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^j_{AR} 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^1_{AR} to the adjusted dataset α_t-F^0_W-F^0_{N1}.
- Fit F^1_W to the adjusted dataset α_t-F^0_W-F^1_{AR}.
- Fit F^1_{N1} to the adjusted dataset α_t-F^1_{AR}-F^1_W.
- Repeat the cycle to generate iteration 2. Continue iterations until the overall fit at the end of the iteration no longer improves 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-maximization algorithm converges to a local maxima of the likelihood, equivalent to a local minima of the square error, we run the algorithm 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± 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. S3*A*), though their turning response to the wall (Fig. S3*B*) is weak and mostly limited to avoiding collisions. (Notice 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 objects, 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 similarities 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 patterns. 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.

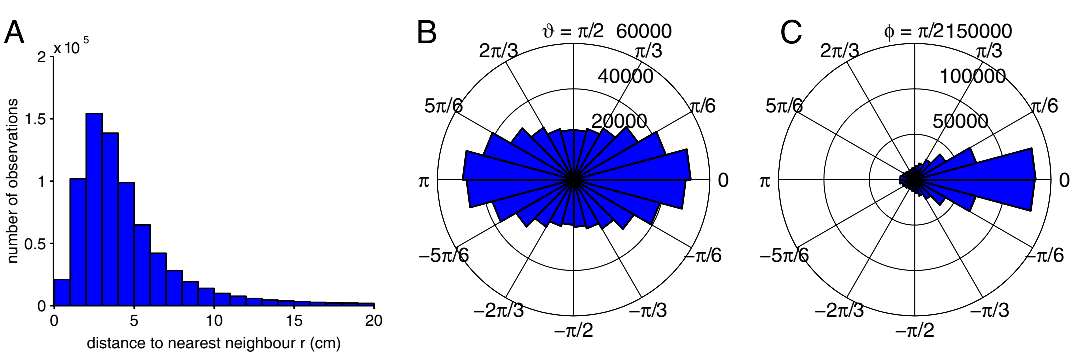


Fig. S1. Descriptive statistics of fish position in relation to other fish. (*A*) Distribution of distances *r* to the nearest neighbor. (*B*) Distribution of positions of the nearest neighbor. (*C*) Distribution of relative orientations of the nearest neighbor.

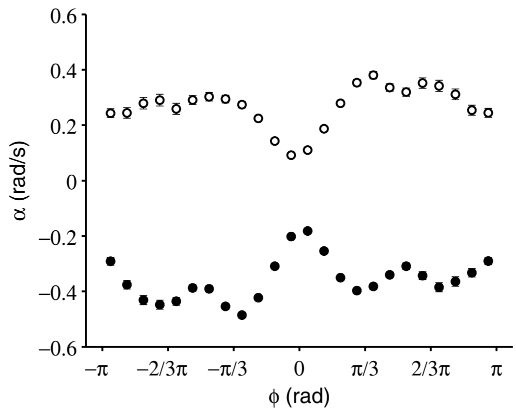


Fig. S2. Average turning angle as a function of the orientation of the neighbor φ. Open symbols: The neighbor is on the left of the focal fish ($\vartheta \geq 0$). Full symbols: The neighbor is on the right of the focal fish ($\vartheta < 0$).

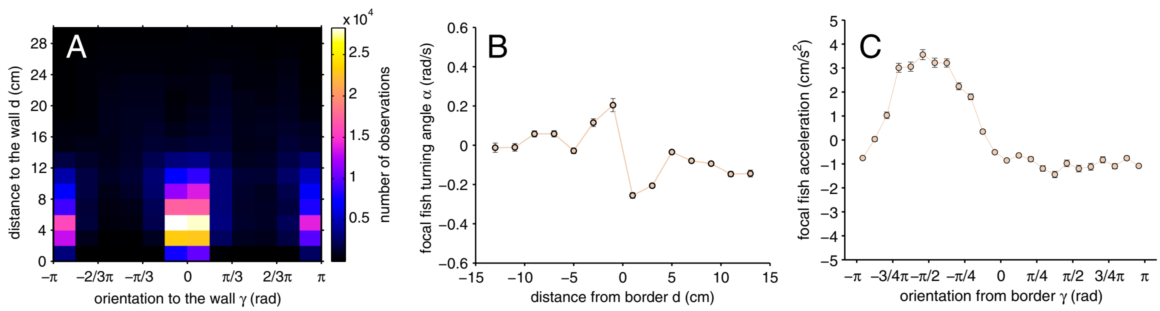


Fig. S3. Descriptive statistics of fish position in relation to the wall. (*A*) Distance *d* and orientation γ of the focal fish with respect to the closest wall of the basin. The colors in the map match the number of observations of a fish in the range of d and γ given by the axes. The higher number 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.

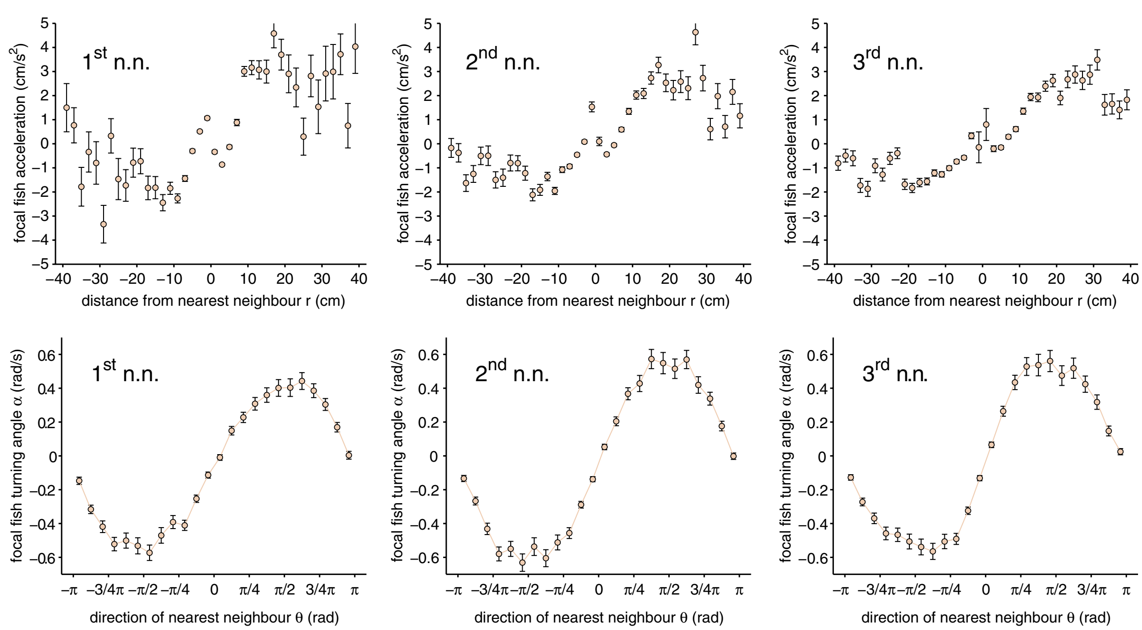


Fig. S4. Average acceleration (top row) and turning angle (bottom plots) of the focal fish vs. distance (*r*) or position (ϑ) of the first, second, and third nearest neighbor (from left to right, respectively). Error bars represent standard errors. The analysis is limited to shoals with four fish. As 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.

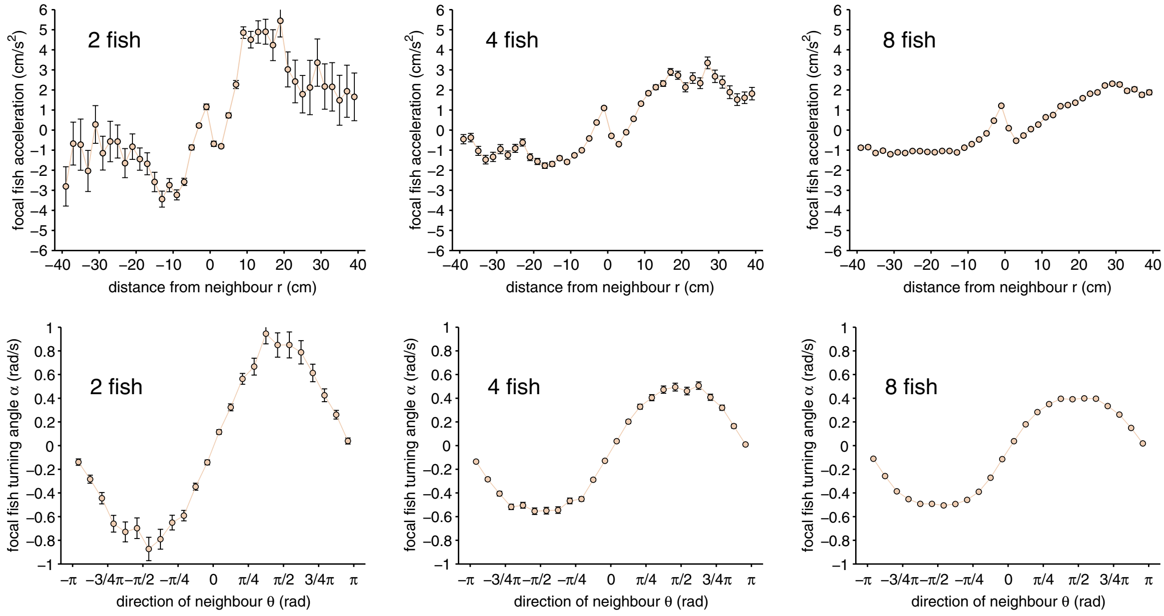


Fig. S5. Effect of different group sizes. Average acceleration (top row) and turning angle (bottom plots) of the focal fish vs. position (ϑ) of neighbors in groups of two, four, and eight fish (from left to right, respectively). Error bars represent standard errors.

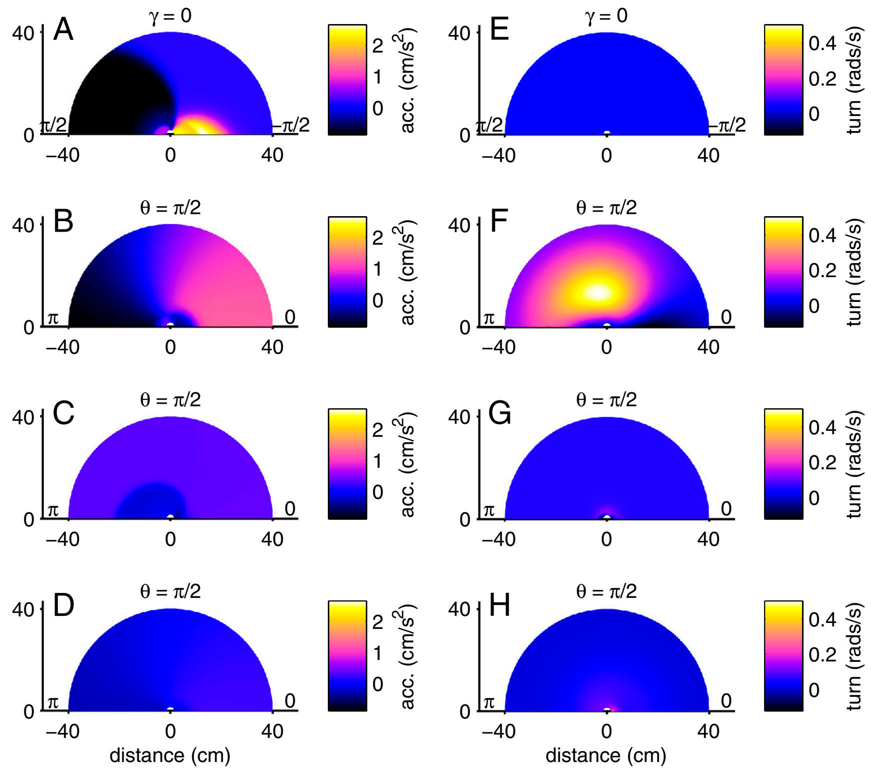
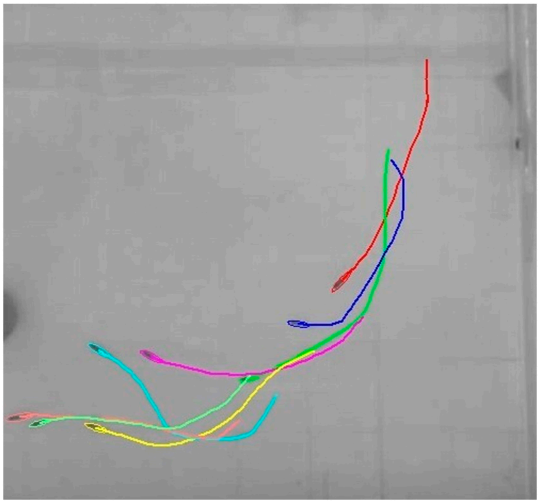


Fig. S6. Effect of the wall, first, second, and third nearest neighbor on the acceleration (*A–D*) and turning angle (*E–H*) of the focal fish. The color scale indicates the size of the acceleration or turning response as a function of the relevant environmental or social variables. Each panel shows 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)