

# Artificial neural networks as models of stimulus control\*

Stefano Ghirlanda and Magnus Enquist

Division of Ethology, Department of Zoology, University of Stockholm

Reprint of December 2, 2006

## Abstract

We evaluate the ability of artificial neural network models (multi-layer perceptrons) to predict stimulus–response relationships. A variety of empirical results are considered, such as generalization, peak-shift (supernormality) and stimulus intensity effects. The networks were trained on the same tasks as the animals in the considered experiments. The subsequent generalization tests on the networks showed that the model replicates correctly the empirical results. It is concluded that these models are valuable tools in the study of animal behaviour.

## 1 Introduction

Artificial neural networks represent an important advance in the modelling of nervous systems and behaviour (see e.g. Churchland & Sejnowski, 1992). During the last 10 or 15 years artificial neural networks have been actively researched in as diverse disciplines as cognitive psychology, neurophysiology, engineering, artificial intelligence and physics. For some reason, however, these models have been more or less ignored by scientists working within the field of animal behaviour. This is surprising since they offer a number of potentials to ethologists. Artificial neural networks can show us how small units like nerve cells can exhibit powerful computational abilities when working together (Hopfield & Tank, 1986; Mezard et al., 1987). They also provide understanding about memory and mental representations (McClelland & Rumelhart, 1985), and about mechanisms such

---

\*Published in *Animal Behaviour*, **56**, 1383–1389. © 1998 The Association for the Study of Animal Behaviour. Minor differences are present compared to the published version. Correspondence (as of December 2, 2006): S. Ghirlanda, stefano.ghirlanda@unibo.it.

as learning (Shanks, 1995) and stimulus control (see e.g. Enquist & Ghirlanda, 1998). In essence, they provide a potential common framework for understanding and modelling behavioural mechanisms relevant for both simple as well as complex organisms. Here we will focus on stimulus control. In his classic text-book *The Study of Instinct*, Tinbergen (1951) devotes a significant part to how external stimuli influence behaviour. Similarly, Hinde (1970) in his attempt to synthesize ethology and animal experimental psychology, dedicates several chapters to stimulus control. Stimulus control has also been and still is a central subject in experimental psychology (see e.g. Terrace, 1966; Mackintosh, 1974; Pearce et al., 1997). One of the fundamental observations on how stimuli control response is referred to as stimulus generalization in psychology (Guttman & Kalish, 1956; Mackintosh, 1974). This refers to the fact that an organism responds in similar ways to many variants of stimulation. For instance, if an animal has been trained to react to a particular stimulus, it will also react to stimuli that are somewhat different. The strength of response is often described by a generalization gradient (over some stimulus dimension) with a maximum of responding usually at or near the training stimulus.

The findings within ethology are very similar, although different concepts have been used and the focus is not on learning but on stimuli that naturally occur in the wild, for example social signals. Ethologists have studied the importance of stimulus components and concluded that certain aspects often are more important than others, and that naturally occurring stimuli can be stripped of many of their components and still can be potent in eliciting the response (Hinde, 1970; Baerends, 1982). In practise this is the same thing as generalization: the animal does not only react to stimuli that occur in nature but also to variants which lack certain aspects or have extra components added.

Both ethologists and experimental psychologists have also shown that certain stimulus variants may even be more efficient than the naturally occurring ones or the training stimuli. Within ethology this is called a supernormal stimulus (Tinbergen, 1951; Hinde, 1970), and psychologists talk about peak-shifts in responding (Hanson, 1959; Mackintosh, 1974).

Empirical studies of stimulus control have resulted in a set of principles (e.g. animals do generalize) which directly reflect observations, but there is very little theory in the strictest sense. One exception is gradient-interaction theory (based on Spence 1936, 1937; Hull 1943), which allows one to calculate response in a complex situation given that one knows how the animal generalizes in all the simple situations that build up the complex case. Note, however, that this model cannot predict the degree of generalization in the individual cases: this has to be obtained in experiments.

Artificial neural networks have several advantages compared with traditional theory of stimulus control. First, these models do not rest on empirical findings on

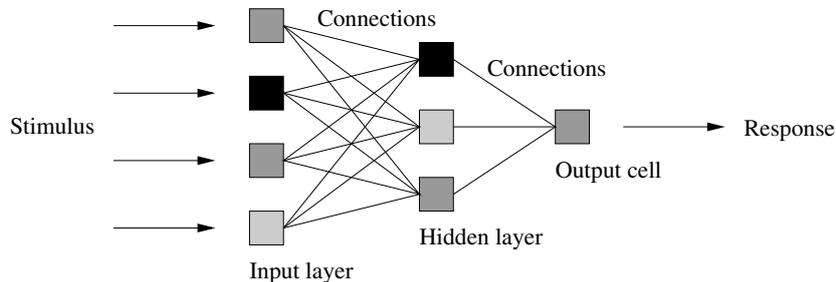


Figure 1: Schematic representation of a three-layer perceptron. Stimulation is assumed to act on the units in the input layer (whose activation is illustrated as the degree of filling), and is then transmitted to the middle layer (called the ‘hidden’ layer) via a system of connections. These can amplify or reduce the signal. The same steps are repeated when the signals travel from the hidden layer to the output cell, whose activation is the response of the network to the stimulation.

stimulus control. Phenomena such as generalization and supernormal stimuli are emergent properties. In addition, they are not ‘black-box models’: they consider the architecture of the nervous system.

The aim of this study is to investigate in some detail how certain popular artificial neural network models generalize and compare these results with empirical data. We shall not be concerned with a quantitative agreement, but rather we want to see if these models are able to reproduce qualitatively a number of generalization phenomena that are important in understanding animal behaviour.

## 2 The Model

We shall use three-layer artificial neural networks with the feed-forward architecture, also called multi-layer perceptrons (see e.g. Haykin, 1994), as shown in figure 1. The input layer can be thought as modelling a perceptual organ, with each unit  $i$  having an activation  $s_i$  in the interval  $[0, 1]$  when the stimulus  $s$  is presented to the network. The output of each hidden-layer cell is a function of the weighted sum of such activations:

$$h_i(s) = \varphi \left( \sum_{j=1}^N w_{ij} s_j \right) \quad (1)$$

where  $N$  is the size of the input layer,  $w$  is the set of connections between the input and hidden layers. The function  $\varphi(\cdot)$  is often referred to as the transfer function; it describes the output of a unit given all its inputs, and it is usually a sigmoid

function. Here we use:

$$\varphi(x) = a \tanh(bx) = a \frac{e^{bx} + e^{-bx}}{e^{bx} - e^{-bx}} \quad (2)$$

with  $a = 1.716$  and  $b = 2/3$ , as suggested by previous works (see e.g. Guyon, 1991). The precise value of these parameters is relevant only to the speed at which networks learn, and does not affect their generalization properties.

The outputs coming from the hidden layer reach a single output cell via a second system  $v$  of connections, and the activation of this cell is taken as the response of the network to the stimulation, according to:

$$o(s) = \psi \left( \sum_{i=1}^M v_i h_i(s) \right) \quad (3)$$

where  $M$  is the number of hidden units and the function  $\psi(\cdot)$  is given by:

$$\psi(x) = \frac{1}{1 + e^{-x}}. \quad (4)$$

This activation function differs from the one pertaining to the hidden layer in that its values are in the interval  $[0, 1]$ , and this enables us to interpret the network output  $o(s)$  as *probability of response* to the stimulus  $s$ . A more detailed model would distinguish between different kinds of response (corresponding to different behaviours), for example by having more output units, but here we rely only on  $o(s)$  as a model of behaviour.

In the present context a learning procedure is just an algorithm that is able to adjust the network outputs to a set of stimuli according to a given criterion, by changing the network connections  $w$  and  $v$ . Here we use the well-known back-propagation algorithm (LeCun, 1985; Parker, 1985; Rumelhart et al., 1986) and a simple model of learning through an evolutionary process (Enquist & Leimar, 1993). In both cases the performance of the network can be quantified by the differences between the present outputs and the target ones, for example by the quantity:

$$e = \sum_{\alpha} (o(s^{\alpha}) - d^{\alpha})^2 \quad (5)$$

where the sum is over the whole training set and  $d^{\alpha}$  is the desired output to stimulus  $s^{\alpha}$ . This performance measure is used by the two algorithms in different ways: in back-propagation a computation is made of how much each connection is contributing to the error equation (5), then changing it a little to diminish the error; in the evolutionary algorithm a new network is created with some mutated connections, and the one is retained that yields the lesser error. In both algorithms the step just sketched is repeated until the network produces the desired outputs



Figure 2: A visual representation of the training stimuli used to train the networks. Each square represents one stimulus unit, and the area filled in black is proportional to the activation of the unit, so that a white square means minimum activation and a completely filled one full activation.

to the training stimuli. A more detailed presentation of the stimuli and targets we have used is given in the next section.

Prior to training each network has to be given some initial set of connections on which the chosen learning procedure will act, and this will affect to some extent its response properties. That this effect exists is both trivial and important, for the initial state of the network has a natural counterpart in the animal’s state before training. This latter state is the result of both genetic and individual history factors, which are of course able to affect behaviour (see e.g. Peterson, 1962; Mason & Reidinger, 1983, for examples relevant to stimulus control). What the initial connections should be to model a real situation is not well understood, and we resort simply on setting weights at random in the interval  $[-0.1, 0.1]$ , with uniform distribution. With these settings the data emerging from the simulations will be best compared with experiments in which no strong pre-existing bias affected generalization performance. On the other hand we can be sure that the observed features, for instance the shape of gradients, are truly due to the training procedure and not to the initial connection values.

### 3 Training and Testing

This section illustrates what stimuli we have used to train the networks, and what tests have been devised to investigate their generalization abilities, relating both of these steps to actual experimental situations; we consider the training first.

#### 3.1 Training Stimuli

To start, we focus for simplicity on just two stimuli (i.e. a single discrimination task), noted  $s^+$  and  $s^-$  (see figure 2). The positive stimulus differs from the negative one in its central region, that we call for short the ‘signal’. The activation of the input units increases towards the centre of the signal. This models the fact that a given stimulus usually activates the receptor cells to different extents, there being cells that are most active, some that are mildly active, and some that are unaffected. The networks are trained to have a high output to  $s^+$  and a low one

to  $s^-$ . The precise value of the target outputs is not important, so long as they are not too close to 0 and 1, to avoid floor or ceiling effects. Similarly, by avoiding choosing extreme values for the units in each stimulus one has the possibility of testing the model with stimuli that are more, or less, intense than the training ones.

The stimulus  $s^-$  can be interpreted as modelling two kinds of real stimuli: an experimental background (the so-called ‘contextual stimuli’), or a stimulus differing from  $s^+$  along a given dimension (e.g. lights of different colours or sounds of different intensities). These two situations are usually referred to as ‘non-differential’ and ‘differential’ training respectively, an important difference being that the contextual stimuli can often be considered as having a very low or zero intensity along the dimension that will be probed in the generalization test, for example a dark key as opposed to an illuminated one when investigating wavelength generalization, or silence in contrast to a tone when looking for sound-frequency generalization.

From these considerations we see that the stimuli in figure 2 are best suited to model a non differential training procedure. A differential discrimination task can be modelled by having two stimuli such as the present  $s^+$ , but differing in the position of the signal, as we will do below.

### 3.2 Testing the Model

In the model, stimulus control has the same definition as in reality: the change in response (network output) coming from a change in stimulation (changes in the values of stimulus units). Thus, in the model, as in real experiments, a generalization test can be given along many dimensions, and we choose here to consider (1) the intensity of the signal in  $s^+$ , that can be varied from 0 to 1 (‘intensity test’), and (2) its position, shifting it from left to right (‘translation test’).

In the intensity test the stimulus change affects the same units throughout the testing, while the translation test can be understood as a rearrangement of the stimulation coming to the network while keeping constant its total intensity. A compound case would be changing the size of the signal in  $s^+$ , involving thus both a change of total intensity and different sets of units involved by different test stimuli. This latter case will not be investigated in this paper (but see Grice & Saltz, 1950; Mednick & Freedman, 1960; Dougherty & Lewis, 1991, for experimental data).

We have chosen the intensity and translation tests because they are readily compared to experiments: intensity generalization gradients have been measured along many dimensions, including sound or light intensity (Razran, 1949; Thomas & Setzer, 1972; Zielinski & Jakubowska, 1977), colour intensity (Czaplicki et al., 1976), taste (Tapper & Halpern, 1968), and click frequency (Weiss & Schindler, 1981); rearrangement of stimulation is implied in studies using such dimensions

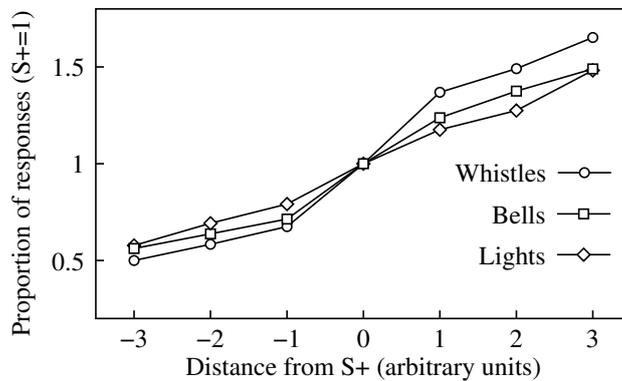


Figure 3: Intensity generalization profiles for three different kinds of stimuli: lights, bells and whistles. On the horizontal axis the intensity of the test stimuli is plotted in arbitrary units as a distance from the conditioning stimulus  $S^+$ , while on the vertical one is shown the proportion of responses emitted for each stimulus in comparison with  $S^+$ . Data from Razran (1949), table II, summarising data from Pavlov's laboratory (Pavlov, 1927).

as sound frequency (Jenkins & Harrison, 1960, 1962), light wavelength (Hanson, 1959; Kalish & Guttman, 1959) and line orientation (Hearst, 1968). A point of interest in distinguishing between these two classes of generalization tests is that they give very different results in real experiments, and are therefore a good first test for any model of stimulus control.

## 4 Results

### 4.1 The Intensity Test

The general appearance of intensity generalization gradients is clearly illustrated in figure 3, where data coming from tests involving three different kinds of stimuli (lights, bells and whistles) are reported (Razran, 1949; Mackintosh, 1974). All three experimental gradients are monotonic, with increased number of responses as the intensity of stimulation increases. All stimuli more intense than  $S^+$  are thus supernormal, eliciting more responses than  $S^+$  itself.

More data are provided in Weiss & Schindler (1981), which we show in figure 4. In this experiment two groups of rats were conditioned for lever pressing in the presence of clicks coming from a loudspeaker, the variable changing in the generalization test being the number of clicks per second (directly related to the amount of stimulation per second). A third group, serving as control, yielded a substantially flat generalization profile (line 1 in figure 4), showing that the intensity of stimulation has no control per se when the considered stimulus is not correlated with any meaningful outcomes. Rats in the first experimental group

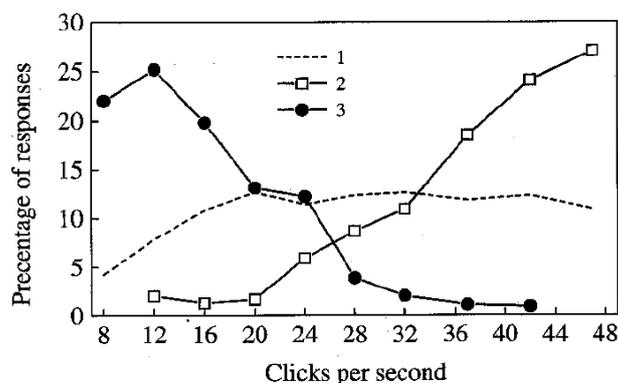


Figure 4: Generalization gradients for rats responding to different click rates. Continuous line: control group, no discrimination training; dashed line:  $S^+$  at 32 Hz and  $S^-$  at 20 Hz; dotted line:  $S^+$  at 20 Hz and  $S^-$  at 32 Hz. Data from Weiss & Schindler (1981, experiment one). On the vertical axis relative percentages of responding are plotted.

were reinforced for responding to click rates of 32 Hz, and not reinforced at a lower click rate, 20 Hz. The resulting gradient (line 2 in figure 4) is similar to those in figure 3, with a strong peak-shift in the direction of increasing click rate and responding dropping well below that of the control group for stimuli less intense than  $S^+$ .

The most interesting feature of this study is the reversed gradient (line 3 in figure 4) obtained from rats trained to respond to the lower click rate of 20 Hz as opposed to the nonrewarded, and more intense, 32-Hz stimulus. From these data we see clearly that it is not intensity in itself that is responsible for increased response rates, but rather the interplay of all experiences that animals are met with along the considered intensity continuum.

All these findings are well reproduced by the network model, as seen in figure 5. The networks exhibit also a consistent negative peak-shift: stimuli that are further than  $s^-$  from  $s^+$  elicit less responding than  $s^-$  itself, not being important which one between  $s^+$  and  $s^-$  is the most intense. Does this occur with animals? In all of the cases in figure 3 the  $S^+$  was rewarded in contrast with the absence of stimulation, i.e. no light, bell or whistle, so that investigation with stimuli less intense than  $S^-$  could not be conducted. A different kind of floor effect, i.e. very low responding to  $S^-$ , prevents us from testing the existence of a negative peak-shift for data in figure 4. The existence of such a phenomenon is, however, well established in other experiments (e.g. Thomas & Setzer, 1972; Zielinski & Jakubowska, 1977), although this effect has not been the object of systematic investigations.

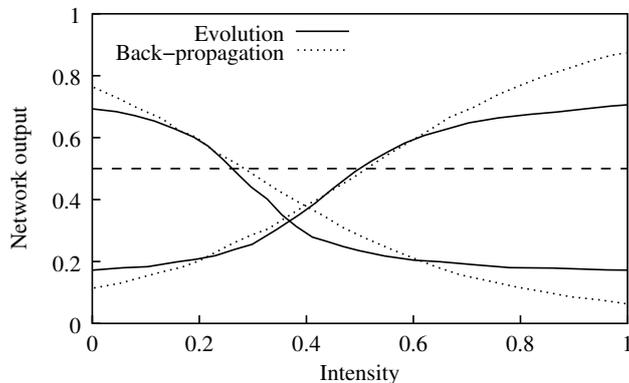


Figure 5: Intensity generalization profiles for networks trained with an evolutionary algorithm and a back-propagation one. The training stimuli, labelled by the intensity of the signal, are shown. The target outputs were 0.2 for  $S^-$  and 0.6 for  $S^+$ . The horizontal line is the response from an untrained network.

## 4.2 The Translation Test

Experimentally, generalization gradients obtained with test stimuli that are a rearrangement of the training ones (as explained in ‘Testing the Model’ above) give very different results from intensity generalization tests. A typical example, relative to wavelength discrimination in pigeons, is given in figure 6, taken from Hanson (1959). We focus first on the control subjects (solid line in the figure), that were trained to peck at a key when it was illuminated with a coloured light and not to respond to the dark key. During the generalization test the wavelength of the light was changed, and the resulting gradient has a maximum located at  $S^+$ , with response decreasing as the test stimulus grows different from the training one.

More generally, the shape of the control gradient in figure 6 is representative of all studies in which, following interdimensional training, the test stimuli are in some way a rearrangement of the training ones (Kalish & Guttman 1959; Jenkins & Harrison 1960, 1962; Marsh 1972; Purtle 1973) although the ways by which rearrangement can be achieved are many (for example, by tilting a line or changing the frequency of a tone maintaining its intensity constant). Theoretical arguments (Ghirlanda & Enquist, unpublished data), as well as the cited empirical evidence, support this conclusion.

The other gradients in figure 6 come from groups of pigeons that, after a pre-training identical to that of the control group, were given discrimination training in which  $S^-$  was a light differing in wavelength from  $S^+$ . These gradients have a similar shape to the control one, but their peaks are shifted from  $S^+$  in the direction away from  $S^-$  (the height difference between the postdiscrimination gradients

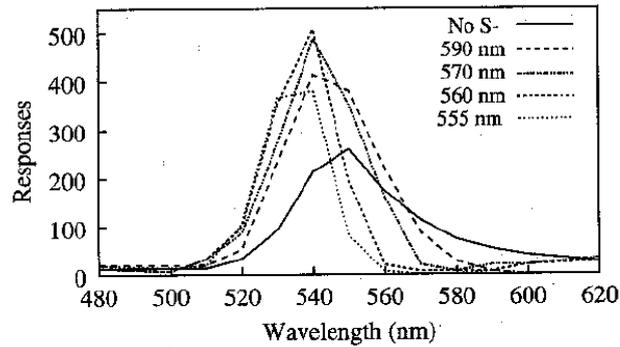


Figure 6: Gradients obtained in a light-wavelength generalization test with pigeons, reproduced from Hanson (1959). The lower gradient is a control one, obtained after training on a 560-nm wavelength. The other gradients result from discrimination training between the 560-nm light and other colours, as indicated in the legend.

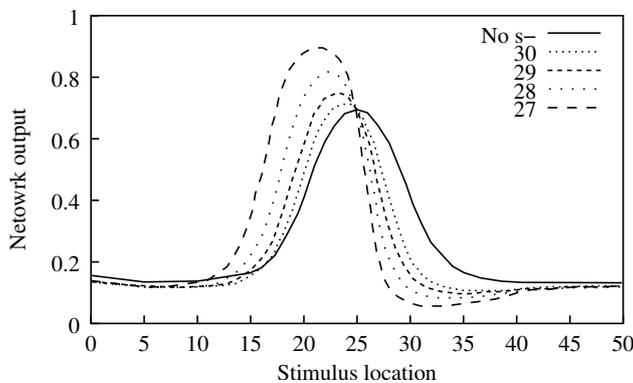


Figure 7: Network generalization gradients coming from the translation test. The  $s^+$  was fixed, and four different negative stimuli were used, consisting in a 'signal' like that in  $s^+$  (see figure 2), but with a displaced position (indicated in the legend above). The evolutionary algorithm gave entirely similar results, but for clarity only data from networks trained with the back-propagation algorithm are shown. The input layer of the networks was made of fifty units, and the positive stimulus was centered at position 30 (with a target output of 0.7) for the no- $s^-$  group and all discrimination groups. The negative stimuli (target output 0.2) were centered at the positions indicated in the legend.

and the control one is due to differences in the amount of training). To model the training of Hanson's pigeons, the networks have been trained in two steps: first in the situation of figure 2 (corresponding to a first conditioning stage to  $S^+$ ), and then an  $s^-$  was introduced consisting of a stimulus similar to the  $s^+$  in figure 2 but with a displaced location (corresponding to discrimination training between two wavelength). Model data relative to the translation test, introduced in section 'Testing the Model' above, are shown in figure 7. The overall shape of the gradient, very different from that characteristic of intensity dimensions, is well reproduced. Moreover, we see that the peak of the gradients is further away from  $s^+$  the closer  $s^-$  comes to it, an effect that can be considered the main characteristic of Hanson's (1959) study. We observe as well that the height of the peak increases both in the model and the empirical data, but a thorough comparison is prevented by the difference in training procedures (the networks are trained to give a fixed response to  $s^+$ , while this variable was not controlled in Hanson's (1959) experiment, as it is clear from figure 6).

## 5 Discussion

In this paper we have evaluated an artificial neural network model's ability to predict stimulus–response relationships. The model was compared with a number of general empirical findings. In summary, the neural network model was very accurate in predicting stimulus–response relationships, appearing, in fact, superior to many psychological and ethological models.

Existing models of stimulus control include many theories based on similarity, in which response is predicted to decline monotonically in all directions from, for example, the training stimulus or a stimulus prototype. The famous bell-shaped generalization curve, with its maximum at the training stimulus, exemplifies this. Stimulus summation theory in ethology (Hinde, 1970) is one example of such theory. If a stimulus contains a number of parts, removing one or several of these is predicted to decrease the rate of responding.

Unfortunately, this kind of 'similarity'-based theories cannot predict peak shifts in maximum responding or supernormality. In this sense the so-called gradient interaction theory is more refined. In addition to generalization it also models the interactions between different experiences: for instance when an  $S^+$  (excitatory stimulus) and an  $S^-$  (inhibitory stimulus) are close in appearance, the maximum responding may not occur at the  $S^+$  but for a stimulus more different from  $S^-$ . However, not all peak shifts or supernormality effects can be predicted by gradient-interaction theory. A major phenomenon is intensity generalization, which does not show the bell-shaped curve but a monotonically increasing rate of response (Mackintosh, 1974).

Both peak shift and intensity generalization are natural properties of the multi-layer perceptron. In addition, the degree of generalization is an emergent property of the model. This is in contrast with ‘similarity’ and gradient-interaction theories in which the degree of generalization has to be assumed.

Two very different learning algorithms, the back-propagation and artificial-selection ones, have been used here, and the networks gave the same results irrespective of which one was used, with only minor quantitative differences. In animals, it is still not clear whether there are any fundamental differences between innate and learned recognition (Baerends, 1982). What the network model suggests is that very similar properties can arise from very different learning mechanisms, given that they act on the same system (the same network in the model, the same nervous system in reality).

Also, it should be pointed out that not all models of neural networks have the properties of the multi-layer perceptron. For instance, it is very difficult to build models of behaviour based on the popular Hopfield model (or derived models), since they will not be able to reproduce any of the particular phenomena of stimulus control considered here. On the other hand, some models that have not been thought originally as artificial neural networks share some properties with multi-layer perceptrons. This is the case, for instance, of the model in Blough (1975), which can be considered a network without hidden layers. We feel anyway that network models can be more readily extended to come closer to real nervous systems, in particular by more accurate modelling of neurons, synaptic dynamics and overall organization of the network. In addition, network models force us to think about what is the more appropriate input that has to be fed to the network in order to accurately model the interaction between a real nervous system and physical stimulation. This reduces the possibility of using in a model abstract representations of stimuli in terms of ill-defined ‘features’ of stimulation. For example, Blough (1975) regarded as a drawback the fact that stimulation had to be cut in small elements before being tractable by his model, because he couldn’t attribute any meaning to such elements. In network models these small elements have the very clear meaning of being the sensory cells’ reactions to stimulation (this point is the object of a forthcoming paper).

To conclude, our results suggest that artificial neural networks can be accurate models of behavioural phenomena, providing a valuable tool for researchers in the field of animal behaviour. In particular, some of the concern that these models have raised seems unjustified (see for instance Dawkins & Guilford, 1995).

## Acknowledgments

We thank Laura Lobasso, Giovanni Ghirlanda, Carl-Adam Wachtmeister and Noél Holmgren for their help. The research was supported by grants from the Swedish Science Research Council and Marianne och Marcus Wallenbergs Stiftelse.

We thank Angela K. Turner at Animal Behaviour and Stuart Moffat at Academic Press for returning us the diskette with the manuscript: our copy was lost in a computer accident.

## References

- Baerends, G. 1982. The hering gull and its egg. Part II. The responsiveness to egg-features. *Behaviour*, **82**, 358–363.
- Churchland, P. & Sejnowski, T. 1992. *The computational brain*. Cambridge, MA: MIT Press.
- Czaplicki, J. A., Borreback, D. E. & Wilcoxon, H. C. 1976. Stimulus generalization of an illness-induced aversion to different intensities of colored water in Japanese quail. *Animal Learning & Behavior*, **4**, 45–48.
- Dawkins, M. & Guilford, T. 1995. An exaggerated preference for simple neural network models of signal evolution? *Proceedings of the Royal Society, Series B*, **261**, 357–360.
- Dougherty, D. M. & Lewis, P. 1991. Stimulus generalization, discrimination learning, and peak shift in horses. *Journal of the Experimental Analysis of Behavior*, **56**, 97–104.
- Enquist, M. & Ghirlanda, S. 1998. The secrets of faces. *Nature*, **394**, 826–827.
- Enquist, M. & Leimar, O. 1993. The evolution of cooperation in mobile organisms. *Animal Behaviour*, **45**, 747–757.
- 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.
- Guttman, N. & Kalish, H. 1956. Discriminability and stimulus generalization. *Journal of Experimental Psychology*, **51**, 79–88.
- Guyon, I. P. 1991. Applications of neural networks to character recognition. *International Journal of Pattern Recognition and Artificial Intelligence*, **5**, 353–382.

- Hanson, H. 1959. Effects of discrimination training on stimulus generalization. *Journal of Experimental Psychology*, **58**, 321–333.
- Haykin, S. 1994. *Neural Networks: A Comprehensive Foundation*. New York: Macmillan, 2nd edition.
- Hearst, E. 1968. Discrimination training as the summation of excitation and inhibition. *Science*, **162**, 1303–1306.
- Hinde, R. A. 1970. *Animal Behaviour*. Tokyo: McGraw-Hill Kogakusha, Ltd, 2nd edition.
- Hopfield, J. J. & Tank, D. W. 1986. Computing with neural circuits: a model. *Science*, **233**, 625–633.
- Jenkins, H. M. & Harrison, R. H. 1960. Effect of discrimination training on auditory generalization. *Journal of Experimental Psychology*, **59**, 246–253.
- Jenkins, H. M. & Harrison, R. H. 1962. Generalization gradients of inhibition following auditory discrimination training. *Journal of the Experimental Analysis of Behavior*, **5**, 435–441.
- Kalish, H. & Guttman, N. 1959. Stimulus generalisation after equal training on three stimuli: a test of the summation hypothesis. *Journal of Experimental Psychology*, **57**, 268–272.
- LeCun, Y. 1985. Une procedure d'apprentissage pour reseau a seuil asymetrique. *Cognitiva*, **85**, 599–604.
- Mackintosh, N. 1974. *The psychology of animal learning*. London: Academic Press.
- Mason, J. R. & Reidinger, R. F. J. 1983. Generalization of and effect of pre-exposure on color avoidance learning by red winged blackbirds (*Agelaius Phoeniceus*). *The Auk*, **100**, 461–468.
- McClelland, J. & Rumelhart, D. 1985. Distributed memory and the representation of general and specific information. *Journal of Experimental Psychology: General*, **114**, 159–188.
- Mednick, S. A. & Freedman, J. L. 1960. Stimulus generalization. *Psychological Bulletin*, **57**, 169–199.
- Mezard, M., Parisi, G. & Virasoro, M. A. 1987. *Spin Glass Theory and Beyond*. Singapore: World Scientific.

- Parker, G. 1985. Models of parent-offspring conflict. V. Effects of the behaviour of the two parents. *Animal Behaviour*, **33**, 519–533.
- Pavlov, I. P. 1927. *Conditioned reflexes*. Oxford: Oxford University Press.
- Pearce, J. M., Aydin, A. & Redhead, E. S. 1997. Configural analysis of summation in autoshaping. *Journal of Experimental Psychology*, **23**, 84–94.
- Peterson, N. 1962. Effect of monochromatic rearing on the control of responding by wavelength. *Science*, **136**, 774–775.
- Razran, G. 1949. Stimulus generalisation of conditioned responses. *Psychological Bulletin*, **46**, 337–365.
- Rumelhart, D. E., Hinton, G. E. & Williams, R. J. 1986. Learning internal representation by back-propagation of errors. *Nature*, **323**, 533–536.
- Shanks, D. S. 1995. *The psychology of associative learning*. Cambridge: Cambridge University Press.
- Tapper, D. N. & Halpern, B. P. 1968. Taste stimuli: a behavioral categorization. *Science*, **161**, 708–709.
- Terrace, H. 1966. Stimulus control. In *Operant Behavior* (Ed. by W. K. Honig), volume 1, pp. 271–344. New York: Appleton Century Crofts.
- Thomas, D. & Setzer, J. 1972. Stimulus generalization gradients for auditory intensity in rats and guinea pigs. *Psychon.Sci*, **28**, 22–24.
- Tinbergen, N. 1951. *The study of instinct*. New York and London: Oxford University Press.
- Weiss, S. J. & Schindler, C. W. 1981. Generalization and peak shift in rats under conditions of positive reinforcement and avoidance. *Journal of the Experimental Analysis of Behavior*, **35**, 175–185.
- Zielinski, K. & Jakubowska, E. 1977. Auditory intensity generalization after CER differentiation training. *Acta Neurobiologiae Experimentalis*, **37**, 191–205.