Behav Ecol Sociobiol (1995) 36: 421–429

David C. Krakauer

Groups confuse predators by exploiting perceptual bottlenecks: a connectionist model of the confusion effect

Received: 1 October 1994/Accepted after revision: 4 February 1993

Abstract Aggregation is a well documented behaviour in a number of animal groups. The "confusion effect" is one mechanism thought to mitigate the success of predators feeding on gregarious prey and hence favour aggregation. An artificial neural network model of prey targeting is developed to explore the advantages prey species might derive through a tendency to group. The network illustrates how an abstract model of the computational mechanisms mediating the perception of prey position is able to show a degradation in performance as group size increases. The relationship between group size and predator confusion has a characteristic decreasing decelerating shape. Prey "oddity" is shown to reduce the impact of the confusion effect, thereby allowing predators to target prey more accurately. Hence shoaling behaviour is most profitable to the prey when prey phenotypes are visually indistinguishable to a predator. Futhermore it is shown that prey "oddity" is relatively more costly in large groups than in small groups and the implications for assortative schooling are discussed. Both the model and the results are intended to make the general point that cognitive constraints will limit the information that a nervous system can process at a number of different levels of neural organization.

Key words Confusion effect · Groups · Oddity
Perceptual Constraints · Shoal

Introduction

There is abundant evidence across a number of taxa which suggests that living in groups is a selectively advantageous strategy for individual organisms (Kenward 1978; Myers et al. 1979; Foster and Treherne 1981). Increased foraging success (Pulliam 1976b; Major 1978) and a reduced predation risk (Hamilton 1971; Calvert et al. 1979; Duncan and Vigne 1979) are currently the most empirically and theoretically compelling hypotheses that account for gregariousness. The "confusion effect" describes the reduced attack-to-kill ratio experienced by a predator resulting from an inability to single out and attack an individual prey in a group (Miller 1922; Milinski 1977a, b; Landeau and Terborgh 1986). While predator attack is thought to be less efficient when confronted with a multitude of similar targets, many predators are highly effective stalkers of solitary or conspicuous prey (Milinski 1977b; Pitcher and Parish 1993).

In this paper I present a neural network model of prey targeting which aims to capture some essential properties of a nervous system involved in the computation of spatial position. Artificial neural networks are models that in some way aim to mimic the operational properties of animal nervous systems. The most critical property of these models are therefore the distributed representation of information among the units of the network, and the parallel processing of data by these units. Units in networks are conceived of as the computational analogues of neurons, while connections between the units correspond to synaptic processes. Inputs presented to the network are processed in parallel through the activities of the units and result in an output value, or list of output values, in the form of a vector. The network can be seen as having implemented a mathematical function mapping the inputs onto a set of outputs.

Marr (1982), in an attempt to define the levels of functional organization in the nervous system, identified three levels at which the visual system in particular might be described: the computational, the algorithmic, and the mechanistic. This paper concentrates

---

D. C. Krakauer
BBSRC NERC Behaviour and Ecology Group, Department of Zoology, University of Oxford, South Parks Road, Oxford OX1 3PS, UK

on a simplified model of the *algorithms* required by a simple network to target on a prey object, circumventing the need for a detailed description of any particular nervous system which would correspond to Marr's *mechanisms*. The network demonstrates how computational principles operating at the algorithmic level are likely to operate in real nervous systems to give rise to processing bottlenecks, thereby restricting the accuracy of the input-output function.

In the model that follows, inputs represent some feature of prey objects stimulating the visual system of a predator, while outputs form an internal representation of these features in the form of a topographic map. This is a map in the sense that neighbouring units of the output are made to correspond to neighbouring units in the input, hence spatial relationships in the input are preserved in the computation of the output. The map therefore provides the predator with adaptively important percepts about its external world by indicating the spatial positions of likely food sources. I assume throughout that an accurate representation of prey position on the internal map, translates into an effective consummatory behaviour. Thus I have located the source of the confusion effect in errors of perception, and not in errors experienced during the execution of a motor response.

If one insists on locating this algorithm in the architecture of real nervous systems, it might plausibly be situated in the higher vertebrate's visual system in the retinal, lateral geniculate nucleus (LGN) to superior colliculus visual pathway (Dowling 1987). Alternatively it might be interpreted as the retina, LGN to optic-tectum projections in non-mammalian vertebrates (Sarnat and Netsky 1981) or the ommatidium-lamina-medulla parallel processing channels in arthropods (Laughlin 1983). However, it is *not* important that the model be thought of as a representation of any particular pathway, the results are intended to be general and the aforementioned structures are merely examples of possible sites where these effects are likely to occur.

It is assumed that the visual system of many predators will have been configured during the course of evolution to identify the spatial position of prey objects passing across the visual field, thus eliciting an appropriate goal-directed movement culminating in a consummatory response. The computational capacity of nervous systems is, however, limited and I suggest that prey species are able to exploit the preferences of predators for small prey distributions, by forming groups that strain the informational capacity of a predator's visual system. Thus it is central to this investigation that there are constraints imposed on the information-processing capacity of a nervous system (Broadbent 1965) by a finite set of algorithms or neural connections. The ability of predators to accurately fix the position of a prey object (analogous to the construction of a "spatial spotlight of attention", Alport 1989) depends upon the elaboration of an attentional mechanism which identifies prey position and excludes vacant positions in space. This ability to target prey objects will be investigated in network models of finite computational ability. This constraint on vision should be distinguished from physical and optical limitations on visual acuity, or on spatial resolving power, and is logically distinct from possible constraints acting on "decisions". More precisely, an animal faced with a choice from a number of options must decide between these options according to some criterion. When confusion results from ambiguity acting on the criterion, rather than ambiguity in perceptual information furnishing this criterion, we are not dealing with the 'confusion' effect but another sort of confusion stemming from a conflict of interest or motivation.

Methods

Architecture

Three models of prey-fixation are explored to demonstrate the contribution of increasing numbers of prey and the conspicuousness of prey on a predator's targeting accuracy. Furthermore, "compaction" or closing-in of an aggregation of prey objects is also explored with respect to its effect on targeting accuracy. The models are a set of artificial neural networks shown in previous studies to exhibit properties common to animal recognition systems (Enquist and Arak 1993, Krakauer 1995). The networks are intended to demonstrate properties assumed to occur at multiple levels of information processing. Objects or features of objects in the visual field, are projected onto an internal representation (i.e. map) from which it is possible to extract environmental spatial relationships. It should be noted that the input units in this model are not synonymous with single cells in the nervous system but potentially groups of such cells. Figure 1 is an illustration of the network showing only some

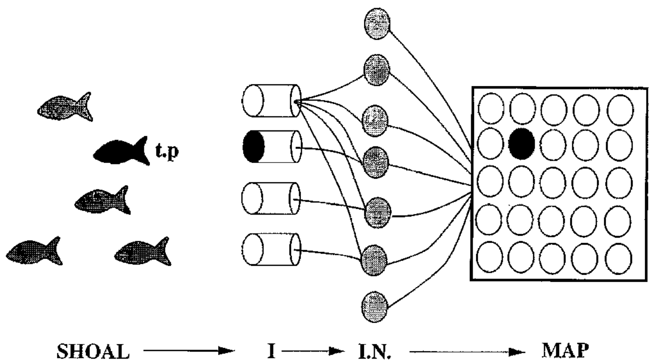

Fig. 1 Schematic illustration of the architecture of the neural network used to perceive a *target* prey (t.p.) in a shoal of fish. The network consists of input units (*I*) responding to features of individual prey , interneurons (*I.N.*) whose connections constitute a receptive field for the network, and a spatially explicit array of cells (*MAP*). These three layers are connected according to the wiring schemes of Fig. 2. The *t.p fish* depicts the *target* prey. The *filled input unit* represents stimulation by the *target* prey and causes an activation of the corrresponding positions on the *MAP*. The *MAP* provides spatial information on the whereabouts of the *target* image in the environment. The *remaining fish* are the distractors

of the connections, which demonstrates the layered architecture proceeding from inputs back via a sequence of connections feeding onto inter neurons (or hidden units) – which define the receptive fields – and then through into further connections that project onto a spatially explicit array or map. The receptive field of one unit is defined as the area of the visual field which upon stimulation elicits a response in this unit. Three 'wiring schemes' are used to configure the connectivities of the network and thereby produce three different types of receptive field. Figure 2a involves non-overlapping visual fields, in which adjacent regions of the inputs are connected exclusively to adjacent regions of the spatial array. Figure 2b involves overlapping visual fields where most connections innervating the map come from corresponding regions of the inputs, but where there is some convergence from more distant regions. In Fig. 2c every input unit connects to each and every map unit, and hence there is no spatial information explicit in the wiring. These connectivities are selected to explore the sensitivity of the confusion effect to anatomical variations in the wiring patterns of sensory channels.

Perceptual systems in most taxa are likely to demonstrate at some points, both the divergence and convergence properties of these networks. We may however be entitled to cariciture the largely parallel connectivities of Fig. 2a with those found in arthropod visual channels (Bullock and Horridge 1965; Laughlin 1983) and the more branching connectivities of Fig. 2b with those found in vertebrates. The divergence and convergence properties of networks can in extreme cases cause information to become distributed over many neural loci or focused at a single neural locus. Divergence maintains signal integrity across many cells but adds little new in terms of power of representation, while convergence increases the power of representation but at a loss of signal integrity. If our caricature is legitimate we might expect the "confusion effect" to play a more significant role in the evasion behaviour of the prey of vertebrates where convergence is high (in contrast to the invertebrates where convergence is often lower, Laughlin 1983). This is a hypothesis seemingly borne out by the loss of the schooling tendency in guppy populations in Trinidad which are heavily predated by freshwater prawns (A.E. Magurran, personal communication).

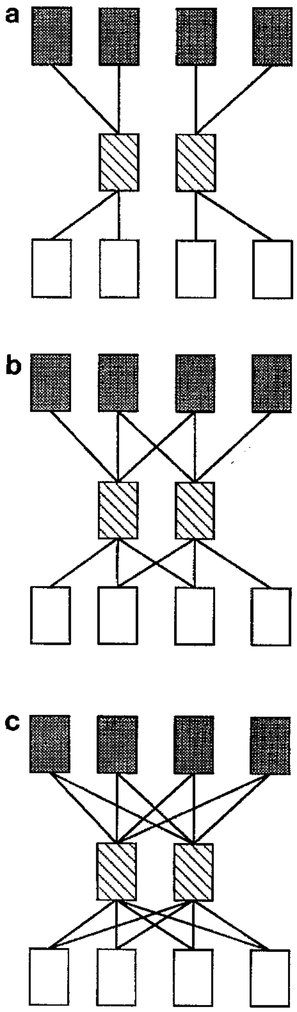

Fig. 2a–c The wiring shemes in two dimensions specifying the connectivity of the networks. The *dark units* make up the MAP, the *cross-hatched units* are interneurons, while the *open units* are the inputs. Only **a** and **b** are true spatial maps where the connectivities preferentially join corresponding positions on the input and MAP units. The inputs are least correlated over the surface of the MAP in **a** and most strongly correlated in **c**

Prey-fixing

The networks are required upon stimulation of an input unit with a *target* prey object to activate the corresponding unit of the spatial array (from here on the map). In other words when a number of input units are set to on, corresponding to a group of prey entering the visual field, the connections from these inputs propagate the stimulation through the hierarchy of network connections from inputs through to outputs culminating with an activation pattern on the map units. The values at the activated map units correspond to the probable spatial position of the input unit occupied by the *target* prey. The connections between the units of the network take particular values referred to as weights which control the influence of the input stimulation on the activation pattern of the map. This series of events is best described algebraically. Let us denote the number of input units and map units as N (16 in the simulations and in Fig. 1) and the number of interneurons as n (4 in the simulations, only 2 shown in Fig. 2). Each one of the input units i can take a value R of 1 if active and -1 if inactive. Interneuron j takes a value T which depends on a sum derived from the input units. Thus

$$T_j = g\left(\sum_{i}^{N} R_i w_{ij} X_{ij}\right) \tag{1}$$

where $g(x)$ is the functions $\tanh(0.5x)$ which constrains the value of T to lie between -1 and 1, w_{ij} are the weights connecting the input units to the interneurons, and X is a matrix that specifies the connectivities or receptive fields of the networks as illustrated in Fig. 2. This matrix is made up of 0s and 1s, with 0s where there are no connections. The values M of the k map units are calculated as follows

$$M_k = g\left(\sum_{j}^{n} T_j v_{jk} X_{jk}\right) \tag{2}$$

where v_{jk} are the weights connecting the interneurons to the map units.

The network weights are modified or trained such that the network preferentially produces activation of the one map unit corresponding to the input unit stimulated with the *target* object. The *target* object is distinguished from non-*target* objects during training by a slightly higher stimulation value to the network at the *target* object position in space. This higher value reflects a choice made by the predator to target a single individual from within the group. The weight training algorithm used is back-propagation (Rumelhart and McClelland 1986), a hill-climbing algorithm which minimises the differences between an observed pattern and a desired pattern by moving the weights of the connections (the matrices w and v) in a direction that minimises this difference. This process is similar to the familiar least-squares procedure used in curve fitting. To be more specific, the expected pattern on the map is simply the stimulation pattern presented to the input unit excluding the non-*target* prey Therefore for each training input pattern, the back propagation algorithm is used to activate the map units corresponding only to the positions of *target* prey in the stimulus pattern. Training is performed with six random configurations of groups with group sizes ranging from one to seven individuals, with 8–15 free spaces on the input. During training the network is therefore exposed to 42 different stimulus configurations. Following training the network is able to identify the position of the *target* prey when presented with a solitary prey or a group of prey objects on the input cells drawn from a testing set up to some given level of accuracy. The M_k vector is scaled such that, $\Sigma_n M_k = 1$, and hence M_k may be interpreted as a probability distribution over the visual field. I stress that the term *target* prey does not imply the recognition of a complete prey form by the predator, simply recognition of a spatially localised target object which might be some feature of the prey (see the Appendix for a full description of the simulation).

The confusion effect

In order to evaluate the performance of the model at the task of prey-fixing during training and testing, a scoring system is employed which provides a measure of confusion. One means of measuring confusion would be simply to plot the probability of identifying the site occupied by the target object M_t, where t is the position of the target on the stimulus pattern, and to do this for different numbers of prey, ignoring whether the predator successfuly identifies the position of other prey that were not the target. This would be a measure of confusion but would not measure the reduced ability of predators to encounter prey *per se.* This is because such a measure might belie a high success rate on non-target prey. For example, neural mechanisms evolved to target specific prey objects might show an improvement in their accuracy at targetting non-specific prey objects, as the number of these objects increases. In such a case, individual prey would not benefit from grouping unless they were consistently the objects of predation in preference to all other individuals. A better measure of accuracy is one that measures success at both encountering the chosen target object and success at encountering objects not selected as the target. In addition to this, one can impose a penalty for targeting vacant spaces in the visual field. Such a measure will reflect the realised success of the network at identifying the positions of all individuals within a group. A predator is rewarded maximally for correctly identifying a chosen target object, slightly less for identifying non-target objects and not punished when correctly excluding vacant positions in space. If we assume that:

a = reward for identifying *target* prey (= 3)
b = reward for identifying *non-target* prey (= 2)
c = penalty for targeting a non-occupied position (= 1)
M_t = probability assigned to the true position of the target object by the predator
M_n = probability assigned to the true position of the non-target objects by the predator
M_v = probability assigned to the true position of the vacant sites by the predator

then we can score the model according to the following expression

$$\text{Accuracy} = \frac{aM_t + \sum_{k \neq k_t} bM_n}{a + \sum_{k \neq k_t, k \neq k_n} cM_v} \qquad (3)$$

This equation provides a maximum score of 1 for correctly identifying the position of *target* prey and eliminating all other sites as likely, a smaller score for identifying non-*target* prey in adddition to the target *prey* and a penalty for targeting empty spaces. This penalty is the right hand term in the denominator of Eq. 3. The a value in the denominator is a scaling constant.

The model is trained to an accuracy score of over 0.75 for each of the connectivities patterns (a, b and c) on each of the 42 stimulus configurations using the back propagation algorithm. To test for prey confusion each of the networks (a, b and c) is presented with a further six random configurations of two, three, four, five, six and seven objects on the input cells where one of these is a target object. The models are required to locate the most probable positions for target objects in increasingly large groups. Note that the specific values assigned to a, b and c are not important, unlike the order of their relative magnitudes which is.

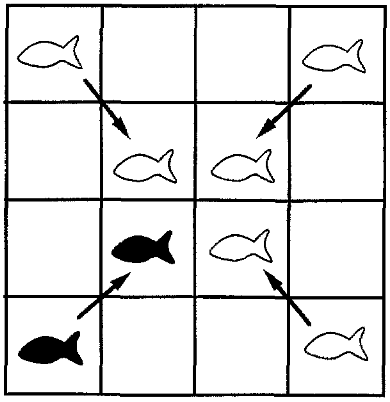

Fig. 3 Prey distribution before and after compaction. A nominal shoal of four individuals with one *target* prey (*solid*) and three non-*target* or distractors (*open*), occupies a widely spaced, diffuse distribution which then closes-in to form a dense, compacted distribution. Both of these distributions are presented to the network to test for the effect of compaction on levels of confusion

Prey oddity

Following the standard training procedure outlined above and in the appendix, to test for an effect of "oddity" the *target* prey is made dissimilar to the other members of the group. This is done by presenting the target prey as an input to the network with a substantially higher signal value than those of the non-target prey. In other words, when the target prey is conspicuous $R_i(\text{target}) - R_i(\text{non-target}) \gg 0$ and $R_i(\text{target}) - R_i(\text{non-target}) \approx > 0$, when the target prey is inconspicuous. The exact nature of this oddity is not specified and hence it may reflect size, morphological irregularity or colour. The networks are scored at each of the five different group sizes where the *target* prey is either conspicuous or inconspicuous.

Compaction

To test whether closing-in behaviour can yield advantages to prey by increasing predator confusion, the networks are scored with four similar prey objects located on the edges of the input array (diffuse) and then with the same number of prey compacted in the centre of the input array. Each network is tested for three diffuse prey distributions and three compacted distributions calculating accuracy scores for each. Figure 3 provides one figurative example of a test configuration before and following compaction.

Results

Effects of aggregation

Predators with all three varieties of wiring schemes demonstrate an increasing inability to accurately target prey objects as the group size increases (Fig. 4a–c). Paired prey objects are more successfully targeted than prey aggregations in excess of two members ($P < 0.05$ in each case, Mann-Whitney U-test: Fig. 4). The advantages accruing to prey through a tendency to aggregate increase with diminishing returns as groups become larger. With the given network architectures, predator performance falls at a decelerating rate to an asymptotic level of accuracy corresponding to approximately random targeting for all three wiring schemes. Wiring scheme a is significantly more accurate at targeting small groups of prey (2, 3 and 4, $P < 0.05$ Mann-Whitney *U*-test: Fig. 4) than wiring schemes b and c.

Effects of oddity

If the *target* prey are made conspicuous in relation to other group members predator accuracy is significantly improved (Fig. 4) at all group sizes.

Oddity versus group size

For wiring schemes *a* and *b* in large groups ($n = 4$) prey oddity reduces the concealing advantage of gregariousness by a greater proportion than in smaller groups ($n = 2$). In other words the probability per group member of predation of an odd individual increases as groups become larger ($P < 0.05$, Mann-Whitney *U*-test: 2a, b).

Effects of compaction

Closing-in behaviour only significantly reduces targeting accuracy in the case where the network employs

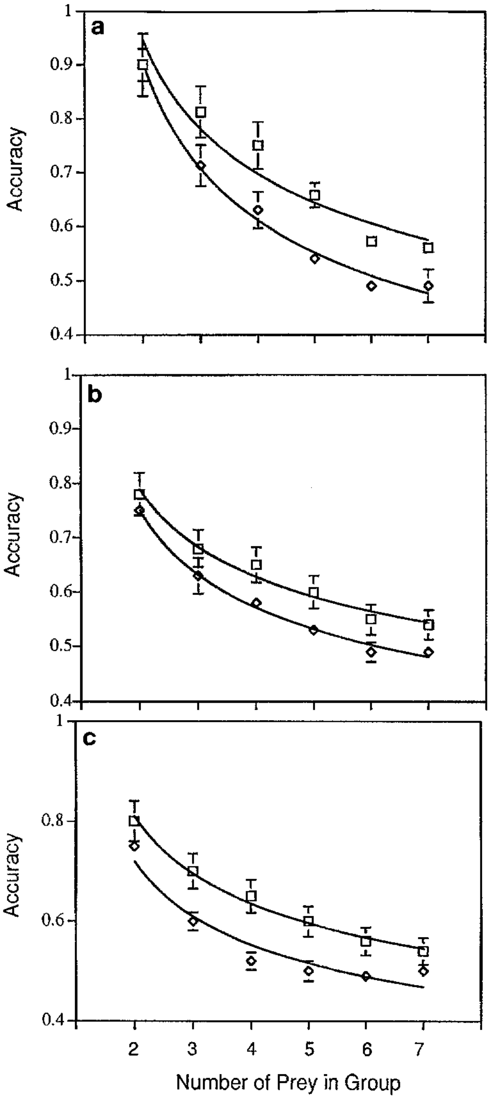

Fig. 4a–c Confusion response curves: the targeting accuracy of a predator presented with groups of increasing numbers of prey. Each of the wiring schemes (**a**, **b**, **c**) shows a reduction in accuracy with an increase in prey number. This accuracy falls off with diminishing returns reaching an asymtotic level of performance above chance levels (*open squares* = conspicuous *target* prey, *open diamonds* = inconspicuous *target* prey). The conspicuous *target* prey are consistently targeted more accurately than the inconspicuous prey. Conspicuous prey are more vulnerable in larger groups for wiring schemes **a** and **b**

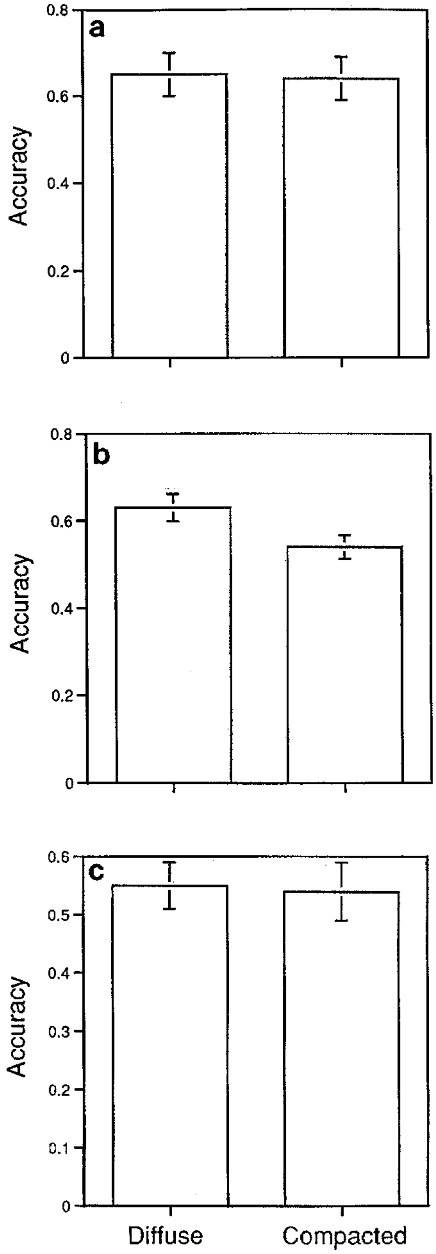

Fig. 5a–c Reduction in targeting accuracy following group compaction for three randomly generated diffuse and compacted prey distributions. Only the wiring scheme **b** shows a reduction in accuracy following compaction through predator confusion

overlapping visual fields ($P < 0.05$, Mann-Whitney U-test: Fig. 5b). Neither the fully connected (Fig. 5c) nor the discrete visual field (Fig. 5a) wiring schemes show a significant reduction in accuracy following compaction. One should note that for Fig. 5c the compacted and diffuse aggregations are spatially congruent with respect to the wiring.

Discussion

Aggregation by prey species is most often accounted for in terms of foraging gains and predator evasion. With the exception of the confusion effect, evasive strategies are behaviours that result in a diminished risk of mortality to select individuals in a group (e.g. marginal predation). Milinski (1977a, b) found that three-spined sticklebacks preferred straying *Daphnia* to aggregations, while Landeau and Terborgh (1986) have provided experimental support for the confusion effect by demonstrating reduced predation for all identical members of a group and increased predation following the introduction of an "odd" individual. An experimental demonstration of the neural principles underlying confusion is hard to achieve and hence this paper has sought to address this difficulty by providing computational evidence in support of "confusion" in an abstract connectionist system (see Appendix). Futhermore, it is suggested that gregariousness be viewed as a form of exploitation (Dawkins and Krebs 1979) where prey exploit the perceptual biases and constraints of predators in order to evade detection.

It has been assumed that a successful predator is one that accurately targets the position of its victim, where targeting is taken as prerequisite for an effective encounter. It is shown that if any predator had an unlimited ability for information processing, groups would not provide a safe haven through the confusion effect, for the confusion effect requires that constraints are explicitly invoked on the part of a predator to account for its inability to attend to multiple targets. To take a neurophysiological example, during visual perception we know that the retina must compress information through a limited number of ganglion cells (Laughlin 1990), a corresponding network would be given a finite number of units in which to store a large amount of information to capture this property. In this model following training on *target* prey, multiple prey are perceived as a collection of objects that compete for representation on cells from which spatial information can be extracted (topographic map). The larger the group, the more information required to pass through the same number of fixed channels to this map. Consequently targeting accuracy drops as more prey enter into the receptive field of a given unit, eventually reaching asymptotic performance above pure chance (Fig. 4). This effect is likely to be very common. An example might be found in an ambush predator such as a pike feeding on a shoal of minnows. One would expect a reduction in targeting accuracy as more minnows crowd the visual field. This is because a large number of prey will increasingly stimulate common receptive fields.

The shape of the confusion response curve demonstrates that the confusion effect is able to provide an immediate advantage to gregariousness (Fig. 4). Groups of three or more prey are better at evading predation than pairs of individuals. Hence this mechanism provides an explanation for the incipient stages of grouping behaviour. It is not clear that few individuals will provide a commensurate advantage via increased vigilance or dilution. Wiring scheme *c* is most effective at reducing confusion when shoal numbers are low. This is because for small numbers of prey, the stimulation is most likely to fall on separate receptive fields. Figure 4 also suggests that there are diminishing returns with increasing group size beyond which additional individuals provide little advantage to the existing members of the group. I should stress here that the precise numbers in this upper group size will depend on the details of the predators visual system. For the sake of analysis this model has been restricted to an artificially small number of units. For most predators there are thousands of cells involved in visual perception and consequently they allow for targeting of many more prey. More exactly, the number of patterns that can be stored in the artificial network scales linearly with the number of units. Hence prey should form aggregations that correspond in some way in their numbers, to the number of cells involved in targeting found in their respective predators. However, the shape of the confusion response curves should remain as they are here.

It is worth labouring the point that confusion tends towards an asymptotic level of accuracy. The existing experimental literature has not yet presented groups sufficiently large to reveal the existence of this ceiling or the decelerating shape of this function. For example, in a hypothetical plot of confusion against swarm density, Milinski (1977a, p. 21 Fig, 11) postulates that confusion should increase exponentially with swarm density. The neural network spells out why this can not be the case: if there are a finite number of input channels, once all of these channels are stimulated, increasing the number of prey will not affect targeting accuracy. In other words the confusion effect is a function of prey number and recognition cell number, not prey number alone. This last point is demonstrated in an experiment on human subjects (Milinski 1990) in which subjects were required to pinpoint dots in artificial two-dimensional swarms. These swarms varied in their densities and it was only in the highest densities (8 dots/cm^2) that confusion set in. At the low and medium densities performance was not hampered by confusion.

Prey oddity reduces the burden on sensory channels by increasing the differences in the incoming signals and therefore provides extra information through intensity. Oddity (stimulus intensity in this model) may therefore be thought of as a signalling dimension in addition to position. Such a result is similar to those of Ohguchi (1981) who placed similar or differently coloured waterfleas in a "confusion machine". This machine shunted a pair of waterfleas in cross-currents between a system of test-tubes visible to a stickleback predator. The stickleback showed a diminished approach rate when the waterfleas were of the same colour and an enhanced response when they were of different colours. Oddity can therefore maintain targeting accuracy above the level found in an equivalently sized phenotypically homogeneous group. This finding is indirectly corroborated by the work of Ranta et al. (1992) who have shown how sticklebacks preferentially join schools which match their own size, a phenomenon they refer to as "assortative schooling". That is not the whole story, however. The results of this paper suggest that an additional component of assortative schooling should be the preference by odd individuals for smaller groups. This is because the relative costs of conspicuousness are lower for an "odd" individual in a small group, than they are for an "odd" individual in a large group (Figs. 4a and b). This result is not found in the fully connected wiring scheme (Fig. 4c). This may be because information is always distributed over the entire map and hence the intensity of this signal is diminished.

Compaction produces this same effect by increasing the correlation between the discrete stimuli, augmenting the number of prey which appear in overlapping visual fields. Schooling fish such as guppies (Seghers 1974), or minnows (Magurran and Pitcher 1987) demonstrate a compaction response when presented by a predatory species. The confusion effect differs from the "selfish herd" hypothesis (Hamilton 1971) in that each individual stands to gain from compaction, and not primarily those at the centre of the huddle. Increased confusion following compaction is not found for wiring scheme 5c because there is no spatial information inherent in the connectivities (it does not preserve spatial information). In Fig. 5a by reducing the overlap between receptive fields, the ratio of information to available channels is similarly negligibly effected following compaction. Only the wiring of Fig. 5b shows compaction to be a beneficial strategy in evading predators.

Groups will only be prevented from growing inexorably if individuals experience other costs associated with gregariousness that reduce the net benefits (Anderson and Wicklund 1978). If there are costs associated with large groups, existing members are expected to close ranks and deny access to aspiring group members. This effect was found by Pulliam and Caraco (1984) who document the attempts made by dominant juncos to evict subordinate birds from a group when resources are threatened.

I have argued that groups provide a means of exploiting some perceptual bottlenecks of a predator implying that prey species are ahead in the evolutionary arms race. Predators are quite likely to have evolved countermeasures in response to their favoured prey forming groups. There are two options available to "confused" predators: surmount the bottleneck with increased cognitive sophistication or disrupting the group to expose the individual. The first of these is difficult to establish, it would require neurobiological data on closely related groups that preferably differ only in the distributions of their respective prey species. The hypothesis tested would be that predators which are required to secondarily predate gregarious prey should manifest neurobiological adaptations specific to this task. The second option is commonly observed in a number of groups: piscine predators feeding on shoaling fish are able to disrupt the structure of schools thereby gaining access to isolated individuals (Major 1978), yellowtail predators cooperate to isolate parts of a school of jack mackerel (Schmidt and Strand 1982), and in an example involving quadrupeds, lionesses cooperate to isolate and hunt down grazing zebras from their herds (Schaller 1972).

These observations do not exclude alternative explanations such as marginal predation. The distinguishing features of the confusion effect and its biological consequences as demonstrated by this model should be stressed; these are:

1. Gains through reduced predation for all members of a group. It is numbers of individuals rather than the relative position within the group which influences predator targeting accuracy.
2. A decelerating drop to an asymptotic level of predator accuracy as group size increases. As numbers increase groups will cease to provide a concealing advantage to existing members, this precise value will depend on the sophistication of the predators nervous system. If there are costs to living in groups, the confusion effect can predict an optimum group size which is a product of confusion benefits with group costs such as interference.
3. Increased gains following compaction. Reducing the free space between individuals increases the information burden placed on a predators sensory channels. This is not the same as dropping below the resolving power of the retina. The centripetal instinct should evolve as a means of evading detection and not merely as a means of evading marginal predation.
4. The confusion effect is most effective when members of the group are all alike. Conspicuous individuals that appear "odd" with respect to the other members of the group do not gain from joining a group.

Hence selection should favour 'assortative schooling' (Ranta et al. 1992).

5. Odd individuals suffer proportionately higher predation in large groups than they do in small groups. We might therefore expect heterospecific aggregations with divergent phenotypes to be more common in small groups than in large groups.

6. Since the principal agency of the confusion effect is numbers and spacing, it should operate with inert/sedentary target species. This suggests that individual species which seek to encourage approach, such as plants with their pollinator's, should space out so as to exploit the perceptual preferences of the pollinators.

If the results of these a models are held to apply in some way to the principles governing nervous activity in animals, the confusion effect offers a substantive explanation for aggregative behaviour. It is quite likely that motion introduces another dimension into predator confusion and the wealth of empirical observations support this assertion (Pitcher and Parish 1993). It is assumed that unpredictable movement by group members will serve to increase confusion but this property of groups has not been examined here. This paper has sought to demonstrate confusion in the simplest case, that of a static assemblage.

While artificial neural networks scarcely resemble networks of biological neurons, it is possible that they share some basic features by virtue of their connectivity and dynamical rules (Churchland and Sejnowski 1994). Small artificial networks might therefore capture some essential properties of neural processing. Recent studies adopting such an approach include those of Hinton and Shallice (1991) who using networks of the sort employed in this paper explore the possible basis of "deep dyslexia". Similarly Smolensky (1986) has employed networks which demonstrate features of procedural memory. Grossberg (1984) has employed a variety of simple networks in order to understand disorders such as Parkinson's disease and schizophrenia, while Enquist and Arak (1993) have employed three-layer networks to reveal perceptual preferences thought common to evolving perceptual systems. These examples are intended to justify the use of models that opt out of physiological verisimilitude. The confusion effect has remained little more than a black-box explanation for a set of behavioural observations. This paper has sought to ground these observation in a simple theory of nervous action.

Acknowledgements Many thanks to Anne Magurran, Alasdair Houston, Mark Pagel, Magnus Enquist, Manfred Milinski and Charlie Paxton for their ideas and comments on the manuscript. All of whose thoughts improved the manuscript significantly. This research was funded by BBSRC and NERC.

Appendix

The network models employed in all the simulations in this paper are N-M-N encoders. There are N inputs and outputs ($N = 16$) and M hidden units ($M = 4$). The connections between the layers of the network are established by the connectivity matrices illustrated in Fig. 2. The network is trained to perform a specific input-output mapping, namely, activate a single unit on the topographic map corresponding to only the *target* unit stimulated in the input pattern. Such a mapping is implemented by configuring the weight values of the network using the back propogation algorithm. The use of such an algorithm is one means of finding acceptable weight values. An evolutionary algorithm of the type employed by (Enquist and Arak 1993) would do just as well.

We select an input pattern R_i (see below for more details of this part) and propogate the signal forward through the network onto the hidden units T_j

$$T_j = g\left(\sum_i^N R_i w_{ij} X_{ij}\right)$$

and then onto the output units M_k

$$Mk = g\left(\sum_j^n T_j v_{jk} X_{jk}\right).$$

We then compute the difference between the computed outputs M_k and the desired output D_k (which is identical to the map except only a single unit is activated corresponding to the *target* units) as a least squares error measure:

$$E = \frac{1}{2}\sum_k\left[D_k - g\left(\sum_j^n g\left(\sum_i^N R_i w_{ij} X_{ij}\right) v_{jk} X_{jk}\right)\right]^2$$

We can then differentiate this function with respect to the weight matrix v_{jk} and thereby move the weights in a downhill direction in the direction of a minimum. The change in the weights is therefore given by:

$$\Delta v_{jk} = -\alpha\,\frac{\partial E}{\partial v_{jk}} = \alpha\sum(D_k - M_k)g'\left(\sum_j^n T_j v_{jk} X_{jk}\right)T_j$$

The change in the input units to hidden unit weights are found by differentiating the error measure with respect to the weight matrix w_{ij}:

$$\Delta v_{jk} = -\alpha\,\frac{\partial E}{\partial w_{jk}} = -\alpha\frac{\partial E}{\partial v_{jk}}\,\frac{\partial v_{jk}}{\partial w_{jk}}$$

$$= \alpha\sum(D_k - M_k)g'\left(\sum_j^n T_j v_{jk} X_{jk}\right)v_{jk}\,g'\left(\sum_i^N R_i w_{ij} X_{ij}\right)R_i$$

where α is a learning rate constant set equal to 0.1.

The input values presented to the input units are 0.6 for the *target* object, 0.5 for the remaining objects are 0.0 for all empty spaces. Each input pattern receives 50 replicates of training on 6 randomly generated configurations for group sizes of one to seven members. That is, the algorithm cycles through the six configurations for each groupsize in turn, moving on to the next larger group size until reaching seven and then returning to groups of one member. This cycle is continued 50 times so that each stimulus configuration will have been exsposed to training a total of 50 times but never in continuous succesion.

References

Alport A (1989) Foundations of cognitive science. MIT Press, Cambridge

Broadbent DE (1965) Information processing in the nervous system. Science 22:457–462

Bullock TH, Horridge GA (1965) Structure and function in the nervous systems of invertebrates, vol 2. Freeman, San Francisco

Calvert WH, Hedrick LE, Brower LP (1979) Mortality of the monarch butterfly, *Danaus plexippus*: avian predation at five overwintering sites in Mexico. Science 204:847–851

Churchland P, Sejnowski T (1994) The computational brain. MIT Press, Cambridge

Dawkins R, Krebs JR (1979) Arms races betweenand within species. Proc R Soc Lond B 205:489–511

Dowling JE (1987) The retina: an approachable part of the brain. Belknap Press, Cambridge

Duncan P, Vigne N (1979) The effect of group size in horses on the rate of attacks by blood-sucking flies. Anim Behav 27:623–625

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

Foster SA, Treherne JE (1981) Evidence for the dilution effect in the selfish herd from fish predation on a marine insect. Nature 295:466–467

Grossberg S (1984) Some normal and abnormal behavioural syndromes due to transmitter gating of opponent processes. Biol Psychiatry 19:1075–1117

Hamilton WD (1971) Geometry of the selfish herd. J Theor Biol 39:295–311

Hinton GE, Shallice T (1991) Lesioning an attractor network: investigations of acquired dyslexia. Psychol Rev 98:74–95

Kenward RE (1978) Hawks and doves: factors effecting success and selection in goshawk attacks on wood-pigeons. J Anim Ecol 47:449–60

Krakauer DC (1995) Simple connectionist models of spatial memory in bees. J Theor Biol 172:149–160

Landeau L, Terborgh J (1986) Oddity and the "confusion effect" in predation. Anim Behav 34:146–153

Laughlin SB (1983) The roles of parallel channels in early visual processing by the arthropod compound eye. In: Ali MA (ed) Photoreception and vision in invertebrates Plenum, New York, pp 457–481

Laughlin SB (1990) Coding effiiciency and visual processing. In: Blakemore C (ed) Vision, coding and efficiency. Cambridge University Press, Cambridge

Magurran AE, Pitcher TJ (1987) Provenance, shoal size and the sociobiology of predator evasion behaviour in minnow shoals. Proc R Soc Lond B 229:439–465

Major PF (1978) Predator-prey interactions in two schooling fishs, *Caranx ignobilis* and *Stolephorus purpureus*. Anim Behav 26:760–77

Marr D (1982) Vision. Freeman, San Francisco

Milinski M (1977a) Do all members of a swarm suffer the same predation? Z Tierpsychol 45:373–388

Milinski M (1977b) Experiments on the selection by predators against spatial oddity of their prey. Z Tierpsychol 43:311–325

Milinski M (1990) Information overload and food selection. In: Hughes RN (ed) Behavioural mechanisms of food selection (NATO ASI Series G: Ecological Sciences, vol 20). Springer, Berlin Heidelberg New York, pp 721–736

Myers JP, Conners PG, Pitelka FA (1981) Territory size in wintering sanderlings: the effects of prey abundance and intruder density. Auk 96:551–561

Ohguchi O (1981) Prey density and selection against oddity by three spined sticklebacks. Adv Ethol 23:1–79

Pitcher TJ, Parish JK (1993) Functions of shoaling behaviour in teleosts. In: Pitcher TJ (ed) Behaviour of teleost fishes. Chapman Hall, London

Pulliam HT (1976) The principle of optimal behaviour and the theory of communities. In: Klopfer PH, Bateson PPG (eds) Perspectives in ethology. pp 311–332. Plenum Press, New York

Pulliam HR, Caraco T (1984) Living in groups: is there an optimal group size? In: Krebs JR, Davis NB (eds) Behavioural ecology: an evolutionary approach, 2nd edn. Blackwell, Oxford, pp 122–147

Ranta E, Lindström K, Peukhuri N (1992) Size matters when three spined sticklebacks go to school. Anim Behav 43:160–162

Rumelhart DE, McClelland JL (1986) Parallel distributed processing: explorations in the microstructure of cognition, vol I. Foundations. MIT Press, Cambridge

Sarnat BH, Netsky MG (1981) Evolution of the nervous system, 2nd edn. Oxford University Press, Oxford

Schaller GB (1972) The Serengeti lion. University of Chicago Press, Chicago

Seghers BH (1974). Schooling behaviour in the guppie (*Poecilia reticulata*): an evolutionary response to predation. Evolution 28:486–489

Schmidt RJ, Strand SW (1982) Cooperative foraging by yellowtail, *Seriola lalandei* (Carangidae), on two species of fish prey. Copeia 714–717

Smolensky P (1986) Harmony theory. In: Rumelhart DE, McClelland JL (eds) Parallel distributed processing, vol 1 MIT Press, Cambridge, pp 194–281

Communicated by P. Pamilo