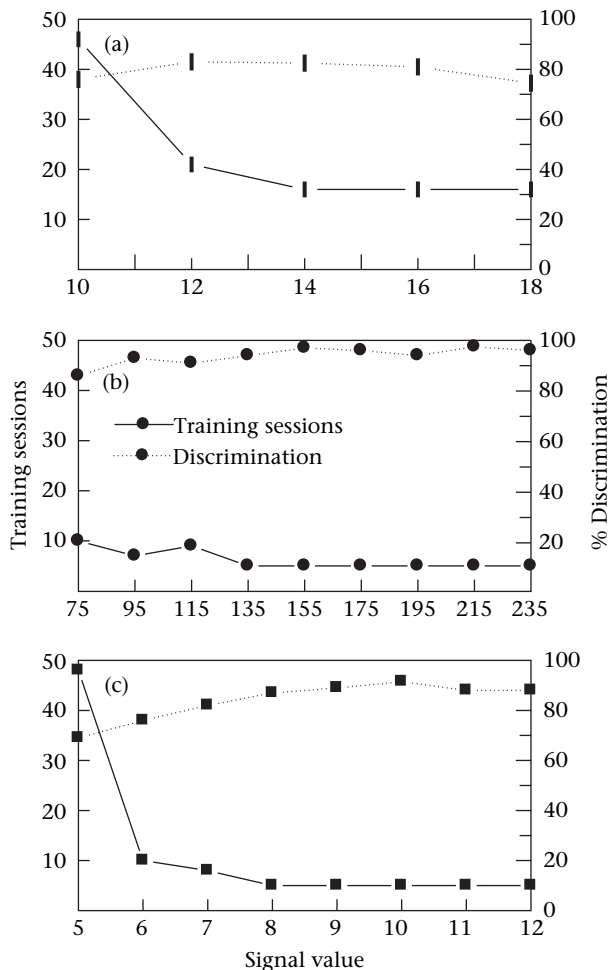
### Statistical Analysis

To establish whether a result of virtual evolution is due to receiver bias we also considered evolution under random choice. We wanted to exclude the possibility that fluctuations in random pecking could explain the result (the null hypothesis). To do this we first derived the probability distribution of signals under random choice after 10 generations (Fig. 3a). The distribution was already very close to the equilibrium distribution. Based on this distribution we calculated the probability after 10 generations of random choice of observing a signal with an index equal to or higher than a particular index (Fig. 3b). This statistical test and its derivation are described in the Appendix.

## RESULTS

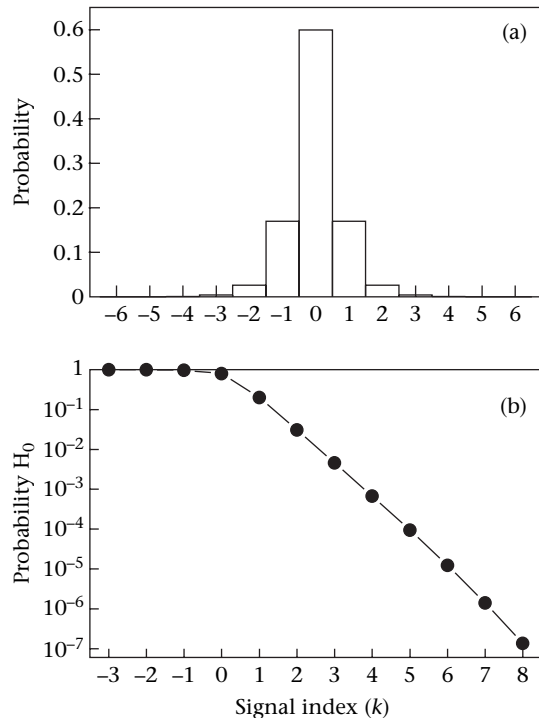
### Evolutionary Trajectories

After 10 generations of virtual evolution in the length experiment, the length had evolved from the original optimum length of 10 pixels to 18 pixels, equal to four steps of evolution (Fig. 4a). The maximum length reached was also 18 pixels. The likelihood of ending up four or more steps above the cost-free length after 10 generations



**Figure 2.** The number of training sessions before the chickens reached the discrimination criterion and subsequent discrimination between the signal and negative stimulus. (a) Length experiment, (b) intensity experiment and (c) area experiment. The discrimination criterion was at least 70% correct pecks in three consecutive sessions. For details of signal values see Table 2. Because the chickens drove the evolution back and forth, more than one chicken was sometimes trained on the same signal value. Some data points are therefore obtained from several chickens. In the length experiment, two chickens were trained on signal value 14, four chickens on 16 and two chickens on value 18. In the intensity experiment, two chickens were trained on signal value 215. In the area experiment, two chickens were trained on signal values 5 10 and 12.

by random choice is small ( $P < 0.0007$ , test described in the Appendix). Thus, the preferences or biases in the chickens must have contributed to this outcome. The results were even clearer in the other two experiments. In the intensity experiment, the signal evolved from a red intensity value of 75 to 215 (RGB value, see Table 2), equal to seven steps of evolution (Fig. 4b,  $P < 0.0001$ , test described in the Appendix). In the area experiment squares evolved from  $5 \times 5$  (25) pixels to  $11 \times 11$  (122) pixels, equal to six steps of evolution (Fig. 4c,  $P < 0.0001$ , test described in the Appendix). The maximum intensity reached was 235 after eight generations and the maximum side of the square was 12 pixels (area of 144 pixels) after nine generations.



**Figure 3.** (a) The distribution of signals after 10 generations of random choice and (b) the statistical significance of possible outcomes of our virtual evolution as a function of the signal index ( $k$ ). For derivation see the Appendix.

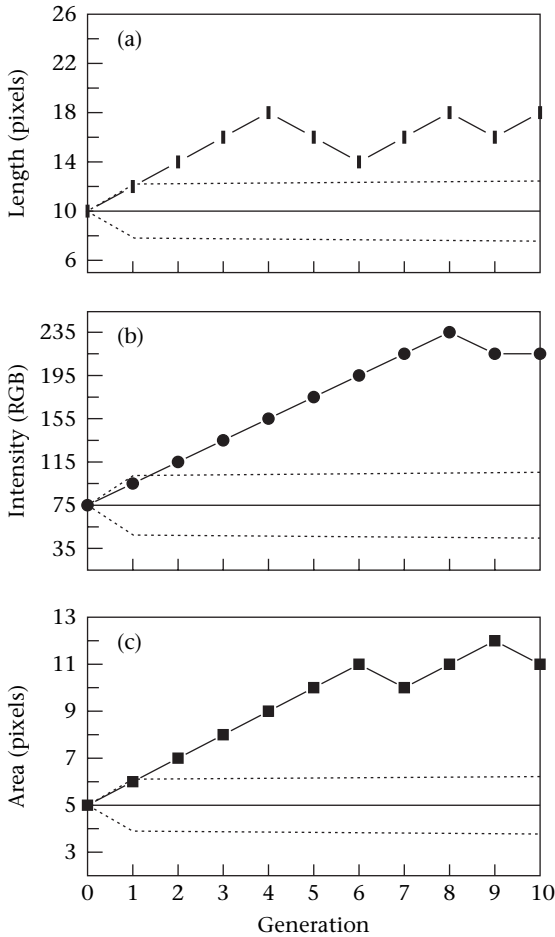
### Fitness

The evolved signals represented a considerable cost to the sender: 0.16 fitness units in the length experiment, 0.32 in the intensity experiment and 0.28 in the area experiment, (Fig. 5). Compared with the absolute fitness, when the common signal variant entailed no cost (the cost-free signal value), this represented a decrease in fitness from 0.5 when the cost-free value was 0.34 in the line experiment, to 0.18 in the intensity experiment and to 0.22 in the area experiment (Fig. 5). However, this is still more than the fitness that the cost-free signal value would have received if introduced as a rare variant in generation 10: 0.25 in the length experiment, 0.06 in the intensity experiment and 0.11 in the area experiment (Fig. 6).

### Biases

In each generation, biases in preferences of signals and stimuli were measured. Figure 7 compares these measures by showing the bias relative to the common signal variant. With a few exceptions the order of preference remained the same, independent of the common signal variant and experiment, and corresponded to a generalization gradient with a peak shift away from the negative stimulus (Figs 7, 8).

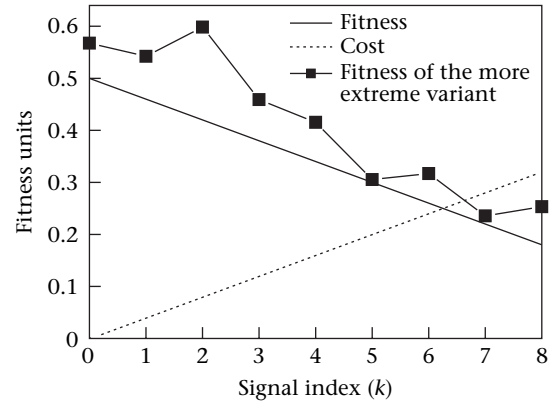
Three major points can be made, even if the results from the length experiment are less clear than in the others (Fig. 6). First, the response towards the negative stimulus



**Figure 4.** The evolution observed in (a) the length experiment, (b) the intensity experiment and (c) the area experiment over the 10 generations. Under random choice at least 99% of the evolutionary trajectories in each generation will occur within the dotted lines. The continuous line and dotted lines indicate the mean  $\pm$  2 SDs under random choice. For the units of intensity see Table 2.

is obviously low since the chickens were trained not to respond to this stimulus. In the intensity and area experiments, at least, the response towards the negative stimulus decreased over the first generations. Second, the relative response to the cost-free signal decreased when the common signal variant increased and approached the response towards the negative stimulus, at least in the intensity and area experiments. Third, the bias in response decreased when the common signal variant became larger. Thus, the selection pressure favouring larger signals seemed to decline when the signal became larger. We can estimate this selection pressure as the relative success of the larger variant minus the relative success of the smaller variant (Fig. 8). In all three experiments, this measure of bias decreased with increasing value of the common signal variant (length experiment: Spearman rank correlation:  $r_s = -0.585$ ,  $N = 10$ ,  $P = 0.075$ ; intensity experiment:  $r_s = -0.829$ ,  $N = 10$ ,  $P = 0.003$ ; area experiment:  $r_s = -0.890$ ,  $N = 10$ ,  $P = 0.0005$ ).

The degree of exaggeration that evolved in our experiments should represent a balance between the cost and



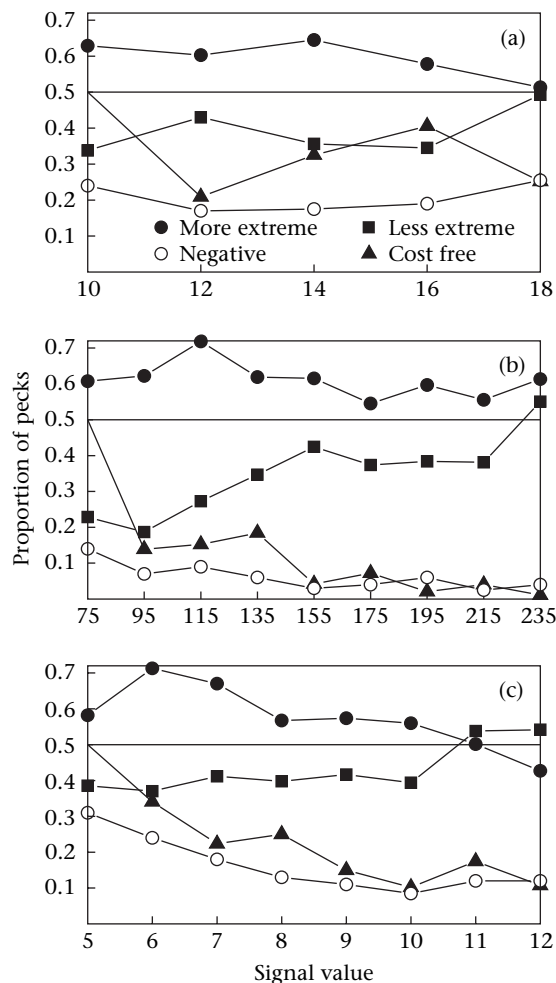
**Figure 5.** Example from the intensity experiment of how the fitness of the signal decreases as it evolves more exaggeration (higher signal index) because the cost increases. Despite being even more costly, the larger variant consistently receives a higher fitness value because the chickens prefer it. This drives the evolution towards more exaggeration (higher index).

benefit of exaggeration, the latter caused by the receiver bias. Since the cost increases linearly with the deviation from the cost-free signal value, this balance will be changed by changes in bias. The result that the signal became most extreme in the intensity experiment, somewhat less extreme in the area experiment and least extreme in the length experiment is consistent with the different declines in biases observed in the different experiments (Fig. 8).

**DISCUSSION**

There are two main results in our study. First, our chickens developed response biases, even though they were never rewarded for selecting anything other than ‘the common signal’. Second, in our virtual evolutionary process these biases were strong enough to drive an evolution towards more exaggerated signals, despite the associated costs for the sender. In all experiments, there was first a rapid increase in exaggeration that later tended to level off, because of a decrease in bias with increasing exaggeration. Since the cost of increasing or decreasing the signal value by one step was always the same, the cost cannot have caused the receding evolution.

The results obtained from the virtual evolution experiment cannot be explained by any major alternative hypothesis for signal evolution. First, the reward of pecking on a signal was the same independent of its exaggeration. Thus, there was no relation between exaggeration and, for instance, reward size that could have driven the evolution. In addition, each receiver was trained only on the common signal variant and thus could not have learnt about any potential benefits of responding to the signal variants. Hence, the result cannot be explained on the basis of a biased response being beneficial or adaptive to the receiver. Second, the bias for exaggeration cannot be explained as a problem of detecting smaller or less intense stimuli since the chickens were never more than 20 cm from the stimuli when facing the

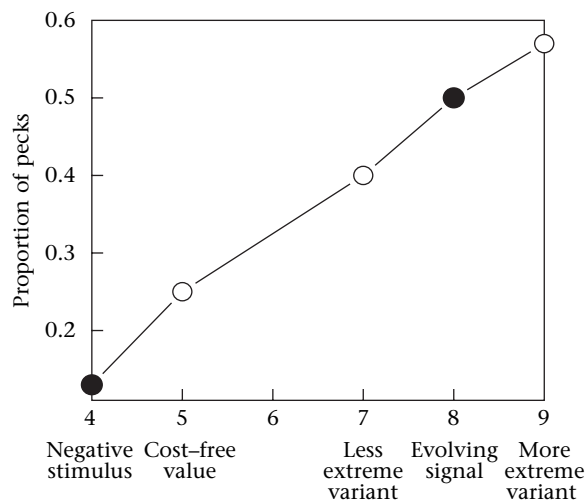


**Figure 6.** Proportion of pecks towards the larger and smaller variant of the signal during the selection trials, the cost-free signal and the negative stimulus as a function of signal value. The data compare the number of pecks on the signal variant relative to the number of pecks on the current appearance of the common signal variant. For some signal values the data were obtained from more than one chicken (see Fig. 2).

computer screen. Furthermore, although less intense stimuli might be harder to perceive, studies have shown that biases develop just as easily for less intensity (e.g. Huff et al. 1975; Jansson & Enquist 2003).

The biases we found are consistent with the general knowledge of stimulus control in the ethological and psychological literature (Tinbergen 1951; Hanson 1959; Kalish 1969; Mackintosh 1974; Eibl-Eibesfeldt 1975; Ghirlanda & Enquist 2003). Generally, the strongest biases are found along size and intensity dimensions (Razran 1949; Grice & Saltz 1950; Magnus 1958; Mackintosh 1974; Ghirlanda & Enquist 2003). Our finding that the bias decreases with increasing distance between the rewarding and the negative stimulus also agrees with general findings (summarized in Ghirlanda & Enquist 2003).

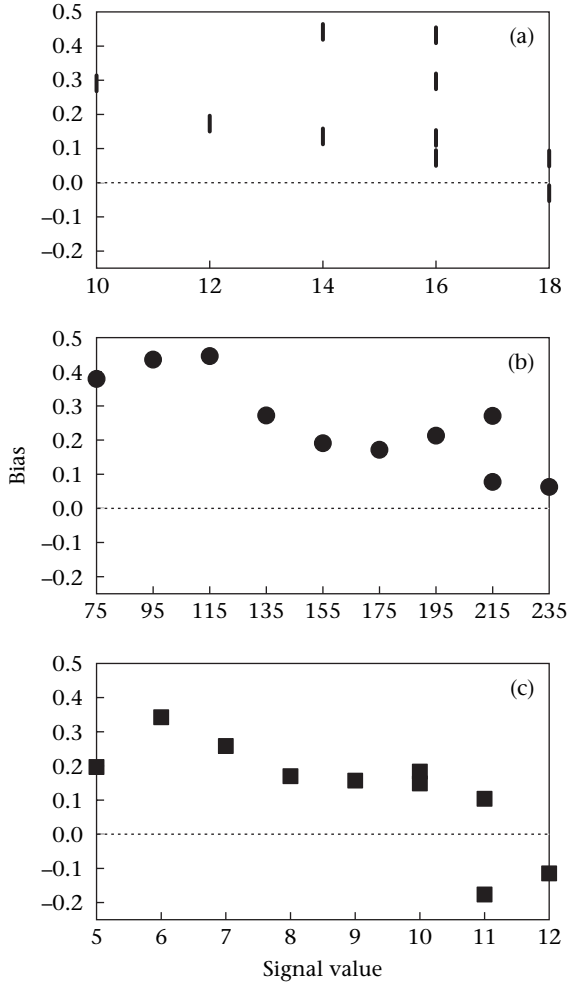
Some differences were found between the experiments. The intensity signal evolved the most by eight steps, followed by the area signal with seven steps, whereas the



**Figure 7.** Example of a generalization gradient. After chickens were trained to peck on the common signal variant but not on the negative stimulus, gradients like this emerged. Note the peak shift; the larger signal variant was chosen more frequently than the common signal variant. The example is from generation 3 of the area experiment when the common signal variant was 8. ●: Training stimuli; ○: Signal variants.

length evolved only four steps. These differences were also reflected in the observed biases. The differences might have been caused by either the nature of the peak shift along the different dimensions or disproportionate evolutionary steps. The strongest result was obtained in the intensity experiment, which is consistent with the finding that intensity generalization gradients are monotonic rather than bell shaped and give rise to responses that are considerably higher than those obtained for the training stimulus (Razran 1949; Ghirlanda & Enquist 2003; Jansson & Enquist 2003). Discrimination along length and area dimensions is also known to generate considerable peak shifts, although not monotonic ones (Weinberg 1973; Ghirlanda & Enquist 2003). That is, response biases for larger size will at specific points, distant from the negative stimulus, eventually peak and become weaker. Why then did the length become less exaggerated? In both the length and the area experiments the size of the signal varied along the dimension, but the form, scale and steps were different. One possibility is that a peak shift is stronger along an area dimension. Another factor might be that the area costs less per pixel. For example, when the area increases from a five- to a six-pixel square this means a total increase of 11 pixels, whereas an increase of the length from 10 to 12 units is a total increase of only four pixels. Thus, if the total number of pixels matters, the area will be more profitable. However, it is also possible that the variants in the length experiment were too similar to the common signal variant. That is, the peak shift was not fully exploited. With signal variants more different from the common signal, the larger variant might have elicited considerably stronger responses by exploiting more of the peak shift. This would have caused the length to become more exaggerated.





**Figure 8.** Selection pressures favouring signals larger than the common signal variants, estimated as the bias in responding to the larger signal variant. (a) Length experiment, (b) intensity experiment and (c) area experiment. The bias was calculated as the relative success of the larger variant minus the relative success of the smaller variant. Thus, a positive value indicates responses in favour of the larger variant and a negative value a preference for the smaller variant.

All three experiments gave results predicted from the receiver bias hypothesis. However, our virtual evolution experiments only crudely resembled reality and the value of such experiments depends, of course, on how reliable they are as a tool for investigating signal evolution. Another powerful experimental approach is given by the virtual studies of Bond & Kamil (1998, 2002). At least three issues seem necessary to consider in these kinds of evolutionary studies: the complexity of the issues in reality, the scale of operation and whether preferences are learned or genetically inherited. Obviously, more factors control the evolution of signals in nature than is feasible to include in an experimental set-up. On the other hand, experiments have the benefit of enabling separation and control of the relevant factors, which may give more detailed information about parts of a process. Our experiment shows in detail how animals respond to several stimuli varying along different dimensions. Hence, even if

the experimental situation is simplified and the stimuli artificial, there is no reason why our results could not be applied to predict the response to signals in nature. Most likely, animals will respond identically, as long as the signals contain the same features and are presented during corresponding conditions. A situation corresponding to our experiment could, for instance, be the presence of two signals, carrying opposite messages, such as edible or nonedible. According to the receiver bias hypothesis and our results, such signals would evolve in opposite directions to each other and thus become exaggerated. This notion is also supported by the prevalent occurrence of exaggerated signals in nature (e.g. Brown 1975; Andersson 1994).

The scales at which real evolutionary processes operate are drastically different from those of our experiment. In reality, selection on signal form involves many interactions between many receivers and many senders in each generation. However, just increasing the number of receivers is unlikely to change any biases, per se, but instead it is likely to remove the most random elements, thus creating a more stable selection pressure. The presence of many signals instead of one may influence the nature of the peak shift. However, very little is known about this (e.g. Scavio & Gormezano 1974; Ghirlanda & Enquist 2003). An ambitious improvement of the experimental design would be to include several subjects in each generation and also to include more signal variation along the dimensions.

### Acknowledgments

This work was supported by Marianne and Marcus Wallenbergs stiftelse, The Bank of Sweden Tercentenary Foundation. We thank Stefano Ghirlanda, Birgitta Tullberg, Ken Cheng, Miguel Rodriguez-Gironés and an anonymous referee for valuable comments and suggestions.

### Appendix: Statistical Test

We derive the probability of observing after  $t \in [0, \infty]$  generations of random choice a signal with index equal or higher than  $k$ . This probability is obtained by the sum  $\sum_{i=k}^{\infty} g_t(i)$ , where  $g_t(k)$  is the probability distribution of  $k$  at time  $t$ . Given an initial distribution  $g_0(k)$  successive distributions can be calculated using the following difference equation:

$$\Delta g(k) = \begin{cases} qg(k-1) + rg(k+1) - (q+r)g(k) & k < -1 \\ qg(-2) + wg(0) - (q+r)g(-1) & k = -1 \\ q(g(-1) + g(1)) - 2wg(0) & k = 0 \\ wg(0) + qg(2) - (q+r)g(1) & k = 1 \\ \sum g(k-1) + qg(k+1) - (q+r)g(k) & k > 1 \end{cases}$$

where  $q$  is the probability of evolving one step towards and  $r$  the probability of evolving one step further away from the signal  $k = 0$ . The probability  $w$  applies to the special case of  $k = 0$ , which is different because changes in both directions go away from  $k = 0$ . Using the fitness

equation described in the [Methods](#) we can calculate  $q$ ,  $r$  and  $w$ .

### Calculation of $r$ and $q$

Consider the case  $k \geq 1$  (the case  $k \leq -1$  is identical and yields the same  $q$  and  $r$ ). Note that  $x_1$  and  $x_2$ , the number of pecks on the variant in selection trials 1 and 2, are replaced with  $X_1$  and  $X_2$ , two random variables having identical binomial distributions ( $n, p = 0.5$ ) describing the distribution of pecks on the variants in selection trials 1 and 2 under random choice. We make the approximation that the total number of pecks  $n$  is always the same (see further below). We first calculate  $r$ :

$$\begin{aligned} r &= \Pr(f(k+1, k) > f(k, k) \text{ AND } f(k+1, k) > f(k-1, k)) \\ &= \Pr\left(\frac{X_2}{n} - (k+1)\lambda > 0.5 - k\lambda \text{ AND } \frac{X_2}{n} - (k+1)\lambda > \frac{X_1}{n} \right. \\ &\quad \left. + (k-1)\lambda\right) = \Pr(X_2 > (0.5 + \lambda)n \text{ AND } X_1 < X_2 - 2\lambda n) \\ &= \sum_{x=a}^n \Pr(X_2 = x) \Pr(X_1 < x - 2\lambda n) \end{aligned}$$

where  $a$  is the smallest whole number greater than  $(0.5 + \lambda)n$ .

To calculate  $q$  we first calculate  $s = 1 - q - r$  (the probability that the current signal will be selected for the next generation) and then calculate  $q$  as  $1 - s - r$ .

$$\begin{aligned} s &= \Pr(f(k, k) \geq f(k-1, k) \text{ AND } f(k) \geq f(k+1)) \\ &= \Pr\left(0.5 - k\lambda \geq \frac{X_1}{n} - (k-1)\lambda \text{ AND } 0.5 - k\lambda \geq \frac{X_2}{n} - (k+1)\lambda\right) \\ &= \Pr(X_1 \leq (0.5 - \lambda)n) \Pr(X_2 \leq (0.5 + \lambda)n) \end{aligned}$$

### Calculation of $w$

Different probabilities apply to the case  $k = 0$ . The probability of changing the signal in the next generation is  $2w$  ( $w$  in each direction). We calculate  $w$  by calculating the probability of not changing the signal:

$$\begin{aligned} 1 - 2w &= \Pr(f(0, 0) \geq f(-1, 0) \text{ AND } f(0, 0) \geq f(1, 0)) \\ &= \Pr\left(0.5 \geq \frac{X_1}{n} - \lambda \text{ AND } 0.5 \geq \frac{X_2}{n} - \lambda\right) \\ &= \Pr(X_1 \leq (0.5 - \lambda)n) \Pr(X_2 \leq (0.5 - \lambda)n) \end{aligned}$$

### Numerical Values for $q$ , $r$ and $w$

On average the hens pecked  $n = 99$  times during each trial type and the lowest number was  $n = 65$ . We use the latter number. This slightly overestimates the probability  $P$  for  $H_0$  (higher  $n$  gives smaller  $P$ ). With  $n = 65$  and  $\lambda = 0.04$  we obtain the following values:  $q = 0.7139$ ,  $r = 0.1098$  and  $w = 0.2024$ .

### The Distribution $g_{10}(k)$

With values of  $q$ ,  $r$ ,  $w$  and the initial distribution  $g_0(0) = 1$  the distribution of  $k$  after 10 generations can be calculated with the equation given at the beginning of this appendix. [Figure 3a](#) shows this probability distribution and [Fig. 3](#) the probability of observing a signal index as high or higher than  $k$  obtained by the sum  $\sum_{i=k}^{\infty} g_{10}(i)$ .

### References

- Andersson, M. 1980. Why are there so many threat displays? *Journal of Theoretical Biology*, **86**, 773–781.
- Andersson, M. 1994. *Sexual Selection*. Princeton, New Jersey: Princeton University Press.
- Arak, A. & Enquist, M. 1995. Conflict, receiver bias and the evolution of signal form. *Philosophical Transactions of the Royal Society of London, Series B*, **349**, 337–344.
- Basolo, A. L. 1990. Female preference predates the evolution of the sword in swordtail fish. *Science*, **250**, 808–810.
- Bond, A. B. & Kamil, A. C. 1998. Apostatic selection by blue jays produces balanced polymorphism in virtual prey. *Nature*, **395**, 594–596.
- Bond, A. B. & Kamil, A. C. 2002. Visual predators select for crypticity and polymorphism in virtual prey. *Nature*, **415**, 609–613.
- Brown, J. L. 1975. *The Evolution of Behavior*. New York: Norton.
- Cott, H. B. 1940. *Adaptive Coloration in Animals*. London: Methuen.
- Dawkins, R. & Krebs, J. R. 1978. Animals signals: information or manipulation? In: *Behavioural Ecology. An Evolutionary Approach* (Ed. by J. R. Krebs & N. B. Davies), pp. 282–309. Oxford: Blackwell Scientific.
- 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: *Cognitive Ecology* (Ed. by R. Dukas), pp. 21–87. Chicago: University of Chicago Press.
- Eibl-Eibesfeldt, I. 1975. *Ethology. The Biology of Behavior*. New York: Holt, Rinehart & Winston.
- Forkman, B. & Enquist, M. 2000. A method for simulating signal evolution using real animals. *Ethology*, **106**, 887–897.
- Ghirlanda, S. & Enquist, M. 2003. A century of generalization. *Animal Behaviour*, **66**, 15–36.
- Grafen, A. 1990. Biological signals as handicaps. *Journal of Theoretical Biology*, **144**, 517–546.
- Grice, G. R. & Saltz, E. 1950. The generalization of an instrumental response to stimuli varying in the size dimension. *Journal of Experimental Psychology*, **40**, 702–708.
- Guilford, T. & Dawkins, M. S. 1991. Receiver psychology and the evolution of animal signals. *Animal Behaviour*, **42**, 1–14.
- Guilford, T. & Dawkins, M. S. 1993. Receiver psychology and the design of animal signals. *Trends in Neurosciences*, **16**, 430–436.
- Hanson, H. 1959. Effects of discrimination training on stimulus generalization. *Journal of Experimental Psychology*, **58**, 321–333.
- Huff, R. C., Sherman, J. E. & Cohn, M. 1975. Some effects of response-independent reinforcement in auditory generalization gradients. *Journal of the Experimental Analysis of Behaviour*, **23**, 81–86.
- Hurd, P. L., Wachtmeister, C. A. & Enquist, M. 1995. Darwin's principle of antithesis revisited: a role for perceptual biases in the evolution of intraspecific signals. *Proceedings of the Royal Society of London, Series B*, **259**, 201–205.
- Jansson, L. & Enquist, M. 2003. Receiver bias for colourful signals. *Animal Behaviour*, **66**, 965–971.

- Kalish, H. I.** 1969. *Learning*. London: Macmillan.
- Mackintosh, N.** 1974. *The Psychology of Animal Learning*. London: Academic Press.
- Magnus, D.** 1958. Experimentelle Untersuchung zur Bionomie und Ethologie des Kaisermantels *Argynnis paphia* L. (Lep. Nymph.) I. Über optische Auslöser von Anfliegereaktion und ihre Bedeutung für das Sichfinden der Geschlechter. *Zeitschrift für Tierpsychologie*, **15**, 397–426.
- Phelps, S. M. & Ryan, M. J.** 1998. Neural networks predict response biases of female tungara frogs. *Proceedings of the Royal Society of London, Series B*, **265**, 279–285.
- Razran, G.** 1949. Stimulus generalization of conditioned responses. *Psychological Bulletin*, **46**, 337–365.
- Ryan, M. J., Fox, J. H., Wilczynski, W. & Rand, S. A.** 1990. Sexual selection for sensory exploitation in the frog, *Physalaemus pustulosus*. *Nature*, **343**, 66–67.
- Scavio, M. J. J. & Gormezano, I.** 1974. CS intensity effects on rabbit nictitating membrane, conditioning, extinction and generalization. *Pavlovian Journal of Biological Science*, **9**, 25–34.
- Tinbergen, N.** 1951. *The Study of Instinct*. Oxford: Clarendon Press.
- Weinberg, S. G.** 1973. Some determinants of the peak shift in stimulus generalization. *Psychological Reports*, **32**, 47–58.
- Weary, D. M., Guilford, T. C. & Weisman, R. G.** 1993. A product of discriminative learning may lead to female preferences for elaborate males. *Evolution*, **47**, 333–336.
- Wiley, R. H.** 1983. The evolution of communication: information and manipulation. In: *Communication* (Ed. by T. R. Halliday & P. J. B. Slater), pp. 156–189. New York: W.H. Freeman.
- Zahavi, A.** 1977. The cost of honesty (further remarks on the handicap principle). *Journal of Theoretical Biology*, **67**, 603–605.