Inferring the rules of interaction of shoaling fish

James E. Herbert-Read[a,1,2], Andrea Perna[b,1], Richard P. Mann[b], Timothy M. Schaerf[a], David J. T. Sumpter[b], and Ashley J. W. Ward[a,3]

[a]School of Biological Sciences, University of Sydney, 2006 Sydney, Australia; and [b]Department of Mathematics, University of Uppsala, Box 480, 75106 Uppsala, Sweden

Collective motion, where large numbers of individuals move synchronously together, is achieved when individuals adopt interaction rules that determine how they respond to their neighbors' movements and positions. These rules determine how group-living animals move, make decisions, and transmit information between individuals. Nonetheless, few studies have explicitly determined these interaction rules in moving groups, and very little is known about the interaction rules of fish. Here, we identify three key rules for the social interactions of mosquitofish (*Gambusia holbrooki*): (*i*) Attraction forces are important in maintaining group cohesion, while we find only weak evidence that fish align with their neighbor's orientation; (*ii*) repulsion is mediated principally by changes in speed; (*iii*) although the positions and directions of all shoal members are highly correlated, individuals only respond to their single nearest neighbor. The last two of these rules are different from the classical models of collective animal motion, raising new questions about how fish and other animals self-organize on the move.

collective animal behavior | fish shoals | group motion | self-propelled particles | self-organization

Collective motion of animal groups occurs when multiple individuals move synchronously, producing large scale "flocking" patterns (1–5). Numerous models have been developed to describe patterns of collective motion in terms of interactions between individuals (6–9). These simulation models usually assume that individuals move at a constant speed and their interactions are mediated through direction changes (1). Often these models use zonal rules, where individuals move away from neighbors at close distances and align and/or move toward neighbors at greater distances. Interactions can be with either all neighbors within some zone (7) or with a set of *n* nearest neighbors (10). These and other models have succeeded in reproducing qualitatively similar patterns to those seen in the collective motion of animal groups in nature.

It remains unclear, however, whether the interaction rules implemented in models are the ones used by animals. Indeed, many collective motion patterns observed in nature can be simulated by models using very different interaction rules (1). We are only now beginning to accumulate evidence about which interaction rules are used. There has been recent identification of zones of repulsion and alignment in surf scoters (11). The structure of starling flocks is consistent with topological interactions between the birds (10). Homing pigeons appear to have hierarchical interactions such that birds with higher route-following fidelity act as leaders (12–15). Partridge showed that lateral line and vision are both important in producing directional alignment in Gadid fish (16). Nonetheless, there remain a large number of open questions about the interactions of fish. For example, do fish adopt attraction and alignment within distinct zones as purported in most models? How many neighbors do fish interact with? Do fish swim at more or less the same speed, with changes in angle mediating a fish's response to conspecifics? Answering these questions requires detailed study of the rules of interaction of particular fish species.

We set out to determine which rules individual fish used by analyzing data on the movements of individuals in groups of mosquitofish (*Gambusia holbrooki*). We first present visualizations of interactions and identify correlations between neighbors in order to determine which (if any) of the above rules are relevant to mosquitofish. In reality, a fish's movement from its current location will be a response to both social cues and environmental topography. These two interactions are hard to separate by correlation alone, yet it is clearly important to do so in order to isolate and understand the rules used in social interactions (17). We isolate the social interactions from other effects using techniques taken from machine learning (18–20). Assuming that the behavioral response of the focal fish can be decomposed into a mixture of the previous behavior of the focal fish, its position in its environment and the position on its *n* nearest neighbors (Fig. 1), we used auto-regression and nonlinear function estimation through artificial neural networks (21) to learn a model that fits the movements of the focal fish as a combination of these variables (22).

Results

We filmed groups of 2, 4, or 8 fish (12 trials of each) (*Gambusia holbrooki*) in a square arena for 5 min (at 15 fps). Using semiautomated tracking software we subsequently analyzed the trajectories of each fish and how they responded to the position and orientation of their neighbors (see *Materials and Methods*). We recorded the following key parameters: ϑ, the relative position of the neighbor with respect to the focal fish; φ, the orientation of the neighbor with respect to the focal fish, r, the distance between the neighbor and the focal fish; and α, the change in direction of the focal fish (Fig. 1).

In terms of acceleration, the fish showed clearly defined zones of repulsion and attraction. Fish accelerated when neighbors were directly behind them (within 6 cm; approximately 2 body lengths) and decelerated when neighboring fish were directly in front of them (Fig. 2 *A* and *B*). At these short distances, the changes to acceleration were more pronounced when individuals were facing each other—i.e., ϑ and φ have opposite signs (Fig. 4*A*). This acceleration response was dependent on the fish's current speed (Fig. 2*C*); fish moving faster decelerated more than fish moving slower.

The focal fish were generally attracted to neighbors that were more than 6 cm from them (Fig. 2 *A*, *B*, and *C*). This was manifest

Fig. 1. Parameters that characterize the position of the focal fish (represented in red) with respect to neighbors and to the walls of the experimental basin. r: distance from the neighbor. d: distance from the closest wall. ϑ: direction to the neighbor. φ: orientation of the neighbor. γ: orientation to the nearest wall. α: change in direction of the focal fish. All angles are computed with reference to the current moving direction of the focal fish.

by fish accelerating toward neighbors in front of them and decelerating in response to neighbors behind them. The transition from repulsion to attraction occurs at about 6 cm, which is close to the mean distance between a fish and its nearest neighbor (mean, median, and modal distance to nearest neighbor are 6.2 cm, 3.8 cm, and 2.6 cm, respectively; Fig. S1*A*). There was no clear relationship between the orientation of neighbors in the attraction zone and the acceleration of the focal fish (Fig. 4*B*). These results suggest that the fish actively regulate the distance to their nearest neighbor, mediated through acceleration, but only respond to orientation in cases where collision is imminent.

The angle to neighbors in the attraction zone had an effect on the turning angle of the focal fish (Fig. 2*E*); fish turned toward the position of their neighbor. This effect increased as a function of ϑ, so that it was greatest when neighbors were at a right angle to the fish's current heading (Fig. 2*F*). This is consistent with the observation that nearest neighbors were more likely to be found directly ahead of or behind each other (Fig. S1*B*). Unlike acceleration, turning angle did not strongly depend on the distance to

neighbors. Thus, the repulsion zone involves fish sharply stopping in response to neighbors within two body lengths, while changing their direction to face toward them.

While fish orientated to face their neighbor's position, alignment with the orientation of neighbors was less important. Fig. 4 *C* and *D* shows a very clear difference in response to direction, ϑ, and less clear pattern in response to orientation, φ (see also Fig. S2). There is possibly a weak effect on the change in angle of a focal fish, whereby when ϑ is approximately equal to $\varphi \pm \pi$, the focal fish produces a stronger turning response toward the neighbor. This would correspond to a fish responding more strongly when neighbors are traveling directly away from (or toward) it.

If fish do not systematically align with their neighbors' orientation, how do groups show strong polarization? One explanation is that fish are attracted more strongly to neighbors in front of them than those behind, producing an asymmetry of leading and following that promotes group alignment (23). Fig. 3 provides empirical support for such leading and following behavior. A fish's direction is maximally correlated with the direction of a fish in front of it after a 1 s time lag, suggesting that the fish behind follows the fish in front, thereby coming into alignment with this neighbor.

The accelerations and turning angles of the fish were also affected by the walls of the tank and the immediate previous behavior of the fish. For example, fish tended to be positioned parallel to and about 5 cm from the wall of the tank (Fig. S3*A*). Moreover, fish generally maintained a cruise speed of 6.9 cms^{-1} (Fig. 2*D*). The influence from these factors cannot be immediately separated from the social interactions we have observed thus far. By assuming that the effect of the fish's past behavior, its interactions with the environmental topography (walls and obstacles), and its social interactions with neighbors are additive, we can estimate each by fitting an additive mixture of response functions. Specifically we model the turning angle and the acceleration of the fish as a sum of five components (see Eq. **1**, *Materials and Methods*). First, a linear auto-regression component mapping past turning angles and accelerations to their current
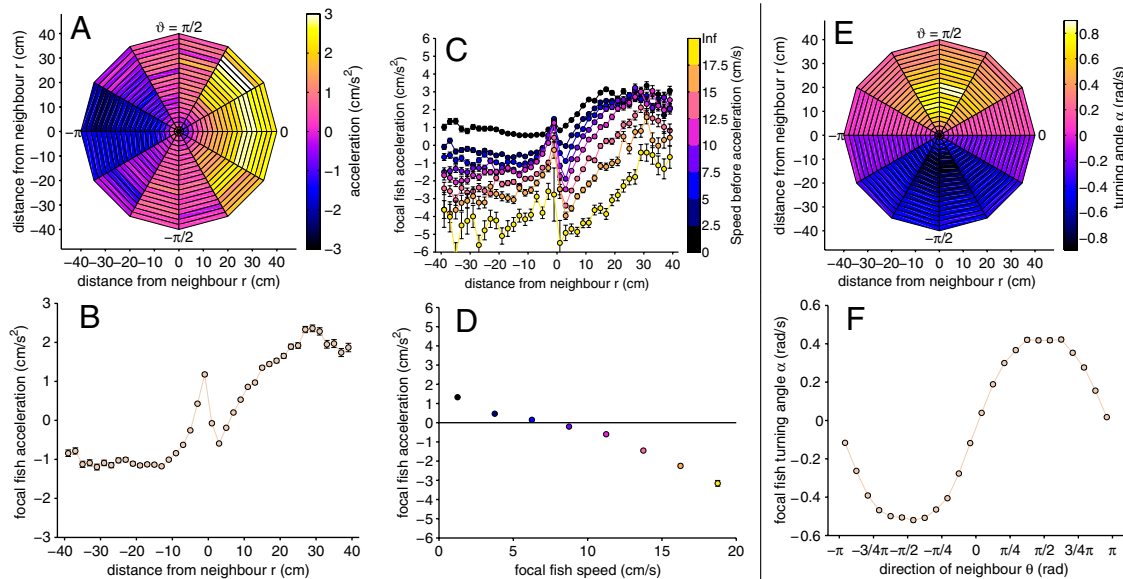
Fig. 2. (*A*) Acceleration of the focal fish vs. distance (r) and position (ϑ) of the neighbors. (In these plots, the focal fish is located at the origin and faces to the right.) (*B*) Acceleration of the focal fish vs. distance of the neighbors. Positive or negative x-axis values indicate the neighbor was respectively in front of or behind the focal fish. (*C*) Same plot as in *B*, but each curve corresponds to a different speed of the focal fish immediately before the acceleration response. (*D*) Average acceleration of the focal fish vs. own current speed. (*E*) Average turning angle of the focal fish as a function of distance and position of all neighbors. Negative values indicate a right turn, positive values indicate a left turn. (*F*) Average turning angle as a function of the position of the neighbor (ϑ). Data are calculated using the position of all neighbors; error bars (often smaller than symbol size) represent standard errors.
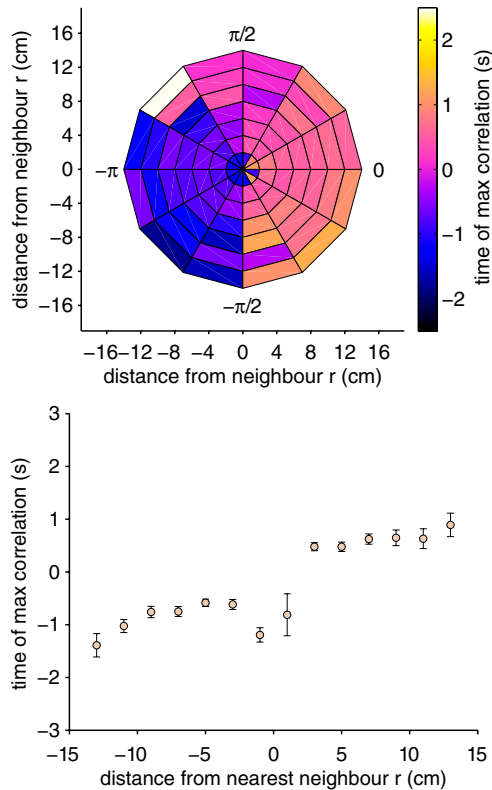
Fig. 3. (*Top*) Directional correlation delay vs. distance (r) and position (ϑ) of the nearest neighbor. In order to avoid spurious correlations introduced by the experimental setup, the analysis is limited to video frames in which the focal fish is far from the corners of the basin (>15 cm). (*Bottom*) Directional correlation delay vs. distance of the nearest neighbor. Positive or negative x-axis values indicate the neighbor was respectively in front or behind the focal fish.

values. This allows us to account for the extended period of time a fish takes to complete its turning or acceleration movements. Second, a nonlinear function of the focal fish's distance d and angle γ to the nearest wall. The third, fourth, and fifth components are functions of the relative position of the first, second, and third nearest neighbors, respectively, specified by the distance r_i and angle ϑ_i to neighbor i. These nonlinear components are fitted using artificial neural networks (see *Materials and Methods*). Maximum-likelihood methods are then used to train the sum of these components. Fig. 2A suggests that the acceleration response to neighbors is symmetric about the axis of the focal fish. Similarly, the turning response should be antisymmetric about the same axis (Fig. 2E). Given these observations, we assume both of these symmetries in our fitting.

The model fitting confirms all our earlier observations and also provides more detail about the components of the interactions. A fish's acceleration was principally controlled by the position of the wall (Fig. 5A) and its nearest neighbor (Fig. 5B). The focal fish tended to accelerate when moving perpendicular and away from the wall ($\gamma = -\pi/2$). The walls had otherwise only weak effects on the fish's acceleration (Fig. 5A) and turning (Fig. 5E). The acceleration toward the nearest neighbor was strongest when it was between 10 and 20 cm and directly in front of the focal fish (Fig. 5B). Deceleration in response to a neighbor behind the focal fish was also present but was weaker than the acceleration response (Fig. 5B). There was also evidence of a repulsion zone close to the focal fish. Model fitting of the turning angle of a fish again confirmed our earlier observations; the turning response to the wall was weak (Fig. 5E), and the focal fish turned most strongly in response to the nearest neighbor when it was 10 cm away and at right angles to the motion of the focal fish (Fig. 5F). The second and third nearest neighbors had a comparatively much weaker effect on both acceleration and turning angle (Fig. 5 C, D, G, and H). For example, the first nearest neighbor 15 cm in front of a focal fish produces an average acceleration five times larger than the second nearest neighbor at an equiva-

Fig. 4. (*A* and *B*) Average acceleration of the focal fish vs. position (ϑ) and orientation (φ) of the neighbors. (*A*) repulsion zone ($r \leq 6$ cm); (*B*) attraction zone ($r > 6$ cm); (*C* and *D*) average turning angle (α) of the focal fish vs. position (ϑ) and orientation (φ) of the neighbors. (*C*) repulsion zone ($r \leq 6$ cm); (*D*) attraction zone ($r > 6$ cm).

Fig. 5. Effect of the wall, first, second and third nearest neighbor on the acceleration (*A–D*) and turning angle (*E–H*) of the focal fish. The color scale indicates the size of the acceleration or turning response as a function of the relevant environmental or social variables. Each panel shows the predicted component of the acceleration or turning angle as a function of the variables indicating the positions of the wall and the nearest neighbors. The total predicted response, either acceleration or turning angle, is the addition of all four of these components. Each panel is plotted as a semicircular arc as we assume left-right symmetry in the acceleration response and antisymmetry in the turning. Only data from groups of four fish are used for producing this figure; see Fig. S6 for similar results from groups of eight fish. See *Materials and Methods* for details of the analysis.

lent position. Additionally, the responses to the second and third nearest neighbor show less spatial structure than those responses to the nearest neighbor.

Discussion

With a combination of high resolution data on group movements and function fitting, we have tested the assumptions made by models of the collective motion of fish. We have identified some key similarities with many model assumptions—namely, that fish responded to the position of their neighbors through short-range repulsion and longer-range attraction rules. These rules have been proposed by a number of models (7, 24, 25, 26). There are also some inconsistencies between the classical models of fish motion and our results. For example, we do not find evidence that fish aligned specifically with their neighbor's orientation; a rule adopted by some models (7, 26). Instead, it appears that group alignment is achieved in some other way, possibly through attraction and repulsion rules and by fish following individuals in front of themselves. Models have already shown that such interactions can produce highly polarized groups (23, 27). However, this is not to say that alignment rules are never adopted by this or other species. Under simulated predation threat, it is clear that alignment with neighbors can allow information to be transmitted rapidly through fish schools (28). Whether fish change their adopted rules depending on context is an intriguing possibility that warrants further investigation.

Another difference is the importance of acceleration in response to a neighbor's position. Unlike many self-propelled particle models, where speed is fixed and interactions are mediated through changes in direction (1, 6), mosquitofish actively changed their speed in order to avoid or move toward neighbors. Other studies on bird flocks and fish schools have also found strong evidence that individuals' speeds fluctuate in groups (29–31). Clearly, therefore, it is important for future models to incorpo-

rate changes and differences in speeds, which is likely to affect a number of group level properties.

Perhaps the most surprising finding of our study is the extent to which the single nearest neighbor dominates social interactions. Previous observations on birds have identified interacting neighbors using correlations in movement (1, 15) or anisotropy of the position of the nearest neighbors (10). These methods have typically identified multiple interacting neighbors. Given that we find only one interacting neighbor using our function fitting method, it would be interesting to see if similar results hold for birds and other species. Indeed, if we apply a purely correlational approach to our data, we find that the turning angle of a focal fish is as correlated with its second and third neighbors as it is with its first, suggesting multiple interacting neighbors (Fig. S4). These correlations likely arise through multiple interactions between the fish over time and do not necessarily reflect their immediate rules of motion.

Previous experiments on mosquitofish have shown that information about the position of predators can be transmitted rapidly between group members (32). A question remains as to whether these nearest neighbor interactions are sufficient to explain these observations or if the fish change their interactions in risky situations. Early models of collective motion suggested that interactions with more than one neighbor are required for collective motion (33). More recent models have shown that single neighbor interactions can produce complex schooling patterns (5). Further modeling work is required to understand the group-level outcome of the rules we have identified. Mosquitofish form relatively small shoals in freshwater, unlike the larger pelagic shoals formed by some marine species, so we would expect models based on the rules we have identified to produce only a limited range of collective patterns.

There is a qualitative difference between paying attention to one neighbor or to multiple individuals. Responding to two or

more neighbors is likely to bear significant costs in information processing time because it requires integrating positional information over the whole visual field (34). Integrating larger quantities of information may be a strategy affordable to animals with larger brains such as birds or mammals but less suitable to animals such as small shoaling fish. However, we are beginning to appreciate that even humans might rely on relatively simple rules to effectively navigate their environment (35). A model on retinal information processing shows that having a threshold mediated response plus sensitivity to particular information can allow individuals to filter out unimportant information from multiple neighbors while responding to strong directed motion of others (36). Such a mechanism is indeed compatible with our findings and may explain how individuals can disregard information from multiple neighbors.

Future studies should attempt to identify the interaction rules between individuals in a comparative way across species. Under different selection pressures different interaction rules are likely to be selectively advantageous (2, 37, 38). Our aim should be to compare the interaction rules adopted by different species, classify the group-level properties generated by these rules, and discuss these in the light of evolution and phylogenetic considerations. If the rules adopted by individuals are similar across distantly related species, then this will show their effectiveness in driving coordinated group movement under different evolutionary pressures. On the other hand, universality in the patterns generated by groups may mean that each species has a very distinct set of rules, each leading to the same set of global patterns (1). Our current study shows us that a detailed analysis of animal interactions of a particular species can provide a different picture of collective motion than we imagine when we concentrate solely on building models.

Materials and Methods

Female mosquitofish (*Gambusia holbrooki*) were collected from Lake Northam (33°53′07, 15°11′35), Sydney, Australia. Fish were kept in 50 L aquaria on a 12:12 hour, dark:light photo period and were fed flake food ad libitum. An experimental arena (730 × 730 × 150 mm; see Movie S1) made of 10-mm acrylic was filled to a depth of 70 mm with dechlorinated water and two liters of water taken from a housing tank containing over 50 mosquitofish. This was done to keep the amount of chemical cues between trials constant. Two circular sandstone cores were placed 240 mm in from two of the arenas corners to add structural complexity to the arena. A frame with an opaque black plastic curtain visually isolated the arena in order to prevent external disturbances to the fish. The arena was lit by fluorescent lights. In one corner of the arena was a transparent and perforated door, which could be remotely lifted with a poly-nylon cord. Fish were held in this holding corner (behind the door) prior to the start of trials. For each trial we placed a group of 2, 4, or 8 individuals in the holding corner for 5 min in order to acclimate the fish to the arena. Following this period, we raised the door, allowing the fish to enter the arena. A webcam (Logitech Pro 9000) placed directly over the arena recorded the movements of the fish for 5 min (at 15 frames per second). Twelve replicates of each group size were done.

Data Collection. Videos were converted from .wmv to .avi format using RAD Video Tools (v 1.9n), uncompressed using VirtualDub (v 1.9.2), and imported to an automated tracking software program, CTrax (v 0.1). This program automatically tracked the position of each fish in each trial, at every frame in the 5-min video, giving 4,500 x and y coordinates of each fish in each video. Data were subsequently imported into MATLAB®, and any errors that the tracking software had made were manually corrected using the associated MATLAB® FixErrors GUI. For a full description of the tracking-software see ref. 39. The x and y coordinates in pixels were converted into cm using a conversion ratio determined by measuring the distance (in pixels) between two corners of the tank (1,032 mm). Coordinates were smoothed using a Savitzky & Sgolay smoothing filter with span 7 frames (approximately 1/2 second) to remove any small spurious changes in position.

Data Analysis. In turn, each fish in the shoal is selected as the "focal fish." At each frame t we measure its instantaneous speed

$$s(t) = \sqrt{(x(t) - x(t-1))^2 + (y(t) - y(t-1))^2}/dt,$$

where $x(t)$ and $y(t)$ are the x and y coordinates of the focal fish at time t and dt is the length of the time interval ($dt = 1/fps$). The direction of movement of the focal fish is

$$\psi(t) = \text{atan2}(y(t) - y(t-1), x(t) - x(t-1)),$$

with respect to the positive x axis. The position and direction of movement of all the other fish in the shoal and the position of the corners and walls of the basin are then expressed in the reference system provided by the focal fish coordinates and direction of movement (Fig. 1). Two parameters characterize the response of the focal fish to the conspecifics and to the environment: Its tangential acceleration

$$a(t) = (s(t) - s(t-1))/dt$$

and its speed of direction change

$$\alpha(t+1) = (\psi(t) - \psi(t-1))/dt,$$

where care is taken to compute the correct angular difference, $\psi(t) - \psi(t-1)$, with regard to the periodicity of $\psi(t)$. We compiled the data from all group sizes together (unless otherwise stated). Analysis revealed no qualitative differences in acceleration or turning angle between group sizes (see Fig. S5).

Function Fitting. Data from the groups of four or eight fish were analyzed separately in this analysis. We assume an additive model that gives the response, α_t, of the focal fish, as a sum of functions of the past responses, the position relative to the nearest wall and the positions of each of the nearest three neighbors,

$$\alpha_t = F_{AR}(\alpha_{t-1},...,\alpha_{t-5}) + F_W(d,\gamma) + \sum_{i=1}^{3} F_{Ni}(r_i,\theta_i) + \epsilon, \quad \textbf{[1]}$$

where F_{AR} is a auto-regressive linear sum of the past five responses, F_W is a nonlinear function of the position of the focal fish relative to the nearest wall, and F_{Ni} is a nonlinear function of the position of the ith nearest neighbor, relative to the focal fish. The stochastic variable ϵ represents Gaussian noise and allows for the intrinsic variability of the response.

Auto-regression was based on five previous time steps, equivalent to one third of a second. Fitting was performed by standard least-squares linear regression. The NETLAB (40) toolbox for MATLAB® was used to construct and optimize artificial neural networks for fitting general nonlinear functions. Each network was initialized with the appropriate input and output data, a hidden layer of 20 nodes, and a linear output function. The maximum-likelihood weights for the network were estimated via the back-propagation algorithm. An expectation-maximization (41) algorithm was used to iteratively fit each function (see SI Text).

Note. While preparing our manuscript we became aware of the excellent work of Katz et al., which addresses similar questions on a different fish species (42).

1. Vicsek T, Zafiris A (2010) Collective motion. arXiv:1010.5017.
2. Sumpter DJT (2010) *Collective Animal Behavior* (Princeton Univ Press, Princeton).
3. Cavagna A, et al. (2008) The starflag handbook on collective animal behaviour: 1. Empirical methods. *Anim Behav* 76:217–236.
4. Giardina I (2008) Collective behavior in animal groups: theoretical models and empirical studies. *HFSP J* 2:205–219.
5. Bode NWF, Franks DW, Wood AJ (2011) Limited interactions in flocks: Relating model simulations to empirical data. *J R Soc Interface* 8:301–304.

6. Vicsek T (1999) Application of statistical mechanics to collective motion in biology. *Physica A* 274:182–189.
7. Couzin ID, Krause J, James R, Ruxton GD, Franks NR (2002) Collective memory and spatial sorting in animal groups. *J Theor Biol* 218:1–11.
8. Huth A, Wissel C (1992) The simulation of the movement of fish schools. *J Theor Biol* 156:365–385.
9. Ginelli F, Chate H (2010) Relevance of metric-free interactions in flocking phenomena. *Phys Rev Lett* 105 168103.
10. Ballerini M, et al. (2008) Interaction ruling animal collective behavior depends on topological rather than metric distance: Evidence from a field study. *Proc Natl Acad Sci USA* 105:1232–1237.
11. Lukeman R, Li YX, Edelstein-Keshet L (2010) Inferring individual rules from collective behavior. *Proc Natl Acad Sci USA* 107:12576–12580.
12. Biro D, Sumpter DJT, Meade J, Guilford T (2006) From compromise to leadership in pigeon homing. *Curr Biol* 16:2123–2128.
13. Sumpter DJT, Buhl J, Biro D, Couzin ID (2008) Information transfer in moving animal groups. *Theory Biosci* 127:177–186.
14. Freeman R, Mann R, Guilford T, Biro D (2011) Group decisions and individual differences: Route fidelity predicts flight leadership in homing pigeons (Columba livia). *Biol Lett* 7:63–66.
15. Nagy M, Akos Z, Biro D, Vicsek T (2010) Hierarchical group dynamics in pigeon flocks. *Nature* 464:890–893.
16. Partridge BL, Pitcher T (1980) The sensory basis of fish schools: Relative roles of lateral line and vision. *J Comp Physiol A* 135:315–325.
17. Eriksson A, Jacobi MN, Nystrom J, Tunstrom K (2010) Determining interaction rules in animal swarms. *Behav Ecol* 21:1106–1111.
18. Vapnik VN (1998) *Statistical Learning Theory* (Wiley-Interscience, New York).
19. Mackay DJC (2003) *Information theory, inference and learning algorithms* (Cambridge Univ Press, Cambridge, UK).
20. Bishop C (2006) *Pattern Recognition and Machine Learning* (Springer, New York).
21. Bishop C (1995) *Neural Networks for Pattern Recognition* (Oxford Univ Press, Oxford).
22. Mann RP (2011) Bayesian inference for identifying interaction rules in moving animal groups. *PLoS ONE* 6:e22827.
23. Strombom D (2011) Collective motion from local attraction. *J Theor Biol*, 10.1016/j.jtbi .2011.05.019.
24. Reynolds CW (1987) Flocks, herds and schools: A distributed behavioral model. *Comput Graph* 21:25–34.
25. Okubo A (1986) Dynamical aspects of animal grouping: Swarms, schools, flocks, and herds. *Adv Biophys* 22:1–94.
26. Vicsek T, Czirók A, Ben-Jacob E, Cohen I, Shochet O (1995) Novel type of phase transition in a system of self-driven particles. *Phys Rev Lett* 75:1226–1229.
27. Romey W (1996) Individual differences make a difference in the trajectories of simulated schools of fish. *Ecol Modell* 92:65–77.
28. Radakov DV (1973) *Schooling in the ecology of fish* (Wiley, New York).
29. Cavagna A, et al. (2010) Scale-free correlations in starling flocks. *Proc Natl Acad Sci USA* 107:11865–11870.
30. Bode NWF, Faria JJ, Franks DW, Krause J, Wood AJ (2010) How perceived threat increases synchronization in collectively moving animal groups. *Proc R Soc Lond B Biol Sci* 277:3065–3070.
31. Hemelrijk C, Hildenbrandt H (2008) Self-organized shape and frontal density of fish schools. *Ethology* 114:245–254.
32. Ward AJW, Herbert-Read JE, Sumpter DJT, Krause J (2011) Fast and accurate decisions through collective vigilance in fish shoals. *Proc Natl Acad Sci USA* 108:2312–2315.
33. Huth A, Wissel C (1992) The simulation of the movement of fish schools. *J Theor Biol* 156:365–385.
34. Dukas R (1998) Cognitive Ecology. *The Evolutionary Ecology of Information Processing and Decision Making*, ed R Dukas (Univ of Chicago Press, Chicago), pp 89–127.
35. Moussaïd M, Helbing D, Theraulaz G (2011) How simple rules determine pedestrian behavior and crowd disasters. *Proc Natl Acad Sci USA* 108:6884–6888.
36. Lemasson B, Anderson J, Goodwin R (2009) Collective motion in animal groups from a neurobiological perspective: The adaptive benefits of dynamic sensory loads and selective attention. *J Theor Biol* 261:501–510.
37. Krause J, Ruxton GD (2002) *Living in groups* (Oxford Univ Press, Oxford).
38. Wood AJ, Ackland GJ (2007) Evolving the selfish herd: Emergence of distinct aggregating strategies in an individual-based model. *Proc R Soc Lond B Biol Sci* 274:1637–1642.
39. Branson K, Robie AA, Bender J, Perona P, Dickinson MH (2009) High-throughput ethomics in large groups of Drosophila. *Nat Methods* 6:451–457.
40. Nabney IT (2002) *NETLAB: Algorithms for Pattern Recognition* (Springer, New York).
41. Neal RM, Hinton GE (1998) *Learning in graphical models*, ed M Jordan (Kluwer Academic Publishers, Dordrecht, the Netherlands), pp 355–367.
42. Katz Y, Tunstrom K, Ioannou CC, Huepe C, Couzin ID (2011) Inferring the structure and dynamics of interactions in schooling fish. *Proc Natl Acad Sci USA*, 10.1073/pnas.1107583108.

ECOLOGY

APPLIED PHYSICAL SCIENCES

Supporting Information

Herbert-Read et al. 10.1073/pnas.1109355108

SI Text

Iterative Expectation-Maximization Algorithm for Function Fitting.
The first three components of the additive model specified by
Eq. 1 (auto-regression, the wall effects, and the first neighbor)
were optimized together using expectation-maximization. This
involves iteratively fitting one function, assuming the current
values of the other two are correct. Let F_{AR}^j be the jth iteration
of the auto-regression function, and similarly for F_W and F_{N1},
then, with the original response data, α_t:

- Initialize each function to a random state, the zeroth iteration.
- Fit F_{AR}^1 to the adjusted dataset α_t-F_W^0-F_{N1}^0.
- Fit F_W^1 to the adjusted dataset α_t-F_W^0-F_{AR}^1.
- Fit F_{N1}^1 to the adjusted dataset α_t-F_{AR}^1-F_W^1.
- Repeat the cycle to generate iteration 2. Continue iterations
  until the overall fit at the end of the iteration no longer im-
  proves beyond a fixed threshold.

After the first three components have been fitted the functions
associated with the positions of the second and third neighbors
are then fitted sequentially to the remaining residual.

- Fit F_{N2} to the adjusted dataset α_t-F_{AR}-F_W-F_{N1}.
- Fit F_{N3} to the adjusted dataset α_t-F_{AR}-F_W-F_{N1}-F_{N2}.

This ordering prioritizes the first neighbor interaction and
represents the prior belief that the fish must interact with the
first neighbor if it interacts with the second and must interact with
the second if it interacts with the third.

Stability of Fitting Algorithm. Because the expectation-maximiza-
tion algorithm converges to a local maxima of the likelihood,
equivalent to a local minima of the square error, we run the algo-
rithm repeatedly from many random initial starting conditions.
Many of these local maxima correspond to solutions with very
weak interactions with no clear pattern. However, those closest
in likelihood and square error to the global maximum resemble
the pattern of interaction shown in this paper, particularly the
range of interaction and the weak or absent interaction with
the second and third neighbors. Randomly removing 10% of
the complete dataset before running the algorithm has negligible
effect on the optimal solution, which suggests that the results are
not due to a small subset of the data.

Additional Descriptive Statistics. Fig. S1 reports the histograms of
the distance d, position ϑ, and orientation φ at which the nearest
neighbor was found with respect to the focal fish. The fish stay
relatively close and well aligned to each other and form elongated
schools, with most neighbors being in front or behind the focal
fish. In spite of the constraints imposed by the experimental setup
(which means fish change direction often), the groups remained
well polarized, with polarization values (measured as in ref. 38 ,
equation 1) of 0.84 ± 0.26, 0.71 ± 0.26, and 0.63 ± 0.25 (mean$\pm$
SD) for groups of two, four, and eight fish, respectively.

Most of the time, the fish are aligned with the closest border
of the basin (Fig. S3A), though their turning response to the wall
(Fig. S3B) is weak and mostly limited to avoiding collisions. (No-
tice that avoiding collision with a neighbor does not necessarily
require a turning response—and indeed we did not find evidence
for a repulsion zone in the turning response—because when one
of the two fish slows down, the other can move away from the
collision zone; on the contrary, avoiding collision with static ob-
jects, such as a wall, always requires a turning response.) The fish
show an acceleration response to the walls, which consists in
speeding up when moving away from the closest wall and slowing
down when approaching the wall.

Correlation Analysis of the Effects of Multiple Neighbors. Fig. S4
shows the response in acceleration and turning angle of a focal
fish to its neighbors, analyzed sequentially per nearest neighbor.
It appears that the focal fish is responding to all of its neighbors as
its acceleration and turning responses show qualitative similari-
ties between neighbor profiles. As our function fitting shows,
however, only the nearest neighbor is necessary in predicting
the direction changes of the focal fish (Fig. 5 and Fig. S6).

Effects of Group Size. Fig. S5 shows the acceleration and turning
angle of a focal fish as a function of the position of its neighbors.
As shown, the three group sizes produce qualitatively similar pat-
terns. As shown, standard error increases in the smaller group
sizes, probably due to less replication (fewer neighbors) in the
smaller group sizes. The amplitude, in both acceleration and
turning angle, decreases as group size increases. This is probably
due to an effect of averaging multiple interactions in the larger
group sizes.
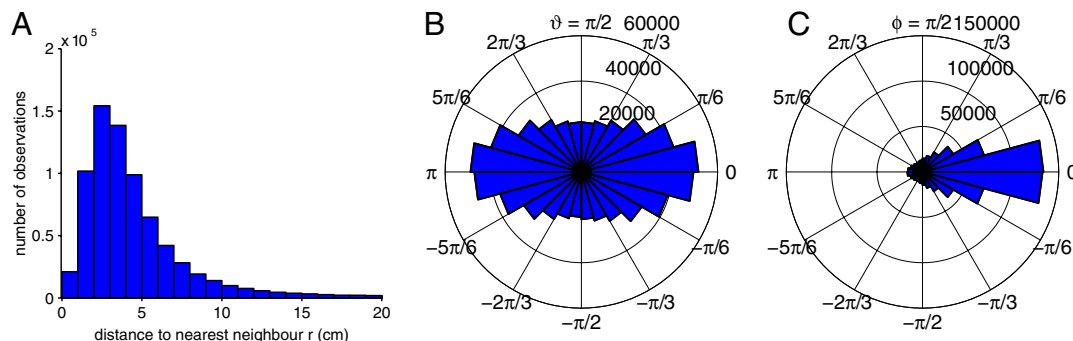
Fig. S1. Descriptive statistics of fish position in relation to other fish. (*A*) Distribution of distances *r* to the nearest neighbor. (*B*) Distribution of positions of the
nearest neighbor. (*C*) Distribution of relative orientations of the nearest neighbor.

Fig. S2. Average turning angle as a function of the orientation of the neighbor φ. Open symbols: The neighbor is on the left of the focal fish ($\vartheta \geq 0$). Full symbols: The neighbor is on the right of the focal fish ($\vartheta < 0$).

Fig. S3. Descriptive statistics of fish position in relation to the wall. (*A*) Distance *d* and orientation γ of the focal fish with respect to the closest wall of the basin. The colors in the map match the number of observations of a fish in the range of d and γ given by the axes. The higher number of observations for values of γ around zero than around $\pm\pi$ indicates a tendency of fish to swim clockwise (i.e., with the wall of the basin on their left). (*B*) Average turning angle as a function of the distance to the wall. Negative (positive) x axis values indicate that the wall was on the right (left) side of the focal fish. (*C*) Average acceleration as a function of the orientation with respect to the closest wall. The analyses are limited to fish far from the corner of the basin (>15 cm). Error bars represent standard errors.

Fig. S4. Average acceleration (top row) and turning angle (bottom plots) of the focal fish vs. distance (*r*) or position (ϑ) of the first, second, and third nearest neighbor (from left to right, respectively). Error bars represent standard errors. The analysis is limited to shoals with four fish. As the positions of all fish in the shoal are highly correlated, the fish appear to turn in the direction of each of the neighbors with similar attraction.

Fig. S5. Effect of different group sizes. Average acceleration (top row) and turning angle (bottom plots) of the focal fish vs. position (ϑ) of neighbors in groups of two, four, and eight fish (from left to right, respectively). Error bars represent standard errors.

Fig. S6. Effect of the wall, first, second, and third nearest neighbor on the acceleration (*A–D*) and turning angle (*E–H*) of the focal fish. The color scale indicates the size of the acceleration or turning response as a function of the relevant environmental or social variables. Each panel shows the predicted component of the acceleration or turning angle as a function of the variables indicating the positions of the wall and the nearest neighbors. The total predicted response, either acceleration or turning angle, is the addition of all four of these components. Each panel is plotted as a semicircular arc as we assume left-right symmetry in the acceleration response and antisymmetry in the turning. Only data from groups of eight fish are used for producing this figure. See *Materials and Methods* for details of the analysis.

Movie S1. A small portion of video from one of the trials with eight fish, showing the experimental arena and tracking of individuals.

Movie S1 (AVI)