Science Advances

AAAS

advances.sciencemag.org/cgi/content/full/3/6/e1603201/DC1

Supplementary Materials for

The effects of external cues on individual and collective behavior of shoaling fish

Timothy M. Schaerf, Peter W. Dillingham, Ashley J. W. Ward

Published 23 June 2017, *Sci. Adv.* **3**, e1603201 (2017)
DOI: 10.1126/sciadv.1603201

This PDF file includes:

- fig. S18. Mean change in speed over time as a function of relative groupmate positions for focal fish traveling at 0 to 20 mm/s (heat maps).
- fig. S19. Mean change in heading as a function of relative groupmate positions for focal fish traveling at 0 to 20 mm/s (heat maps).
- fig. S20. Observation frequency of groupmates and changes in velocity as a function of relative groupmate positions for focal fish traveling at 60 to 80 mm/s (projections).
- fig. S21. Observation frequency of groupmates as a function of relative groupmate positions for focal fish traveling at 60 to 80 mm/s (heat maps).
- fig. S22. Relative directions of motion of groupmates as a function of relative groupmate positions for focal fish traveling at 60 to 80 mm/s (arrow and heat maps).
- fig. S23. Mean change in speed over time as a function of relative groupmate positions for focal fish traveling at 60 to 80 mm/s (heat maps).
- fig. S24. Mean change in heading as a function of relative groupmate positions for focal fish traveling at 60 to 80 mm/s (heat maps).
- fig. S25. Observation frequency of groupmates and changes in velocity as a function of relative groupmate positions for focal fish traveling at 120 to 140 mm/s (projections).
- fig. S26. Observation frequency of groupmates as a function of relative groupmate positions for focal fish traveling at 120 to 140 mm/s (heat maps).
- fig. S27. Relative directions of motion of groupmates as a function of relative groupmate positions for focal fish traveling at 120 to 140 mm/s (arrow and heat maps).
- fig. S28. Mean change in speed over time as a function of relative groupmate positions for focal fish traveling at 120 to 140 mm/s (heat maps).
- fig. S29. Mean change in heading as a function of relative groupmate positions for focal fish traveling at 120 to 140 mm/s (heat maps).
- fig. S30. Relative frequency of partner observation and mean speed of focal fish for bins centered about the x or y axes.
- fig. S31. Changes in velocity of focal fish for bins centered about the x or y axes.
- fig. S32. Observation frequency of groupmates and changes in velocity as a function of relative groupmate positions for focal fish traveling at 0 to 20 mm/s for bins centered about the x or y axes.
- fig. S33. Observation frequency of groupmates and changes in velocity as a function of relative groupmate positions for focal fish traveling at 60 to 80 mm/s for bins centered about the x or y axes.
- fig. S34. Observation frequency of groupmates and changes in velocity as a function of relative groupmate positions for focal fish traveling at 120 to 140 mm/s for bins centered about the x or y axes.
- fig. S35. Entropy, mutual information, and entropy rate associated with changes in displacement.
- fig. S36. Entropy, mutual information, and entropy rate associated with changes in velocity.
- fig. S37. Distributions of estimated areas sighted by individual fish.

- fig. S38. Distributions of estimated area sighted by all group members.
- fig. S39. Distributions of estimated number of groupmates seen by each individual.
- fig. S40. Survival functions for durations between changes in sense of motion around the arena.
- fig. S41. Distributions of observed isolation events per minute.
- fig. S42. Survival functions for individual durations in isolation.
- fig. S43. Distributions of the number of swaps between group exterior and interior.
- fig. S44. Survival functions for durations spent on the group exterior.
- fig. S45. Distributions of the group fraction facing outward from the nearest wall.
- fig. S46. Survival functions for durations spent within one body length of the outer boundary.
- fig. S47. Survival functions for durations spent within two body lengths of the outer boundary.
- fig. S48. Survival functions for durations spent within one body length of the inner boundary.
- fig. S49. Survival functions for durations spent within two body lengths of the inner boundary.
- table S1. States associated with changes in displacement.
- table S2. States associated with changes in velocity.
- table S3. Bootstrap confidence intervals for median and SD change in speed over time.
- table S4. Bootstrap confidence intervals for median and SD distances from individual fish to the group centroid, nearest and mean neighbor distances, and group expanse and polarization (in direction of motion or facing direction).
- table S5. Bootstrap confidence intervals for median and SD distances to boundaries.
- table S6. Bootstrap confidence intervals for median and SD estimates for areas and number of groupmates sighted.
- table S7. Log-rank test comparison of survival functions for durations between individual changes in sense of motion around the arena.
- table S8. Bootstrap confidence intervals associated with time lag to maximum mean correlation in direction of motion.
- table S9. Bootstrap confidence intervals associated with the analysis of rapid turns.
- table S10. Bootstrap confidence intervals for median and SD number of isolation events per fish per minute.
- table. S11. Log-rank test comparison of survival functions for durations in isolation.
- table S12. Bootstrap confidence intervals for mean and SD number of interior-exterior group position swaps per minute.
- table S13. Log-tank test comparison of durations spent on the group exterior.
- table S14. Bootstrap confidence intervals for the mean and SD group fraction facing outward with their right eye.

section S1. Analysis

All analysis was performed using custom code developed in MATLAB (The Mathworks Inc., Natick, Massachusetts, USA).

S1.1 Labelling of time intervals, automatic tracking output, boundary identification and conversion of tracking output to standard units

As noted in the materials and methods, fish in each trial were filmed for a total of 15 minutes at 25 frames per second. We converted the films to .avi format using VirtualDub and then extracted six one minute sections from each film. We took three sections of film from before the introduction of the cues, beginning 2 minutes, 4 minutes and 6 minutes after the conclusion of the acclimation period. We also took three sections of film from after the introduction of the cues, beginning at 10 minutes, 12 minutes and 14 minutes following the conclusion of the acclimation period. It was guaranteed that fish would have been exposed to the external cues for the initial section of film from the post cue period (10 minutes after the conclusion of the acclimation period) as cues were introduced gradually from 7 to 9 minutes after the acclimation period, and our dye tests suggested that cues would spread throughout the entire annulus within this 2 minute time frame. For each section of film, we then used the Ctrax automated tracking software to extract the (x, y) trajectories of each fish (*39*). Tracking work was performed blind, with the person operating the automatic tracking software unaware of the hypothesis being tested. We labelled the sections of film (and corresponding derived data) for the period before the release of cues as A1, A2 and A3, and the sections of film for the period after the release of cues as P1, P2 and P3. When referring to data from A1–A3 combined or P1–P3 combined we refer to the *ante* (A) time interval or *post* (P) time interval respectively.

Preliminary analysis showed little difference between the behaviour of the fish across the three different ante-cue film sections, or between the behaviour of the fish across the three different post-cue film sections. Hence, for much of our analysis we pooled the data to provide a single measure for the ante-cue treatment and a single measure for the post-cue treatment.

Output from Ctrax that we used included each fish's (x, y) coordinates as a function of time (initially recorded in pixels), two parameters, e_x and e_y, that corresponded to half the length of semi-major and semi-minor axes of an ellipse that was fitted over the image of each fish for each time step (also initially recorded in pixels) and the facing direction of each fish, θ, relative to the positive x-axis (recorded in radians). We smoothed the (x, y) coordinates of each fish's trajectory using a Savitzky-Golay filter implemented through MATLAB's intrinsic *smooth* function with span 5 and degree 2.

We used a still image taken from the first frame tracked in each trial to identify the approximate location of the inner and outer boundaries of the tank (bounding the water) and the outer edge of the tank, see fig. S1 A. Each of the inner boundary, outer boundary

and outer edge of the tank were circular (or very close to circular). We developed a MATLAB script to manually identify three points on each boundary ((x_1, y_1), (x_2, y_2) and (x_3, y_3) (in pixels), input by mouse clicks on the still image of the first frame tracked), and subsequently determined the equation of the circle that passed through these points. The centre, (x_c, y_c), and radius, r_c, of the circle that passes through (x_1, y_1), (x_2, y_2) and (x_3, y_3) can be found by solving the following equations simultaneously

$$(x_1 - x_c)^2 + (y_1 - y_c)^2 = r_c^2 \quad \text{(S1)}$$
$$(x_2 - x_c)^2 + (y_2 - y_c)^2 = r_c^2 \quad \text{(S2)}$$
$$(x_3 - x_c)^2 + (y_3 - y_c)^2 = r_c^2 \quad \text{(S3)}$$

Provided that $x_1 \neq x_2 \neq x_3$ and $y_1 \neq y_2 \neq y_3$ (which we enforced manually by our choice of points), the solution to the above system of equations can be written as

$$y_c = \frac{(x_3(x_2^2 - x_1^2 + y_2^2 - y_1^2) + x_2(x_1^2 - x_3^2 + y_1^2 - y_3^2) + x_1(x_3^2 - x_2^2 + y_3^2 - y_2^2))}{2\left((y_2 - y_1)(x_3 - x_2) - (y_3 - y_2)(x_2 - x_1)\right)} \quad \text{(S4)}$$

$$x_c = \frac{x_3^2 - x_2^2 + y_3^2 - y_2^2 - 2y_c(y_3 - y_2)}{2(x_3 - x_2)} \quad \text{(S5)}$$

$$r_c = \sqrt{(x_1 - x_c)^2 + (y_1 - y_c)^2} \quad \text{(S6)}$$

We superimposed each fitted circle onto the image of the experimental arena, and if the fit seemed poor we re-selected points until there was a visually acceptable fit. fig. S1 B illustrates the circles fitted for one experimental trial. We used the diameter of the circle fitted to the outer edge of the tank, and the known diameter of the tank (730 mm) to determine the conversion ratio from pixels to millimetres ($= 2r_c/730$ (pixels/mm)) for each trial. The coordinates of all objects including the (x, y) coordinates of fish, the coordinates of the inner and outer boundaries, the coordinates of the outer edge of the tank and the length measurements associated with ellipses fitted over the images of the fish were then converted to millimetres by dividing by the pixels/mm ratio. Additionally, y-coordinates (in pixels) derived from Ctrax increase from the top of an image to the bottom, whereas y-coordinates are traditionally plotted in ascending value from the bottom of a plot to the top. To maintain consistency between what we saw in our videos and plots derived from our trajectory data we multiplied our y-coordinate converted to millimetres by -1. We also multiplied θ (the facing direction of fish in radians) by -1 for similar reasons. Neither of the changes in sign to y or θ had an effect on any of the analysis that we performed, but it made visualising our data more convenient.

S1.2 Basic measures of individual locomotion

We determined each fish's velocity, speed, acceleration, magnitude of acceleration, change in speed over time and turning speed (change in angle of motion over time) directly from the distance calibrated and smoothed tracking data using the following series of calculations.

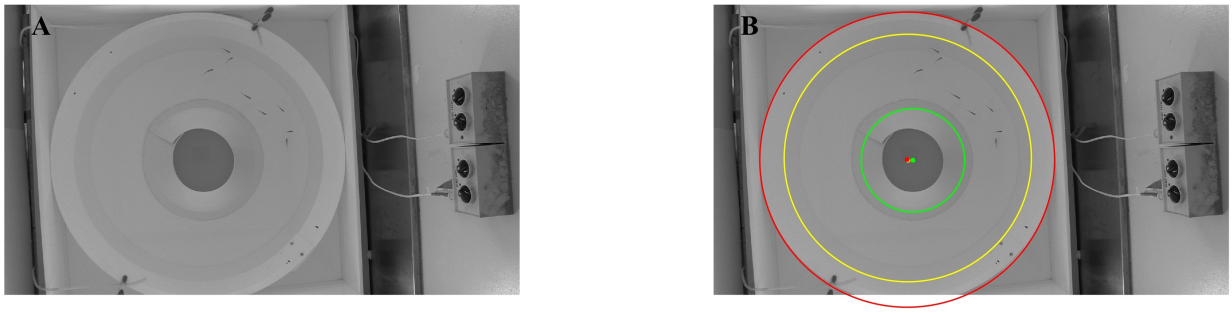

fig. S1. A – The annular circular arena where all experiments were performed (this is the first frame from the first trial, which was a control group). B – Circles fitted to the inner (green) and outer (yellow) boundaries of the water, and to the edge of the tank (red) for distance calibration purposes. The centres of each fitted circle are plotted as green, yellow and red dots respectively.

Writing $(x_i(t), y_i(t))$ as the coordinates of fish i at time t we determined the x and y components of a fish's velocity using the *standard* forward-difference approximations

$$u_i(t) = \frac{x_i(t+\Delta t) - x_i(t)}{\Delta t} \qquad \text{and} \qquad v_i(t) = \frac{y_i(t+\Delta t) - y_i(t)}{\Delta t} \tag{S7}$$

where $\Delta t = 1/25$ s was the constant duration between consecutive video frames. A fish's speed at time t was then approximated as

$$s_i(t) = \sqrt{(u_i(t))^2 + (v_i(t))^2} \tag{S8}$$

Following immediately from this calculation we determined the change in a fish's speed over time via

$$\frac{\Delta s_i}{\Delta t}(t) = \frac{s_i(t+\Delta t) - s_i(t)}{\Delta t} \tag{S9}$$

The above measure is referred to as tangential acceleration in (*4*), although the quantity differs from the component of an individual's acceleration vector in the direction of its velocity vector except in the limit of infinite frames per second ($\Delta t \to 0$). The measure in equation (S9) differs from both the acceleration of a fish (a vector), and the magnitude of acceleration. $\frac{\Delta s}{\Delta t}$ can take negative values (representing deceleration).
We determined the x and y components of a fish's acceleration respectively using the centred difference approximations

$$\dot{u}_i(t) = \frac{x_i(t+\Delta t) - 2x_i(t) + x_i(t-\Delta t)}{(\Delta t)^2} \qquad \text{and} \qquad \dot{v}_i(t) = \frac{y_i(t+\Delta t) - 2y_i(t) + y_i(t-\Delta t)}{(\Delta t)^2} \tag{S10}$$

and thus the magnitude of a fish's acceleration was determined by

$$a_i(t) = \sqrt{(\dot{u}_i(t))^2 + (\dot{v}_i(t))^2} \tag{S11}$$

We estimated a fish's turning speed at time t based on the direction of its velocity vector at times t and $t + \Delta t$. To do this we constructed unit vectors in the direction of each fish's velocity vector, with components

$$\hat{u}_i(t) = \frac{u_i(t)}{s_i(t)} \qquad \text{and} \qquad \hat{v}_i(t) = \frac{v_i(t)}{s_i(t)} \tag{S12}$$

The internal angle between the unit vectors for a given fish's direction of motion at at times t and $t + \Delta t$ was then determined using the dot product; we then divided this angle by the duration between consecutive frames to estimate turning speed. Compactly, the formula for calculating a fish's turning speed (in radians/s) can be written as

$$\alpha_i(t) = \frac{\cos^{-1}(\hat{u}_i(t)\hat{u}_i(t + \Delta t) + \hat{v}_i(t)\hat{v}_i(t + \Delta t))}{\Delta t} \tag{S13}$$

We pooled all observed individual speed, magnitude of acceleration, change in speed over time and turning speed values for each treatment divided into A and P time intervals, and then generated histograms to illustrate the distributions of each of these quantities. Results of analysis of the quantities described in this section appear in the main text, Table 1, and here in the supplementary information in section S2.1, table S3 and figs. S3 to S6.

S1.3 Basic measures of group configuration

We determined the coordinates for the group centre, $(c_x(t), c_y(t))$, for each time step (given by the mean x- and y-coordinates of all group members for a given frame). We then determined the distance between each fish and the group centroid for all time steps

$$d_{i,c}(t) = \sqrt{(x_i(t) - c_x(t))^2 + (y_i(t) - c_y(t))^2} \tag{S14}$$

For each time step we determined the distance between each fish i and all their neighbours $j \neq i$ (individually), given by

$$d_{i,j}(t) = \sqrt{(x_j(t) - x_i(t))^2 + (y_j(t) - y_i(t))^2} \tag{S15}$$

We thus determined the distance from each fish i to their nearest neighbour, $d_{i,nn}(t)$, and the mean distance from fish i to all their neighbours, $d_{i,mn}(t)$ for each time t.
We measured the spread of the group about the group centroid as a function of time using the expanse. Expanse is the square root of the mean distance from each fish to the group centroid squared

$$E(t) = \sqrt{\frac{\sum_{i=1}^{n}(d_{i,c}(t))^2}{n}} \tag{S16}$$

The idea behind calculating the mean of the distance squared is to give greater weighting in the average to distances further away from the centroid. The square root then returns

the units of the measure to distance (rather than distance squared). Both (*14*) and (*37*) make use of expanse, but the formulae presented are inconsistent with each other. As far as we can tell the measure reported by (*37*) does not make use of the square of distances between individual fish and the centroid, but rather examines the square of the difference in the distance from the origin to each fish and the distance from the origin to the group centroid – as reported we do not think that is a particularly informative measure. The notation used by (*14*) is ambiguous and the formula that they report may not measure expanse in the way that they intended either.

We examined the polarisation of both the direction of motion of group members and the facing direction of group members separately.

Polarisation in the directions of motion of group members at time t was calculated via

$$R_{\text{motion}}(t) = \frac{\sqrt{U(t)^2 + V(t)^2}}{n} \tag{S17}$$

where $U(t) = \sum_{i=1}^{n} \hat{u}_i(t)$, $V(t) = \sum_{i=1}^{n} \hat{v}(t)$ and n was the number of fish tracked in each group.

Polarisation in the facing directions of group members at time t was determined with

$$R_{\text{facing}}(t) = \frac{\sqrt{C(t)^2 + S(t)^2}}{n} \tag{S18}$$

where $C(t) = \sum_{i=1}^{n} \cos(\theta_i(t))$, $S(t) = \sum_{i=1}^{n} \sin(\theta_i(t))$, $\theta_i(t)$ was the facing direction of fish i at time t relative to the positive x-axis and n was the number of fish tracked in a given group.

We pooled all observed values of the distance of each fish from the group centroid, $d_{i,c}(t)$, distance from each fish to their nearest neighbour, $d_{i,nn}(t)$, the mean distance from each fish to all their neighbours, $d_{i,mn}(t)$, expanse, $E(t)$, polarisation in direction of motion, $R_{\text{motion}}(t)$ and polarisation in facing direction, $R_{\text{facing}}(t)$, for all the A and P time intervals observed in each treatment. As with the basic measures of locomotion, we then generated histograms to examine the distribution of these quantities.

Results of analysis of the quantities described in this section appear in the main text, Table 2, and here in the supplementary information in section S2.2, table S4 and figs. S7 to S12.

S1.4 Basic interactions with arena boundaries

We determined the points on each of the inner and outer boundaries of the tank closest to the location of each fish, and in the process also determined the shortest distance from each fish to either boundary. In general, the shortest distance between a boundary and a point is the length of the straight line segment from the point to the boundary such that the straight line segment is perpendicular to the boundary. (See for example (*40*), where we determined perpendicular distances from straight boundary edges in an arena with five edges.) A straight line that intersects a circle at right angles must coincide with a radii

of the circle, and therefore must also pass through the centre of the circle. We used this fact to construct a straight line that both passed through a fish's position, $(x_i(t), y_i(t))$, and that was perpendicular to (and passed through) either the inner or outer boundary. Writing the centre of the circular boundary of interest (inner or outer) as (x_c, y_c), and the radius of the circle as r_c, the straight line that passed through both $(x_i(t), y_i(t))$ and (x_c, y_c) was

$$\begin{aligned} x &= x_c \quad \text{if } x_i(t) = x_c \\ y &= mx + b \quad \text{otherwise} \end{aligned} \tag{S19}$$

where $m = (y_i(t) - y_c)/(x_i(t) - x_c)$ and $b = y_c - mx_c$. We then determined the two points of intersection of the straight line described in equation (S19) with the circular boundary given by

$$(x - x_c)^2 + (y - y_c)^2 = r_c^2 \tag{S20}$$

by solving equations (S19) and (S20) simultaneously. (Since (S19) coincided with a radii of the circle, it was guaranteed that there would be two solutions.) If $x_i(t) = x_c$, then the two points of intersection were $(X_1, Y_1) = (x_c, y_c + r_c)$ and $(X_2, Y_2) = (x_c, y_c - r_c)$. Otherwise, the two points of intersection could be written as

$$X_{1,2} = \frac{-k_2 \pm \sqrt{k_2^2 - 4k_1k_3}}{2k_1} \tag{S21}$$

$$Y_{1,2} = mX_{1,2} + b \tag{S22}$$

where

$$k_1 = m^2 + 1 \tag{S23}$$

$$k_2 = 2bm - 2x_c - 2my_c \tag{S24}$$

$$k_3 = b^2 + x_c^2 + y_c^2 - 2by_c - r_c^2 \tag{S25}$$

We then determined the distances from the location of the fish, $(x_i(t), y_i(t))$, to both (X_1, Y_1) and (X_2, Y_2). We identified the least of the two distances and stored that distance as the minimum distance to the given boundary for fish i at time t, denoted $d_{i,b}(t)$.
We pooled all observed values of the distances from the fish to the closest points on the inner and outer boundaries, $d_{i,b_{\text{inner}}}(t)$ and $d_{i,b_{\text{outer}}}(t)$ for all the A and P time intervals observed in each treatment. We then generated histograms to examine the distribution of these quantities.
Results of analysis of the quantities described in this section appear in section S2.3, table S5 and figs. S13 and S14.

S1.5 Speed, relative alignment and rules of interaction as a function of relative partner displacement and focal fish speed

We determined the relative frequency that each fish observed their group mates at given relative coordinates, as well as the mean speed, mean relative alignment in direction of

motion, mean change in speed over time and mean change in angle of motion over time of fish as functions of the relative locations of their group mates. We rendered these functions as heat maps in two dimensions, as well as examining each quantity as a function of the relative x or y-coordinates of partner fish in isolation. Further, we examined relative frequency of observations of partner fish, mean alignment in directions of motion and changes in speed and direction of motion as a function of both the relative locations of all group mates and speed of focal fish.

The first step in making each heat map was to determine the distance between each pair of fish for all times t, as given by equation (S15). Next we calculated the angle between the direction of motion of each *focal* fish, i, (given in component form by equation (S12)) and the directed straight line segment from the location of fish i to the location of all other group mates j in turn, termed *partner* fish here in the context of the pairwise calculations required, for all time steps t. The calculations that follow in this description were repeated treating all fish i in turn as a focal fish, and all other group mates as partners (for $j \neq i$). To aid in the calculation of the angle of motion of each focal fish and the line segment to each of their group mates/partners, we constructed a unit vector in the direction of the straight line segment from fish i to fish j, with components

$$\hat{x}_{ij}(t) = \frac{x_j(t) - x_i(t)}{d_{i,j}(t)} \qquad \text{and} \qquad \hat{y}_{ij}(t) = \frac{y_j(t) - y_i(t)}{d_{i,j}(t)} \tag{S26}$$

The internal angle between the unit vectors representing the direction of motion of fish i (equation (S12)) and the direction from fish i to fish j (equation (S26)) can be determined using a dot product (similar to equation (S13))

$$\phi_{ij}(t) = \frac{180}{\pi} \cos^{-1}(\hat{u}_i(t)\hat{x}_{ij}(t) + \hat{v}_i(t)\hat{y}_{ij}(t)) \tag{S27}$$

Using equation (S27) will determine an angle constrained so that $0 \leq \phi_{ij} \leq 180°$. An additional calculation is required to determine if fish j is either to the left or the right of fish i. Relative to the direction of motion of fish i, fish j lies to the left (right) of fish i if the sign of the following equation is positive (negative)

$$\lambda_{ij}(t) = \operatorname{sgn}\left(\hat{u}_i(t)\hat{y}_{ij}(t) - \hat{v}_i(t)\hat{x}_{ij}(t)\right) \tag{S28}$$

The term in the parentheses of equation (S28) is the vertical component of the cross-product of the unit vector pointing in the direction of motion of fish i with the unit vector pointing from fish i to fish j. We defined the signed angle between the direction of motion of fish i and the relative location of fish j as

$$\varphi_{ij}(t) = \begin{cases} \lambda_{ij}(t)\phi_{ij}(t) & \text{if } \lambda_{ij}(t) \neq 0 \\ \phi_{ij}(t) & \text{if } \lambda_{ij}(t) = 0 \end{cases}$$

Each heat map was constructed in Cartesian coordinates (x, y), where $-120 \leq x \leq 120$ (mm) and $-120 < y \leq 120$; a spatial extent of approximately four body lengths in both

the x- and y-directions. Focal fish (i) were treated as being located at the origin, moving to the right (parallel to the x-axis). A separate map was produced for the A and P time interval for each treatment and each quantity of interest. (Here we discuss calculations relating to speed by means of example, but the method is identical for other quantities.) We converted the relative locations of all partner fish from the polar form described by $(d_{i,j}(t), \varphi_{ij}(t))$ to Cartesian coordinates via

$$x_{ij,\text{relative}}(t) = d_{i,j}(t)\cos\left(\varphi_{ij}(t)\right) \tag{S29}$$

$$y_{ij,\text{relative}}(t) = d_{i,j}(t)\sin\left(\varphi_{ij}(t)\right) \tag{S30}$$

We divided the domain centred on each focal fish into a set of overlapping bins such that the left edges of the bins were located at $x_{l,\text{left}} = -120, -116, -112, -108, \ldots, 104$ (mm), the right edges of the bins were located at $x_{l,\text{right}} = -104, -100, -96, -92, \ldots, 120$ (mm), the bottom edges of the bins were located at $y_{k,\text{bottom}} = -120, -116, -112, -108, \ldots, 104$ (mm) and the top edges of the bins were located at $y_{k,\text{top}} = -104, -100, -96, -92, \ldots, 120$ (mm). That is, bins extend 16 mm in both the x and y directions (approximately half a body length), and were separated by 4 mm in both x and y directions. The biological reason behind using such smoothing is that it is reasonable to assume that small changes in the relative position of partner fish should not result in dramatically different behaviour of focal fish (on average).

For each fish i in a given set, and each time-step, fish i's speed at time t was included in bin (l, k) if $x_{l,\text{left}} < x_{ij,\text{relative}}(t) \leq x_{l,\text{right}}$ and $y_{k,\text{bottom}} < y_{ij,\text{relative}}(t) \leq y_{k,\text{top}}$ (for all partner fish j at time t). Once data corresponding to all fish and time steps were allocated to bins, we calculated the mean of the finite entries in each bin, and rendered the results with the help of MATLAB's intrinsic *surf* function. In the case where alignment was the quantity of interest, we determined the mean angle between the facing direction of the focal fish and their partners using standard methods of circular statistics (*41*) (plotted as arrows in the relevant plots), along with R, which is a measure of the scatter of all the angles in a set. For reference, the mean, $\bar{\vartheta}$, of a set of angles, ϑ_i, is given by

$$\bar{\vartheta} = \tan^{-1}\left(\frac{Y}{X}\right) \tag{S31}$$

where $X = \sum_{i=1}^{n}\cos\vartheta_i$, $Y = \sum_{i=1}^{n}\sin\vartheta_i$, and

$$R = \frac{\sqrt{X^2 + Y^2}}{n} \tag{S32}$$

In practice we evaluated equation (S31) using MATLAB's intrinsic *atan2* function such that $\bar{\vartheta}$ was located in the correct quadrant and $-\pi < \bar{\vartheta} \leq \pi$. In surface plots of alignment, colours corresponded to the R value in each bin, rather than a mean.

In addition to the magnitude of turning speed given by equation (S13), we required information about the sense of rotation of fish (clockwise or anti-clockwise) to construct appropriate plots of turning behaviour. This sense of rotation was determined by examining

the vertical component of the cross product of unit velocity vectors for a fish at times t and $t+\Delta t$, similar to the calculations for $\lambda_{ij}(t)$ in equation (S28). We refer to the quantity that combines sense of rotation and magnitude of turning speed as change in angle of motion over time, or change in angle over time, denoted $\frac{\Delta\theta}{\Delta t}$.

In addition to surface plots, we produced line-graphs of the proportion of observations of neighbour fish, mean speed of focal fish, mean change in speed over time of focal fish and mean change in angle of motion over time of focal fish by projecting data contained in the square bins (described above) onto both the x and y-axes. Data was projected onto the x-axis by combining all data that satisfied $x_{l,\text{left}} < x_{ij,\text{relative}}(t) \leq x_{l,\text{right}}$ into bin l, irrespective of the y-coordinate associated with each data point. Similarly, data was projected onto the y-axis by combining all elements that satisfied $y_{k,\text{bottom}} < y_{ij,\text{relative}}(t) \leq y_{k,\text{top}}$ into bin k. As well as calculating means for the line-graphs, we determined the standard deviation of values contained in each bin, and hence standard errors (based on a sample size equal to the number of elements contained in a given bin). Further, we examined each quantity (relative frequency of neighbour observations, speed, change in speed over time and change in angle of motion over time) only across bins centred along the lines $x = 0$ and $y = 0$; the relevant bins satisfied either $-8 < x \leq 8$ or $-8 < y \leq 8$ mm.

Finally, we repeated the above procedure to produce both heat and line graphs with an additional dimension added to take into account the speed of each focal fish i. In effect, we produced heat and line graphs of the above form using subsets of data where the speed of focal fish lay between $s_{i,\text{low}} = 0, 10, 20, \ldots, 130$ (mm/s) and $s_{i,\text{high}} = 20, 30, 40, \ldots, 150$ (mm/s).

The use of the type of calculations described in this section to examine anything other than the distribution of neighbours about focal individuals is relatively new to studies of collective motion (*1, 4, 5*). There are both pros and cons to the approach we adopted for this study that are connected to binning data for every neighbour (as opposed to a single neighbour) for each time step of our data set. The main advantages of our approach include that it allows us to make use of our full data set and to summarise our results with single plots for each treatment group and time interval. To the best of our knowledge, (*1*) used the same approach to binning data that we used in this study (based on the description of the calculations perfomed in (*1*) to produce plots comparing the alignment of interior members in groups of surf scoters), and we applied the same approach to binning data for groups of 2, 4 or 8 eastern mosquitofish in (*4*), albeit using a polar coordinate based grid, rather than the Cartesian coordinate based grid used here. The main drawbacks of our approach include that given that data is binned based on the relative positions of 7 other group mates (as opposed to a single group mate) there is some ambiguity in interpreting the response of focal fish to the positions of their neighbours (our plots essentially show the result of combined interactions across all partners/neighbours), and that our plots may not accurately predict changes in velocity (or speed and relative alignment) for groups of sizes other than 8 (the shoal size in this study). (*5*) began work to try to address these sorts of problems by examining velocity changes in pairs and trios of golden shiners. In particular for groups of 3 (*5*) examined velocity changes of focal individuals as a function of the relative coordinates of both their partners in two dimensions (that is, velocity changes

of focal individuals as a function of four variables); subsequent analysis then suggested that the interactions in a group of three were not an average of pairwise interactions (determined from analysis of groups of 2). If we were to adopt a similar approach here, then our initial calculations would involve estimating speed, relative alignment or velocity changes of focal individuals as a function of 14 variables (2 for the relative locations of each partner) – there are unresolved issues associated with such calculations, including how to associate each partner with a set of relative coordinates (for example, is it more appropriate to associate the nth nearest neighbour with the nth pair of coordinates, or is a better approach to associate a particular partner with the nth pair of coordinates irrespective of the topological relationship between focal individual and partner). Given that the work of (*5*) suggests that simple averaging of pairwise interactions does not predict the changes in velocity seen in groups of 3, a suitable method for visualising the results of any interactions is then also complicated. Another viable alternative could be to generate different plots based on the number of partners in front of, or behind, each focal fish (which would result in a sequence of 8 plots for each of the cases of 0 to 7 group mates in front of a focal fish). At the heart of all of this is the important question of how best to measure and visualise the rules real animals use to adjust their velocity based on the relative locations and behaviour (such as velocity, acceleration) of their group mates from tracking data with a view to then using these measurements to make predictions about collective motion in other contexts. This is a vitally important problem to address in collective behaviour, but is also beyond the scope of our study of the effect of food and alarm cues here, and should be examined as part of a dedicated study.
Results of the above analysis appear in the main text, Figs. 1 to 7, and here in the supplementary information in section S2.4 and figs. S15 to S34.

S1.6 Predictability of changes in displacement and velocity

We used measures from information theory to examine the predictability of elements of motion of fish subject to each treatment. In particular we determined the conditional entropy, mutual information and entropy rate associated with changes in displacement and velocity over short time intervals. These calculations were made under the assumption that such changes in displacement and velocity could be treated as part of a Markov process, where the next state depends only on the current state of the system.
We first re-formulated observed changes in displacement and velocity as Markov chains for each individual, i, using the method outlined in (*9*). We defined 9 different states based on differences in displacement between equally spaced sample times. Writing differences in displacement between frames at times t and $t + L\Delta t$ as

$$dx_i(t) = x_i(t + L\Delta t) - x_i(t) \qquad \text{and} \qquad dy_i(t) = y_i(t + L\Delta t) - y_i(t) \tag{S33}$$

where L was an arbitrary relatively small integer chosen to try to reduce possible effects of noise, we defined a fish's state $D_i(t)$ at time t dependent on the sign and magnitude of $dx_i(t)$ and $dy_i(t)$ relative to a threshold, K, as detailed in table S1.

table S1. States associated with changes in displacement.

$dx_i(t)$	$dy_i(t)$	State, $D_i(t)$
$\leq -K$	$\geq K$	1
$-K < dx_i(t) < K$	$\geq K$	2
$\geq K$	$\geq K$	3
$\leq -K$	$-K < dy_i(t) < K$	4
$-K < dx_i(t) < K$	$-K < dy_i(t) < K$	5
$\geq K$	$-K < dy_i(t) < K$	6
$\leq -K$	$\leq -K$	7
$-K < dx_i(t) < K$	$\leq -K$	8
$\geq K$	$\leq -K$	9

Similarly, we defined a fish's state dependent on changes in velocity, $U_i(t)$, as detailed in table S2, where $du_i(t)$ and $dv_i(t)$ were given by

$$du_i(t) = \frac{dx_i(t + L\Delta t) - dx_i(t)}{L\Delta t} \qquad \text{and} \qquad dv_i(t) = \frac{dy_i(t + L\Delta t) - dy_i(t)}{L\Delta t} \tag{S34}$$

table S2. States associated with changes in velocity.

$du_i(t)$	$dv_i(t)$	State, $U_i(t)$
$\leq -K$	$\geq K$	1
$-K < du_i(t) < K$	$\geq K$	2
$\geq K$	$\geq K$	3
$\leq -K$	$-K < dv_i(t) < K$	4
$-K < du_i(t) < K$	$-K < dv_i(t) < K$	5
$\geq K$	$-K < dv_i(t) < K$	6
$\leq -K$	$\leq -K$	7
$-K < du_i(t) < K$	$\leq -K$	8
$\geq K$	$\leq -K$	9

If the threshold parameter, K, was set to a relatively small value (close to zero) then the states in tables S1 and S2 would essentially describe if a fish made any change to their position or velocity over a short duration. If the threshold was set to a high value, then the states would describe if a fish made large jumps in position or velocity between sample frames. In (*9*) we set L to 10 frames (which corresponded to a separation in sample frames of approximately 0.67 seconds at 15 frames per second); we set $L = 20$ for our calculations here (a slightly larger gap of 0.80 seconds at 25 frames per second) and further examined the sensitivity of our results to small changes in L by performing additional calculations

for $L = 19$ and $L = 21$. We chose values of K to examine predictability of nominally small, medium and large changes in displacement and velocity for each treatment group. Small changes were associated with $K \in \{0.7, 0.8, 0.9\}$, medium changes were associated with $K \in \{7, 8, 9\}$ and large changes were associated with $K \in \{17, 18, 19\}$. Three values were chosen for each set to again examine sensitivity of our results (this time to small changes in K). With $K = 0.8$ and $L = 20$ the threshold for changes in displacement corresponded to fish moving at 1 mm/s in either or both the x and y directions, $K = 8$ and $L = 20$ corresponded to fish moving at 10 mm/s and $K = 18$, $L = 20$ corresponded to fish moving at 22.5 mm/s.

We determined the frequency that each fish was observed to transition between all 81 possible pairs of states in each time interval for both changes in displacement and changes in velocity and each pairing of parameters K and L. We then pooled these frequencies for fish from all trials in the same treatment over the larger A or P time intervals. Writing $F_{i,j}$ as the observed frequency of transitions from state i at one sample time to state j at the next sample time, we then estimated the conditional probability of moving from state i to state j as $P_{i,j} = P(j|i) = F_{i,j}/(\sum_{j=1}^{9} F_{i,j})$. Additionally, we estimated the probabilities that state i was the observed starting state at time t, given by $P_i = (\sum_{j=1}^{9} F_{i,j})/(\sum_{i=1}^{9}\sum_{j=1}^{9} F_{i,j})$, and that state j was the state transitioned to at time $t + L\Delta t$, given by $P_j = (\sum_{i=1}^{9} F_{i,j})/(\sum_{i=1}^{9}\sum_{j=1}^{9} F_{i,j})$.

We determined the conditional entropy of observed changes in displacement and velocity via

$$H(J|I) = -\frac{\sum_{i=1}^{9} P_i \left(\sum_{j=1}^{9} P_{i,j} \log_e P_{i,j}\right)}{\log_e(2)} \tag{S35}$$

measured in bits (see for example (*38*)). Entropy is a measure of uncertainty of a random variable.

We calculated the information about state J given by state I in bits via

$$\begin{aligned} \mathcal{I}(I;J) &= H(J) - H(J|I) && \text{(S36)} \\ &= -\frac{\sum_{j=1}^{9} P_j \log_e P_j}{\log_e(2)} + \frac{\sum_{i=1}^{9} P_i \left(\sum_{j=1}^{9} P_{i,j} \log_e P_{i,j}\right)}{\log_e(2)} && \text{(S37)} \end{aligned}$$

Information in the above equation, $\mathcal{I}$, is referred to as mutual information and is a measure of the reduction of uncertainty of J due to the knowledge of I.

Finally, we calculated the entropy rate associated with changes in displacement or velocity for each treatment, before and after application of cues via

$$\mathcal{H}(J) = -\frac{\sum_{i=1}^{9}\sum_{j=1}^{9} \pi_i P_{i,j} \log_e P_{i,j}}{\log_e(2)} \tag{S38}$$

where the π_i terms are the elements of the stationary distribution of the time-homogeneous Markov chain described by the $P_{i,j}$ terms. The elements of the stationary distribution

satisfy the conditions $\pi_j = \sum_{i=1}^{9} \pi_i P_{i,j}$, $\sum_{j=1}^{9} \pi_j = 1$, and $0 \leq \pi_j \leq 1$. In practice we determined the stationary distribution by solving the matrix equation

$$\begin{bmatrix} 1 & 1 & \dots & 1 \\ P_{1,1}-1 & P_{2,1} & \dots & P_{9,1} \\ P_{2,1} & P_{2,2}-1 & \dots & P_{9,2} \\ \vdots & \vdots & \ddots & \vdots \\ P_{9,1} & P_{9,2} & \dots & P_{9,9}-1 \end{bmatrix} \begin{bmatrix} \pi_1 \\ \pi_2 \\ \vdots \\ \pi_9 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

using MATLAB's intrinsic *mldivide* function. Entropy rate is the rate at which entropy grows as a sequence of random variables (such as we treated changes in displacement or velocity) also grows. Entropy rate is also associated with the freedom of options in changing state, with higher entropy rates corresponding to greater freedom (see for example, Example 4.3.1 in (*38*) which relates to random walks on chess boards by pieces with different restrictions in movement).
Results of the above analysis appear in the main text, Fig. 8, and here in the supplementary information in section S2.5 and figs. S35 and S36.

S1.7 Approximate area under vigilance by individuals and groups, and number of group mates visible to individuals

We sought measures for what was visible to fish, both at the individual and group levels, subject to different treatments. The motivations for this analysis were the apparent changes in group configuration suggested by differences in neighbour distances, polarisation and expanse. In particular, polarisation, measures of neighbour distances and expanse diminished in groups subject to the alarm cue – the reduction in all these measures suggested that alarmed groups tended to be more closely grouped, but that the directions faced by group members were more different to each other than under other treatments. The greater dispersion in facing/movement directions observed in alarmed groups could result in a greater area covered by the eyes of the group (due to their vision of a greater portion of 360° compared to a more polarised group), but the more tightly packed group might result in obstructed vision for individuals. The effect of these competing factors on what could be seen by individuals and the group was not obvious *a priori*.
Inspired by the work of (*10, 11*) we used a modified ray-casting method to estimate the area visible to each fish. A ray-casting method is a geometric method for determining what objects can be seen when an individual is in a particular location; such methods are often used in video games to identify which objects in a two or three-dimensional environment a player should be able to see, and which objects are hidden/obstructed by other features of the environment. The basic idea of a ray-casting method is to draw a series of straight line segments (the rays) from the approximate location of the eyes of an individual and then identify the objects closest to the individual intersected by each ray. The surfaces of the closest objects intersected by each ray are then treated as being visible to the individual. The calculations as described in (*10, 11*) were extremely intensive, with 2000 rays cast per

eye per fish for 70 fish per trial in (*10*) and 1000 rays per eye per fish in (*11*) (implied by the angular separation between rays). The points of intersection with other fish and the boundaries of the arena would then have to be determined for every ray cast from both eyes of every fish for the entire time series (or relevant portion) of tracking data. Initially we planned on making use of an efficient ray-casting method (where instead of casting a fixed number of rays with equal angular separation, rays would be cast from their origin to the vertices of all other objects, with efficiency coming from a reduction in the overall number of rays used) to ultimately determine the area that individual fish could see (which would be described by a set of triangles). Once all individual areas sighted were determined for a given time step, we then planned on overlaying these areas on a single fine grid of square boxes to approximately determine the area seen by the entire group. However, a reasonably large number of points (over 300 points) were required to produce relatively accurate discrete representations of the inner and outer circular boundaries of the arena, vastly reducing the benefit of using the basic efficient ray-casting method. Further, we did not have the computational resources to perform calculations of the same scale performed as those in (*10, 11*), so instead we adopted a compromise that still made use of ray-casting ideas, but without the very fine scale accuracy achieved by the two previous studies. In part, we reduced the overall number of calculations that needed to be performed by transferring information about the location of fish, the rays from their eyes and the boundaries of the arena to a course grid comprised of 146 $\times$ 146 squares with dimensions 5 mm $\times$ 5 mm (prior to determining individual areas sighted, rather than after calculating the individual areas, as we originally planned). We then determined the first object intersected by each ray from each fish's eye on the grid. As with the analysis presented by (*10, 11*), we made no prior assumptions about the angular extents of the blind and visible zones of each fish. Here we estimated the blind angle behind each fish in each frame based on geometric arguments by approximating the bodies of the fish with the ellipses provided by Ctrax, and assuming that the fish could not see through their own bodies. We also assumed that the range at which the x-ray tetras could resolve objects was greater than the diameter of the arena (660 mm). Recent work by (*42*) has estimated the visual coverage and the distances at which golden shiners and zebrafish (*Danio rerio*) can resolve con-specifics; once such information has also been deduced for x-ray tetras we will be able to further refine the methods described here to better estimate areas under vigil by both individuals and groups.

The essential first step of our procedure was to estimate the position of each fish's eyes using the coordinates of each fish, $(x_i(t), y_i(t))$, the facing direction of each fish relative to the positive x-axis, $\theta_i(t)$, and measurements of the semi-major and semi-minor axis of the ellipse fitted to the image of each fish during automated tracking ($2e_x$ and $2e_y$). For simplicity, we made the reasonable assumption that each fish's eyes were located 90% of the distance from the end of the ellipse closest to the fish's tail to the end of the ellipse closest to the fish's snout, along the edge of the ellipse. The standard Cartesian equation

for an ellipse derived from our tracking data was

$$\frac{x^2}{(2e_x)^2} + \frac{y^2}{(2e_y)^2} = 1 \tag{S39}$$

(with the ellipse centred at the origin, and its major axis parallel to the x-axis). In such a standard form, the x-coordinates of both eyes of the fish would be at $x = 2e_x q$, with the parameter $q = 4/5$ to ensure that the eyes are located 9/10ths of the way along the major axis (4/5ths of the distance from the origin to the right edge of the ellipse (with $y = 0$) corresponds to 9/10ths of the distance from the left most edge of the ellipse to the rightmost edge of the ellipse (both at $y = 0$)). Substituting $x = 2e_x q$ into equation (S39) and rearranging then gave the y-coordinates of both eyes, at $y = \pm 2e_y\sqrt{(1-q^2)}$. For each time step we then tested that the approximate locations of all the fish's eyes remained within the outer boundary of the arena (there was a possibility that small errors in manually fitting the boundary to an image, as well as tracking and the placement of ellipses therein could result in parts of an ellipse extruding through our approximation for the location of the outer boundary). The test was simply to determine the distance from each eye to the centre of the outer boundary; if any such distance exceeded the radius of the outer boundary, then the eye was located outside the boundary. We excluded any time steps where any fish's eyes lay outside the boundary from subsequent analysis (there were 321 out of 270000 frames excluded from our analysis for this reason).

We assumed that in free space the range of vision from each eye of a solo fish was only obstructed by the fish's own body (which was approximated by the fitted ellipse). Therefore, the angular range of rays from each eye would be bounded by rays tangential to the ellipse that passed through the coordinates of each eye. In the configuration described by equation (S39), the slopes of the ellipse at the fish's left eye (where $x = 2e_x q$ and $y > 0$) and the fish's right eye (where $x = 2e_x q$ and $y < 0$) were

$$\frac{dy}{dx} = -\frac{e_y q}{e_x}\left(1-q^2\right)^{-\frac{1}{2}} \qquad \text{and} \qquad \frac{dy}{dx} = \frac{e_y q}{e_x}\left(1-q^2\right)^{-\frac{1}{2}} \tag{S40}$$

respectively. We then explicitly determined the angles between the tangents to the ellipse that passed through the left eye or right eyes respectively and the positive x-axis via

$$\alpha_{\text{left}} = \text{atan2}\left(-\frac{e_y q}{e_x}\left(1-q^2\right)^{-\frac{1}{2}}, 1\right) \qquad \text{and} \qquad \alpha_{\text{right}} = \text{atan2}\left(\frac{e_y q}{e_x}\left(1-q^2\right)^{-\frac{1}{2}}, 1\right) \tag{S41}$$

where $\text{atan2}(Y, X)$ is the four quadrant inverse tangent of X and Y as implemented by MATLAB, such that $-\pi < \text{atan2}(Y, X) \leq \pi$. Rays emanating from the left eye then had a possible angular range from α_{left} to $\pi - \alpha_{\text{left}}$ (radians, increasing in the anti-clockwise direction), and rays from the right eye had a possible angular range from $-\pi + \alpha_{\text{right}}$ to α_{right} radians (relative to the positive x-axis with the ellipse in the configuration described by equation (S39)). We chose the angular separation between consecutive rays so that consecutive rays would not be separated by more than the side length of the coarse grid

(5 mm) to which we ultimately transferred all coordinate data. The diameter of the outer boundary of the annulus was 660 mm; we thus treated the maximum possible distance from the eye of a fish to another object in the environment as 660 mm (aware that this was likely an over-estimate). Using the cosine rule ($c^2 = a^2 + b^2 - 2ab\cos C$), we determined the angular separation between two straight line segments with the same starting point that were separated by 5 mm at a distance of 660 mm along each segment (≈ 0.007576 radians). Therefore $n_{\text{rays}} = 415$ rays per eye (equally spaced in angle) were required to achieve the minimum separation in rays that we desired. Initially, we extended each ray from a given eye to a distance that guaranteed the ray would end outside the arena (we used the corner to corner distance of the square containing the annular arena, $\sqrt{730^2 + 730^2} \approx 1032$ mm). Finally, with the ellipse still in the standard configuration described by (S39), we identified the coordinates of an ordered set of discrete points around the ellipse, ($x_e(\gamma_k) = 2e_x\cos(\gamma_k), y_e(\gamma_k) = 2e_y\sin(\gamma_k)$) with $\gamma_k \in \left\{0, \frac{\pi}{180}, \frac{2\pi}{180}, \ldots, 2\pi - \frac{pi}{180}\right\}$ (radians), that we would ultimately use to form a grid based representation of the location of each fish. We used these coordinates to form a closed polygon (which approximated each ellipse).

We rotated the coordinates of the approximate location of each fish's eyes, the points on the border of the ellipse and the end points of all the rays cast from each eye $\theta_i(t)$ radians about the origin, and then shifted all coordinates $+x_i(t)$ mm in the x-direction and $+y_i(t)$ mm in the y-direction (so that the centre and orientation of the ellipse matched that of fish i at time t). We had to ensure that the coordinates of all objects, including the rays, were wholly contained within a square that bounded the outer edge of the annular arena before converting all objects to a grid based format. To enforce such a condition on the rays, we first explicitly calculated the points of intersection of each ray from each fish's eyes with the inner wall (closest to the water) of the outer boundary of the tank, since fish would not be able to see anything beyond that boundary and it would also be guaranteed that the ends of the rays would not extrude outside the domain to be converted to gridded format. Provided that the ray was not vertical, the points of intersection were the solution to the equations

$$r_c^2 = (x - x_c)^2 + (y - y_c)^2 \tag{S42}$$

$$y = m'x + b' \tag{S43}$$

where the slope $m' = (y_{r2} - y_{r1})/(x_{r2} - x_{r1})$ and intercept $b' = y_{r1} - m'x_{r1}$ of each ray was determined using the coordinates of the two end points of the ray ((x_{r1}, y_{r1}) (at the fish's eye) and (x_{r2}, y_{r2}) (at the far end of the line segment)). The two x-coordinates that satisfied equations (S42) and (S43) can be written as

$$x = \frac{(x_c - m'b' + m'y_c) \pm \sqrt{(x_c - m'b' + m'y_c)^2 - ((m')^2 + 1)(x_c^2 - 2b'y_c + (b')^2 + y_c^2 - r_c^2)}}{(m')^2 + 1} \tag{S44}$$

We then determined which of the two solutions (corresponding to the plus or minus square root term) in equation (S44) correctly identified the single crossing point between the ray and the outer boundary of the arena. The appropriate solution would satisfy the condition

that $\mu \in [0,1]$ where $\mu = \frac{x-x_{r1}}{x_{r2}-x_{r1}}$ (derived from the x-coordinate of the parametric form of the straight line segment from (x_{r1}, y_{r1}) to (x_{r2}, y_{r2}) – $x = x_{r1} + \mu(x_{r2} - x_{r1})$). We then determined the y-coordinate of the crossing point by substitution into equation (S44). If the ray was vertical, then the x-coordinate of the possible intersections between ray and boundary was $x = x_{r1} = x_{r2}$, with corresponding y-coordinates

$$y = y_c \pm \sqrt{r_c^2 - (x_{r1} - x_c)^2} \tag{S45}$$

Similar to a non-vertical ray, the appropriate solution branch ($\pm_{\surd}$) satisfied $\mu \in [0,1]$ where $\mu = \frac{y-y_{r1}}{y_{r2}-y_{r1}}$ (this time derived from the y-coordinate of the parametric form of the straight line segment from (x_{r1}, y_{r1}) to (x_{r2}, y_{r2})). Some numerical errors occured during our calculations that involved lines that were not exactly vertical, but that had very steep slopes (with gradients in excess of 1000000). We treated such rays as being 'approximately vertical', and then used equation (S45) to determine an approximate point of intersection of the ray with the outer boundary.

Once the coordinates of our discrete approximation to all ellipses fitted to the image of each fish and the rays from each fish's eyes (contrived to terminate on the interior side of the rim of the arena) were determined we proceeded to construct a coarser image of each object on an $n_g \times n_g = 146 \times 146$ grid of 5 mm $\times$ 5 mm squares. To do this we modified a method that was originally developed to construct gridded versions of contours from simulations of fluid motion, (*43,44*), and that we used more recently to quantify the area covered by the trajectories of eastern mosquitofish (*Gambusia holbrooki*) and crimson spotted rainbowfish (*Melanotaenia duboulayi*), (*40, 45*). Similar to the ellipses representing each fish, we constructed discrete representations of the inner and outer water level boundaries of the arena (once and for all for each trial). The coordinates on each circular boundary were given by $x_b(\gamma_k) = r_c \cos(\gamma_k) + x_c$, $y_b(\gamma_k) = r_c \sin(\gamma_k) + y_c$ for $\gamma_k = 0, \frac{\pi}{720}, \frac{2\pi}{720}, \ldots, \frac{1439\pi}{720}$ – we then used these coordinates to construct a closed polygon approximating each boundary. We then identified the extreme x- and y-coordinates of the discrete approximation to the outer circular edge of the tank, denoted $x_{\min}$, $x_{\max}$, $y_{\min}$ and $y_{\max}$. For each trial we then performed the following calculations for the inner circular boundary of the arena, and within each trial we performed the same calculations for the ellipses representing each fish for each time step. Writing the coordinates of the kth discrete point on either the circular boundary or ellipse as $(\check{x}_k(t), \check{y}_k(t))$, we then shifted and scaled the coordinates of each object via

$$\tilde{x}_k(t) = \text{floor}\left(\left(\frac{\check{x}_k(t) - x_{\min}}{x_{\max} - x_{\min}}\right)(n_g - 1) + 1\right) \quad \text{and} \quad \tilde{y}_k(t) = \text{floor}\left(\left(\frac{\check{y}_k(t) - y_{\min}}{y_{\max} - y_{\min}}\right)(n_g - 1) + 1\right) \tag{S46}$$

where the argument of floor$(\cdot)$ is rounded down towards zero. Such a transformation guaranteed that $1 \leq \tilde{x}_k(t) \leq n_g$ and $1 \leq \tilde{y}_k(t) \leq n_g$.

We used separate $n_g \times n_g$ matrices, $T_i(t)$, to store each of the grid based representations of either the inner circular boundary of the arena, or the ellipse fitted to each fish (with the locations of fish stored in $T_i(t)$ for $i = 1, \ldots, 8$ and $i = 9$ corresponding to the inner boundary). Initially all entries in each matrix $T_i(t)$ were set to 0. We then cycled through

the ordered series of transformed coordinates, $(\tilde{x}_k(t), \tilde{y}_k(t))$, for each object i. If the entry in row $\tilde{y}_k(t)$, column $\tilde{x}_k(t)$ equalled zero, we would then set that same entry to 1. It was possible that points k and $k+1$ on a given object did not occupy adjacent grid boxes, and thus we needed some method to fill in the intermediate boxes to produce a reasonable approximation to the ellipse or circle on the grid. We made use of Bresenham's line algorithm, (*46*), to identify grid boxes that lay on the straight line segment from point k to $k+1$. We then set entries in $T_i(t)$ corresponding to the intermediate points to 1. Next, we identified all grid boxes that lay inside or on the boundary of the polygon with coordinates given by equation (S46) using MATLAB's intrinsic *inpolygon* function. We set the corresponding entries in $T_i(t)$ to 1 for these interior points as well, effectively producing a solid grid based version of the fish or the inner circular section of the arena. (The reason for producing these solid grid based images, as opposed to just converting the edges of each object to the grid was connected to the next part of our method where the Bresenham algorithm was used to produce grid based versions of each of the rays from a fish's eye. We found that in some cases in practice the rays could cross the edge of objects on the grid without ever passing through a grid box where the corresponding entry of $T_i(t)$ was 1. Making the objects solid on the grid mitigated this problem.)
For each time step for each fish i we then summed all the matrices $T_j(t)$ excluding the matrix containing fish i's grid based representation (that is, we formed the sum $T_{\text{superimposed}}(t) = \sum_{j \neq i} T_j(t)$). We then determined the sign of $T_{\text{superimposed}}(t)$, which resulted in a matrix with entries equal to 1 where other fish or the circular boundary were located, or entries equal to zero otherwise. For each ray k from fish i's eye we initialised another $n_g \times n_g$ matrix of zeros, $\Omega_{i,k}(t)$. We shifted and scaled the coordinates of the two endpoints of the ray using equation (S46), and then set the entry in $\Omega_{i,k}(t)$ corresponding to the location of the fish's eye to 1. Next, we used the Bresenham algorithm to fill in intermediate points on the gridded representation of the line segment (and thus set appropriate entries in $\Omega_{i,k}(t)$ to 1), starting at the endpoint located at the fish's eye. For each point on the gridded version of the line (including the start point at the fish's eye) we checked the corresponding entry in $\text{sgn}(T_{\text{superimposed}}(t))$ – if the entry was 1, then we truncated the ray at that point in $\Omega_{i,k}(t)$ and examined the corresponding entries in all the $T_i(t)$ to determine which object (specific fish or the inner circular boundary) the gridded version of the ray first touched. If a ray was not terminated before the other endpoint was reached, we assumed the ray touched the outer boundary of the arena (since this was the endpoint of the ray that we had previously defined). We determined the sign of the sum of all the matrices $\Omega_{i,k}(t)$ corresponding to all the rays from a fish's eyes to produce another matrix $\bar{\Omega}_i(t)$ that contained what was effectively a coarse grained picture (drawn with a set of 1s) of the area covered by all the rays from a fish's eyes. We summed all the elements in $\bar{\Omega}_i(t)$ and then multiplied the result by the area of a grid box (25 mm^2) to obtain an estimate of the area seen by fish i at time t, denoted $V_i(t)$. Additionally, we determined the sign of the sum of all the matrices $\bar{\Omega}_i(t)$ to produce a coarse picture of the area covered by the rays from all fish's eyes at time t, denoted $\check{\Omega}(t)$. As with the individual gridded representations, we then summed all the elements in $\check{\Omega}(t)$ and multiplied by the area of a grid box to obtain an estimate for the area visible to the group at time t, $V_{\text{group}}(t)$. Finally,

we used information about the objects touched by each ray to determine the number of partners that each fish could see at time t, denoted $G_i(t)$.
Due to the relative coarse grid, it was possible that one fish's eyes (and thus the start point for the rays from its eyes) occupied the same grid box as a part of another fish. The consequence of such an arrangement was that all gridded versions of the rays from the fish's eyes terminated in the same grid box as they started, therefore only one element in $\bar{\Omega}_i(t)$ would be set to 1 (with all other entries zero) and $V_i(t)$ would equal 25 mm^2. We treated any time step where $V_i(t) = 25$ for any fish as a time where a likely large underestimate had been made of the area that fish could see, with the error also likely affecting our approximation for the area seen by all fish. We therefore excluded any time steps where $V_i(t) = 25$ for any fish from further analysis (a total of 34963 out of 270000 frames had to be ignored for this reason, in addition to the 321 frames excluded due to obvious errors in the identification of the approximate location of fish's eyes – ultimately we made use of 86.93% of our video data for the analysis described in this section).
We generated relative frequency histograms of the area sighted by individuals, the area sighted by the group and the number of group mates seen by each fish derived from data pooled within each treatment and larger A or P time interval across all individuals (where applicable) and time steps.
Results of analysis of the quantities described in this section appear in section S2.6, table S6 and figs. S37 to S39.

S1.8 Individual durations between swaps in sense of motion (clockwise or anticlockwise motion about approximate centre of tank)

We determined the duration between swaps in senses of direction (either anticlockwise or clockwise) about the approximate centre of the tank (defined to be the centre of the inner circular boundary of the tank) by individual fish. Another way this could be thought of is that we determined the durations that individual fish had uninterrupted travel either clockwise or anticlockwise around the tank.
For each time step we constructed a vector (in three dimensions) from the centre of the inner circular boundary of the tank (x_c, y_c) to the location of each fish. The x and y components of such a vector respectively were

$$\rho_{i,x}(t) = x_i(t) - x_c \qquad \text{and} \qquad \rho_{i,y}(t) = y_i(t) - y_c \tag{S47}$$

We then formed the cross-product of the above vector ($\boldsymbol{\rho}_i(t) = (\rho_{i,x}(t), \rho_{i,y}(t), 0)$) with the unit vector in the fish's current direction of motion ($\hat{\mathbf{u}}_i(t) = (\hat{u}_i(t), \hat{v}_i(t), 0)$), denoted $\mathbf{S}_i(t) = \boldsymbol{\rho}_i(t) \times \hat{\mathbf{u}}_i(t)$. If the vertical component of $\mathbf{S}_i(t)$ (the component of $\mathbf{S}_i(t)$ in the positive z-direction) was positive, then that indicated that the fish was travelling anticlockwise about the centre of the inner circular boundary. If the component of $\mathbf{S}_i(t)$ in the z-direction was negative, then the fish was moving clockwise about the centre of the inner circle (when viewed from above). (The components of $\mathbf{S}_i(t)$ in the x and y directions

were guaranteed to be zero.) We recorded the sign of the vertical component of $\mathbf{S}_i(t)$ as $\mathrm{AC}_i(t)$. We determined the start and end frames of each uninterrupted duration of motion in either the clockwise or anticlockwise senses by examining sign changes in $\mathrm{AC}_i(t)$; motion in a particular sense was concluded by either a swap from clockwise to anticlockwise motion, a swap from anticlockwise to clockwise motion, or by a fish stopping movement completely (hence having zero velocity, and therefore $\mathrm{AC}_i(t) = 0$). For initial examination of the data we excluded any durations that started at the first frame of a given time interval, or concluded on the final frame of a given time interval (effectively ignoring any censored data). We then pooled all durations of uninterrupted motion (in either sense) for all fish across all three ante (A) and all three post (P) time intervals for each treatment. We constructed relative frequency histograms to then examine the distribution of observed durations (see fig. S2). These histograms were characterised by much of the data being allocated to short duration bins, with relatively fewer entries allocated to higher duration bins. Given the observed shape of the data we opted to use survival analysis to compare the durations of uninterrupted motion in a given sense. See section S1.17 for more details of our survival analysis calculations, which we applied to the durations between changes in sense of motion, as well as durations spent within a threshold distance of either the inner or outer boundaries of the water in the arena.
Results of survival analysis of the quantities described in this section appear in section S2.7, table S7 and fig. S40

S1.9 Correlation in direction of motion as a function of time delay

We examined the tendency of fish to match the direction of motion adopted by their partners at other times within ±4 seconds (*2, 5*), dependent on if partner fish were located in front or behind a focal fish, relative to the facing direction of the focal fish.
For all time steps and pairs of fish, i and j, we first determined the magnitude of the relative angle between the facing direction of fish i, $\theta_i(t)$, and the straight line segment from the coordinates of fish i, $(x_i(t), y_i(t))$, to the location of their partner j, $(x_j(t), y_j(t))$. This angle was given by

$$\beta_{i,j}(t) = \cos^{-1}\left(\frac{\cos(\theta_i(t))(x_j(t) - x_i(t)) + \sin(\theta_i(t))(y_j(t) - y_i(t))}{\sqrt{(x_j(t) - x_i(t))^2 + (y_j(t) - y_i(t))^2}}\right) \tag{S48}$$

If $\beta_{i,j}(t) \leq 90°$ then fish j was in front of fish i relative to the facing direction of fish i at time t, otherwise (when $\beta_{i,j}(t) > 90°$) fish j was behind fish i.
We next determined unit vectors in the direction of motion of each fish for each time step, which had components as given by equation (S12). For each pair of fish, i and j, we then constructed two sets of time series of each fish's direction of motion (implicit in equation (S12)). The first of these time series left all entries where fish j was in the frontmost position blank (that is, it only contained information for the time steps when fish i was in front), and the second time series left all entries where fish i was in front

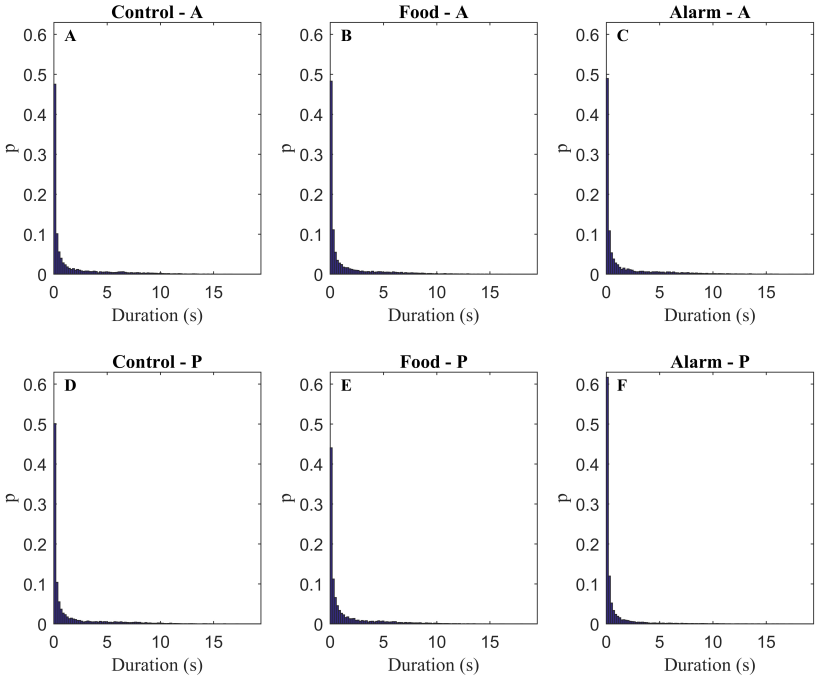

fig. S2. Observed distribution of the durations of unbroken travel by individual fish either in the clockwise or anticlockwise sense about the centre of the inner circular boundary of the arena over the full range of observed values for the treatments and time intervals: A – control treatment ante cue, B – food treatment ante cue, C – alarm treatment ante cue, D – control treatment post cue, E – food treatment post cue, F – alarm treatment post cue.

blank. For each set of time series (corresponding to fish i in front or fish j in front), we determined the directional correlation

$$C_{ij}(\tau) = \langle \hat{u}_i(t)\hat{u}_j(t+\tau) + \hat{v}_i(t)\hat{v}_j(t+\tau) \rangle \tag{S49}$$

where $\tau = \tau_n \Delta t$ was time-lag in seconds, $\tau_n \in \{-100, -109, \ldots, 100\}$ was the number of frames corresponding to a given time-lag and $\langle \cdot \rangle$ represented the mean taken over all t. (The term inside the angle brackets is the dot/inner product of the direction of motion of fish i at time t and the direction of motion of fish j at time $t + \tau$.) We then identified the maximum value of $C_{ij}(\tau)$ and the value of τ that corresponds to this maximum, denoted τ^*_{ij}. Provided that $C_{ij}(\tau^*_{ij})$ was large enough to suggest that there was reasonable correlation in the directions of motion of the two fish, a positive value of τ^*_{ij} suggested that fish j adjusted its direction of motion to match that adopted by fish i at an earlier time (that is, fish j was following the direction of fish i), whereas a negative value of τ^*_{ij} suggested that fish i was following fish j on average.

For each fish i we determined the mean values of τ_{ij}^* and $C_{ij}(\tau_{ij}^*)$ across all partner fish j for both instances when partner fish were in front of or behind fish i (relative to the facing direction of fish i). We further averaged these values across all fish in a trial for each one minute time interval, and then again averaged these values across all A or P time intervals for each treatment.
Results of the analysis of directional correlation appear in section S2.8 and table S8.

S1.10 Alignment responses associated with rapid turns

In addition to examining the average tendency of fish to align with partners after a short time delay, we also examined the tendency of fish to align with rapidly turning partners on a case-by-case basis.
We first pooled all observed turning speeds exhibited by all fish in all treatment groups prior to the application of any cue (as given by equation (S13)). We then defined a rapid turning speed as a turning speed greater than or equal to the mean plus two standard deviations of the pooled turning speeds (denoted α_{rapid}, approximately 912.01 degrees/s). For each fish, i, we identified the ordered set of indices of all frames where the fish's turning speed was greater than or equal to α_{rapid}. We then formed differences of all consecutive pairs of the ordered set of indices. We used these differences to identify separate instances of rapid turning; we defined separate rapid turning events as any where the associated separation in frames where there was rapid turning speed was greater than 10 frames (0.400 seconds). Ultimately we identified the time steps, t_r, which indicated the start of a rapid turn for each fish i, and tallied the number of rapid turns produced by each individual in each one minute time interval. We then cycled through all partner fish $j \neq i$. If fish j could see i according to the calculations described in section S1.7 at time t_r and fish i and j were separated by 90 mm (approximately three body lengths) or less, we calculated the set of correlations in direction

$$\tilde{C}_{ij}(\tau) = \hat{u}_i(t_r)\hat{u}_j(t_r + \tau) + \hat{v}_i(t_r)\hat{v}_j(t_r + \tau) \tag{S50}$$

where $\tau = \tau_n \Delta t$ and $\tau_n \in \{-75, -74, \ldots, 75\}$. That is, we calculated the tendency for j to align with the rapid turner i for time delays between -3 and 3 seconds; unlike equation (S49), we did not average the values of $\tilde{C}_{ij}(\tau)$ over time. We identified the maximum value of $\tilde{C}_{ij}$ for each instance of rapid turning and the corresponding time delay, τ. Denoting the maximum correlation as a function of time delay as $\tilde{C}_{ij}^*$, we treated fish j as closely aligning with the rapid turn of fish i if $\tilde{C}_{ij}^* \geq 0.9848$, which corresponded to alignment in directions of motion of 10° or less, and the time delay corresponding with maximum correlation was positive.
We pooled the time delays of maximum alignment with other fish's rapid turns for each fish in a group for each time interval. We determined the mean and standard deviation of each fish's set of time delays within a given one minute time interval, as well as the number of rapid turns preformed by each fish within the same time interval; we then calculated the mean of these means and standard deviations across all members in the group. We further averaged the measures again across all three A or P time intervals.

Results of the analysis of rapid turns appear in section S2.9 and table S9.

S1.11 Number of times isolated from group, and duration of individual periods in isolation

We used the algorithm outlined in (*45*) to classify fish as being part of a distinct subgroup based on the distances between individual fish. In the case of the tetras here, the algorithm identified a distinct subgroup of fish as a set of fish where no fish in the set was more than three body lengths (90 mm) from any other fish in the set at time t. For completeness, we again outline the basic algorithm here.

For each time t, we assigned the first fish in a given group to subgroup 1. We then identified all fish that were less than or equal to 90 mm from fish 1 with the help of equation (S15), and assigned any such fish to subgroup 1. For each newly assigned fish i, we then identified any other fish j not already assigned to a subgroup for which $d_{i,j}(t) \leq 90$ mm. If such fish were found, then they were assigned to subgroup 1, and then the process of cycling through newly assigned fish to find any other unassigned fish for which $d_{i,j}(t) \leq 90$ mm was repeated. If no fish for which $d_{i,j}(t) \leq 90$ mm were found during a cycle, then a new subgroup was started, with the first fish not already assigned to another subgroup identified as the first member of the new subgroup. The process of finding all fish less than or equal to 90 mm from the fish identified as the founding member of a new subgroup was then repeated as described above. The algorithm terminated when all fish in a group had been assigned to a subgroup for time t. We stored the indices of all fish that appeared in each distinct subgroup for all times t.

We identified and tallied all frames where each fish i was the only member of their particular subgroup. We then treated each series of consecutive time steps where a fish was the only subgroup member as a distinct instance of isolation. We tallied the number of distinct instances of isolation by each fish, and determined the duration of each distinct instance of isolation. We pooled the number of isolation events observed for each fish within a one minute time interval across all A or P time intervals for all trials subject to a given treatment (so that there were three measurements of the number of isolation events per minute for each fish included in the pooled data). We then used the pooled data to construct relative frequency histograms for the number of observed isolation events per fish per minute. We applied survival analysis, as outlined in section S1.17, to examine if any differences in individual periods spent in isolation was evident across treatments or larger (A or P) time intervals (using isolation duration data pooled from all individuals within each treatment and A or P time intervals).

Results of the analysis of isolation events appear in section S2.10, tables S10 and S11, and figs. S41 and S42.

S1.12 Durations spent on the edge of the group, and frequency of swapping between interior and exterior positions

We sought evidence that fish adjusted the duration that they spent on the edge of the group (as opposed to the group's interior) under different external cues. To do this, we determined the convex hull for the set of fish's coordinates for each time step (in each trial) using MATLAB's intrinsic *convhull* function. In two dimensions, a convex hull is the smallest convex set that contains all the points in a given set of coordinates. Also in two dimensions, a convex set is a region such that if two arbitrary points are chosen from the region, then every point on the straight line segment joining the two arbitrary points must also be contained within, or lie on the edge of, the convex set/region. If a fish was on an edge or corner of the convex set, then we treated it as also being on the edge of the group. Fish that were not on the edge of the convex set were identified as occupying the interior of the group for a given time step. We constructed a time series for each fish that identified when the fish was on the edge of the group and when the fish was in the interior. We tallied the number of times each fish swapped position (from the interior to the exterior of the group, or from the exterior of the group to the interior) in each minute long time interval, and then pooled the number of swaps made by all fish across all A and P time intervals within a given treatment (so that three counts were added to the pool for each fish, similar to the pooling of instances of isolation in section S1.11). We then constructed relative frequency histograms for each treatment and larger time interval (A or P). We determined each unbroken duration that fish spent on the exterior of the group, and then examined if there were any differences in durations spent on the exterior using the survival analysis outlined in section S1.17.

Results of the analysis of individual occupancy of interior and exterior group positions appear in section S2.11, tables S12 and S13, and figs. S43 and S44.

S1.13 An eye for trouble? Do x-ray tetras prefer a particular eye when scanning for food or danger?

We sought evidence that fish may orient themselves so that a particular eye pointed inwards or outwards relative to the group centre under different treatments. A fish was oriented anticlockwise relative to the group centre, $(c_x(t), c_y(t))$, and thus had their right eye pointing outwards if $\operatorname{sgn}((\mathbf{F}(t) \times \mathbf{f}(t)) \cdot \mathbf{k}) = 1$; a fish was oriented clockwise/had their left eye pointing outwards if $\operatorname{sgn}((\mathbf{F}(t) \times \mathbf{f}(t)) \cdot \mathbf{k}) = -1$ where $\mathbf{F}(t) = (x_i(t) - c_x(t))\mathbf{i} + (y_i(t) - c_y(t))\mathbf{j}$ was the vector from the group centre to the location of fish i, $\mathbf{f}(t) = \cos(\theta_i(t))\mathbf{i} + \sin(\theta_i(t))\mathbf{j}$ was a unit vector pointing in the facing direction of fish i, $(x_i, y_i(t))$ were the coordinates of fish i at time t, $\theta_i(t)$ was the facing direction of fish i relative to the positive x-axis, $\mathbf{i}$, $\mathbf{j}$ and $\mathbf{k}$ were unit vectors pointing in the x, y and z directions respectively, $\cdot$ denoted the dot product and $\times$ denoted the cross product. We then calculated the fraction of the group with their right eye pointing outwards, $\text{frac}_{\text{right}}(t)$ for each time step.

Results of analysis of the fraction of the group with their right eye pointing outwards appear in the supplementary information in section S2.12 and table S14.

S1.14 Tendency to face towards or away nearest walls

We determined the fraction of individuals facing outward from the wall that was on average closest to all group members in a given time step.
The mean distance of individual fish from the inner boundary at time t was determined directly from the mean across individuals, i, of $d_{i,b_{\text{inner}}}(t)$ (refer to section S1.3). Similarly the mean distance of individuals from the outer boundary at time t was the mean of $d_{i,b_{\text{outer}}}(t)$ across all i. Thus it was possible to determine which wall was closest to group members on average for each time step t.
We classified fish as facing towards or away from a boundary based on the magnitude of the angle between a unit vector pointing in the fish's current facing direction and a unit vector perpendicular to the closest wall to the group on average, at the point on that wall closest to the fish's location, $(x_{i,b}(t), y_{i,b}(t))$ (see section S1.3).
For either boundary, the components of a vector normal to the boundary at the point closest to a given fish were calculated via

$$n_{i,x}(t) = x_{i,b}(t) - x_c \qquad \text{and} \qquad n_{i,y}(t) = y_{i,b}(t) - y_c \tag{S51}$$

where (x_c, y_c) was the coordinates of the centre of the boundary of interest.
The vector with components given by (S51) either pointed towards the water-filled interior of the tank if determined relative to the inner boundary, or away from the water-filled interior if determined relative to the outer boundary. We then determined a unit vector pointing in the direction of the vector with components given by (S51), writing the components of the unit vector in the x and y directions as $\hat{n}_{i,x}(t)$ and $\hat{n}_{i,y}(t)$ respectively. We also determined unit vectors in the facing direction of each fish, with components

$$f_{i,x}(t) = \cos(\theta_i(t)) \qquad \text{and} \qquad f_{i,y}(t) = \sin(\theta_i(t)) \tag{S52}$$

We then determined the cosine of the angle, $\beta_i(t)$, between the unit vector perpendicular to the wall and the unit vector in the facing direction of each fish i

$$\cos(\beta_i(t)) = \hat{n}_{i,x}(t) f_{i,x}(t) + \hat{n}_{i,y}(t) f_{i,y}(t). \tag{S53}$$

If $0^\circ < \beta_i(t) < 90^\circ$ with respect to the inner wall, then we classified fish i as facing away from that wall. For comparisons between facing directions and vectors perpendicular to the outer wall, $90^\circ < \beta_i(t) < 180^\circ$ indicated that fish were facing away from the outer wall.
We determined the fraction of group members facing away from the mean nearest wall for each time step, $F_o(t)$. We constructed relative frequency histograms of the observed values of $F_o(t)$ within each treatment and A or P time intervals.
Results of the analysis of wall facing appear in section S2.13, table S15 and fig. S45.

S1.15 Durations spent within a threshold distance of the walls

We used the time series of distances from each fish to either the outer or inner boundaries of the arena ($d_{i,b_{\text{outer}}}(t)$ or $d_{i,b_{\text{inner}}}(t)$ respectively, as calculated in section S1.3) to identify all

time steps where each fish was less than or equal to a threshold distance, d_{thresh}, from the given boundary (we performed separate calculations for $d_{\text{thresh}} = 30, 60$ mm; approximately one or two body lengths). We identified a visit to a boundary region as a set of consecutive frames where a fish was within the threshold of the wall, and hence deduced the duration of each visit in seconds using the formula ((index of last frame of visit) - (index of first frame of visit) + 1)/(frames/s). We pooled all durations of individual visits to each boundary region for each treatment and larger time interval (A or P). We then applied the survival analysis outlined in section S1.17 to determine if different treatments affected the duration spent close to either of the arena's boundaries.
Results of the analysis of durations spent close to the boundaries appear in section S2.14, tables S16 to S18, and figs. S46 and S49.

S1.16 Bootstrap analysis

We collapsed our data relating to many of the measures described in the previous sections down to single values for the A and P time intervals for each group according to the following procedures (resulting in two values per group).
We determined the median and standard deviation over time of the speed, turning speed, magnitude of acceleration, change in speed over time, distance to group centroid, nearest neighbour distance, mean neighbour distance, distance to the closest point on the inner wall, distance to the closest point on the outer wall, approximate area of free space visible and number of group mates visible for each individual fish within each of the one minute A1 to P3 time intervals. We then calculated the mean across group members of each of the medians and standard deviations within each time interval (A1 to P3), and then further averaged these values across all three A or P time intervals to obtain a single number representation of each quantity relating to individual behaviour in the larger A and P time intervals for each group.
For each of the smaller A1 to P3 time intervals we determined the mean and standard deviation (across group members) of the number of rapid turns, the number of times isolated from any group member by more than 90 mm and the number of times fish changed between positions on the interior or exterior of the group during each one minute interval. As above, we then averaged each of these values across all three A or P time intervals.
Values for the mean time delay associated with maximum average alignment with partner fish in front or behind focal fish as well as the mean maximum average alignment and the mean and standard deviation in delay before alignment with rapid turns of visible partners were calculated as described in sections S1.9 and S1.10.
We determined the median and standard deviation over time of group polarisation based on the direction of motion of fish, group polarisation based on the facing direction of fish, the fraction of 360° viewed by each group, group expanse, the fraction of the group facing outwards from the section of wall that was at the least mean distance from all group members and the area sighted by each group for each one minute (A1 to P3) time interval. We also determined the mean and standard deviation over time of the fraction

of the group that had their right eye pointing outwards relative to the group centroid for the A1 to P3 time intervals. As with all the individual measures, we then further averaged each summary statistic of group properties across all three A or P time intervals for each group.

We then applied bootstrap analysis to determine if each of the above measures changed from interval A to P. Here we describe the bootstrap specifically applied to the mean median speed of individual fish, but the procedure was analogous for all other measures examined. We assumed that there was an effect of the duration in the arena on observed behaviour (an effect of time from interval A to P for each group) and that there was an effect of group composition on observed behaviour (that is, variation between groups was a reasonable expectation). We determined the mean across all 10 trials within each treatment and time interval of the mean median speeds, denoted $\bar{s}_{CA}$ and $\bar{s}_{CP}$ and for control groups during time intervals A and P respectively, $\bar{s}_{FA}$ and $\bar{s}_{FP}$ for food groups during intervals A and P respectively, and $\bar{s}_{AA}$ and $\bar{s}_{AP}$ for alarm groups during intervals A and P respectively. From these means we calculated the reference statistics

$$\begin{aligned} \Delta\bar{s}_C &= \bar{s}_{CP} - \bar{s}_{CA} \\ \Delta\bar{s}_F &= \bar{s}_{FP} - \bar{s}_{FA} - \Delta\bar{s}_C \quad \text{and} \\ \Delta\bar{s}_A &= \bar{s}_{AP} - \bar{s}_{AA} - \Delta\bar{s}_C \end{aligned}$$

In the above formulation, the reference statistics $\Delta\bar{s}_F$ and $\Delta\bar{s}_A$ were essentially corrected for underlying time effects under control conditions by subtracting any difference in mean median speeds observed for control groups. We proceeded with our bootstrap by randomly selecting mean median speed data from 10 trials within each treatment with replacement. Selecting data for a given trial automatically meant selecting data from both A and P time intervals (to account for any underlying difference due to group composition). We determined the means of each pool of randomly selected data, $\bar{s}^*_{CA,i}$, $\bar{s}^*_{CP,i}$, $\bar{s}^*_{FA,i}$, $\bar{s}^*_{FP,i}$, $\bar{s}^*_{AA,i}$ and $\bar{s}^*_{AP,i}$, and from these calculated the statistics

$$\begin{aligned} \Delta\bar{s}^*_{C,i} &= \bar{s}^*_{CP,i} - \bar{s}^*_{CA,i} \\ \Delta\bar{s}^*_{F,i} &= \bar{s}^*_{FP,i} - \bar{s}^*_{FA,i} - \Delta\bar{s}^*_{C,i} \quad \text{and} \\ \Delta\bar{s}^*_{A,i} &= \bar{s}^*_{AP,i} - \bar{s}^*_{AA,i} - \Delta\bar{s}^*_{C,i} \end{aligned}$$

We repeated the above process of random sampling with replacement and subsequent calculation of $\Delta\bar{s}^*_{C,i}$, $\Delta\bar{s}^*_{F,i}$ and $\Delta\bar{s}^*_{A,i}$ 9999 times. We then identified the 2.5th and 97.5th percentiles of each set of bootstrap samples (denoted $\hat{q}_{C,0.025}$ and $\hat{q}_{C,0.975}$ for $\Delta\bar{s}^*_{C,i}$, $\hat{q}_{F,0.025}$ and $\hat{q}_{F,0.975}$ for $\Delta\bar{s}^*_{F,i}$, and $\hat{q}_{A,0.025}$ and $\hat{q}_{A,0.975}$ for $\Delta\bar{s}^*_{A,i}$). Finally , we used the basic method to estimate 95% bootstrap confidence intervals for differences in mean mean median speeds (see for example (*36*)). These confidence intervals were given by $(2\Delta\bar{s}_C - \hat{q}_{C,0.975}, 2\Delta\bar{s}_C - \hat{q}_{C,0.025})$ for control groups, $(2\Delta\bar{s}_F - \hat{q}_{F,0.975}, 2\Delta\bar{s}_F - \hat{q}_{F,0.025})$ for food groups and $(2\Delta\bar{s}_A - \hat{q}_{A,0.975}, 2\Delta\bar{s}_A - \hat{q}_{A,0.025})$ for alarm groups.

If a confidence interval lay entirely above or below 0 for a given treatment, this indicated a significant change in the underlying measure (such as mean mean median speed). A

confidence interval entirely above 0 indicated that the underlying measure had increased from interval A to P, whereas a confidence interval entirely below 0 indicated a decrease in the underlying measure from interval A to P (taking into account any time effects, as observed in control groups).

S1.17 Survival analysis

We used standard methods of survival analysis (*47*) to examine changes in duration between swaps in the sense of motion of individual fish, durations spent in isolation, durations spent by individuals on the edge of their group and durations spent by individuals within a relatively small threshold distance of the arena's inner and outer boundaries, and durations spent by individuals on the edge of their group (see sections S1.8, S1.11, S1.12 and S1.15). We treated durations associated with each behaviour of interest as right censored (that is, the duration of the behaviour was known to be at least that observed, but the actual duration of the behaviour was unknown) if the behaviour was observed to start on the first frame of a given time interval, end on the last frame of a given time interval, or both. All other durations were treated as uncensored. We then formed separate pools of uncensored and censored durations for each of the larger A and P time intervals for each treatment. We then constructed Kaplan-Meier estimates of the survival functions associated with the sets of durations for each of the A and P time intervals for each treatment following the method outlined in (*47*). Each survival function, $S(t) = P(T > t)$, represented the probability that a fish exhibited a given behaviour for a duration greater than t seconds. In addition, for visualisation purposes we determined approximate bounds for the 95% confidence interval for each survival function.

We performed a log-rank test to determine the probability that at least one survival function relating to a given behaviour differed from the others (as detailed in (*47*). If there was a significant difference (at significance level $\alpha_{\text{sig}} = 0.05$), we then performed pairwise comparisons of all possible pairs of survival functions (15 pairs in total for pairs derived from 6 survival functions) using additional log-rank tests to determine which pairs differed. Once the p-value was determined for each pairwise comparison, we sorted the results of all tests in ascending order of p-value and identified significant results after applying a Holm-Bonferroni correction (*48*). Subsequently, we only retained results for more meaningful pairwise comparisons (where comparisons were made within the same treatment across time intervals, or across treatments within the same time interval).

section S2. Results

S2.1 Basic measures of individual locomotion

Figures S3 to S6 illustrate the observed distributions of the speed, magnitude of acceleration, change in speed over time and turning speed of individual fish during the ante (A) and post (P) time intervals of each of the control, food and alarm treatments. Table 1 in the main text summarises the results of bootstrap analysis of the median and standard deviation of speed, turning speed and magnitude of acceleration; table S3 tabulates results for the median and standard deviation of the change in speed over time.

The mean median speed of fish subject to alarm cues decreased after cues were released into the arena according to our bootstrap analysis. The most dramatic change in the distribution of observed speeds occured for fish subject to alarm cues during the post cue time interval, with an obvious shift in the distribution such that fish were more frequently observed swimming at low speed (fig. S3 F).

The mean median magnitude of acceleration decreased from interval A to interval P for fish in control and alarm groups, but increased for fish subject to food cues, according to bootstrap analysis. Changes in distributions of the magnitude of acceleration from A to P time intervals were evident for fish subject to all treatments, including control groups, with the most marked changes being a greater proportion of higher magnitudes of acceleration for fish in food groups, and a greater proportion of lower magnitudes of acceleration for fish in alarm groups (fig. S4).

The mean median change in speed over time increased from interval A to P for fish in alarm groups and decreased for fish in food groups according to bootstrap analysis. The distributions of changes in speed over time changed from A to P time intervals such that a lesser proportion of changes in speed close to zero were observed for fish subject to food cues, and a greater proportion of changes in speed close to zero were observed for fish subject to alarm cues (fig. S5).

The mean median and mean standard deviation turning speed of fish subject to alarm cues increased after cues were released according to bootstrap analysis. Visible differences in the distributions of observed turning speeds were not easy to see overall, but the proportion of low turning speeds (close to 0 rad/s) diminished from A to P time intervals for fish subject to alarm cues, so there must necessarily have been a greater proportion of turning speeds observed above the bottom end of the scale (compare fig. S6 panels C and F).

Overall, fish subject to alarm cues exhibited lower median speeds, greater median and standard deviations in turning speeds, decreased median magnitudes of acceleration and increased median changes in speed over time. Fish subject to food cues exhibited greater median magnitudes of acceleration and decreased median changes in speed over time. The only significant effect from ante to post cue time intervals for fish in control groups was a decrease in median magnitude of acceleration.

table S3. Basic 95% confidence intervals for test statistics derived from the median and SD change in speed over time, $\frac{\Delta s_i}{\Delta t}(t)$, of individual fish. Statistically significant effects are marked with an asterisk (*) (ie. confidence intervals for the test statistics that lie entirely above or below 0). If a confidence interval lies entirely below zero, then the associated quantity decreased from interval A to P; if a confidence interval lies entirely above zero then the associated quantity increased from A to P.

Variable	Control	Food	Alarm
median $\frac{\Delta s_i}{\Delta t}(t)$ (mm/s^2)	$(-4.53, 8.48)$	$(-22.84, -4.35)^*$	$(11.32, 30.49)^*$
std $\frac{\Delta s_i}{\Delta t}(t)$ (mm/s^2)	$(-33.67, 30.51)$	$(-1.42, 83.75)$	$(-64.18, 16.92)$

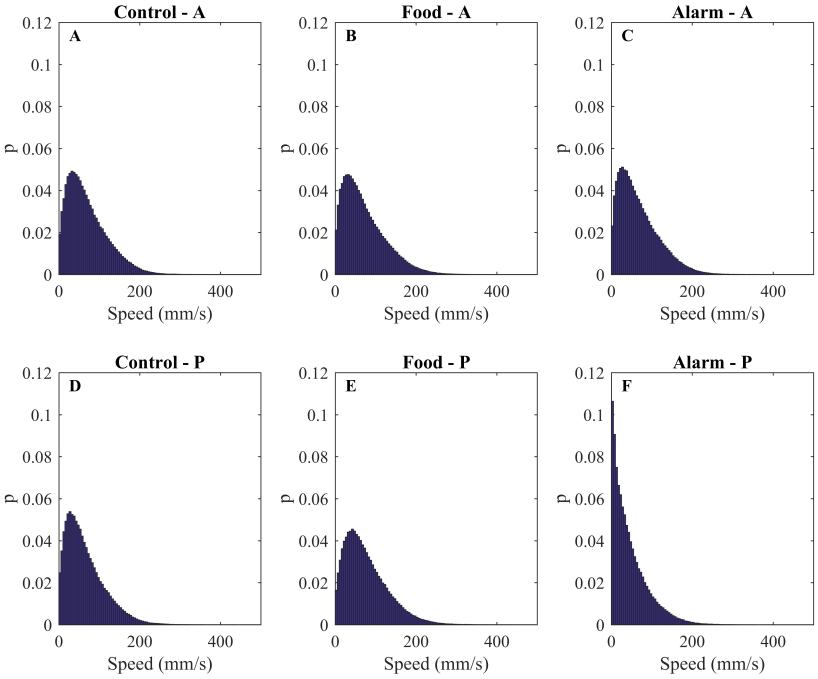

fig. S3. Observed distribution of the speed of individual fish (in mm/s) over the range 0 to 500 mm/s for the treatments and time intervals: A – control treatment ante cue, B – food treatment ante cue, C – alarm treatment ante cue, D – control treatment post cue, E – food treatment post cue, F – alarm treatment post cue. Details of the calculations used to generate these plots are given in section S1.2.

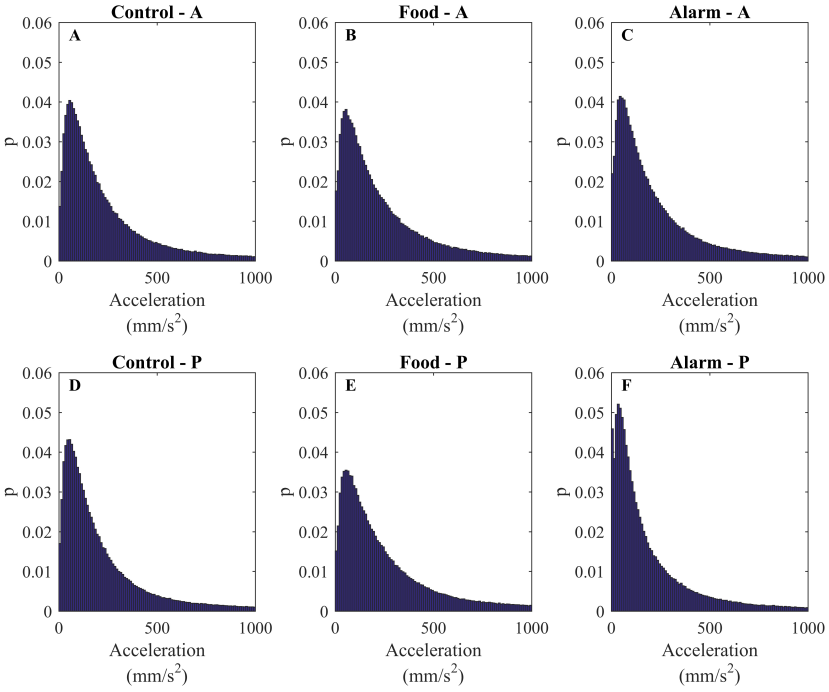

fig. S4. Observed distribution of the magnitude of acceleration of individual fish (in mm/s^2) over the range 0 to 1000 mm/s^2 for the treatments and time intervals: A – control treatment ante cue, B – food treatment ante cue, C – alarm treatment ante cue, D – control treatment post cue, E – food treatment post cue, F – alarm treatment post cue. Details of the calculations used to generate these plots are given in section S1.2.

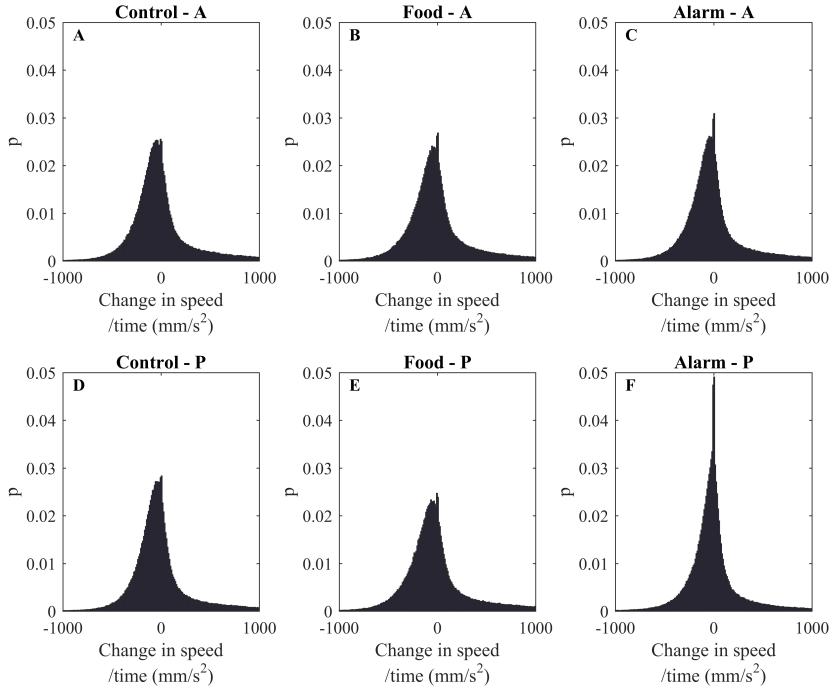

fig. S5. Observed distribution of the change in speed over time of individual fish (in mm/s^2) over the range -1000 to 1000 mm/s^2 for the treatments and time intervals: A – control treatment ante cue, B – food treatment ante cue, C – alarm treatment ante cue, D – control treatment post cue, E – food treatment post cue, F – alarm treatment post cue. Details of the calculations used to generate these plots are given in section S1.2.

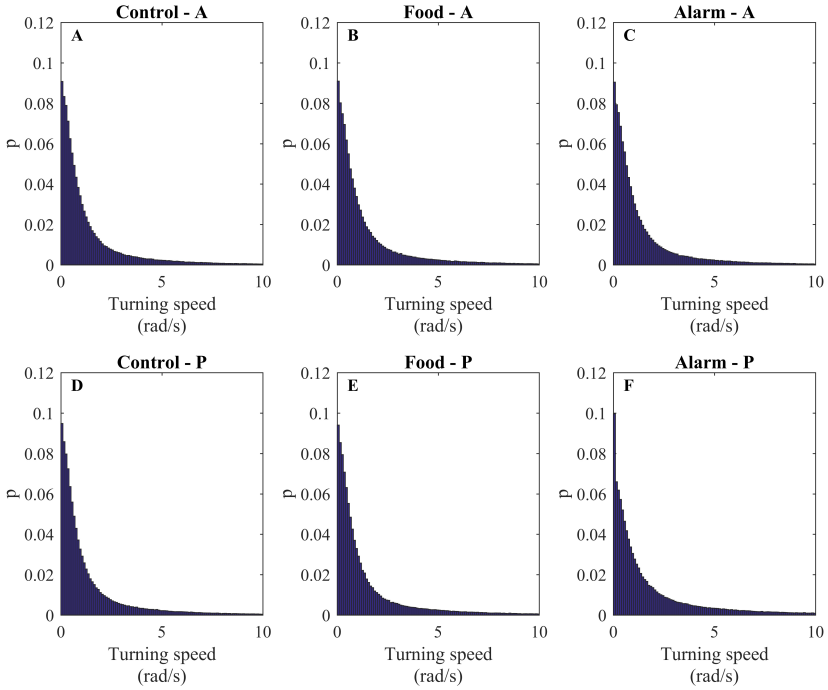

fig. S6. Observed distribution of the turning speed of individual fish (in radians/s) over the range 0 to 10 radians/s for the treatments and time intervals: A – control treatment ante cue, B – food treatment ante cue, C – alarm treatment ante cue, D – control treatment post cue, E – food treatment post cue, F – alarm treatment post cue. Details of the calculations used to generate these plots are given in section S1.2.

S2.2 Basic measures of group configuration

Figures S7 to S12 illustrate the observed distributions of the distance of individual fish from the group centroid, distance to nearest neighbours, mean distance to all neighbours at a given time, group expanse, polarisation in direction of motion and polarisation in facing direction. Table S4 summarises the results of bootstrap analysis of the median and standard deviation of each of the above measures.

The mean median and standard deviation in distance from individual fish to the group centroid decreased from interval A to P for fish subject to alarm cues according to bootstrap analysis. Changes in the observed distribution of distances of individuals to the group centre were most evident for fish subject to alarm cues, with a greater proportion of distances closer to zero observed after deployment of a cue (fig. S7 F).

The mean median and standard deviation distance from fish to their nearest neighbour decreased post cue for fish in alarm groups according to bootstrap analysis. There were shifts in the observed distributions of nearest neighbour distances that suggested that a greater proportion of smaller distances were observed for fish subject to alarm cues (fig. S8). Additionally, mean median mean neighbour distance decreased for fish subject to alarm cues according to bootstrap analysis. As with changes in the distributions of distances of individuals to group centroids and nearest neighbour distances, there were a greater proportion of small distances observed for fish subject to alarm cues (fig. S9).

Median group expanse decreased for fish subject to alarm cues according to bootstrap analysis, and the standard deviation in group expanse decreased for fish subject to food cues. Changes in the distribution of observed values of group expanse from A to P time intervals were largely consistent with what was indicated by the analysis of individual spacing (individual distances to group centroid, nearest neighbours and mean neighbour distances). A greater proportion of smaller expanse values were observed for fish in alarm groups post cue (fig. S10 C and F).

All the above analysis strongly suggests that fish subject to alarm cues tended to bunch up/decrease the distance to their neighbours (according to four different, albeit similar measures of spacing between fish). Additionally spacing between fish became less variable post deployment of alarm cues. The only other significant effect was a decrease in the standard deviation of group expanse for fish subject to alarm cues.

The median and standard deviation of polarisation of direction of motion decreased for fish subject to alarm cues according to bootstrap analysis. Examination of the distributions of $R_{\text{motion}}(t)$, suggests in particular that after application of alarm cues, low levels (below 0.5) of polarisation in direction of motion were observed more frequently (fig. S11 C and F). The standard deviation of polarisation of facing direction also decreased from interval A to P for fish subject to alarm cues according to bootstrap analysis.

Overall the above analysis suggests that fish subject to alarm cues tended to form more compact groups, with smaller distances to their neighbours, and that group members tended to be less aligned in their direction of motion than under other observed circumstances (food or control treatments). However, there was no significant effect of alarm treatment on polarisation in the facing direction of group members. This is possibly be-

cause fish do not always travel in the direction that they are facing, and may imply that fish in alarm groups in particular engaged in more sideways and backwards swimming post application of cues (although we did not quantify such behaviour).

table S4. Basic 95% confidence intervals for test statistics derived from the median and SD distances of individual fish from the group centroid, $d_{i,c}(t)$, distances of individual fish to their nearest neighbour, $d_{i,nn}(t)$, the mean distance of all neighbours from individual fish, $d_{i,mn}(t)$, group expanse, $E(t)$, polarisation of the direction of motion of group members, $R_{\text{motion}}(t)$ and polarisation of the facing direction of group members, $R_{\text{facing}}(t)$. Statistically significant effects are marked with an asterisk (*) (ie. confidence intervals for the test statistics that lie entirely above or below 0). If a confidence interval lies entirely below zero, then the associated quantity decreased from interval A to P; if a confidence interval lies entirely above zero then the associated quantity increased from A to P.

Variable	Control	Food	Alarm
median $d_{i,c}(t)$ (mm)	$(-3.88, 4.37)$	$(-6.81, 9.04)$	$(-13.91, -3.45)^*$
std $d_{i,c}(t)$ (mm)	$(-1.94, 1.49)$	$(-7.41, 2.86)$	$(-6.40, -0.42)^*$
median $d_{i,nn}(t)$ (mm)	$(-1.93, 2.96)$	$(-2.28, 7.21)$	$(-8.42, -2.09)^*$
std $d_{i,nn}(t)$ (mm)	$(-1.15, 0.68)$	$(-1.96, 2.45)$	$(-4.78, -0.61)^*$
median $d_{i,mn}(t)$ (mm)	$(-5.11, 6.09)$	$(-5.87, 15.97)$	$(-20.03, -5.43)^*$
std $d_{i,mn}(t)$ (mm)	$(-1.83, 1.51)$	$(-10.48, 1.33)$	$(-6.64, 0.22)$
median $E(t)$ (mm)	$(-4.08, 4.46)$	$(-5.03, 10.73)$	$(-15.39, -3.82)^*$
std $E(t)$ (mm)	$(-1.21, 1.28)$	$(-10.92, -0.15)^*$	$(-6.19, 0.75)$
median $R_{\text{motion}}(t)$	$(-0.03, 0.04)$	$(-0.05, 0.07)$	$(-0.18, -0.01)^*$
std $R_{\text{motion}}(t)$	$(-0.04, 0.003)$	$(-0.04, 0.02)$	$(0.01, 0.08)^*$
median $R_{\text{facing}}(t)$	$(-0.03, 0.04)$	$(-0.05, 0.06)$	$(-0.16, 0.01)$
std $R_{\text{facing}}(t)$	$(-0.04, 0.01)$	$(-0.04, 0.02)$	$(0.001, 0.08)^*$

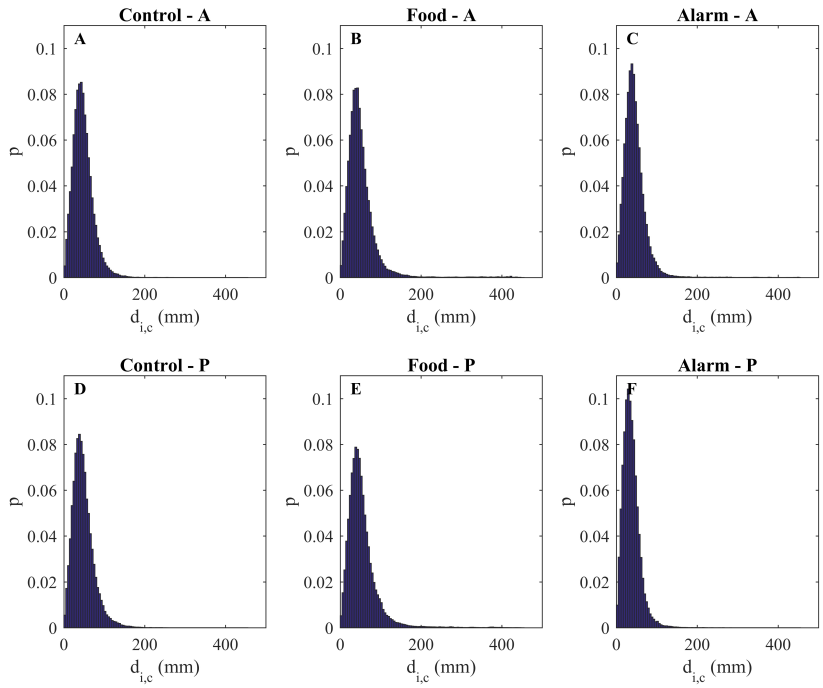

fig. S7. Observed distribution of the distance of individual fish from the group centroid (in mm) over the range 0 to 500 mm for the treatments and time intervals: A – control treatment ante cue, B – food treatment ante cue, C – alarm treatment ante cue, D – control treatment post cue, E – food treatment post cue, F – alarm treatment post cue. Details of the calculations used to generate these plots are given in section S1.3.

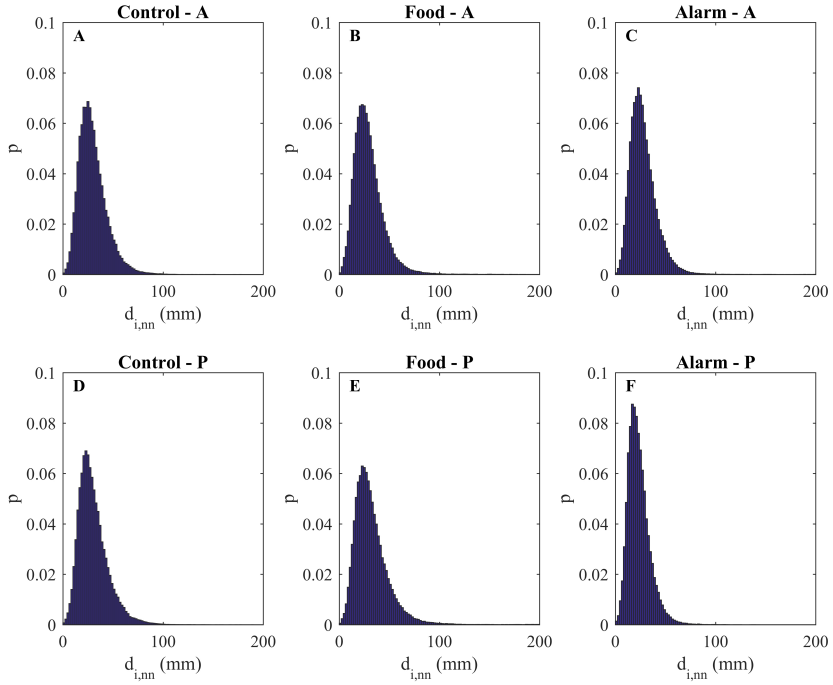

fig. S8. Observed distribution of the distance of individual fish from their nearest neighbour (in mm) over the range 0 to 200 mm for the treatments and time intervals: A – control treatment ante cue, B – food treatment ante cue, C – alarm treatment ante cue, D – control treatment post cue, E – food treatment post cue, F – alarm treatment post cue. Details of the calculations used to generate these plots are given in section S1.3.

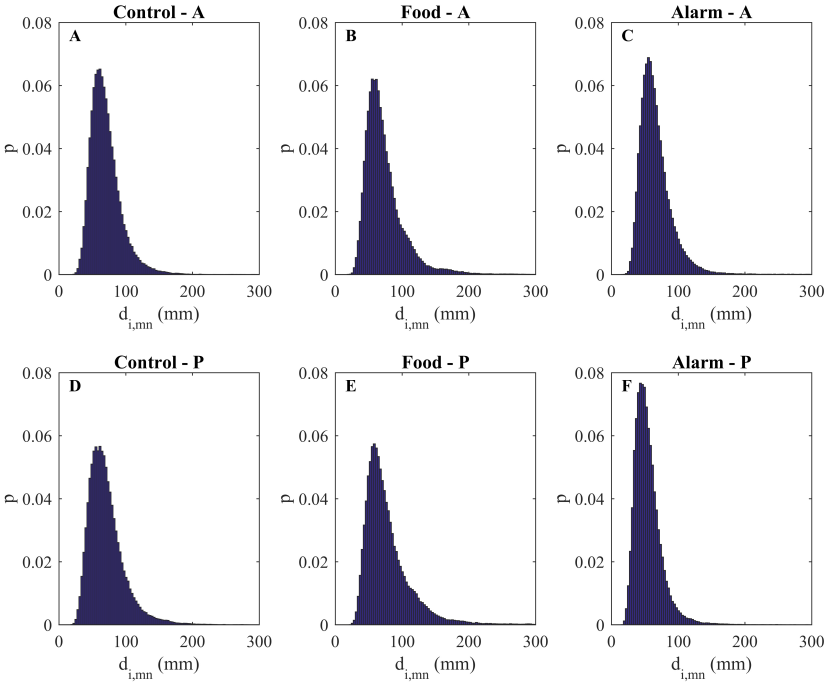

fig. S9. Observed distribution of the mean distance of individual fish from all their neighbours (in mm) over the range 0 to 300 mm for the treatments and time intervals: A – control treatment ante cue, B – food treatment ante cue, C – alarm treatment ante cue, D – control treatment post cue, E – food treatment post cue, F – alarm treatment post cue. Details of the calculations used to generate these plots are given in section S1.3.

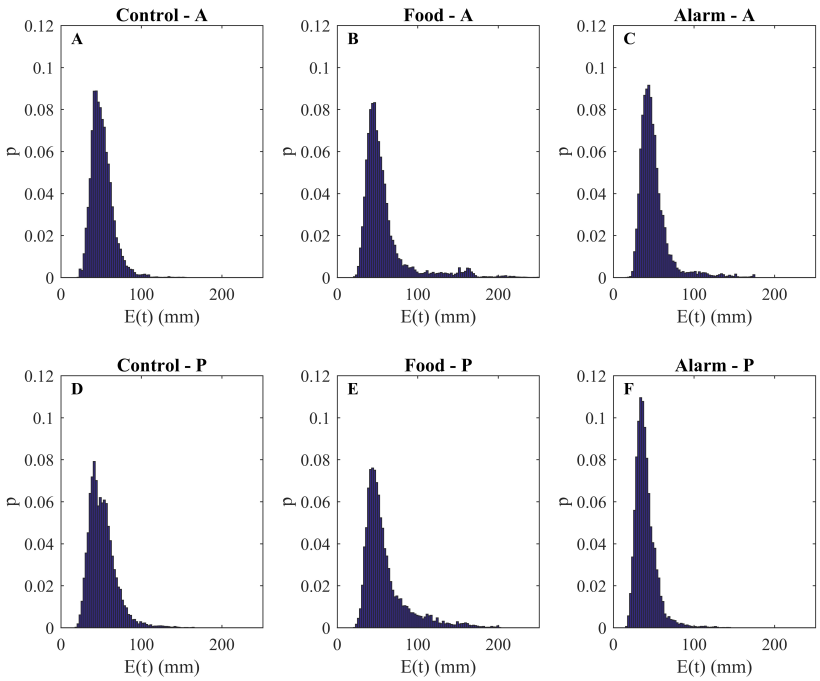

fig. S10. Observed distribution of group expanse over the range from 0 to 268.2 mm for the treatments and time intervals: A – control treatment ante cue, B – food treatment ante cue, C – alarm treatment ante cue, D – control treatment post cue, E – food treatment post cue, F – alarm treatment post cue. Details of the calculations used to generate these plots are given in section S1.3.

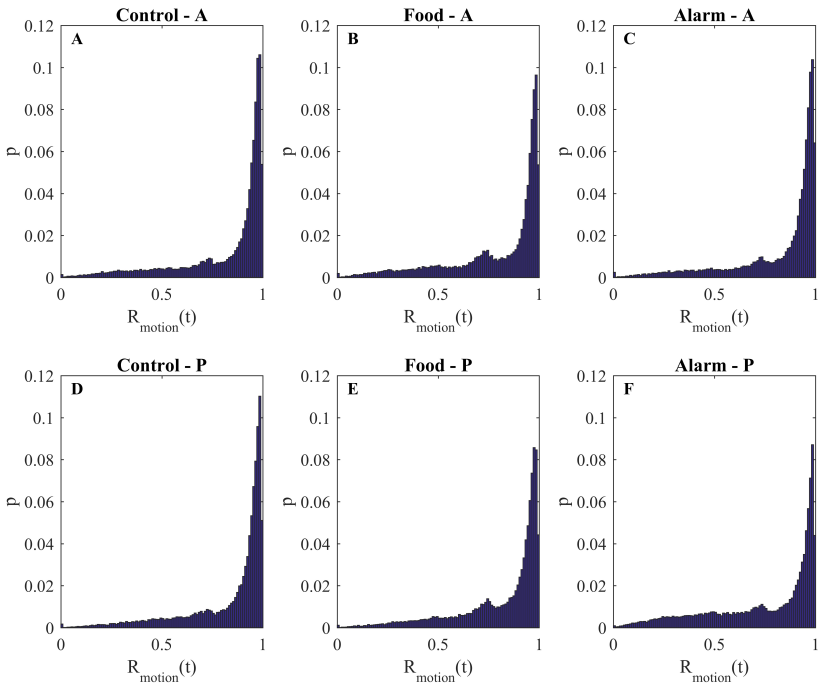

fig. S11. Observed distribution of polarisation based on direction of motion for the treatments and time intervals: A – control treatment ante cue, B – food treatment ante cue, C – alarm treatment ante cue, D – control treatment post cue, E – food treatment post cue, F – alarm treatment post cue. Details of the calculations used to generate these plots are given in section S1.3.

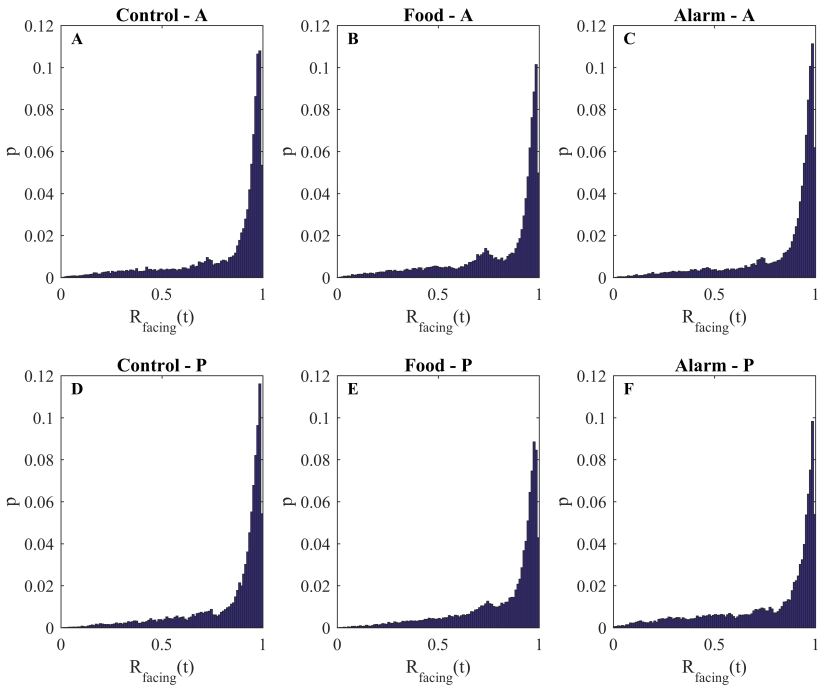

fig. S12. Observed distribution of polarisation based on the facing direction of fish for the treatments and time intervals: A – control treatment ante cue, B – food treatment ante cue, C – alarm treatment ante cue, D – control treatment post cue, E – food treatment post cue, F – alarm treatment post cue. Details of the calculations used to generate these plots are given in section S1.3.

S2.3 Basic interactions with arena boundaries

Figures S13 and S14 illustrate the observed distributions of individual distances from the inner and outer boundaries of the arena respectively. Table S5 summarises the results of bootstrap analysis of the median and standard deviation of each of these distance measures. The mean median distance between individual fish and the inner boundary of the arena decreased from interval A to P for fish in control groups and increased from interval A to P for fish subject to alarm cues. The mean standard deviation of individual distances from the inner boundary decreased from interval A to P for fish subject to alarm cues. Complimenting the statistically significant decrease in distance to the inner boundary of the arena observed for fish in control groups, the mean median distance between individual fish and the outer boundary increased from interval A to P for fish in control groups. However, there was no significant decrease in mean median distances to the outer wall for fish subject to alarm cues. The mean standard deviation of individual distances to the outer wall decreased for fish subject to alarm cues. Examination of the distributions of distances of individuals from the inner (fig. S13) and outer (fig. S14) walls further suggests that fish in alarm groups did indeed move further away from the inner boundary post deployment of cue, and tended to exhibit reduced spread in their distances from inner and outer boundaries.

table S5. Basic 95% confidence intervals for test statistics derived from the median and SD distances of individual fish from the inner, $d_{i,b_{\text{inner}}}(t)$, and outer, $d_{i,b_{\text{outer}}}(t)$, boundaries of the annular arena. Statistically significant effects are marked with an asterisk (*) (ie. confidence intervals for the test statistics that lie entirely above or below 0). If a confidence interval lies entirely below zero, then the associated quantity decreased from interval A to P; if a confidence interval lies entirely above zero then the associated quantity increased from A to P.

Variable	Control	Food	Alarm
median $d_{i,b_{\text{inner}}}(t)$ (mm)	$(-13.07, -0.95)^*$	$(-10.26, 4.66)$	$(0.46, 19.06)^*$
std $d_{i,b_{\text{inner}}}(t)$ (mm)	$(-2.15, 2.74)$	$(-0.54, 5.53)$	$(-7.74, -0.38)^*$
median $d_{i,b_{\text{outer}}}(t)$ (mm)	$(1.39, 12.63)^*$	$(-4.58, 10.05)$	$(-18.60, 0.24)$
std $d_{i,b_{\text{outer}}}(t)$ (mm)	$(-2.05, 2.96)$	$(-0.57, 5.54)$	$(-8.10, -0.30)^*$

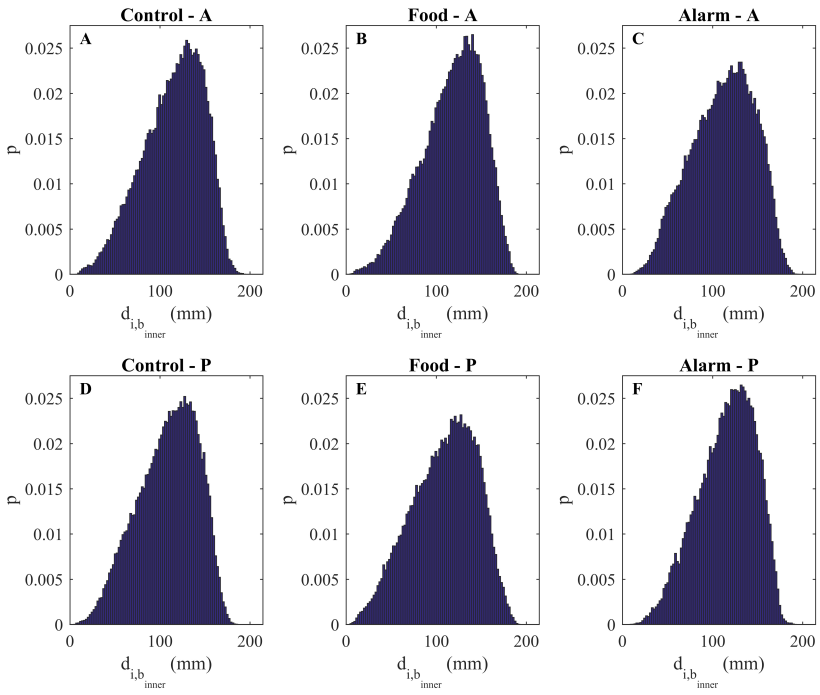

fig. S13. Observed distribution of the distance of individual fish from the inner circular boundary (in mm) over the range 0 to 210.9 mm for the treatments and time intervals: A – control treatment ante cue, B – food treatment ante cue, C – alarm treatment ante cue, D – control treatment post cue, E – food treatment post cue, F – alarm treatment post cue. Details of the calculations used to generate these plots are given in section S1.4.

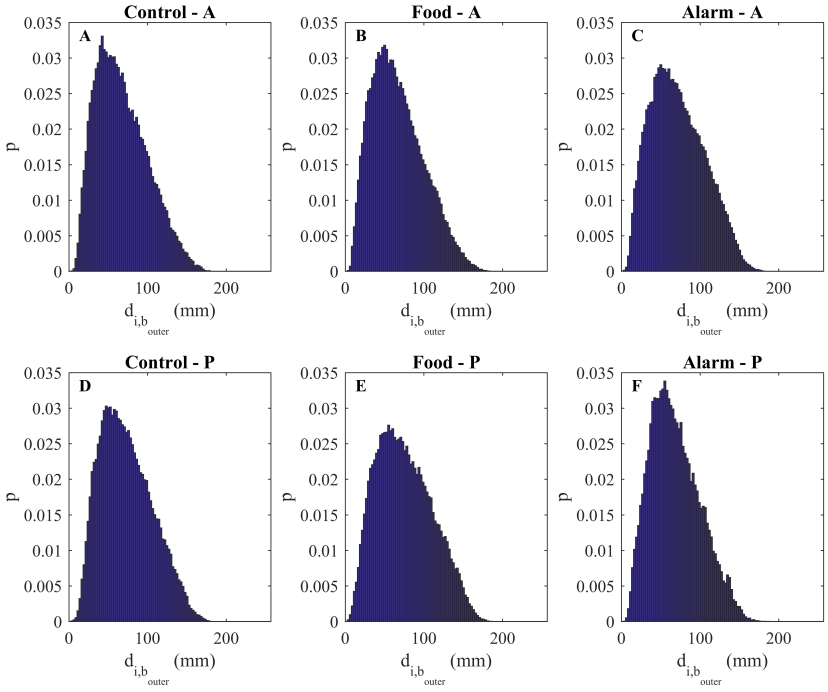

fig. S14. Observed distribution of the distance of individual fish from the outer circular boundary (in mm) over the range 0 to 274.6 mm for the treatments and time intervals: A – control treatment ante cue, B – food treatment ante cue, C – alarm treatment ante cue, D – control treatment post cue, E – food treatment post cue, F – alarm treatment post cue. Details of the calculations used to generate these plots are given in section S1.4.

S2.4 Speed, relative alignment and rules of interaction as a function of relative partner displacement and focal fish speed

The results section and Figs. 1 to 7 in the main text detail how introduction of food or alarm cues affected the relative frequency that group members were observed by focal fish in different relative locations, as well as the speed, relative alignment, change in speed over time and change in angle of motion over time of focal fish as a funtion of the relative positions occupied by all other group members.

Figure S15 illustrates the probability of observing partner fish at given x- or y-coordinates when focal fish were swimming at 0 to 20 mm/s, along with the mean change in speed over time and mean change in angle over time of focal fish swimming at 0 to 20 mm/s as a function of the relative x- or y-coordinates of partner fish. Figures S16 to S19 illustrate via heat maps the probability of observing partner fish at given (x, y) coordinates, the relative orientation of partner fish at given (x, y) coordinates, the mean change in speed over time and the mean change in angle of motion over time of focal fish as a function of the relative (x, y) coordinates of partner fish with focal fish swimming between 0 and 20 mm/s. Figures S20 to S24, and S25 to S29 illustrate the same quantities as in figs. S15 to S19, but with focal fish swimming between 60 and 80 mm/s or between 120 and 140 mm/s respectively. We chose the above speed divisions to report in detail, as they were reasonable representations of the mean behaviour of individuals travelling at lower, medium and higher speeds relative to the range of speeds for which we determined rules of interaction as a function of relative partner location and focal fish speed.

As expected due to the general decrease in speed exhibited by fish in alarm treatment groups, the probability of observing partners at given (x, y) coordinates while focal fish travelled at between 0 and 20 mm/s increased markedly post deployment of the alarm cue (fig. S15 A & B, fig. S16 E & F). The probability of observing control group partner fish at given (x, y) coordinates with focal fish travelling at 0 to 20 mm/s also increased, but not to the extent observed for alarm group fish (fig. S15 A & B, fig. S16 A & B). There was a decrease in the associated probabilities for fish subject to food cues, driven by the increase in speed observed for these fish. There tended to be greater variation in the relative directions of motion of partner fish when focal fish travelled at lower speeds (evidenced by lower values of R in fig. S17), but the mean directions of motion of partner fish still tended to be similar to that of focal fish, particularly for partners in the range $-50 \leq y \leq 50$. Focal fish tended to increase their speed whilst in the lowest speed range, irrespective of the relative location of their partners (fig. S15 C & D and fig. S18). The magnitude of the increase in speed of focal fish travelling at between 0 and 20 mm/s as a function of relative partner location was least for fish subject to alarm cues. Further, at low speeds the potential zone of repulsion where fish increased speed when partners were close behind, or decreased speed when partners were close in front was not evident (contrast fig. S18 with Fig. 2). At low speed, fish tended to turn towards their partners (fig. S15 F and fig. S19). The largest magnitudes of the mean changes in angle of motion associated with these turns were again observed for fish subject to alarm cues (fig. S15 F and fig. S19 F).

The probability of observing partner fish at given (x, y) coordinates when focal fish were travelling between 60 and 80 mm/s diminished for fish subject to alarm cues (fig. S20 A & B). This reduced probability was consistent with the tendency for fish subject to alarm cues to swim at lower speeds. There was less variation in the relative directions of motion of partner fish when focal fish were travelling between 60 and 80 mm/s (fig. S22) compared to when fish travelled at low speeds (0 to 20 mm/s, see fig. S17). The repulsion like zone evident in Fig. 2 was apparent when focal fish swam at between 60 and 80 mm/s (fig. S23). Fish subject to food cues tended to exhibit greater, and positive, mean changes in speed over time when their partners were located outside their potential zone of repulsion (fig. S20 C & D and fig. S23 C & D). Only fish subject to alarm cues had a clear region of relative partner locations where focal fish travelling between 60 and 80 mm/s would slow down on average, for $x < -30$ mm approximately, outside of the potential zone of repulsion (fig. S20 C and fig. S23 F). On average, focal fish travelling at 60 to 80 mm/s tended to turn towards their partners, except for in a small region near focal fish (approximately $-30 \leq x \leq 30$, $-10 \leq y \leq 10$ mm) where the focal fish tended to turn away from their partner (another short range repulsive effect), (fig. S24). The magnitudes of the change in angle of motion over time associated with turns towards partners tended to be greatest for fish subject to alarm cues and travelling between 60 and 80 mm/s, as was the case for fish swimming at low speeds (fig. S20 F and fig. S24 F).

The probability of observing partner fish at given x or y coordinates when focal fish were travelling at 120 to 140 mm/s increased slightly for fish subject to food cues, and diminished for fish subject to alarm cues (fig. S25 A & B). These changes in probability were consistent with the general reduction in speed of fish subject to alarm cues. At higher speeds of focal fish, the variation in relative directions of motion of partner fish further reduced as compared to when focal fish were travelling at 60 to 80 mm/s (fig. S27). There may have been greater scatter in the relative directions of motion of partner fish for fish subject to alarm cues at higher speeds of focal fish (fig. S27 F); such an observation is consistent with the reduction in polarisation in direction of motion suggested by our bootstrap analysis. Local repulsion zones remained evident in heat plots of the mean change in speed over time of focal individuals that were travelling at 120 to 140 mm/s, except for fish subject to alarm cues where a general tendency to slow down when partner fish were located behind seemed to overwhelm any small region where focal fish would speed up if their partner was close and directly behind them (fig. S28). At high speeds, focal fish subject to alarm cues generally exhibited changes in speed over time below those of other fish (fig. S25 C & D), and these changes in speed tended to be negative over a larger region than for fish subject to other treatments (fig. S28). As at lower speeds, focal fish travelling at high speeds still tended to turn away from very near partners (faintly visible) and towards partners at greater distances (fig. S25 F and fig. S29). The magnitude of the mean change in angle of motion over time of fish subject to alarm cues was still greatest at higher speeds, compared to fish subject to other cues (fig. S25 F and fig. S29 F).

Ultimately the analysis in this section suggests that fish subject to food cues tended to travel at higher speeds on average and that fish subject to alarm cues tended to travel at lower speeds on average. Our bootstrap analysis also suggested a decrease in speed for

alarmed fish, but there was no significant change in mean median or standard deviation of speed for fish in food groups. There was also some evidence that fish subject to alarm cues tended to be less aligned with their partners, especially when focal fish were travelling at relative high speeds, which further supports the results of our analysis of median group polarisation that suggested a reduction in polarisation for alarm treatment groups. Further, plots of the relative frequency that partner fish were located at given relative (x, y) coordinates suggested that fish subject to food cues tended to spread out, whereas fish subject to alarm cues reduced the distance to their group mates. These results relating to relative partner locations support our bootstrap analysis of median individual distances to the group centroid, median nearest neighbour distances, median mean neighbour distances and median group expanse – all of which suggested that fish subject to alarm cues bunched up. Again, there were no significant changes in median measures of spacing for fish in food treatment groups according to bootstrap analysis though, so any trend might be more subtle.

Finally, we include plots of the relative frequency of observations of neighbours and speed of focal individuals along the bins centered on $y = 0$ or $x = 0$ in fig. S30, the change in speed over time and change in angle of motion over time of focal individuals along the bins centred on $y = 0$ or $x = 0$ in fig. S31, and plots of relative observation frequency of neighbours, changes in speed and changes in heading along the bins centred on $y = 0$ or $x = 0$ for focal fish speeds of 0 to 20 mm/s, 60 to 80 mm/s or 120 to 140 mm/s in figs. S32 to S34 respectively. (These plots are cross sections of the corresponding heat maps for each quantity along the lines $y = 0$ or $x = 0$.)

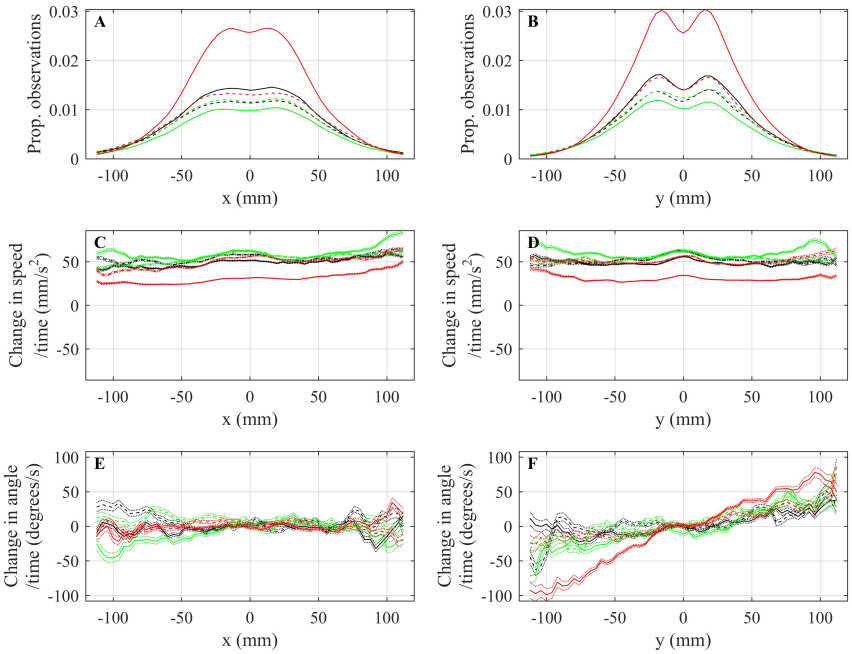

fig. S15. Relative frequency of observations of partner fish (A and B), mean change in speed over time of focal fish (C and D) and mean change in angle of motion over time of focal fish (E and F) as a function of the relative x- (A, C and E) or y-coordinates (B, D, and F) of partner fish for focal fish travelling with speeds between 0 and 20 mm/s. Curves corresponding to control trials are plotted in black, curves corresponding to food treatments are plotted in green and curves corresponding to alarm cue treatments are plotted in red. Data before application of a cue is plotted as a dashed line; post cue data is plotted as solid lines. Dotted lines are plotted one standard error above and below each curve. Details of the generation of these plots are given in section S1.5.

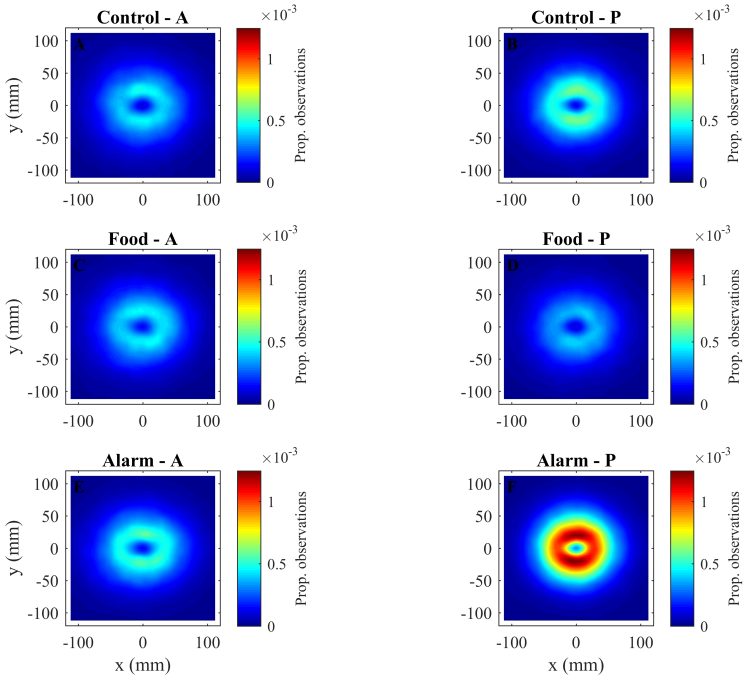

fig. S16. Relative frequency of observations of partner fish before (A, C and E) and after (B, D and F) application of cues for focal fish located at the origin travelling at 0 to 20 mm/s, subject to control (A and B), food (C and D) or alarm (E and F) treatments. The direction of motion of focal fish in each plot is parallel to the positive x-axis. Details of the generation of these plots are given in section S1.5.

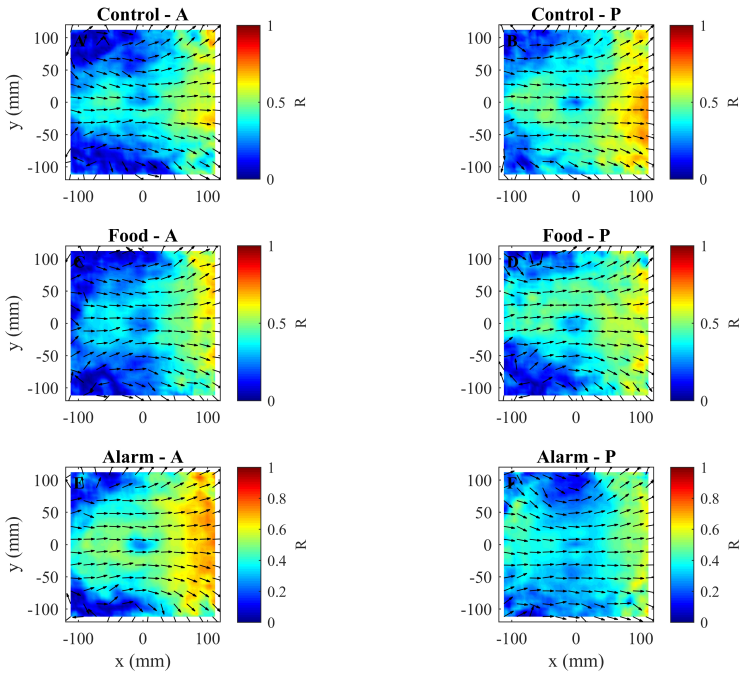

fig. S17. The mean direction of motion of partner fish (indicated by arrows) relative to focal fish travelling at 0 to 20 mm/s located at the origin before (A, C and E) and after (B, D and F) application of cues for focal fish located at the origin, for fish subject to control (A and B), food (C and D) or alarm (E and F) treatments. The direction of motion of focal fish in each plot is parallel to the positive x-axis. The heat portion of these plots indicates the polarisation of angles contained in each bin, R. Details of the generation of these plots are given in section S1.5.

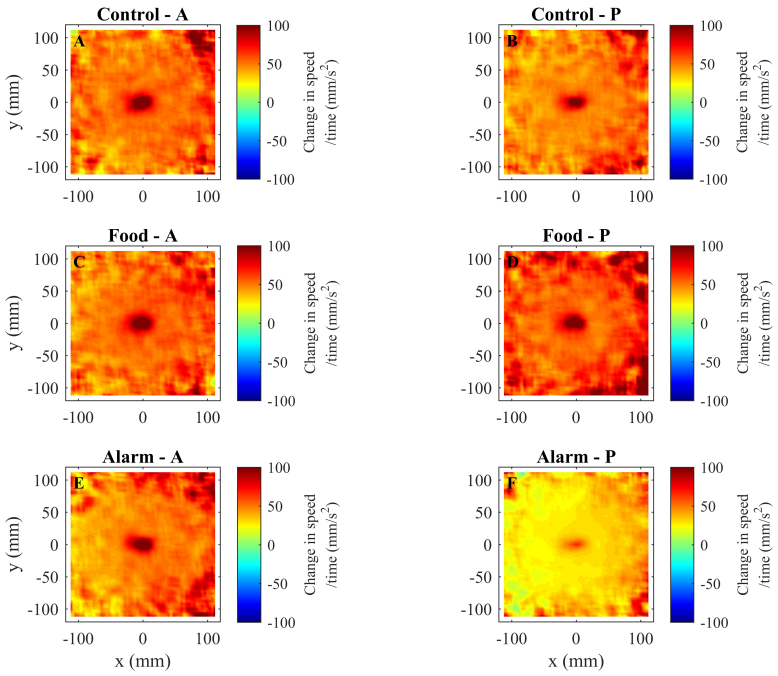

fig. S18. Mean change in speed over time of focal fish travelling at 0 to 20 mm/s as a function of relative partner location before (A, C and E) and after (B, D and F) application of cues for focal fish located at the origin, for fish subject to control (A and B), food (C and D) or alarm (E and F) treatments. The direction of motion of focal fish in each plot is parallel to the positive x-axis. Extreme changes in speed have been truncated at ± 100 mm/s^2 in these plots for visualisation purposes. Details of the generation of these plots are given in section S1.5.

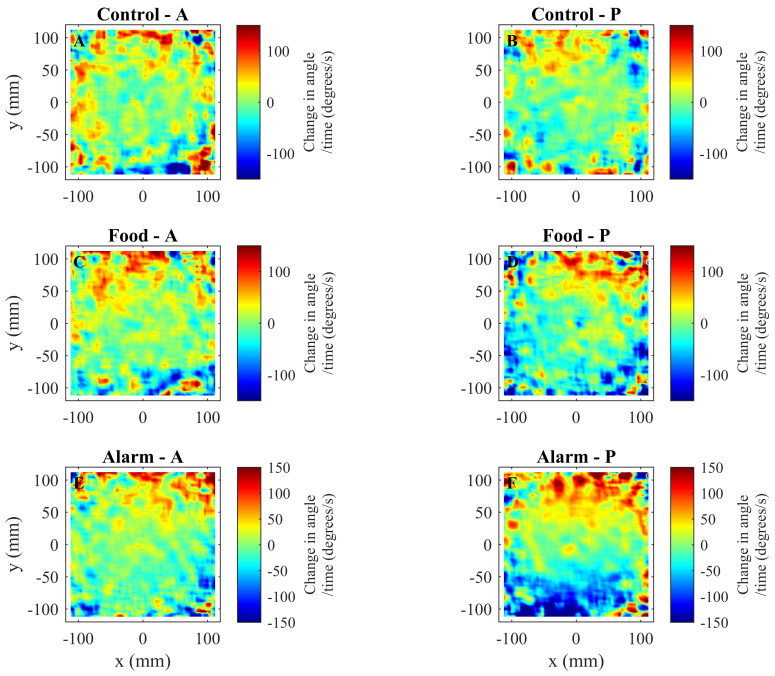

fig. S19. Mean change in angle of motion over time of focal fish travelling at 0 to 20 mm/s as a function of relative partner location before (A, C and E) and after (B, D and F) application of cues for focal fish located at the origin, for fish subject to control (A and B), food (C and D) or alarm (E and F) treatments. The direction of motion of focal fish in each plot is parallel to the positive x-axis. Extreme changes in angle have been truncated at ± 150 degrees/s in these plots for visualisation purposes. Details of the generation of these plots are given in section S1.5.

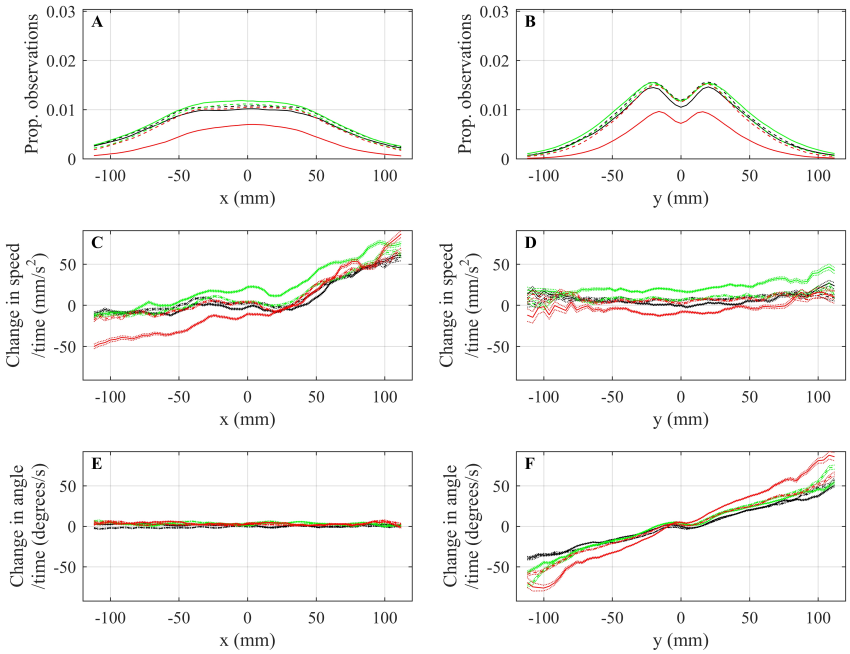

fig. S20. Relative frequency of observations of partner fish (A and B), mean change in speed over time of focal fish (C and D) and mean change in angle of motion over time of focal fish (E and F) as a function of the relative x- (A, C and E) or y-coordinates (B, D, and F) of partner fish for focal fish travelling with speeds between 60 and 80 mm/s. Curves corresponding to control trials are plotted in black, curves corresponding to food treatments are plotted in green and curves corresponding to alarm cue treatments are plotted in red. Data before application of a cue is plotted as a dashed line; post cue data is plotted as solid lines. Dotted lines are plotted one standard error above and below each curve. Details of the generation of these plots are given in section S1.5.

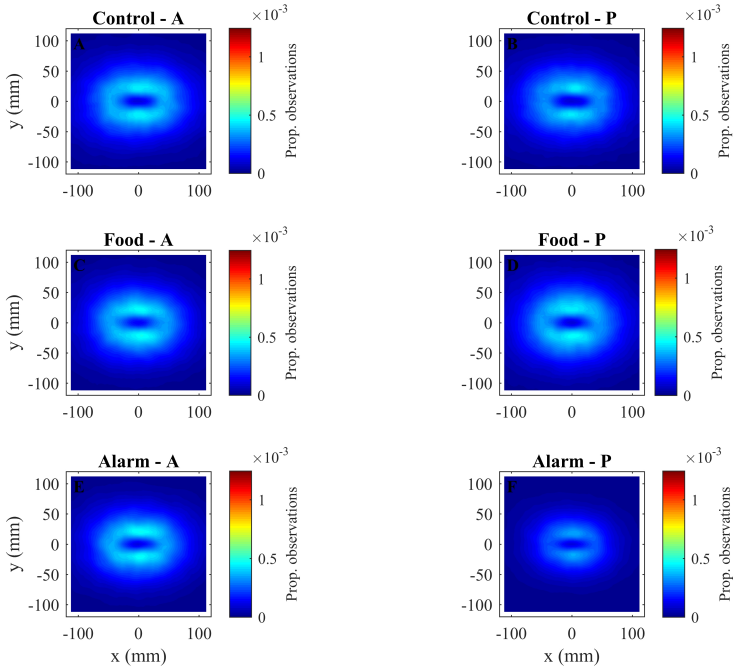

fig. S21. Relative frequency of observations of partner fish before (A, C and E) and after (B, D and F) application of cues for focal fish located at the origin travelling at 60 to 80 mm/s, subject to control (A and B), food (C and D) or alarm (E and F) treatments. The direction of motion of focal fish in each plot is parallel to the positive x-axis. Details of the generation of these plots are given in section S1.5.

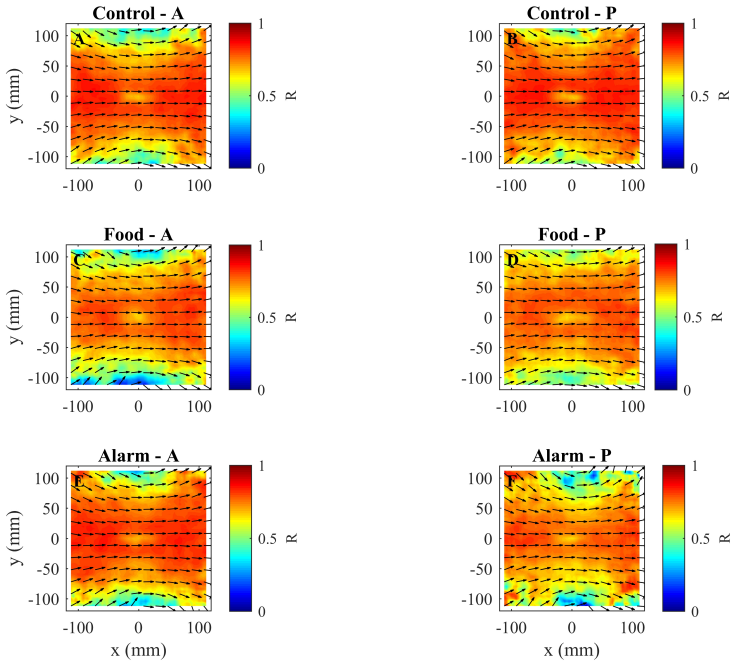

fig. S22. The mean direction of motion of partner fish (indicated by arrows) relative to focal fish travelling at 60 to 80 mm/s located at the origin before (A, C and E) and after (B, D and F) application of cues for focal fish located at the origin, for fish subject to control (A and B), food (C and D) or alarm (E and F) treatments. The direction of motion of focal fish in each plot is parallel to the positive x-axis. The heat portion of these plots indicates the polarisation of angles contained in each bin, R. Details of the generation of these plots are given in section S1.5.

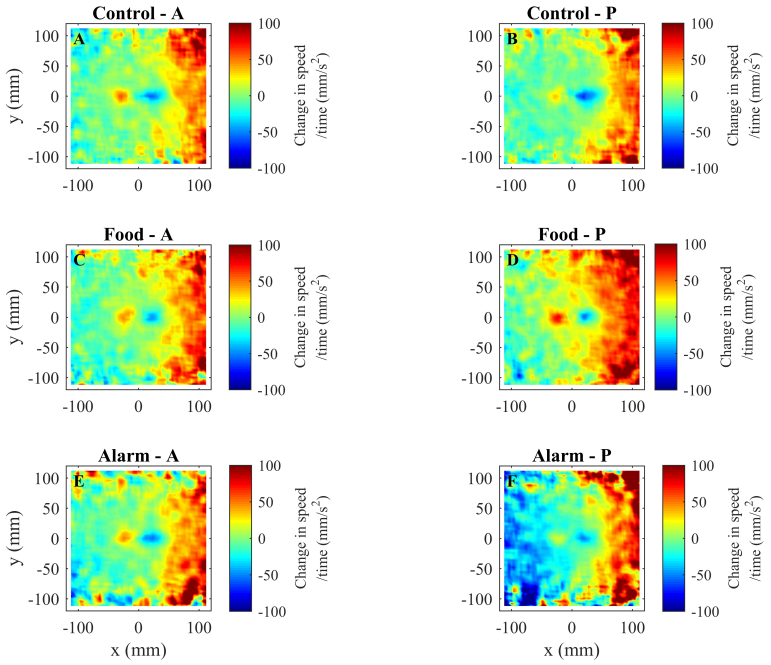

fig. S23. Mean change in speed over time of focal fish travelling at 60 to 80 mm/s as a function of relative partner location before (A, C and E) and after (B, D and F) application of cues for focal fish located at the origin, for fish subject to control (A and B), food (C and D) or alarm (E and F) treatments. The direction of motion of focal fish in each plot is parallel to the positive x-axis. Extreme changes in speed have been truncated at ± 100 mm/s^2 in these plots for visualisation purposes. Details of the generation of these plots are given in section S1.5.

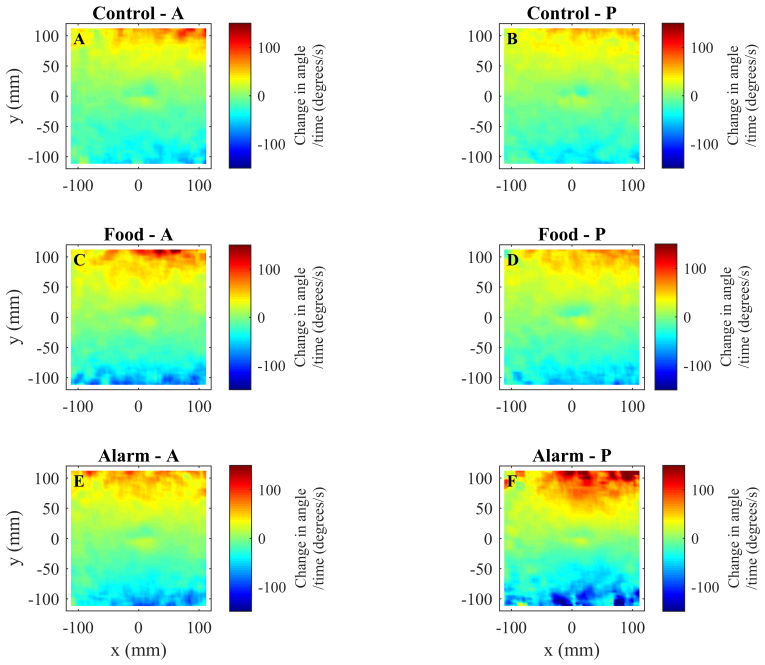

fig. S24. Mean change in angle of motion over time of focal fish travelling at 60 to 80 mm/s as a function of relative partner location before (A, C and E) and after (B, D and F) application of cues for focal fish located at the origin, for fish subject to control (A and B), food (C and D) or alarm (E and F) treatments. The direction of motion of focal fish in each plot is parallel to the positive x-axis. Extreme changes in angle have been truncated at ± 150 degrees/s in these plots for visualisation purposes. Details of the generation of these plots are given in section S1.5.

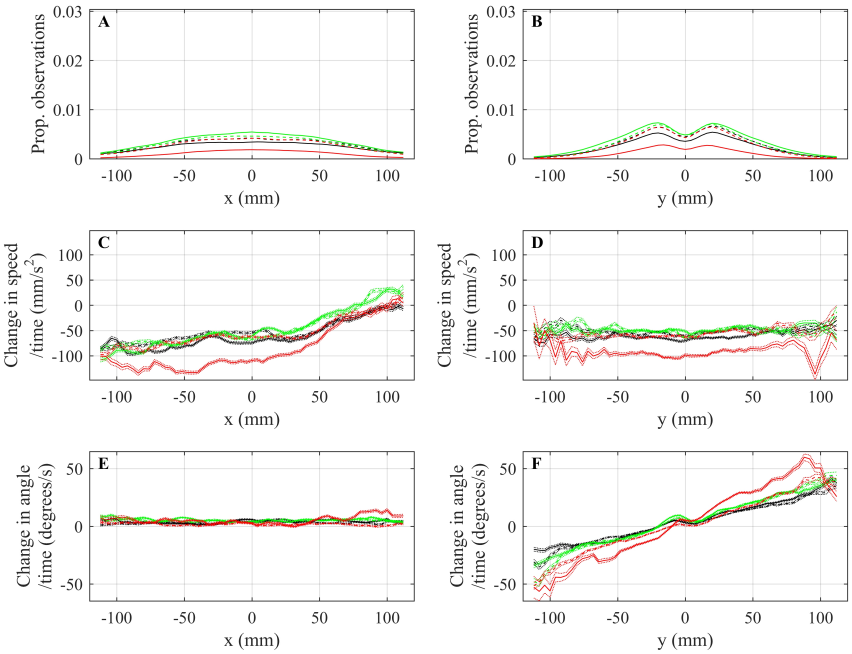

fig. S25. Relative frequency of observations of partner fish (A and B), mean change in speed over time of focal fish (C and D) and mean change in angle of motion over time of focal fish (E and F) as a function of the relative x- (A, C and E) or y-coordinates (B, D, and F) of partner fish for focal fish travelling with speeds between 120 and 140 mm/s. Curves corresponding to control trials are plotted in black, curves corresponding to food treatments are plotted in green and curves corresponding to alarm cue treatments are plotted in red. Data before application of a cue is plotted as a dashed line; post cue data is plotted as solid lines. Dotted lines are plotted one standard error above and below each curve. Details of the generation of these plots are given in section S1.5.

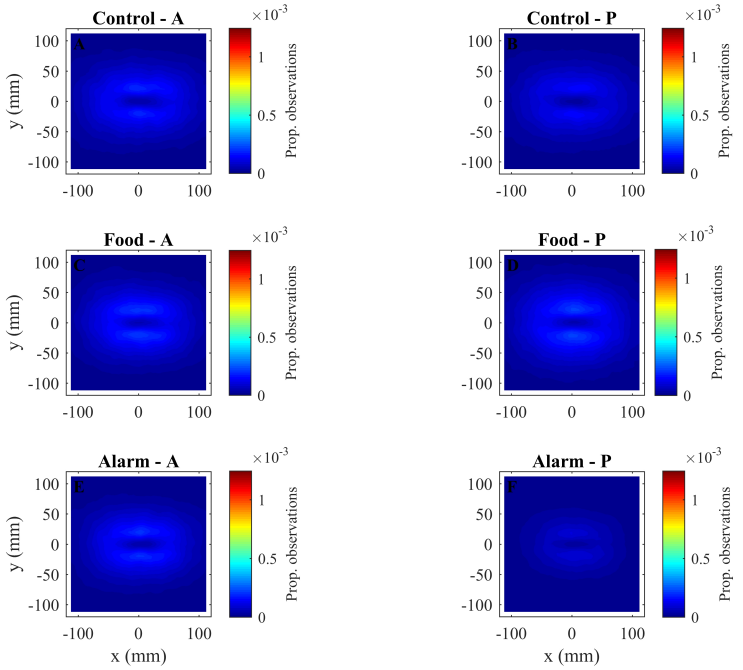

fig. S26. Relative frequency of observations of partner fish before (A, C and E) and after (B, D and F) application of cues for focal fish located at the origin travelling at 120 to 140 mm/s, subject to control (A and B), food (C and D) or alarm (E and F) treatments. The direction of motion of focal fish in each plot is parallel to the positive x-axis. Details of the generation of these plots are given in section S1.5.

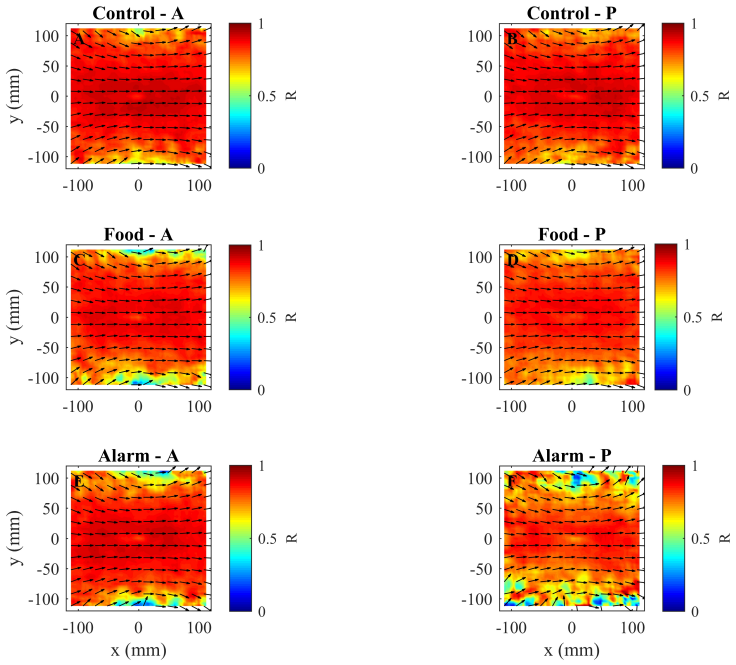

fig. S27. The mean direction of motion of partner fish (indicated by arrows) relative to focal fish travelling at 120 to 140 mm/s located at the origin before (A, C and E) and after (B, D and F) application of cues for focal fish located at the origin, for fish subject to control (A and B), food (C and D) or alarm (E and F) treatments. The direction of motion of focal fish in each plot is parallel to the positive x-axis. The heat portion of these plots indicates the polarisation of angles contained in each bin, R. Details of the generation of these plots are given in section S1.5.

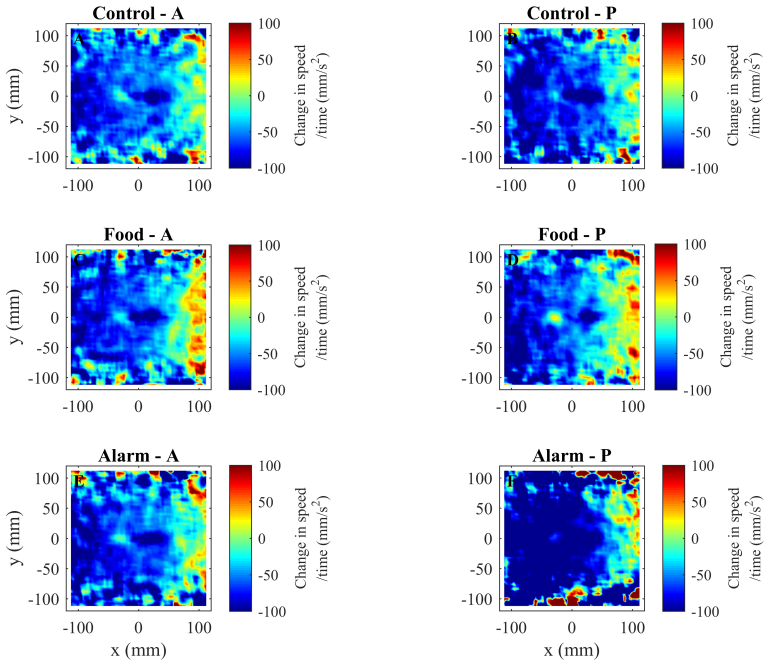

fig. S28. Mean change in speed over time of focal fish travelling at 120 to 140 mm/s as a function of relative partner location before (A, C and E) and after (B, D and F) application of cues for focal fish located at the origin, for fish subject to control (A and B), food (C and D) or alarm (E and F) treatments. The direction of motion of focal fish in each plot is parallel to the positive x-axis. Extreme changes in speed have been truncated at ± 100 mm/s^2 in these plots for visualisation purposes. Details of the generation of these plots are given in section S1.5.

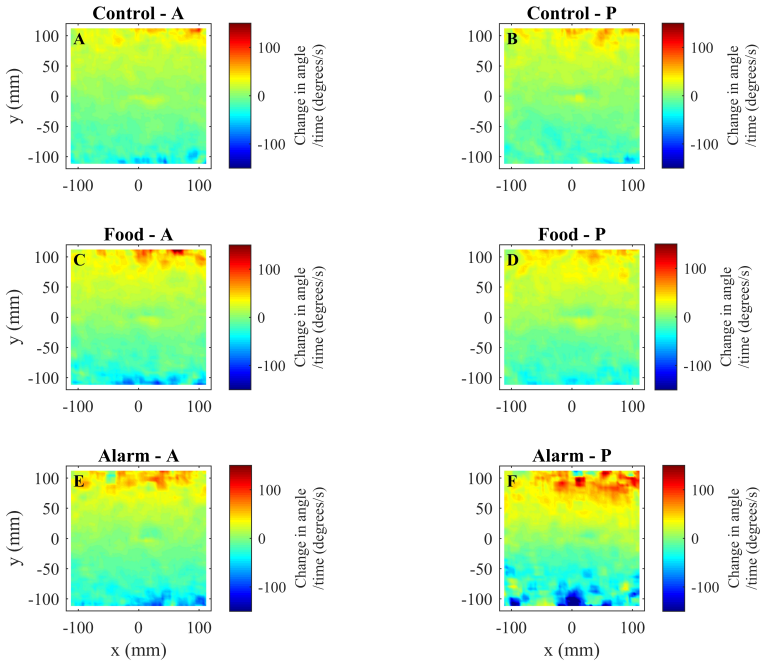

fig. S29. Mean change in angle of motion over time of focal fish travelling at 120 to 140 mm/s as a function of relative partner location before (A, C and E) and after (B, D and F) application of cues for focal fish located at the origin, for fish subject to control (A and B), food (C and D) or alarm (E and F) treatments. The direction of motion of focal fish in each plot is parallel to the positive x-axis. Extreme changes in angle have been truncated at ± 150 degrees/s in these plots for visualisation purposes. Details of the generation of these plots are given in section S1.5.

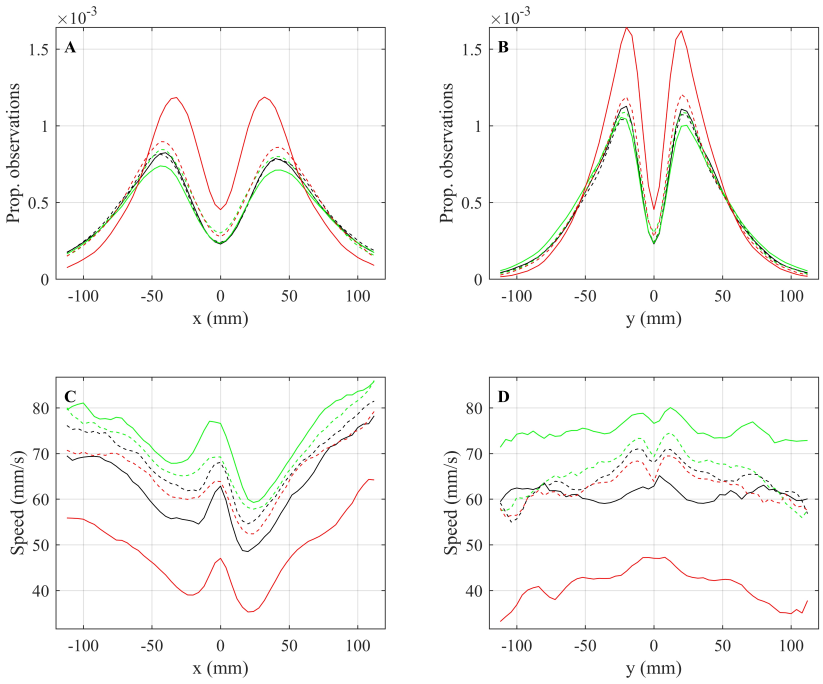

fig. S30. Relative frequency that partner fish were observed in bins centred about the x-axis (A) or y-axis (B) relative to focal fish located at the origin, and the mean speed of focal fish in bins centred about the x-axis (C) or y-axis (D). Curves corresponding to control trials are plotted in black, curves corresponding to food treatments are plotted in green and curves corresponding to alarm cue treatments are plotted in red. Data before application of a cue is plotted as a dashed line; post cue data is plotted as solid lines. Details of the generation of these plots are given in section S1.5.

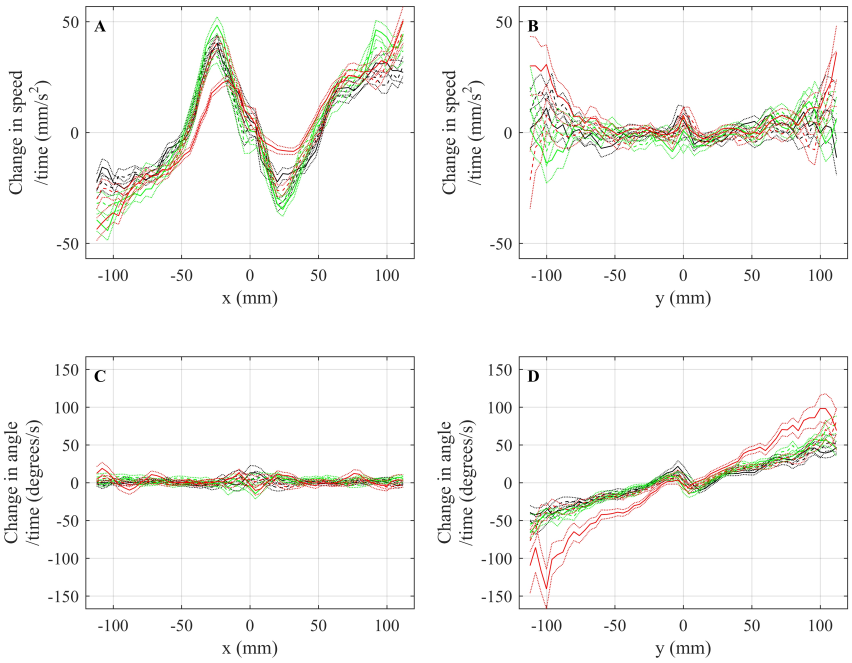

fig. S31. Mean change in speed over time (A and B) and mean change in angle of motion over time (turning speed) (C and D) of focal fish (located at the origin, moving parallel to the positive x-axis) as a function of the relative x-coordinates (A and C) or y-coordinates (B and D) of partner fish for bins centred about the x-axis (A and B) or y-axis (C and D). Curves corresponding to control trials are plotted in black, curves corresponding to food treatments are plotted in green and curves corresponding to alarm cue treatments are plotted in red. Data before application of a cue is plotted as a dashed line; post cue data is plotted as solid lines. Dotted lines are plotted one standard error above and below each curve. Details of the generation of these plots are given in section S1.5.

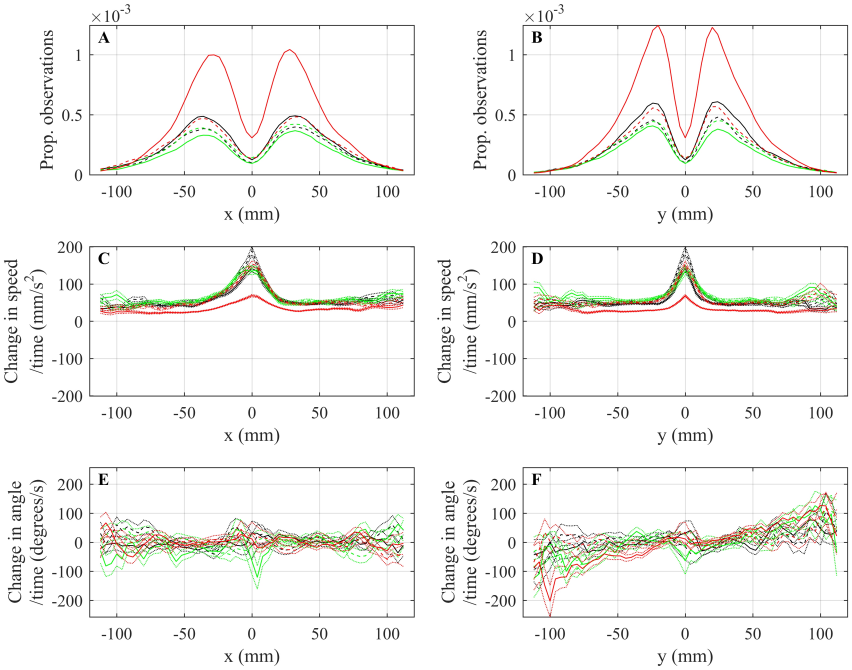

fig. S32. Relative frequency of observations of partner fish (A and B), mean change in speed over time of focal fish (C and D) and mean change in angle of motion over time of focal fish (E and F) as a function of the relative x- (A, C and E) or y-coordinates (B, D, and F) of partner fish for focal fish travelling with speeds between 0 and 20 mm/s, and bins centred about the x-axis (A, C and E) or the y-axis (B, D and F). Curves corresponding to control trials are plotted in black, curves corresponding to food treatments are plotted in green and curves corresponding to alarm cue treatments are plotted in red. Data before application of a cue is plotted as a dashed line; post cue data is plotted as solid lines. Dotted lines are plotted one standard error above and below each curve. Details of the generation of these plots are given in section S1.5.

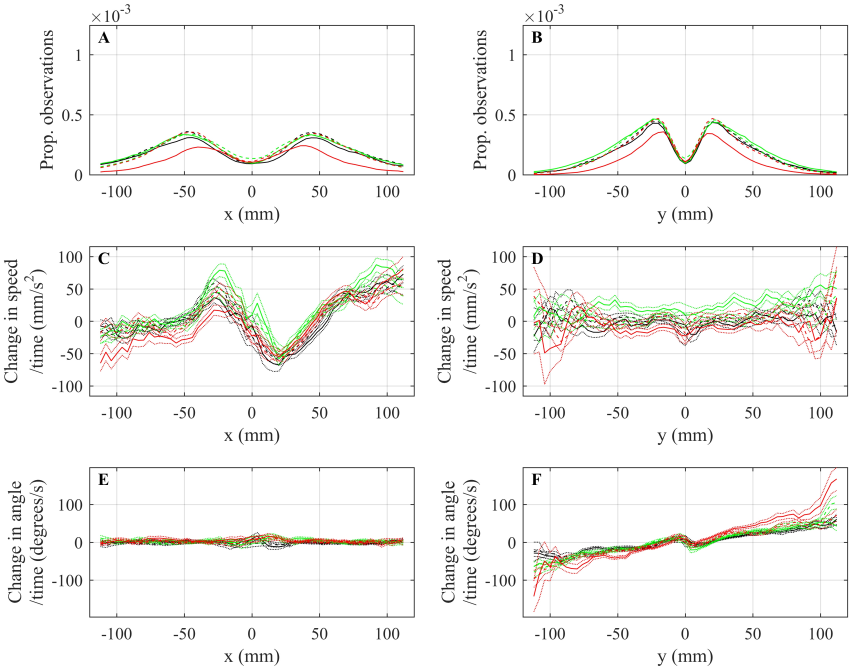

fig. S33. Relative frequency of observations of partner fish (A and B), mean change in speed over time of focal fish (C and D) and mean change in angle of motion over time of focal fish (E and F) as a function of the relative x- (A, C and E) or y-coordinates (B, D, and F) of partner fish for focal fish travelling with speeds between 60 and 180 mm/s, and bins centred about the x-axis (A, C and E) or the y-axis (B, D and F). Curves corresponding to control trials are plotted in black, curves corresponding to food treatments are plotted in green and curves corresponding to alarm cue treatments are plotted in red. Data before application of a cue is plotted as a dashed line; post cue data is plotted as solid lines. Dotted lines are plotted one standard error above and below each curve. Details of the generation of these plots are given in section S1.5.

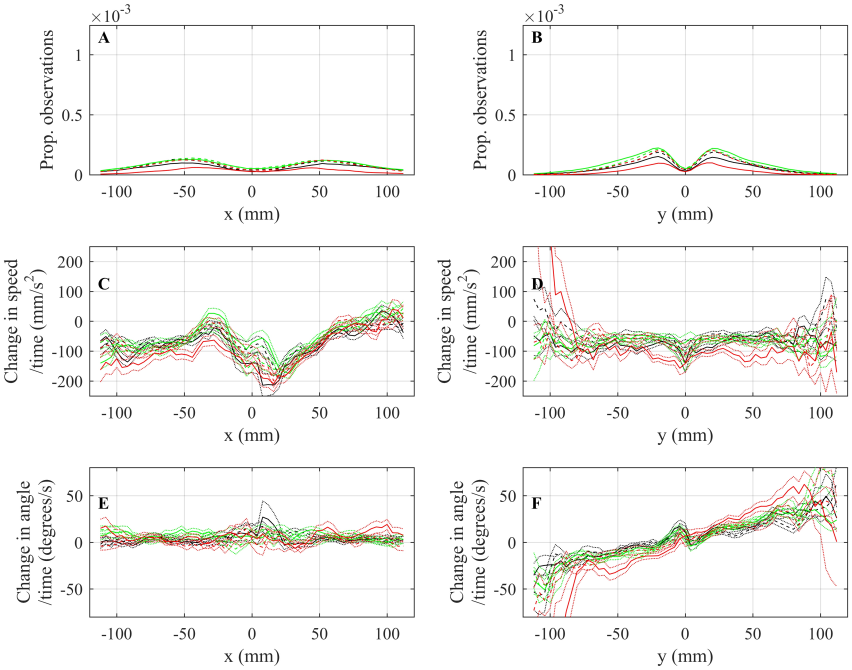

fig. S34. Relative frequency of observations of partner fish (A and B), mean change in speed over time of focal fish (C and D) and mean change in angle of motion over time of focal fish (E and F) as a function of the relative x- (A, C and E) or y-coordinates (B, D, and F) of partner fish for focal fish travelling with speeds between 120 and 140 mm/s, and bins centred about the x-axis (A, C and E) or the y-axis (B, D and F). Curves corresponding to control trials are plotted in black, curves corresponding to food treatments are plotted in green and curves corresponding to alarm cue treatments are plotted in red. Data before application of a cue is plotted as a dashed line; post cue data is plotted as solid lines. Dotted lines are plotted one standard error above and below each curve. Details of the generation of these plots are given in section S1.5.

S2.5 Predictability of changes in displacement and velocity

There were limited variations in the results of our calculations for parameter values within the range of small ($K \in \{0.7, 0.8, 0.9\}$), medium ($K \in \{7, 8, 9\}$) and large ($K \in \{17, 18, 19\}$) thresholds for changes in displacement or velocity. There was also limited sensitivity of our results to the specific value of $L \in \{19, 20, 21\}$. As our results were not particularly sensitive to small perturbations in K or L, we discuss the results for $(K, L) = (0.8, 20)$, $(K, L) = (8, 20)$ and $(K, L) = (18, 20)$ in detail below as typical examples of the results observed across all parameters within small, medium or large threshold groups respectively.
Entropy related to small ($K = 0.8$) changes in displacement increased from ante to post cue time intervals for fish in all treatments (fig. S35 A). The most marked increase in entropy at the small threshold occur
Information given by the current state about the next change in displacement decreased for fish subject fish subject to food and alarm cues, whereas there was an increase in this mutual information from ante to post cue time intervals for fish in control groups for both small and medium thresholds (fig. S35 D & E). Mutual information decreased post cue for all three treatments at large thresholds (fig. S35 F), with the largest decrease observed for alarmed fish. For all thresholds and treatments, the information about the next change in displacement given by the current change in displacement was in the vicinity of 1 bit – only sufficient information to unambiguously distinguish between two states.
Entropy rate related to small changes in displacement increased for fish in control and alarm treatment groups, whereas there was a decrease in entropy rate for fish subject to food cues (fig. S35 G). The largest magnitude change in entropy rate occur
Entropy related to small changes in velocity increased post cue for fish in control and alarm groups, but decreased slightly for fish in food groups (fig. S36 A). The magnitude of change in entropy post cue was largest for fish in alarm groups at the small threshold. Entropy decreased post cue for fish subject to food and alarm treatments at medium thresholds for changes in velocity, whereas entropy increased slightly for fish in control groups (fig. S36 B). Entropy related to large changes in velocity decreased most markedly for fish subject to alarm cues; there was also a decrease in entropy for control group fish (fig. S36 C). Entropy related to large changes in velocity increased for fish subject to food cues.
Little information was given by the change in velocity at sample time t about the change in velocity at the next sample time $t + L\Delta t$ at small, medium and large thresholds for fish in all three treatments (less than 0.15 bits in all cases, see fig. S36 D, E & F). Mutual information increased post cue for fish subject to alarm cues at all three thresholds, whereas there was a decrease in mutual information for fish in control and food cue groups (this decrease was greatest for fish subject to food cues). The greatest magnitude changes in mutual information occurred for fish in alarm groups at small, medium and large thresholds.
Entropy rate associated with small changes in velocity increased for control and alarm group fish, and was steady for food group fish (fig. S36 G). At medium thresholds, entropy rate associated with changes in velocity decreased slightly for fish in food and alarm groups and increased slightly for fish in control groups (fig. S36 H). Finally, entropy

rate associated with large changes in velocity decreased for control and alarm group fish, and increased slightly for food group fish post application of cue (fig. S36 I).

What do all these measures tell us about the predictability of different elements of the movement of the tetras under the influence of different external cues? The most marked effects relating to changes in displacement tended to appear for fish subject to alarm cues. In terms of making a change in position of magnitude greater than a small (or even medium) threshold, fish subject to alarm cues became less predictable in their movements after alarm cues were present in the water (as evidenced by an increase in entropy, a decrease in the mutual information that a previous movement gave about the next movement and an increased entropy rate, indicating a greater freedom in the movement choices utilised). However, entropy associated with large changes in position actually decreased for fish subject to alarm cues. This is most likely because fish subject to alarm cues tended to travel at lower speeds in general (see section S2.1), and thus it was more predictable that they would not be observed making large magnitude changes in displacement (due to their low speed).

In terms of changes in velocity, fish subject to all treatments gave very little mutual information about their next change in velocity, but fish subject to alarm cues tended to give the most away. Alarmed fish were less predictable in terms of making any sort of adjustment (above a small threshold) to their velocity according to measures of entropy and entropy rate. However alarmed fish were more predictable in terms of changes in velocity relative to medium and large thresholds, again as measured via entropy and entropy rate. In general fish subject to alarm cues showed lower magnitude changes in speed (see fig. S5 F in section S2.1) compared to fish in control and food treatments post cue; reduced magnitudes of changes in speed of alarmed fish speed seemed present irrespective of the speed of focal fish (see figs. S15 (C and D), S20 (C and D) and S25 (C and D) in section S2.4). Thus the increased predictability associated with making changes in velocity above a medium (or larger) threshold likely reflects the fact that alarmed fish tended not to adjust their speed rapidly compared to fish in control and food groups.

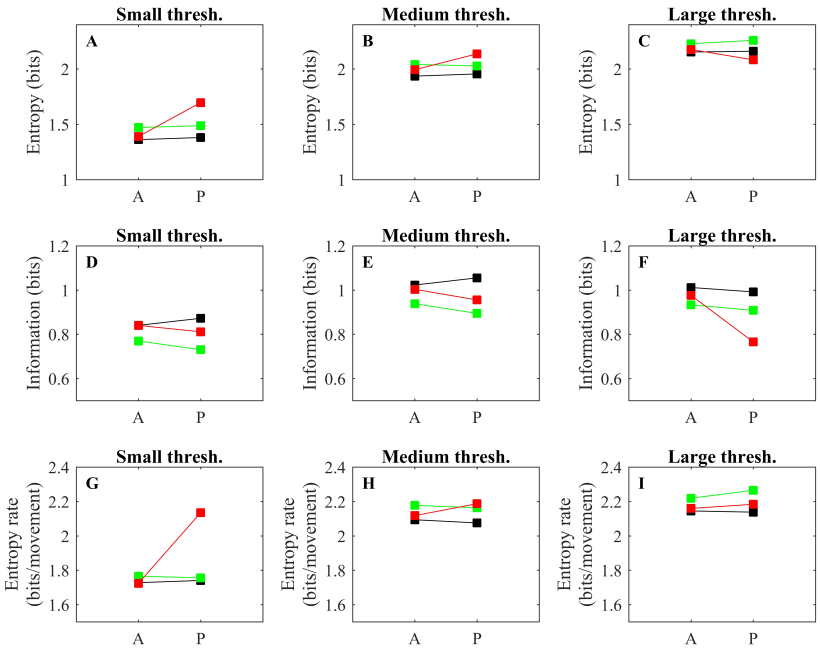

fig. S35. Entropy, mutual information and entropy rate associated with small ($K = 0.8$), medium ($K = 8$), and large ($K = 18$) changes in displacement over $L = 20$ frames for control (black), food (green) and alarm (red) treatments before (A) and after (P) deployment of the cue. Details of associated calculations are given in section S1.6.

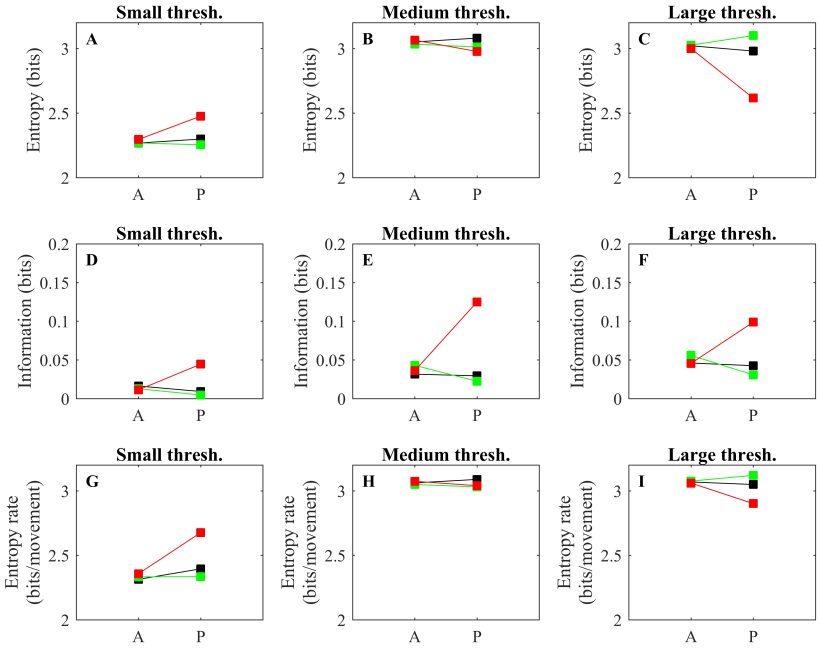

fig. S36 Entropy, mutual information and entropy rate associated with small ($K = 0.8$), medium ($K = 8$), and large ($K = 18$) changes in velocity over $L = 20$ frames for control (black), food (green) and alarm (red) treatments before (A) and after (P) deployment of the cue. Details of associated calculations are given in section S1.6.

S2.6 Approximate area under vigilance by individuals and groups, and number of group mates visible to individuals

The mean standard deviation of the area sighted by individual fish, $V_i(t)$, increased from interval A to P for fish subject to alarm cues (table S6). This change in the standard deviation of $V_i(t)$ somewhat reflects evident changes in the distribution of $V_i(t)$, as illustrated in fig. S37. Peaks in the relative frequency of observations occur

The mean standard deviation of the area sighted by groups, $V_{\text{group}}(t)$, increased from interval A to interval P for control groups. Relative frequency histograms of the approximate area covered by all group members eyes appear in fig. S38.

The mean median number of other group members sighted by individual fish, $G_i(t)$, decreased from interval A to P for fish subject to alarm cues. The most evident shifts in the distributions of $G_i(t)$ seemed to occur for alarm groups, post alarm cue, with the relative frequency that $G_i(t) \leq 4$ increasing during the P time interval (fig. S39 C and F).

Thus, changes in group configuration for fish subject to alarm cues, such as spacing between group mates and polarisation, seemed to have an effect on what individual fish could see. In particular, there was greater variability in the area sighted by individual fish, and the number of group mates visible at a given time decreased for alarmed fish. These changes for alarmed fish could reasonably be attributed to the vision of alarmed fish being more likely to be obstructed by their group mates, who tended to be closer on average according to measures of individual and group spacing examined in section S2.2. In spite of possible reductions in the area sighted by individuals, the union of areas sighted by all fish did not seem to change appreciably from A to P time intervals for fish subject to alarm cues (with the exception of fewer extreme areas sighted, compare fig. S38 C and F). Thus, at the group level roughly the same amount of open space remained under vigil for all treatments, but there was greater variability in open space under vigil during the second half of trials for fish in control groups. The details of what could be seen (for example more area in front or behind the group) could have differed between treatments though. A reduction in the number of group mates visible also means a reduction in the number of indirect sources of information about potential danger for alarmed fish. Perhaps alarmed fish traded indirect sources of information of potential danger from group mates for the improved safety of being closer to group mates?

table S6. Basic 95% confidence intervals for test statistics derived from the median and SD of the area sighted by individual fish, $V_i(t)$, the union of the areas sighted by all group members, $V_{\text{group}}(t)$, and the number of group members visible to individual fish, $G_i(t)$. Statistically significant effects are marked with an asterisk (*) (ie. confidence intervals for the test statistics that lie entirely above or below 0). If a confidence interval lies entirely below zero, then the associated quantity decreased from interval A to P; if a confidence interval lies entirely above zero then the associated quantity increased from A to P.

Variable	Control	Food	Alarm
median $V_i(t)$ (mm^2)	$(-4852.37, 1132.83)$	$(-3425.01, 5126.55)$	$(-7052.59, 1084.98)$
std $V_i(t)$ (mm^2)	$(-4440.99, 593.67)$	$(-4322.09, 2115.44)$	$(1456.74, 9158.62)^*$
median $V_{\text{group}}(t)$ (mm^2)	$(-3229.58, 1843.83)$	$(-3862.56, 4880.28)$	$(-6950.45, 1012.45)$
std $V_{\text{group}}(t)$ (mm^2)	$(97.50, 1359.48)^*$	$(-4350.78, 259.39)$	$(-3968.65, 538.30)$
median $G_i(t)$	$(-0.12, 0.13)$	$(-0.06, 0.30)$	$(-0.60, -0.15)^*$
std $G_i(t)$	$(-0.06, 0.02)$	$(-0.08, 0.02)$	$(-0.01, 0.11)$

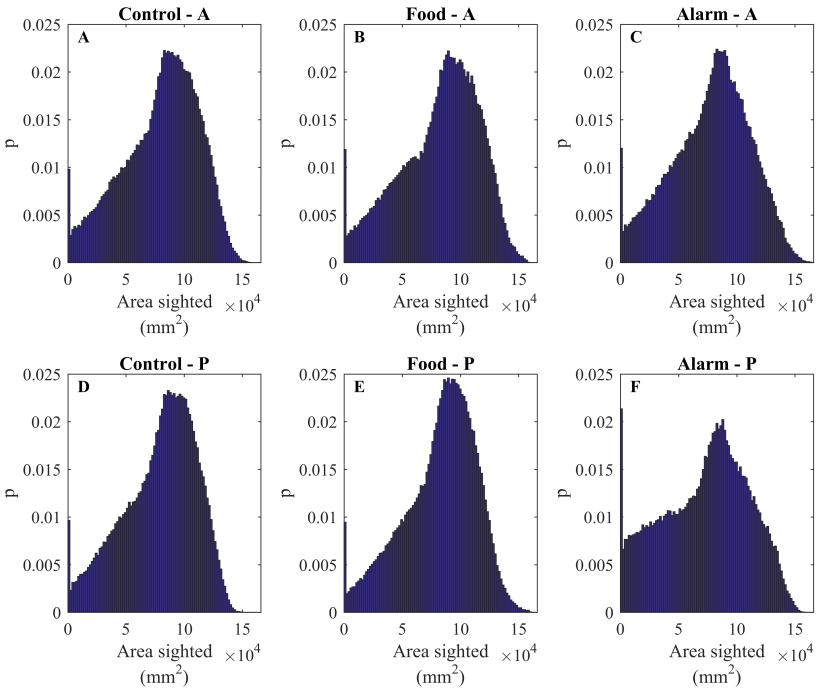

fig. S37. Relative frequency histograms of the approximate area visible to individual fish (in mm^2) for: A - control treatment ante cue, B - food treatment ante cue, C - alarm treatment ante cue, D - control treatment post cue, E - food treatment post cue, F - alarm treatment post cue. Details of associated calculations are given in section S1.7.

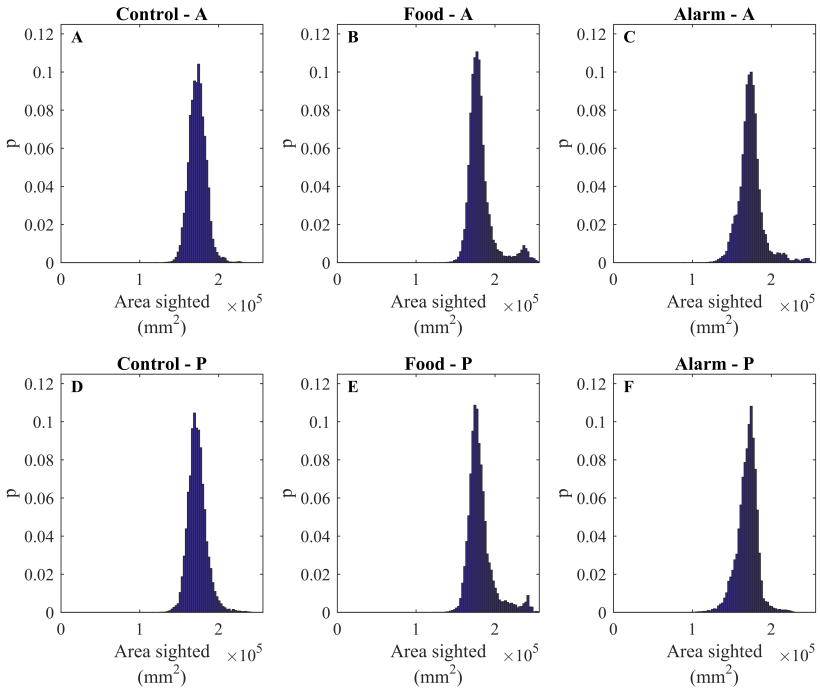

fig. S38. Relative frequency histograms of the approximate area covered by all group members eyes (in mm^2) for: A - control treatment ante cue, B - food treatment ante cue, C - alarm treatment ante cue, D - control treatment post cue, E - food treatment post cue, F - alarm treatment post cue. Details of associated calculations are given in section S1.7.

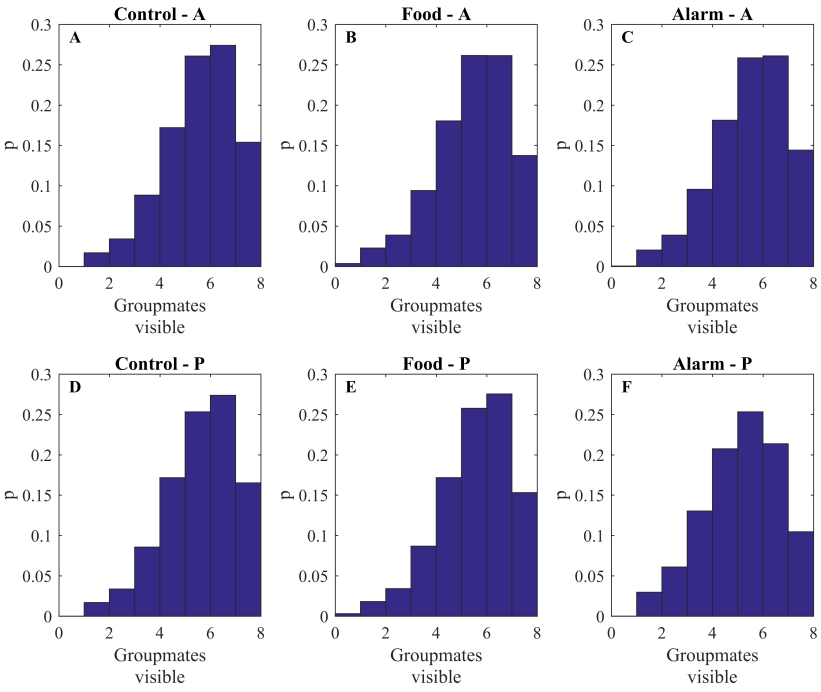

fig. S39. Relative frequency histograms of the approximate number of group mates visible to individual fish for: A - control treatment ante cue, B - food treatment ante cue, C - alarm treatment ante cue, D - control treatment post cue, E - food treatment post cue, F - alarm treatment post cue. Details of associated calculations are given in section S1.7.

S2.7 Individual durations between swaps in sense of motion (clockwise or anticlockwise motion about approximate centre of tank)

At least one estimated survival function for the duration between changes in sense of motion about the arena centre differed from the others ($p \approx 0$, log-rank test, $DF = 5$, test-statistic ≈ 2058). Subsequent comparisons of every pair of survival curves (again using the log-rank test, (*47*)) revealed that the Alarm P survival function differed from Alarm A, Control P and Food P, and the Control A survival function differed from Control P, Food A and Alarm A, see table S7 for details.

table S7. Summary of pairwise comparison of survival functions for unbroken durations of motion in either the clockwise or anti-clockwise sense about the approximate centre of the annular arena. $DF = 1$ for each pairwise comparison, with the significance level (α_{sig}) for rejection or acceptance of the null hypothesis determined via a Holm-Bonferroni correction.

Pair	Test statistic	p	α_{sig}	H_0 (0) or H_1 (1)
Food P & Alarm P	1079	≈ 0	0.0038	1
Alarm A & Alarm P	894.2	≈ 0	0.0042	1
Control P & Alarm P	827.7	≈ 0	0.0045	1
Control A & Food A	15.48	8.34×10^{-5}	0.0050	1
Control A & Control P	12.48	4.11×10^{-4}	0.0056	1
Control A & Alarm A	8.643	0.0033	0.0063	1
Food A & Food P	2.938	0.087	0.0083	0
Control P & Food P	2.009	0.1563	0.0100	0
Food A & Alarm A	0.8148	0.3667	0.0125	0

Even though statistical differences were found between six pairs of survival curves, the practical difference between the values of these curves were small (see fig. S40) with the exception that $S(t)$ was markedly lower for alarmed individuals post application of alarm cue (red dashed line in fig. S40). Re-examination of the pairwise calculations in table S7 reinforces the fact that the most dramatic statistical differences are between the alarm P survival function and everything else. This suggests that the alarmed fish genuinely tended to swap sense of motion about the centre of the tank (or stop completely) more frequently then fish in the control or food treatments, and that the swaps made by alarmed fish were more frequent post cue than ante cue.

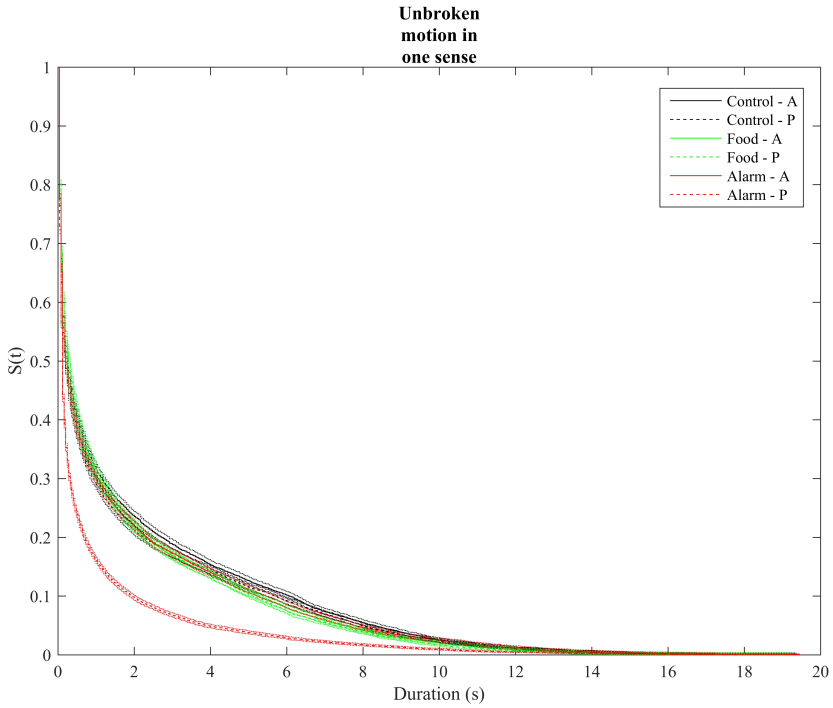

fig. S40. Kaplan-Meier survival function estimates, $S(t)$, for durations of unbroken motion in either the clockwise or anticlockwise sense about the centre of the arena's inner circular boundary. Survival curves associated with control groups are plotted in black, curves associated with food treatment groups are plotted in green and curves associated with the alarm treatment are plotted in red. Solid curves are associated with data pooled from the ante (A) time intervals, dashed curves are associated with the post (P) time intervals. Approximate 95% confidence intervals for each of the curves are bounded by dotted lines. Details of associated calculations are given in section S1.8.

S2.8 Correlation in direction of motion as a function of time delay

In most instances the time delay associated with maximum C_{ij} was negative when partner fish (j) were located in front of focal fish (i), and positive when partner fish we located behind focal fish. This indicates that on average fish tended to adjust their direction of motion to match that previously adopted by partners in front of them. Thus, on average, tetras led from the front and followed from behind. In general it is thought that fish direct group movement from the front-most positions of groups, and there is empirical evidence to support this hypothesis for other species of fish (specifically for golden shiners, (*5*), and eastern mosquitofish (*Gambusia holbrooki*, (*49*)). However, the assumption that changes in direction flow from the front to the back of groups on average is something that should be tested on a species by species basis. Outside fish, there is evidence that domestic pigeons (*Columba livia*) direct group motion when at the front of groups, (*2*). On the other hand, honey bees adopt a different mechanism of guidance ('streaking') where bees thought to be guiding a group travel through the upper portions of a swarm at relatively high speed in their target direction rather than solely occupying the most forward portions of the group, (*23-25, 50-52*).

The mean maximum mean delayed correlation in direction of motion decreased from interval A to P for fish subject to alarm cues when partner fish were located in front of, or behind, focal fish (table S8). These reductions in maximum time delayed correlation may reflect decreases in the instantaneous alignment of fish's motion when subject to alarm cues (as noted in section S2.2 with respect to polarisation of the direction of motion of group members). There were no significant effects on the time delay associated with reaching maximum delayed correlation, so on average external cues did not affect how rapidly fish adjusted to their partners' direction of motion, just the extent to which the directions of motion matched.

table S8. Basic 95% confidence intervals for test statistics derived from the time lag to maximum mean correlation in direction of motion, τ^*_{ij}, and the associated maximum mean correlation, $C_{ij}(\tau^*_{ij})$, when partner fish (j) were located in front of (PIF) or behind (PB) each focal fish (i). Statistically significant effects are marked with an asterisk (*) (ie. confidence intervals for the test statistics that lie entirely above or below 0). If a confidence interval lies entirely below zero, then the associated quantity decreased from interval A to P; if a confidence interval lies entirely above zero then the associated quantity increased from A to P.

Variable	Control	Food	Alarm
mean τ^*_{ij} PIF (s)	$(-0.03, 0.13)$	$(-0.18, 0.13)$	$(-0.28, 0.06)$
mean $C_{ij}(\tau^*_{ij})$ PIF	$(-0.02, 0.02)$	$(-0.05, 0.03)$	$(-0.16, -0.04)^*$
mean τ^*_{ij} PB (s)	$(-0.18, 0.14)$	$(-0.31, 0.12)$	$(-0.17, 0.33)$
mean $C_{ij}(\tau^*_{ij})$ PB	$(-0.03, 0.02)$	$(-0.01, 0.08)$	$(-0.19, -0.05)^*$

S2.9 Alignment responses associated with rapid turns

The mean number of rapid turns per minute performed by individual fish increased from interval A to P for fish subject to alarm cues (table S9). The mean delay to align with visible rapidly turning near neighbours decreased from interval A to P for fish subject to food cues; in other words fish subject to food cues responded more rapidly to fast turning partners post cue according to our chosen measure.

table S9. Basic 95% confidence intervals for test statistics derived from the mean and SD of the number of rapid turns performed by individual fish per minute, and the time delay to align with rapidly turning, visible, near neighbours. Statistically significant effects are marked with an asterisk (*) (ie. confidence intervals for the test statistics that lie entirely above or below 0). If a confidence interval lies entirely below zero, then the associated quantity decreased from interval A to P; if a confidence interval lies entirely above zero then the associated quantity increased from A to P.

Variable	Control	Food	Alarm
mean rapid turns	$(-1.78, 2.16)$	$(-2.22, 2.62)$	$(3.83, 9.78)^*$
std rapid turns	$(-0.07, 1.00)$	$(-0.68, 0.75)$	$(-0.72, 0.63)$
mean delay (s)	$(-0.05, 0.20)$	$(-0.32, -0.03)^*$	$(-0.11, 0.17)$
std delay (s)	$(-0.01, 0.16)$	$(-0.17, 0.03)$	$(-0.15, 0.03)$

S2.10 Number of times isolated from group, and duration of individual periods in isolation

There were no significant changes from interval A to P in the mean or standard deviation of the number of times per minute that fish became isolated (were $>$ 90 mm from any other group member) (table S10). Figure S41 contains relative frequency histograms of the number of times per minute that fish became isolated subject to each treatment.

table S10. Basic 95% confidence intervals for test statistics derived from the mean and SD of the number of times per minute that fish became isolated from a larger group (were more than three body lengths from any other fish).

Variable	Control	Food	Alarm
mean no. isolations	$(-0.13, 0.25)$	$(-0.27, 0.69)$	$(-0.37, 0.07)$
std no. isolations	$(-0.20, 0.13)$	$(-0.01, 0.72)$	$(-0.47, 0.03)$

At least one estimated survival function for durations spent in isolation differed from the others ($p = 7.8461 \times 10^{-5}$, log-rank test, $DF = 5$, test-statistic $= 26.2880$). Subsequent

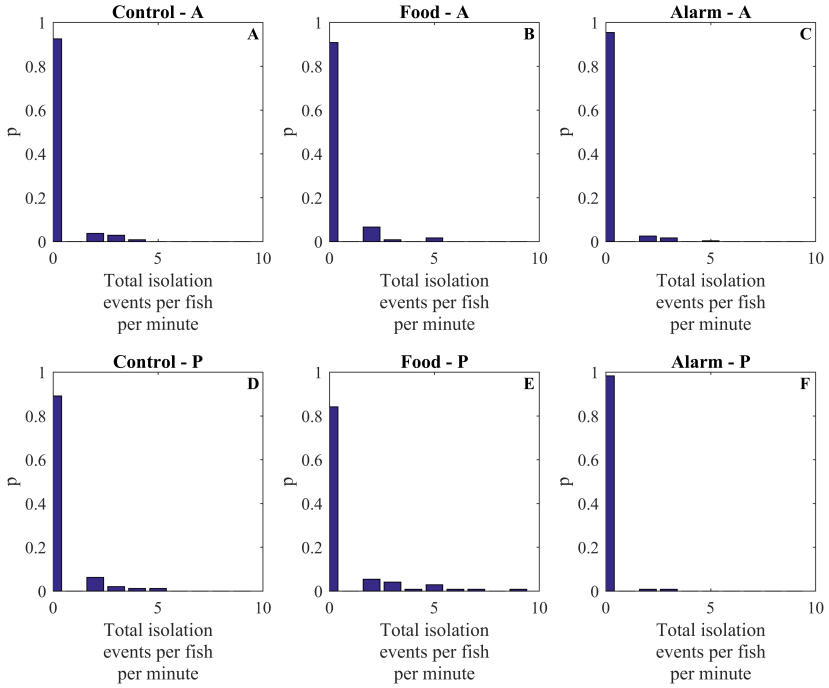

fig. S41. Relative frequency histograms of the number of times each fish became isolated from the group (was > 90 mm from any other member of the group) during 1-minute time intervals ante (A) or post (P) application of cues to the water. Details of associated calculations are given in section S1.11.

comparisons of every pair of survival curves suggested that the control P survival function differed from the food P survival function (table S11), with fish subject to the food cue tending to spend a greater period in isolation when isolated from the group (green dashed curve, fig. S42).

S2.11 Durations spent on the edge of the group, and frequency of swapping between interior and exterior positions

There were no significant changes in the mean or standard deviation of the rate that individuals swapped between positions on the group exterior and interior (table S12). Distributions of the observed number of swaps per minute between group exterior and interior are plotted in fig. S43

At least one estimated survival function for durations spent by individuals on their group's exterior differed from the others ($p = 8.1046 \times 10^{-15}$, log-rank test, $DF = 5$, test-statistic = 75.30). Pairwise comparisons of survival curves suggested that the control P survival function differed from the food P and alarm P survival functions, the food

table S11. Summary of pairwise comparison of survival functions for periods in isolation (separation of more than 90 mm/approximately three body lengths from any other group member). $DF = 1$ for each pairwise comparison, with the significance level (α_{sig}) for rejection or acceptance of the null hypothesis determined via a Holm-Bonferroni correction.

Pair	Test statistic	p	α_{sig}	H_0 (0) or H_1 (1)
Control P & Food P	18.4740	1.7224×10^{-5}	0.0033	1
Control A & Food A	5.3604	0.0206	0.0042	0
Control A & Control P	2.8353	0.0922	0.0050	0
Control A & Alarm A	2.6609	0.1028	0.0056	0
Control P & Alarm P	1.3339	0.2481	0.0063	0
Food P & Alarm P	0.6795	0.4098	0.0071	0
Alarm A & Alarm P	0.4384	0.5079	0.0083	0
Food A & Food P	0.2687	0.6042	0.0125	0
Food A & Alarm A	0.0056	0.9402	0.0500	0

table S12. Basic 95% confidence intervals for test statistics derived from the mean and SD of the number of times individuals swapped between positions on the group exterior and interior (defined with respect to the convex hull for the set of coordinates for each frame of data) during one minute intervals.

Variable	Control	Food	Alarm
mean no. swaps	$(-2.52, 1.07)$	$(-0.76, 4.30)$	$(-2.04, 3.72)$
std no. swaps	$(-0.56, 1.23)$	$(-1.86, 0.61)$	$(-1.93, 0.73)$

P survival function differed from the alarm P survival function and the food A survival function differed from the control A survival and alarm A survival functions (table S13). There were no differences in survival functions for control A versus control P, food A versus food P or alarm A versus alarm P, so the statistical differences may have been a function of the individuals in each treatment group, rather than due to the cues that were deployed. In general the survival functions for the control groups (black curves in fig. S44) tended to sit above the survival functions for both alarm groups (red curves) and food groups (green curves), with the survival functions for the alarm groups sitting in the middle. This indicates that fish in the control groups tended to spend longer on the edge of the group, with fish in alarm groups spending less time on the group's edges and fish in food treatment groups spending the least time on their groups' edges, irrespective of if cues had been deployed or not.

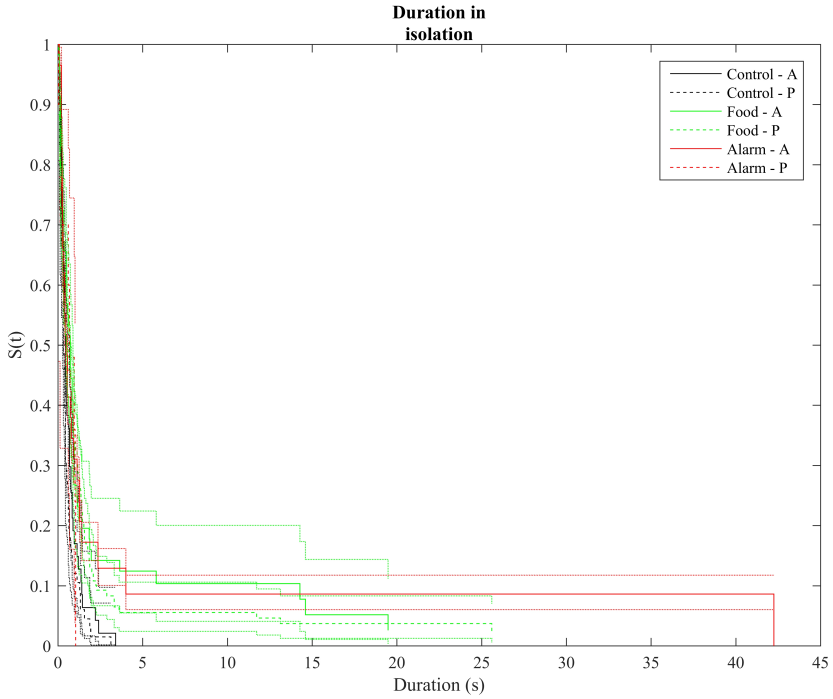

fig. S42. Kaplan-Meier survival function estimates, $S(t)$, for durations of in isolation (separation of more than 90 mm/approximately three body lengths from any other group member). Survival curves associated with control groups are plotted in black, curves associated with food treatment groups are plotted in green and curves associated with the alarm treatment are plotted in red. Solid curves are associated with data pooled from the ante (A) time intervals, dashed curves are associated with the post (P) time intervals. Approximate 95% confidence intervals for each of the curves are bounded by dotted lines. Details of associated calculations are given in section S1.11.

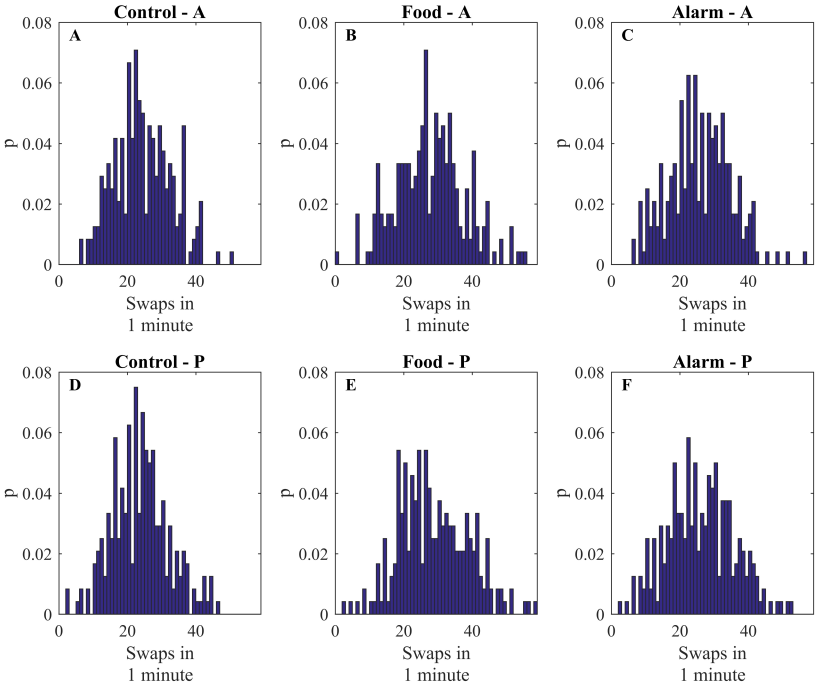

fig. S43. Relative frequency histograms of the number of swaps from group exterior to group interior or vice versa made by each fish per minute. Details of associated calculations are given in section S1.12.

table S13. Summary of pairwise comparison of survival functions for durations spent by individuals on their group's exterior. $DF = 1$ for each pairwise comparison, with the significance level (α_{sig}) for rejection or acceptance of the null hypothesis determined via a Holm-Bonferroni correction.

Pair	Test statistic	p	α_{sig}	H_0 (0) or H_1 (1)
Control P & Food P	46.3792	9.7445×10^{-12}	0.0033	1
Control A & Food A	27.0474	1.9853×10^{-7}	0.0042	1
Food P & Alarm P	13.2129	2.7802×10^{-4}	0.0050	1
Control P & Alarm P	10.2776	0.0013	0.0056	1
Food A & Alarm A	8.6453	0.0033	0.0063	1
Control A & Alarm A	4.9395	0.0262	0.0125	0
Food A & Food P	1.9519	0.1624	0.0167	0
Alarm A & Alarm P	0.4469	0.5038	0.0250	0
Control A & Control P	0.1475	0.7010	0.0500	0

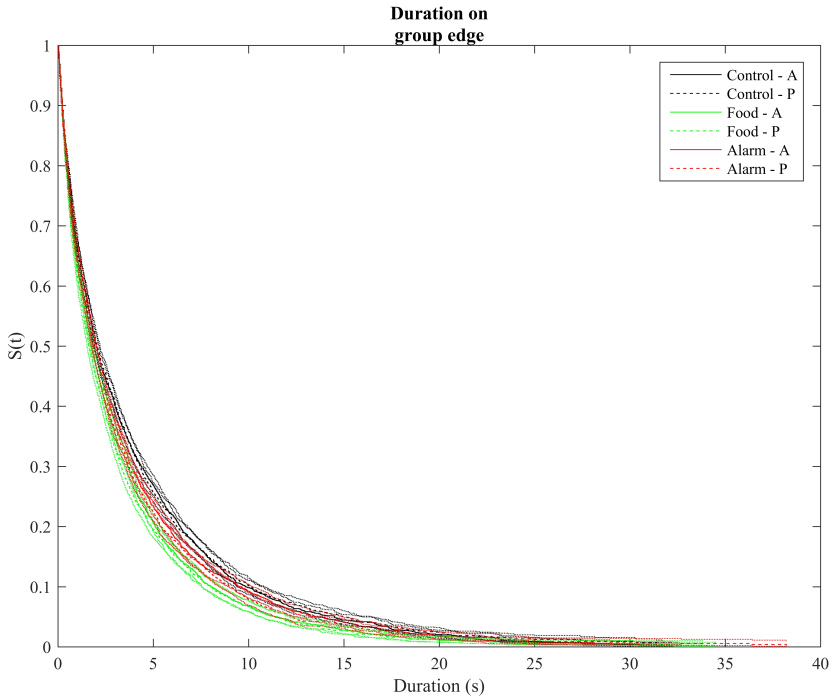

fig. S44. Kaplan-Meier survival function estimates, $S(t)$, for durations spent by individuals on their group's exterior. Survival curves associated with control groups are plotted in black, curves associated with food treatment groups are plotted in green and curves associated with the alarm treatment are plotted in red. Solid curves are associated with data pooled from the ante (A) time intervals, dashed curves are associated with the post (P) time intervals. Approximate 95% confidence intervals for each of the curves are bounded by dotted lines. Details of associated calculations are given in section S1.12.

S2.12 An eye for trouble? Do x-ray tetras prefer a particular eye when scanning for food or danger?

The standard deviation of the fraction of a group with their right eye pointing outwards from the group centre decreased from interval A to P for fish in control groups (table S14).

table S14. Basic 95% confidence intervals for test statistics derived from the mean and SD of the fraction of group members that had their right eye pointing outwards relative to the group centroid. Statistically significant effects are marked with an asterisk (*) (ie. confidence intervals for the test statistics that lie entirely above or below 0). If a confidence interval lies entirely below zero, then the associated quantity decreased from interval A to P; if a confidence interval lies entirely above zero then the associated quantity increased from A to P.

Variable	Control	Food	Alarm
mean $\text{frac}_{\text{right}}(t)$	$(-0.02, 0.02)$	$(-0.04, 0.01)$	$(-0.02, 0.02)$
std $\text{frac}_{\text{right}}(t)$	$(-0.02, -0.004)^*$	$(-0.02, 0.02)$	$(-0.004, 0.03)$

S2.13 Tendency to face towards or away nearest walls

There were no significant changes from interval A to P to the median or standard deviation of the fraction of group members facing outwards from the wall that was on average closest to all group members during a given video frame for any treatment (table S15). Relative frequency histograms of the fraction of group members facing outwards from the boundary appear in fig. S45.

table S15. Basic 95% confidence intervals for test statistics derived from the median and SD of the fraction of group members facing outwards from the wall that was closest on average during a given time step, $F_o(t)$.

Variable	Control	Food	Alarm
median $F_o(t)$	$(-0.05, 0.02)$	$(-0.06, 0.03)$	$(-0.05, 0.04)$
std $F_o(t)$	$(-0.01, 0.04)$	$(-0.04, 0.02)$	$(-0.05, 0.001)$

S2.14 Durations spent within a threshold distance of the walls

At least one estimated survival function for durations spent within one body length (30 mm) of the outer water level boundary of the arena differed from the others ($p \approx 0$, log-rank test, $DF = 5$, test-statistic $= 127.8225$). Subsequent pairwise comparisons of survival

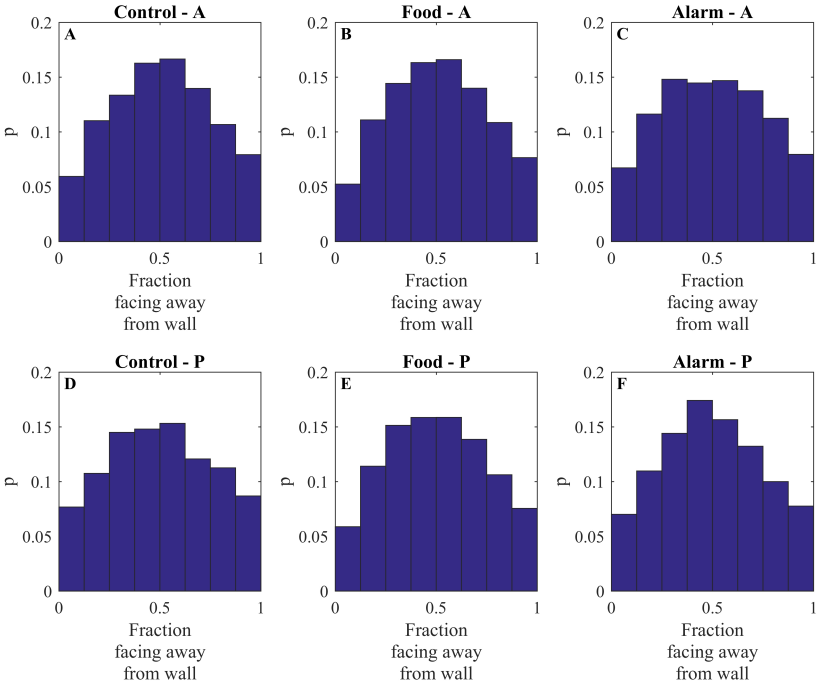

fig. S45. Relative frequency histograms of the fraction of the group facing away from the boundary that was the least mean distance from all group members for a given time step. Details of associated calculations are given in section S1.14.

functions suggested that the alarm P survival function differed from the food P and control P survival functions, the food A and food P survival functions differed from one another, the control A and control P survival functions differed from one another, and the alarm A survival function differed from the control A survival function (table S16) at the one body length threshold. Inspection of graphs of the Kaplan-Meier estimates of the survival functions as shown in fig. S46 suggests that fish subject to alarm cues (red dashed line) tended to stay within one body length of the outer boundary for greater durations then fish subject to control (black dashed line) or food (green dashed line) treatments during the post cue (P) time interval.

At least one estimated survival function for durations spent within two body lengths (60 mm) of the outer water level boundary of the arena differed from the others ($p \approx 0$, log-rank test, $DF = 5$, test-statistic $= 210.9055$). Subsequent pairwise comparisons suggested that there was no difference between all possible pairs of survival functions associated with any of the treatments during the ante time interval, but that the food P survival function differed from the food A, control P and alarm P survival functions, the alarm P survival function differed from the alarm A and control P survival functions and the control P survival function differed from the control A survival function (table S17). Examination

of the Kaplan-Meier estimates of each of the survival functions (illustrated in fig. S47) suggested that post cue, fish subject to food cues tended to spend the least duration during distinct periods within two body lengths of the outer boundary, whereas fish subject to alarm cues tended to spend the greatest duration per visit to the region within two body lengths of the outer boundary. The survival curve for fish in control groups post cue tended to lie in-between the food P survival curve (which lay below the control P curve) and the alarm P curve (which lay above the control P curve).

At least one estimated survival function for durations spent within one body length (30 mm) of the inner water level boundary of the arena differed from the others ($p = 5.1238 \times 10^{-5}$, log-rank test, $DF = 5$, test-statistic $= 27.2392$). Subsequent pairwise comparisons of survival curves suggested that the control A and alarm A survival curves differed from each other, and that the food P and alarm P survival curves differed from each other (table S18). Examination of the Kaplan-Meier estimates of the corresponding survival curves (illustrated in fig. S48) suggests that fish in control groups tended to spend greater durations within one body length of the inner boundary before application of cues (solid black line) than fish in alarm groups ante cue. Further, fish subject to food cues tended to spend greater durations within one body length of the inner wall post cue (green dashed line) than fish subject to alarm cues (red dashed line).

No estimated survival function for durations spent within two body lengths (60 mm) of the inner water level boundary of the arena differed from any others ($p = 0.5747$, log-rank test, $DF = 5$, test-statistic $= 3.8265$). Kaplan-Meier estimates of the corresponding survival functions are plotted in fig. S49.

Overall, the most marked effect seemed to be in changes to durations spent near the outer boundary by fish subject to both food and alarm cues, particularly at the tested threshold distance of 60 mm (approximately two body lengths). Fish subject to alarm cues tended to spend greater durations within the threshold distance of the outer boundary, whereas fish subject to food cues tended to spend lesser durations near the outer boundary. These results seem to be consistent with what was suggested by the analysis of spatial positioning with respect to the boundaries in section S2.3, where fish subject to alarm cues increased their median distance to the inner boundary (although there was no significant change in distance to the outer boundary) and fish subject to food cues increased their median distance from the outer boundary.

table S16. Summary of pairwise comparison of survival functions for durations spent within one body length (30 mm) of the outer water level boundary of the arena. $DF = 1$ for each pairwise comparison, with the significance level (α_{sig}) for rejection or acceptance of the null hypothesis determined via a Holm-Bonferroni correction.

Pair	Test statistic	p	α_{sig}	H_0 (0) or H_1 (1)
Food P & Alarm P	80.7386	≈ 0	0.0033	1
Control P & Alarm P	52.4872	4.3310×10^{-13}	0.0038	1
Food A & Food P	37.0891	1.1285×10^{-9}	0.0042	1
Control A & Control P	14.9094	1.1280×10^{-4}	0.0071	1
Control A & Alarm A	6.6719	0.0098	0.0100	1
Food A & Alarm A	5.1487	0.0233	0.0125	0
Control P & Food P	3.1759	0.0747	0.0167	0
Alarm A & Alarm P	3.0113	0.0827	0.0250	0
Control A & Food A	0.0835	0.7725	0.0500	0

table S17. Summary of pairwise comparison of survival functions for durations spent within two body lengths (60 mm) of the outer water level boundary of the arena. $DF = 1$ for each pairwise comparison, with the significance level (α_{sig}) for rejection or acceptance of the null hypothesis determined via a Holm-Bonferroni correction.

Pair	Test statistic	p	α_{sig}	H_0 (0) or H_1 (1)
Food P & Alarm P	167.1704	≈ 0	0.0038	1
Control P & Alarm P	67.7317	3.8858×10^{-15}	0.0042	1
Food A & Food P	60.1090	8.9928×10^{-15}	0.0045	1
Control P & Food P	29.3994	5.8895×10^{-8}	0.0056	1
Alarm A & Alarm P	23.4631	1.2733×10^{-6}	0.0063	1
Control A & Control P	18.8691	1.4000×10^{-5}	0.0071	1
Control A & Food A	5.6688	0.0173	0.0125	0
Food A & Alarm A	1.4853	0.2229	0.0250	0
Control A & Alarm A	0.9384	0.3327	0.0500	0

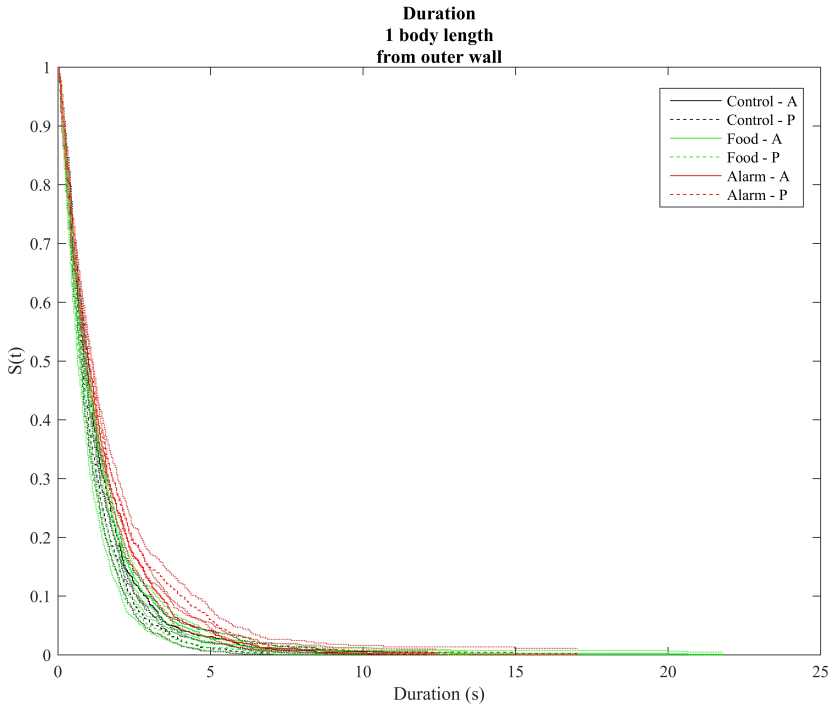

fig. S46. Kaplan-Meier survival function estimates, $S(t)$, for durations spent within one body length (30 mm) of the outer water level boundary of the arena. Survival curves associated with control groups are plotted in black, curves associated with food treatment groups are plotted in green and curves associated with the alarm treatment are plotted in red. Solid curves are associated with data pooled from the ante (A) time intervals, dashed curves are associated with the post (P) time intervals. Approximate 95% confidence intervals for each of the curves are bounded by dotted lines. Details of associated calculations are given in section S1.15.

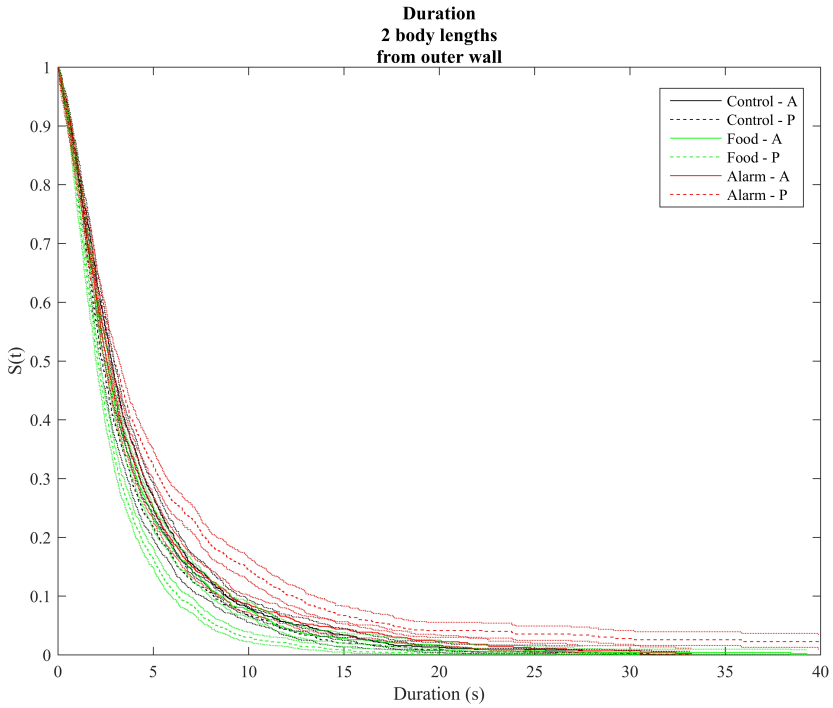

fig. S47. Kaplan-Meier survival function estimates, $S(t)$, for durations spent within two body lengths (60 mm) of the outer water level boundary of the arena. Survival curves associated with control groups are plotted in black, curves associated with food treatment groups are plotted in green and curves associated with the alarm treatment are plotted in red. Solid curves are associated with data pooled from the ante (A) time intervals, dashed curves are associated with the post (P) time intervals. Approximate 95% confidence intervals for each of the curves are bounded by dotted lines. Details of associated calculations are given in section S1.15.

table S18. Summary of pairwise comparison of survival functions for durations spent within one body length (30 mm) of the inner water level boundary of the arena. $DF = 1$ for each pairwise comparison, with the significance level (α_{sig}) for rejection or acceptance of the null hypothesis determined via a Holm-Bonferroni correction.

Pair	Test statistic	p	α_{sig}	H_0 (0) or H_1 (1)
Control A & Alarm A	12.0808	5.0943×10^{-4}	0.0036	1
Food P & Alarm P	8.6488	0.0033	0.0038	1
Food A & Alarm A	4.6667	0.0308	0.0050	0
Control P & Food P	4.3192	0.0377	0.0056	0
Control P & Alarm P	3.1212	0.0773	0.0063	0
Control A & Control P	1.8709	0.1714	0.0071	0
Food A & Food P	0.7250	0.3350	0.0100	0
Control A & Food A	0.2658	0.3562	0.0125	0
Alarm A & Alarm P	0.1833	0.3891	0.0167	0

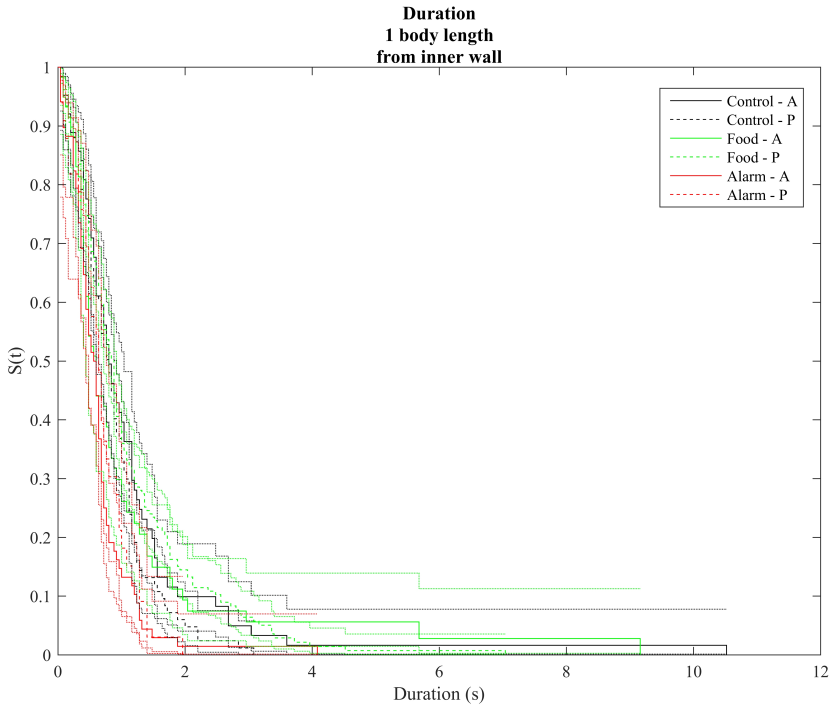

fig. S48. Kaplan-Meier survival function estimates, $S(t)$, for durations spent within one body length (30 mm) of the inner water level boundary of the arena. Survival curves associated with control groups are plotted in black, curves associated with food treatment groups are plotted in green and curves associated with the alarm treatment are plotted in red. Solid curves are associated with data pooled from the ante (A) time intervals, dashed curves are associated with the post (P) time intervals. Approximate 95% confidence intervals for each of the curves are bounded by dotted lines. Details of associated calculations are given in section S1.15.

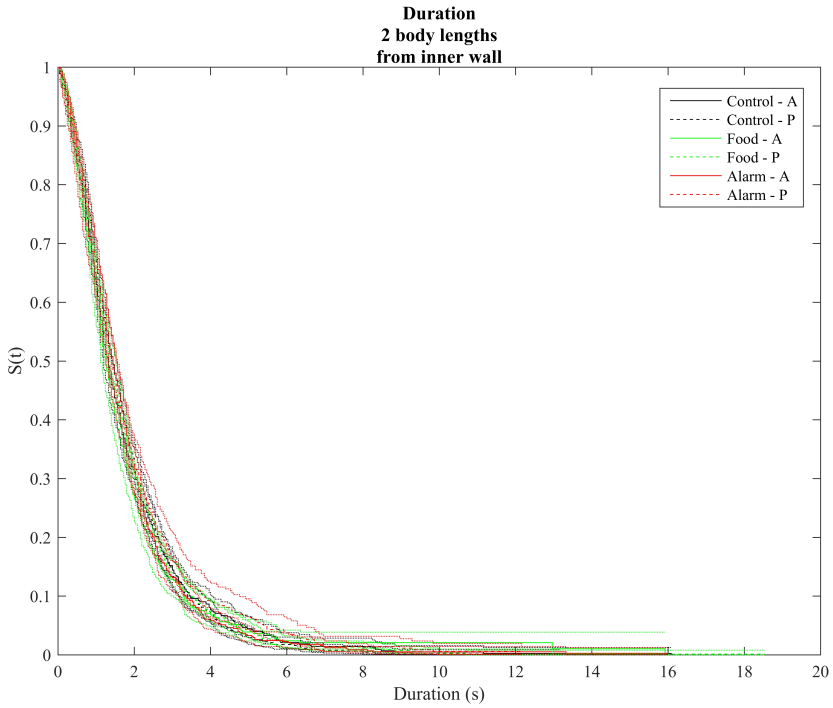

fig. S49. Kaplan-Meier survival function estimates, $S(t)$, for durations spent within two body lengths (60 mm) of the inner water level boundary of the arena. Survival curves associated with control groups are plotted in black, curves associated with food treatment groups are plotted in green and curves associated with the alarm treatment are plotted in red. Solid curves are associated with data pooled from the ante (A) time intervals, dashed curves are associated with the post (P) time intervals. Approximate 95% confidence intervals for each of the curves are bounded by dotted lines. Details of associated calculations are given in section S1.15.