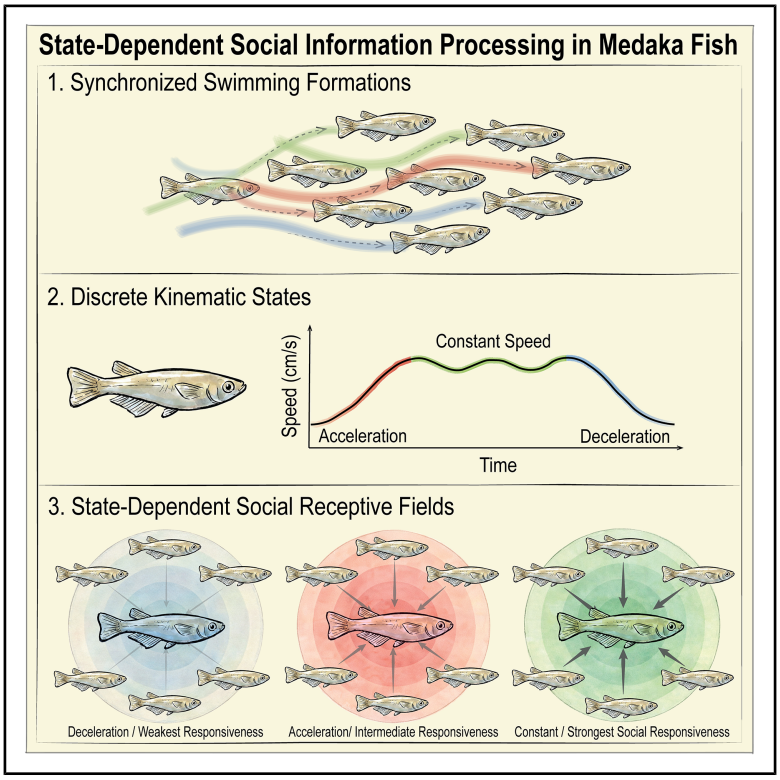
Report

Current Biology

Social interactions in medaka fish depend on discrete kinematic states of swimming behavior

Graphical abstract



Highlights

- Synchronized swimming formations in medaka develop early and stabilize rapidly
- Continuous swimming behavior can be segmented into discrete kinematic states
- Social information processing depends on the kinematic states
- Social responsiveness varies systematically across states

Authors

Roy Harpaz, Pedro Piquet, Yasuko Isoe, Mark C. Fishman, Florian Engert

Correspondence

harpazone@gmail.com

In brief

Harpaz et al. analyze collective swimming in medaka and show that continuous locomotion can be segmented into discrete kinematic states. State-dependent computational modeling reveals distinct social information processing in each state, with social responsiveness strongest during prolonged constant-speed swimming.

Harpaz et al., 2026, Current Biology *36*, 1–10
March 9, 2026
https://doi.org/10.1016/j.cub.2026.01.039

CellPress

Please cite this article in press as: Harpaz et al., Social interactions in medaka fish depend on discrete kinematic states of swimming behavior, Current Biology (2026), https://doi.org/10.1016/j.cub.2026.01.039

Current Biology

CellPress
OPEN ACCESS

Report

Social interactions in medaka fish depend on discrete kinematic states of swimming behavior

Roy Harpaz,[1,5,*] Pedro Piquet,[2] Yasuko Isoe,[1] Mark C. Fishman,[4] and Florian Engert[1,3]
[1]Department of Molecular and Cellular Biology, Harvard University, Cambridge, MA 02138, USA
[2]Department of Biomedical Engineering, McGill University, Montréal, QC H2S 3H, Canada
[3]Center for Brain Science, Harvard University, Cambridge, MA 02138, USA
[4]Department of Stem Cell and Regenerative Biology, Harvard University, Cambridge, MA 02138, USA
[5]Lead contact
*Correspondence: harpazone@gmail.com
https://doi.org/10.1016/j.cub.2026.01.039

SUMMARY

Complex collective behaviors such as schooling are believed to emerge from simple, individual-level computations that translate incoming information from conspecifics into actions.[1–14] Recently, it has been proposed that discrete behavioral modes, or internal states, may modulate these computations, affecting the resulting collective behaviors.[15–18] Direct evidence for such hierarchical control remains limited due to challenges in inferring hidden perception-action computations and uncovering discrete behavioral modes from continuous behaviors.[18–22] To address this, we analyzed swimming behaviors of medaka fish (*Oryzias latipes*) throughout development.[23–25] At the group level, medaka exhibit synchronized swimming formations that develop early, emerging around 2 weeks of age and stabilizing within 1 month. Unlike many teleost species that use burst-and-coast swim patterns,[19,26–35] medaka exhibit continuous tail and body undulations.[31] We show that this continuous behavior can be segmented into three distinct kinematic states: acceleration, deceleration, and prolonged constant-speed swimming. Using state-dependent computational models,[15,18] we tested how medaka translate social information from neighbors into actions across these kinematic states. The models revealed distinct computations governing social information processing and movement responses in each state. Moreover, social responsiveness varied significantly between states: it was strongest during constant-speed epochs, intermediate during accelerations, and lowest during decelerations. These findings highlight discrete behavioral modes as key modulators of social interaction computations underlying collective behavior.

RESULTS

Collective behavior in medaka develops early and saturates rapidly

To study individual and collective behavior in continuous undulatory swimmers, we analyzed the swimming behavior of groups of the Japanese rice fish (*Oryzias latipes*), known as medaka[23–25,36] (Figures 1A and 1B; Video S1). We tracked groups through development and measured their behavior at the ages of 10, 16, 24, and 34 days post fertilization (dpf) and at the adult stage (>60 dpf) (Figures S1A and S1B). We estimated each fish's position, heading orientation, and the midline of the body and calculated the relative distances, azimuths, and heading directions of its neighbors (Figure 1B, inset; STAR Methods).

We find that social behavior develops early in these fish, with signs of attraction emerging already at 16 dpf (8 days post hatching; STAR Methods), and monotonically increasing as the fish mature (group attraction = 0.01 ± 0.08 [−0.03, 0.04], 0.06 ± 0.08 [0.02, 0.09], 0.34 ± 0.16 [0.27, 0.41], 0.43 ± 0.18 [0.35 0.51], and 0.42 ± 0.16 [0.32, 0.51]; mean ± SD, 95% confidence interval [CI] in brackets, for ages 10, 16, 24, 34, and >60 dpf) (Figures 1C and 1E). Alignment with neighbors follows a similar developmental trajectory (group alignment = 0.41 ± 0.02 [0.4, 0.42], 0.44 ± 0.04 [0.43, 0.46], 0.55 ± 0.06 [0.52, 0.58], 0.56 ± 0.07 [0.53, 0.6], and 0.58 ± 0.05 [0.55, 0.61]) (Figures 1D and 1F; STAR Methods), and both properties are close to the adult values at the juvenile age of 34 dpf. Compared with adult zebrafish, a similarly sized teleost commonly studied for collective behaviors,[8,15,37,38] medaka groups exhibit more strongly aligned formations but reduced group cohesiveness (Figures 1C and 1D), indicating a stronger tendency toward schooling rather than shoaling behaviors.[39]

Analyzing group structure at the age of 60 dpf, when social behavior is fully developed, revealed the unique spatial organization of medaka groups: neighbors are more likely to be found directly to the side of a focal fish at intermediate distances (~2.5 cm) but are less likely to be found directly in front or behind the focal fish (Figure 1G). Alignment of heading direction is similarly high in areas corresponding to the most likely positions of neighbors, and also directly in front of the focal fish, emphasizing the anisotropic structure of the group (Figure 1H).

These results confirm that medaka show distinct and robust social behaviors in a group setting, which are fully developed at the age of 60 dpf, allowing us to study the

Please cite this article in press as: Harpaz et al., Social interactions in medaka fish depend on discrete kinematic states of swimming behavior, Current Biology (2026), https://doi.org/10.1016/j.cub.2026.01.039

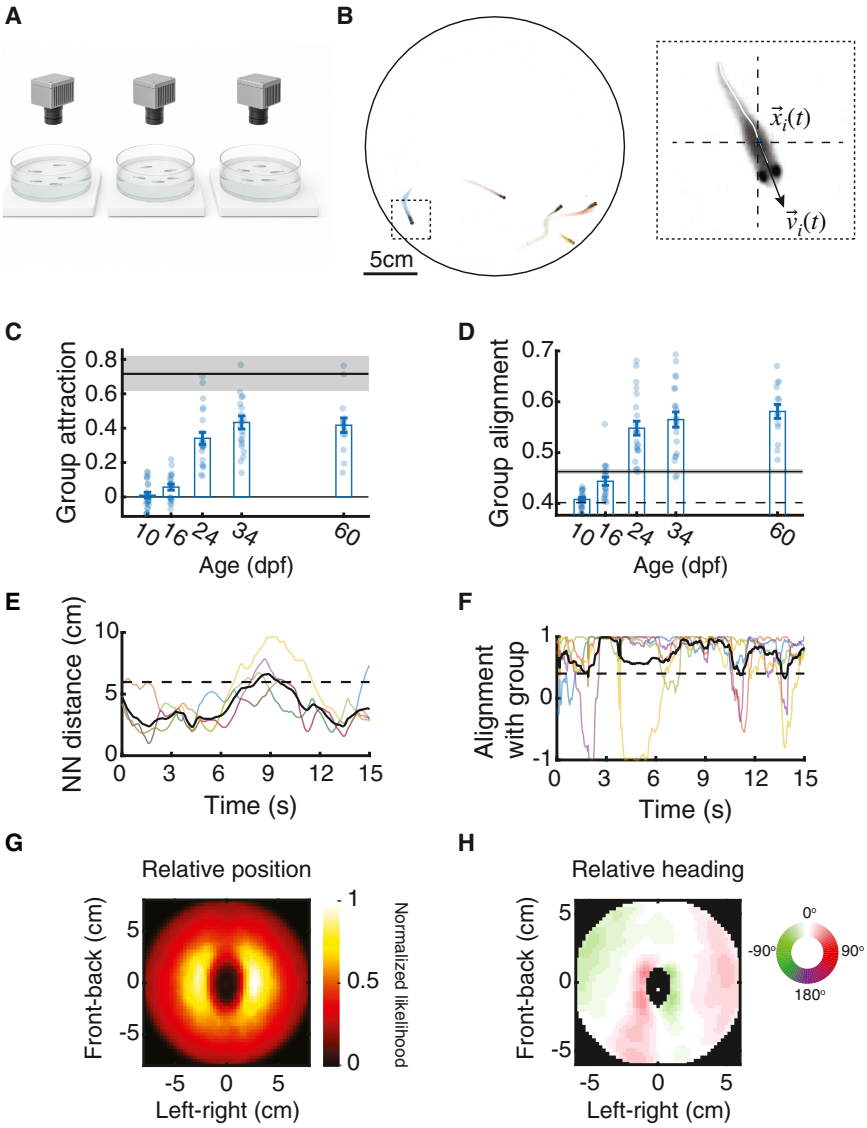


Figure 1. Social behavior in medaka develops early and rapidly

(A) Sketch of the experimental setups used for behavioral tracking of medaka groups (STAR Methods).

(B) A snapshot of a group of 5 adult medaka (>60 dpf) swimming in a large arena (diameter = 26 cm). Different colors represent different fish in the group. Inset: zoom-in view of a single fish *i*, its center of mass position $\vec{x}_i(t)$, velocity vector $\vec{v}_i(t)$, and its midline (solid white line).

(C and D) Average attraction (C) and alignment (D) values of medaka groups (5 fish in a group) at different developmental stages. Attraction = $-\log\left(\frac{NNd}{NNd^{Shuffle}}\right)^{55}$, where *NNd* is the average nearest neighbor distance, and $NNd^{Shuffle}$ is the expected value for shuffled groups (STAR Methods). Alignment = $|\sum_i \vec{d}_i|/n$, where $\vec{d}_i$ denotes the unit heading vector of fish *i* and *n* is the number of fish in the group. Dotted line is the expected value for shuffled groups (STAR Methods). Solid horizontal black lines and shaded areas are the mean and SEM of attraction and alignment values for adult zebrafish groups (5 fish in a group, $n = 6$ groups; STAR Methods).

(E) Example traces of the NN distances for 5 adult medaka (>60 dpf) swimming in a group (colored lines), the average of the group (solid black line), and the expected value of shuffled groups (dotted black line).

(F) Example traces of the alignment of individual fish with the heading direction of the group (colored lines; STAR Methods) and the alignment of the group (solid black line). Same group as in (E).

(G) Heatmap showing the likelihood of observing a neighboring fish at different spatial positions relative to a focal fish located at (0,0) pointing north (STAR Methods); groups of 5 fish (age > 60 dpf).

(H) Average relative heading of neighbors at different spatial locations with respect to a focal fish located at (0,0) pointing north (STAR Methods). White colors represent no deviation or perfect alignment; same groups as in (G).

In (C) and (D), dots represent average values of different groups, and error bars are mean ± SEM. Number of medaka groups tested at each age: $n_{10\ dpf} = 19$, $n_{16\ dpf} = 20$, $n_{24\ dpf} = 22$, $n_{34\ dpf} = 22$, and $n_{60\ dpf} = 15$.
See also Figure S1 and Video S1.

social interaction computations underlying collective behaviors at this age.

Medaka display distinct kinematic swim states

We next analyzed the swim kinematics of adult medaka (age > 60 dpf). As reported before,[31] medaka swim with continuous tail and body undulations (Figure 2A; Video S2), as opposed to the punctuated burst-and-coast kinematics utilized, for example, by zebrafish (Figure S2A; Video S3). Importantly, we were able to naturally segment the speed profile of medaka into three distinct kinematic states: acceleration and deceleration phases where fish modulate their swimming speed upward or downward, and a constant-speed state in which fish maintain constant velocity magnitudes (Figures 2B and S2B; STAR Methods). To confirm that these three states are distinct, we demonstrate that speed profiles in each state are characterized by a unique family of functions. Accelerations and decelerations are best described by sigmoidal functions with positive and negative slopes, respectively ($R_{Acc}^2 = 0.99 \pm 0.001$, $R_{Dec}^2 = 0.98 \pm 0.001$; mean ± SD), and the constant-speed epochs by an oscillating sine wave ($R^2 = 0.74 \pm 0.03$) (Figure 2B). Each function family provided a uniquely good fit for only one state and failed to describe the other states (Figure S2C). Mechanistically, accelerations were characterized by a range of time constants $\tau_{Acc} = 0.044 \pm 0.013$ (mean ± SD) that were independent of the amplitude of the swim epoch (Figure 2C). Decelerations were characterized by longer time constants $\tau_{Dec} = 0.055 \pm 0.019$, with large amplitude epochs having longer time constants (Spearman's correlation = 0.27; $p < 0.0001$, permutation test). During constant-speed epochs, speed oscillated at 3.34 ± 1.34 Hz, slower than

Please cite this article in press as: Harpaz et al., Social interactions in medaka fish depend on discrete kinematic states of swimming behavior, Current Biology (2026), https://doi.org/10.1016/j.cub.2026.01.039

CellPress
OPEN ACCESS

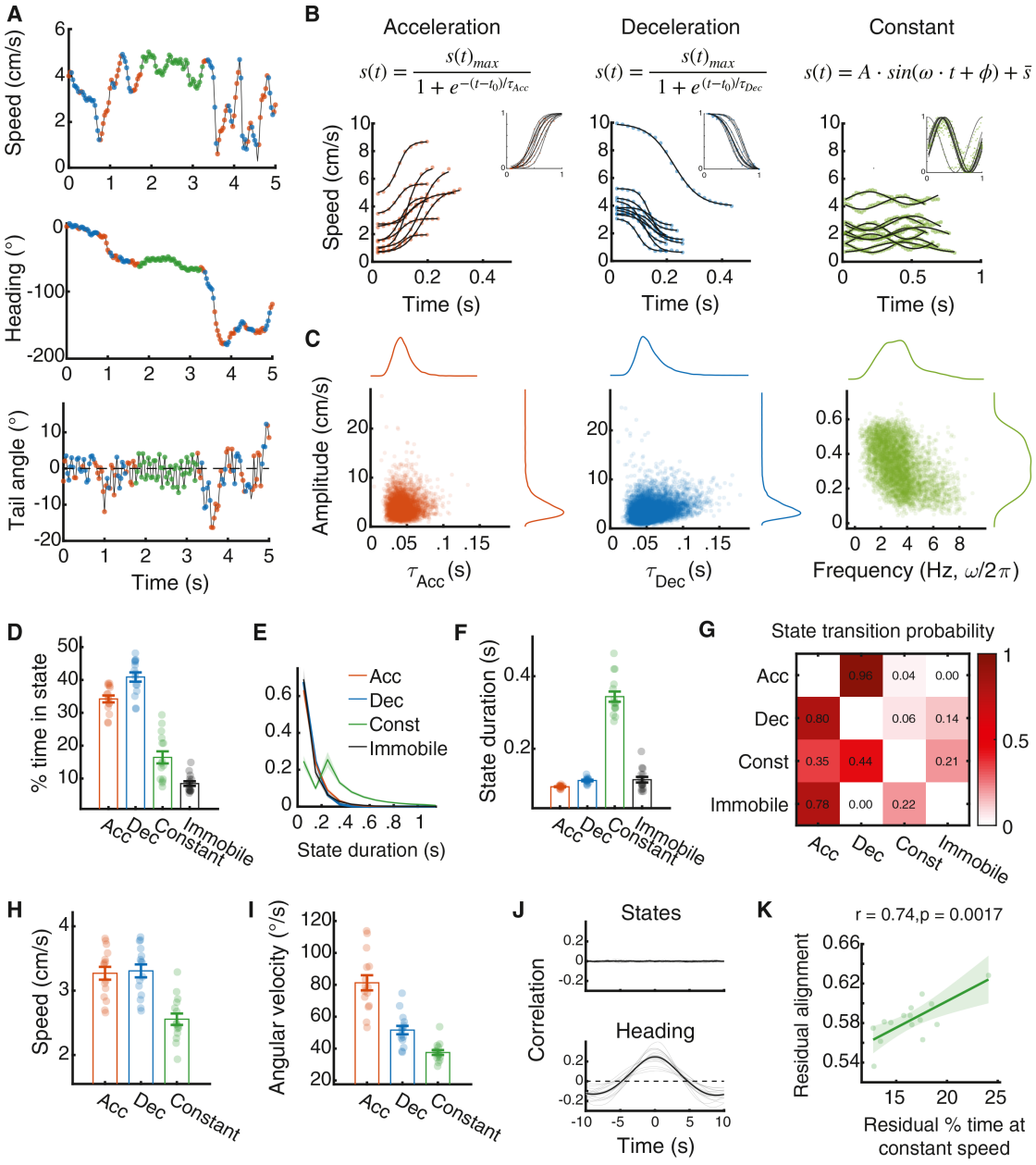


Figure 2. Medaka utilize distinct kinematic swim states

(A) Examples of the speed, heading direction and tail angle deviation from the body axis of an adult medaka swimming in a group of five fish. Red, blue, and green colors represent acceleration, deceleration, and constant-speed epochs (Figure S2B; STAR Methods).

(B) Examples of speed traces showing acceleration, deceleration, and constant-speed epochs (colored dots) along with their best fits (black lines) when using increasing sigmoidal, decreasing sigmoidal, and sine wave functions respectively. Inset: same fits as shown in the main figures, but with time and speed magnitude normalized between 0 and 1. Normalized sine waves were also inverted to align the direction of their initial oscillation.

(C) Speed amplitude ($|S_{max} - S_{min}|$) plotted vs. function parameters inferred from fitted segmented speed epochs (as in B). Colored lines show the marginal distributions of the parameters.

(D) Mean percent time spent in each of the segmented states for groups of adult medaka.

(E) Distribution of state durations for each of the kinematic states. Solid lines are averages over fish; shaded areas are standard deviations. Bin width = 0.1 s.

(F) Average duration of the segmented states (same groups as in D).

(G) Transition matrix showing mean probability of transitions from a given state (rows) to each of the other states (columns). Probabilities are calculated over rows; averages are over groups.

(H and I) Speed (H) and angular velocity (I) of adult medaka in the three kinematic states.

(J) Mean pairwise cross correlation of kinematic states (top) and heading direction (bottom) of pairs of adult fish, swimming in groups of 5 fish. Lines represent averages of all pairs in a group; bold lines and shaded areas are the mean ± SEM values across groups.

(legend continued on next page)

Please cite this article in press as: Harpaz et al., Social interactions in medaka fish depend on discrete kinematic states of swimming behavior, Current Biology (2026), https://doi.org/10.1016/j.cub.2026.01.039

the measured tail beat frequencies (Figure S2D), with narrow amplitude variations around the mean speed (0.35 ± 0.13) (Figures 2C and 2H). Higher oscillation frequencies tended to have lower amplitudes (Spearman's correlation = −0.48, $p < 0.0001$, permutation test).

Next, we analyzed how these discrete states shape the swimming behavior of the fish. Adult medaka spent 16.4% ± 7.1% [12.5, 20.4] (mean ± SD, 95% CI in brackets) of their swim time in the constant-speed state, compared with 34.2% ± 4.1% [32, 36.5] and 40.9% ± 5.5% [37.9, 43.9] in the acceleration and deceleration states (fish are considered immobile if speed is <0.6 cm/s) (Figure 2D). The average duration of constant-speed epochs, however, was ∼3.6 and 3 times longer than acceleration and deceleration events respectively (Figures 2E and 2F). In contrast, adult zebrafish, which are known to utilize "burst-and-coast" kinematics, are rarely found in a similar constant-speed state (<3%) (Figures S2A and S2E; Video S3). Analyzing the transitions between states, we find that medaka frequently alternate between acceleration and deceleration states, transitioning less frequently into the constant-speed state, which was sustained for longer durations (Figures 2G and S2F). Average swimming speed was similar across acceleration and deceleration events ($speed_{Acc}$ = 3.27 ± 0.39 cm/s [3.06, 3.48], $speed_{Dec}$ = 3.31 ± 0.39 cm/s [3.09, 3.52]) (Figure 2H). Yet speed was significantly slower in the constant-speed state ($speed_{Const}$ = 2.55 ± 0.34 cm/s [2.37, 2.75]), indicating that fish maintain slower speeds for longer durations. Turning was most strongly associated with accelerations and weakest during constant-speed epochs ($turning_{Acc}$ = 81 ± 18°/s [71.1, 91.4], $turning_{dec}$ = 52 ± 10°/s [46, 57.2], and $turning_{Const}$ = 38 ± 6°/s [34.3, 40.8]) (Figure 2I). We did not observe any differences in distances and angles to the walls of the arena between the three states (Figure S2G). Interestingly, while the heading direction between pairs of fish in the group is correlated, the timing of state transitions was not (Figure 2J).

To study how individual swim statistics are related to medaka's tendency to swim in aligned formations, we regressed group alignment values on mean speed, mean nearest-neighbor distance, and the fraction of time spent at constant speed. We find that all predictors contributed significantly to the model: higher alignment was associated with faster swimming (β_{speed}= 0.12 [0.06, 0.17]) closer spacing (β_{NNd} = −0.05 [−0.08, −0.03]), and more time spent at the constant-speed state ($\beta_{\%\ time\ in\ constant\ speed}$ = 0.005 [0.002, 0.009]) (see STAR Methods for complete model details). Notably, the association between time at constant speed and alignment remained strong after accounting for both speed and nearest neighbor distances (r = 0.74, p = 0.003, partial correlation coefficient) (Figures 2K and S2H).

Together, these findings demonstrate that while medaka utilize continuous body and tail undulations for swimming, they employ at least three discrete kinematic states to regulate their swimming behaviors. Next, we assessed how these kinematic states also influence the inter-individual interactions underlying collective behaviors.

Distinct swim states imply varying levels of social responsiveness

Next, we study the relation between the distinct kinematic states of medaka and their responses to social information from neighbors. Based on the measured spatial structure of neighbor positions and relative headings (Figures 1G and 1H), we opted to utilize a previously suggested family of models in which the change in velocity $\Delta\vec{v}_i$ of fish i depends on the spatiotemporal properties of its neighbors.[15] Specifically, we decompose the change in velocity of fish i at time t into two components (Figure 3A):

$$\Delta\vec{v}_i(t) = \Delta\vec{v}_i^{\,Active}(t) + \Delta\vec{v}_i^{\,Passive}(t). \quad \text{(Equation 1)}$$

The passive component represents velocity decay due to water friction, which is defined as:

$$\Delta\vec{v}_i^{\,Passive}(t) = -\eta\vec{v}_i(t-\tau_c), \quad \text{(Equation 2)}$$

where η is the friction coefficient taken as the inverse of the decay time constant of exponential fits to deceleration segments $1/\langle\tau_{Dec}\rangle$ (STAR Methods), and τ_c is a fixed decay timescale, estimated from the autocorrelation function of the speed time series (Figure S3A; STAR Methods).

The active component $\Delta\vec{v}_i^{\,Active}(t)$ represents the response of the fish to social and sensory stimuli and is defined here as a spatiotemporal filter (or "receptive field") of the velocities of neighboring fish and the direction of the closest walls of the arena:

$$\Delta\vec{v}_i^{\,Active}(t) = \sum_{j,k}\beta_j(k)\cdot\vec{v}_j(t-\Delta t_k) + \sum_{l,k}\beta_l(k)\cdot\vec{d}_l(t-\Delta t_k), \quad \text{(Equation 3)}$$

where $\vec{v}_j$ is the average velocity of neighbors in the spatiotemporal bin j at time $t-\Delta t_k$, and $\vec{d}_l$ is the direction tangent to the wall closest to the fish. $\beta_j(k)$ and $\beta_l(k)$ are the weights assigned to each spatiotemporal bin for both social and sensory information respectively (Figure 3B). In addition, we assume that a unique set of model weights $\beta_j(k)^{state}$ and $\beta_l(k)^{state}$ define each of the kinematic states of the fish—acceleration, deceleration, and constant speed.

To test the ability of the models to predict fish active swimming responses, and to test for differences in social interactions between the states, we train a unique model for each group in each of the 3 kinematic states and test how well the predicted motion vectors $\Delta\vec{v}_i^{\,model}(t)$ describe the active changes in fish swim velocities $\Delta\vec{v}_i^{\,Active}(t)$(Figures 3C and 3D; STAR Methods). While all states show above-chance prediction accuracies (defined as the Pearson correlation between real and predicted velocities), we find clear differences between the state-dependent models: deceleration epochs showed the lowest prediction accuracies (0.31 ± 0.02 [0.29, 0.33]; mean ± SD, 95% CI in brackets), acceleration events were predicted with higher accuracies (0.44 ± 0.02 [0.43, 0.46]), and the constant-speed epochs had the highest

(K) Relationship between group alignment and percentage time at constant speed, after controlling for variation in speed and nearest neighbor distances. Dots represent groups, solid line is the best fit linear model, and shaded area is the 95% CI of the model (STAR Methods).
In (D), (F), (H), and (I), dots represent average values of different groups, and error bars are mean ± SEM.
See also Figure S2 and Videos S2 and S3.

Please cite this article in press as: Harpaz et al., Social interactions in medaka fish depend on discrete kinematic states of swimming behavior, Current Biology (2026), https://doi.org/10.1016/j.cub.2026.01.039

CellPress
OPEN ACCESS

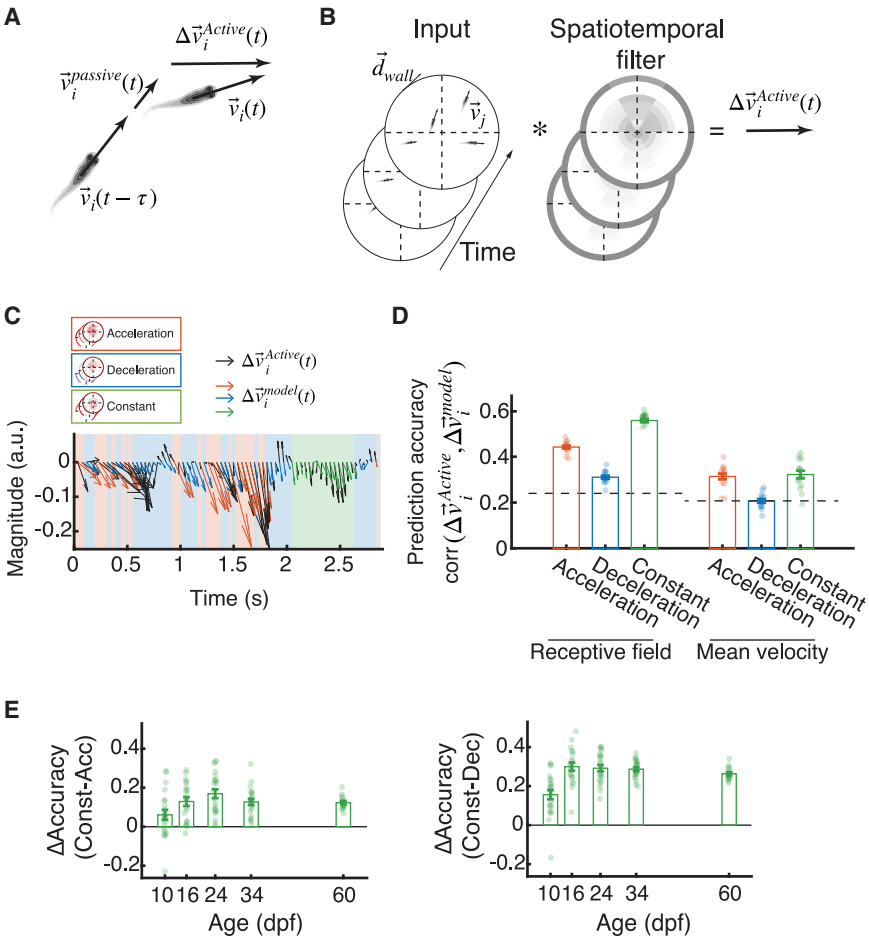


Figure 3. Distinct swim states imply varying levels of social responsiveness

(A) The velocity of fish *i* is decomposed into passive and active components (Equations 1 and 2).

(B) The active motion component $\Delta\vec{v}_i^{\,Active}$ is modeled as the weighted spatiotemporal integration of neighboring fish velocities $\vec{v}_j$ and the direction of the closest walls of the arena $\vec{d}_{wall}$ (Equation 3). Grayscale values represent different weights for different spatiotemporal bins; the outer ring represents weights for arena walls.

(C) Active responses of an example fish—$\Delta\vec{v}_i^{\,Active}(t)$ (black arrows) for different kinematic states (colored background) and the predicted velocities $\Delta\vec{v}_i^{\,Model}(t)$. At each time point, the model used for prediction matches the kinematic state of the fish (colored arrows).

(D) Prediction accuracy measured as the correlation between the active motion component of the fish $\Delta\vec{v}_i^{\,Active}$ and the predicted active component by the models $\Delta\vec{v}_i^{\,model}$. Models are learned separately for each state (see Figure 4A) and accuracy is calculated using held-out test data not used for training. "Receptive field" is the full spatiotemporal model presented in (B), and "mean velocity" is a simpler model that ignores the spatial positions of neighbors (Equation 5; STAR Methods). Black dashed lines represent the accuracy of models fitted to shuffled data (STAR Methods).

(E) Difference in prediction accuracy between the constant-speed state and acceleration (left) and deceleration (right) states for different ages when using the full spatiotemporal models for prediction.

In (D and E), dots represent groups and error bars are mean ± SEM.

See also Figure S3.

prediction accuracy of the three states (0.56 ± 0.03 [0.54, 0.57]) (Figure 3D, left—receptive field). Similarly, the relative contribution of the active component over the passive one was highest in the constant-speed state, lower in the acceleration state, and lowest for decelerations (Figure S3B).

To test the importance of spatial information in the fish's computations, we compared these accuracies to a baseline model where we predict the change in velocity of fish *i* as a simple average of all neighbor velocities (and wall direction), regardless of their relative spatial positions (Equation 5; STAR Methods). This "mean velocity" model had markedly lower prediction accuracies compared with the full receptive field model ($C_{acceleration}$= 0.31 ± 0.05 [0.29, 0.33], $C_{deceleration}$ = 0.2 ± 0.034 [0.19, 0.21], and C_{const}= 0.31 ± 0.06 [0.29, 0.33]), and no difference in accuracy between acceleration and constant-speed epochs (Figure 3D, right—mean velocity), emphasizing the importance of spatiotemporal information in the fish's social interaction computations.

Importantly, segmenting fish swimming into only two states (acceleration and deceleration only; STAR Methods) in the context of the full spatiotemporal model, "averages out" the improvements in accuracy we observed in the constant-speed state, resulting in prediction accuracies that are similar to those found for deceleration and acceleration events in the 3 state models (Figure S3C). A two-state model that treats acceleration/decelerations as a single mode and constant speed as a second mode is also less accurate than the full, three-state model (Figure S3C). Training a single model for the entire dataset again results in intermediate accuracy values, which are expected due to averaging (Figure S3C), emphasizing the importance of the different social interaction states.

To confirm that prediction accuracy indeed depends on state-specific social interactions in the different kinematic states, and not on inherent differences in complexity of the executed swims, we fitted our models to two control datasets. First, we created a "shuffled identity" dataset that retains the individual swim kinematics of the fish, yet randomizes available social information, by mixing fish identities from different groups (STAR Methods). We find that models trained on this shuffled dataset yield prediction accuracies that are at chance levels with almost no noticeable differences between the different kinematic states (Figure S3D). Second, we created for every group two complexity-matched datasets by subsampling the acceleration and deceleration states such that the distributions of $\Delta\vec{v}_i^{\,Active}(t)$ is similar across the 3 states (Figure S3E, left). We find that models trained on these complexity-matched datasets give very similar prediction accuracies to those from the original datasets, with similar differences in accuracies between states (Figure S3E, right). Inspecting the ontogeny of state-dependent social interaction

Please cite this article in press as: Harpaz et al., Social interactions in medaka fish depend on discrete kinematic states of swimming behavior, Current Biology (2026), https://doi.org/10.1016/j.cub.2026.01.039

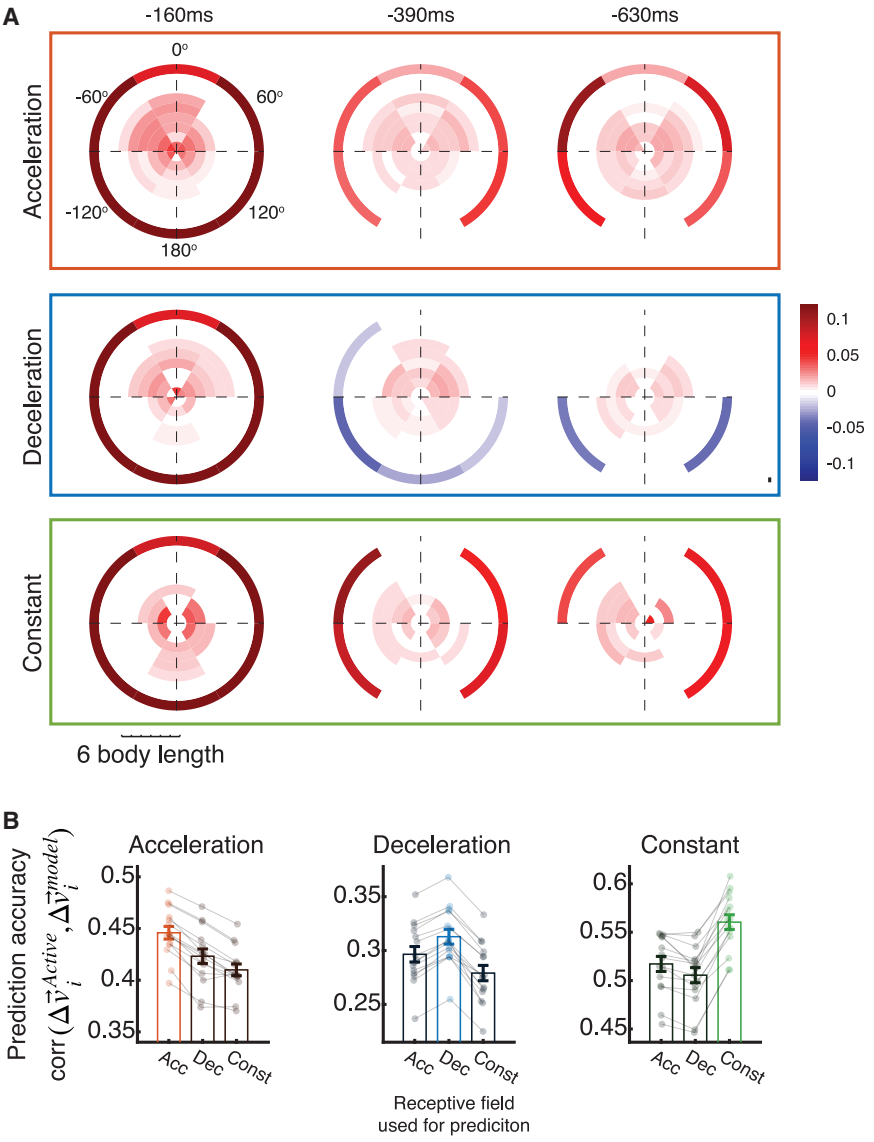


Figure 4. Fish utilize different spatiotemporal filters in each of the kinematic states

(A) Learned spatiotemporal weights for the three kinematic states (rows). Each circular map represents the weights assigned to neighbor velocities at a specific distance and angle from a focal fish situated in the center of the map pointing north. Columns show maps at increasing Δt offsets. The outer ring in each map represents the weights assigned to the wall of the arena at different angles. We only show weights that are significantly different from 0 ($p < 0.05$; t test; STAR Methods). (B) Prediction accuracy for each state (separate panels) using models trained in-state (light-colored bars) or out-of-state (dark-colored bars). For example, in the acceleration panel (left), the light-colored bar shows prediction accuracy for held-out velocities using the RF model trained on acceleration data, whereas the dark bars show accuracy for out-of-state predictions using RF models trained on the other two states. In-state predictions are consistently and significantly more accurate, regardless of the specific group or state analyzed, demonstrating the state-specificity of learned spatiotemporal filters. $p = 0.00006$ for all six in-state vs. out-of-state comparisons (Wilcoxon's sign rank test); effect sizes: Cohen's $d_{Acceleration} = 0.89$ and 1.57 (vs. deceleration and constant), Cohen's $d_{decelerations} = 0.6$ and 1.25 (vs. acceleration and constant), and Cohen's $d_{constant} = 1.45$, and 1.83 (vs. acceleration and deceleration) (STAR Methods). In all images, dots represent groups and error bars are mean ± SEM. See also Figure S4.

computations, we found that the improved prediction accuracy in the constant-speed state compared to the acceleration and deceleration states was consistent throughout development. Specifically, models fitted for individual groups, at different ages (Figure S3F), showed an increase in accuracy in the constant-speed state over the deceleration state ($\Delta Accuracy_{Const-Dec}$ = 0.16 ± 0.11 [0.11, 0.21], 0.3 ± 0.09 [0.26, 0.34], 0.29 ± 0.07 [0.26, 0.33], 0.29 ± 0.05 [0.27, 0.31], and 0.26 ± 0.03 [0.24, 0.28] for ages 10, 16, 24, 34, and 60 dpf, respectively), and over the acceleration state ($\Delta Accuracy_{Const-Acc}$ = 0.06 ± 0.12 [0.01, 0.11], 0.13 ± 0.1 [0.08, 0.18], 0.17 ± 0.11 [0.12, 0.22], 0.13 ± 0.07 [0.09, 0.16], and 0.12 ± 0.04 [0.10, 0.14]) from very early in development and became stable around 24 dpf onward (Figure 3E).

Taken together, these results indicate that the three kinematic swim states are indicative of different social information processing states, with the continuous constant-speed state showing the strongest social interaction tendencies.

Fish utilize different spatiotemporal filters in each of the kinematic states

We next assessed the structure of the spatiotemporal filters (receptive fields) fish use in each of the kinematic states (Figure 4A). In all states, the magnitude of the weights assigned to neighbor velocities was strongest at short temporal delays and gradually decreased as the temporal offset between neighbor velocities and focal fish responses increased (compare delays of −160, −390, and −630 ms in Figure 4A). Spatially, the learned weights in the acceleration state extended up to 6 body lengths directly in front of the fish (Figure 4A, top row, first column, 0° sector) but were weaker behind it (sectors 180°, 120°, and −120°). By contrast, in the constant-speed state, high-magnitude weights were tightly localized, reaching only about 2–3 body lengths in front of and to the sides of the focal fish but extended farther behind it (Figure 4A, bottom row, left column). Deceleration events showed a spatial structure similar to that of acceleration events at short temporal delays (−160 ms) but appeared to decay more quickly at longer delays (−390 and −630 ms) (Figure 4A, middle row). Responses to arena walls (Figure 4A, colored rings surrounding the receptive field) were strong in short temporal delays in all states, consistent with previous reports in other fish species.[15,40] These differences in weights,

Please cite this article in press as: Harpaz et al., Social interactions in medaka fish depend on discrete kinematic states of swimming behavior, Current Biology (2026), https://doi.org/10.1016/j.cub.2026.01.039

together with the fact that group structure did not markedly differ between states (Figure S4A), suggest that fish differentially attend to their social surroundings in each kinematic state.

To confirm that fish indeed employ unique spatiotemporal filters in each kinematic state and to quantify their effect on behavior, we compared the accuracy of model predictions when using each of the three learned models shown in Figure 4A to predict changes in velocity $\Delta \overrightarrow{v}_i^{Active}(t)$ in each of the states (Figures 4B and S4B). We show that model accuracy is significantly higher (and prediction error is consistently lower) when the model used for prediction matches the kinematic state of the target velocity vector, as evaluated on held-out data across all states (Figures 4B and S4B, see statistics in figure legend). For example, predicting target velocities in the constant-speed epochs using models learned from acceleration or deceleration epochs, resulted in a 0.043 ± 0.01 [0.037, 0.046] and 0.055 ± 0.02 [0.043, 0.055] (mean ± SD, 95% CI in brackets) decrease in accuracy, respectively, compared with the predictions of models trained on constant-speed data (Figure 4B, right). This pattern is consistent across all kinematic states and experimental groups (Figures 4B and S4B). Moreover, the spatial structure of the filters predicts that model accuracy should increase when group structure is tight and neighbors are close by, and that this effect should be most pronounced for the constant-speed state given the filter's tight spatial structure. To test these hypotheses, we separated our data to instances where fish swim in close proximity to neighbors and to instances when they are far away (bottom and top 20% of NNd distributions; STAR Methods). We find that the swimming behavior of the fish changed in response to neighbor proximity: fish swam slower and turned more when in close proximity to their neighbors, and that speed changes were most pronounced in the constant-speed state, while turning differences were strongest during acceleration and decelerations (Figure S4C). Fitting our models to the different states in these two proximity scenarios we find, as predicted from the filters' structure, that proximity to neighbors improved model accuracy, and that this improvement was strongest during the constant-speed state (Figure S4D).

In summary, the distinct models learned for each state confirm that fish process and respond to spatiotemporal social information in a state-dependent manner.

DISCUSSION

In the current study, we demonstrate that the continuous undulatory swimmer, medaka, exhibits robust schooling behaviors that develop early, around 2 weeks of age, and rapidly stabilize, reaching adult-level values by the juvenile age of 1 month. These findings establish medaka as a valuable model species for studying the development and mechanisms of collective social behaviors.[23–25,41,42] Continuous swimming in medaka can be naturally segmented into discrete acceleration, deceleration, and constant-speed states, each associated with distinct social interaction computations. The tendency to respond to social information was stronger during accelerations than decelerations but was strongest during constant-speed epochs. At the group level, we find that the social interaction modes are temporally uncorrelated among individuals; yet, group alignment levels correlate with the relative frequency with which fish occupy each state. These findings reveal state-dependent social responses of an individual to its conspecifics, which can now be leveraged in future studies to explain collective behaviors in this species.

While previous results suggest that executed behavioral responses occur predominantly during accelerations, and less so during decelerations,[12,15,18] the marked increase in responsiveness to social information in the constant-speed state has not been described previously. We hypothesize that maintaining slow, constant swim speeds enhances responsiveness to social cues, possibly due to enhanced sensorimotor processing during prolonged steady swimming. Future studies can test whether this improved responsiveness is specific to social information or represents a general enhancement of sensory processing, also to non-social stimuli.[43–47]

Furthermore, additional behavioral states beyond the three identified in this study likely influence the social behaviors of this species.[19,47,48] For example, fish may utilize different behavioral rules, or modify the usage of these kinematic states in specific contexts such as during foraging[19,31,49] or in the presence of predators.[37,50] Moreover, current or previous social densities and social experience may modulate behavioral states of individuals either in a discrete or continuous manner.[16,51] We also note that while we focused on quasi-2D behaviors, depth-dependent effects might also modulate social behaviors in madeka.[52,53] Future research can now systematically explore how these various factors shape medaka collective behaviors.

The ontogeny of collective social behavior in medaka can be directly compared with previous findings in zebrafish, another freshwater fish species of a similar size, which utilizes intermittent kinematics.[8,12,38,39,54]

Zebrafish larvae exhibit predominantly repulsive interactions at a young age of ~7 dpf, developing weak attraction tendencies at 14 dpf, which grow stronger at 21 dpf.[8,12,38,54,55] Adult zebrafish utilize burst-and coast kinematics that can be parsed into "active" and "passive" social interaction states,[15] with spatiotemporal interactions strongly dependent on the instantaneous speed magnitude of the fish.[56]

Social behavior in medaka follows a different trajectory—larvae show no signs of purely repulsive tendencies at early ages, with robust social alignment developing already at 16 dpf (Figure 1D). Medaka's continuous kinematics can be segmented into three states, all classified as active, even decelerations, and social interactions show strong dependence on these prolonged states but are less dependent on absolute speed magnitude, as described in zebrafish.[56] Importantly, since behavior in these species diverges already at the young larval stages, when brain-wide calcium imaging and optogenetics circuit dissection are widely available,[57–60] our findings provide an opportunity to study differences and commonalities in the underlying circuit mechanisms.[61,62]

Notably, we found that individual swim kinematics and group-level collective patterns in medaka follow distinct developmental trajectories. As fish mature, the fraction of time spent at constant speed decreases, whereas group aggregation and alignment increase. This dissociation suggests that kinematic states serve multiple functions, beyond social interactions, such as energy economy,[32,34] oxygen uptake,[63] and possibly increased

sensitivity to non-social stimuli.[64,65] Consistent with this view, state-transition timing is largely uncorrelated across individuals, yet group aggregation and alignment still emerge. Thus, collective organization can arise from noisy, uncoordinated individual dynamics, in line with prior findings in adult zebrafish[15] and with theoretical models.[6]

To analyze the social interaction computations underlying collective behaviors in medaka, we adapted a family of receptive field models previously validated in zebrafish.[15] Despite considerable behavioral differences at both individual and collective levels, we found that this modeling framework effectively captures medaka social responses, emphasizing the general applicability of these models in studying collective behaviors. The fitted models allowed us to characterize how fish utilize spatiotemporal information from neighbors, revealing distinct receptive fields for each kinematic state. These insights can inform future theoretical studies investigating how diverse social interaction states shape collective behaviors in continuous undulatory swimmers, as was previously done for fish utilizing burst-and-coast swimming.[9,35,66–68]

Recent technological advances have positioned medaka as a promising model system for linking neural circuit mechanisms and emergent adaptive behaviors.[61,69–72] Our results pave the way for studying such mechanisms in the context of social and collective behaviors in this important model system.

RESOURCE AVAILABILITY

Lead contact

Requests for further information and resources should be directed to the lead contact, Roy Harpaz (harpazone@gmail.com).

Materials availability

This study did not generate new materials.

Data and code availability

- All data used in this study are publicly available at https://dataverse.harvard.edu/dataset.xhtml?persistentId=doi:10.7910/DVN/QTNPJL.
- This study did not generate new code.
- Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

ACKNOWLEDGMENTS

We thank all members of the Engert and Fishman labs for their support and advice throughout this project. We particularly thank Kumaresh Krishnan and Cristina Santoriello for their valuable assistance in establishing the behavioral setups and Emily Chan for her help with medaka breeding. We also thank Prof. Elad Schneidman for valuable discussions and insights. Finally, we thank Ed Soucy, Brett Graham, and Yuwei Li from the Neuroengineering Core Facility at Harvard's Center for Brain Science for their technical support. R.H. received funding from Harvard Minds Brain and Behavior Initiative. F.E. received funding from the National Institutes of Health (U19NS104653 and 1R01NS124017-01), and the Simons Foundation (SCGB 542973 and NC-GB-CULM-00003241-02).

AUTHOR CONTRIBUTIONS

R.H., P.P., Y.I., M.C.F., and F.E. designed the research; R.H., P.P., and Y.I. performed experiments; R.H., P.P., and F.E. analyzed the data; and R.H., P.P., Y.I., M.C.F., and F.E. wrote the manuscript.

DECLARATION OF INTERESTS

The authors declare no competing interests.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS
  - Fish Husbandry
- METHOD DETAILS
  - Behavioral experiments
  - Tracking
  - Midline extraction and tail tracking
  - Individual and group level swim properties
  - Neighbor Maps
  - Segmenting fish speed profiles into discrete states
  - Function fitting to kinematic states
  - Operational definition of social responsiveness
  - 'Receptive field' models to predict active velocity changes of individuals
  - 'Mean velocity' model to predict active velocity changes of individuals
  - Surrogate data of shuffled groups
  - Complexity matched datasets
  - Model parameters
  - Weight map plots
- QUANTIFICATION AND STATISTICAL ANALYSIS
  - Linear models linking group properties to kinematic features
  - Sample sizes, trial numbers and power estimation
  - Statistical testing

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.cub.2026.01.039.

Received: July 25, 2025
Revised: December 4, 2025
Accepted: January 16, 2026

REFERENCES

1. Couzin, I.D., Krause, J., James, R., Ruxton, G.D., and Franks, N.R. (2002). Collective memory and spatial sorting in animal groups. J. Theor. Biol. *218*, 1–11. https://doi.org/10.1006/jtbi.2002.3065.
2. Aoki, I. (1982). A simulation study on the schooling mechanism in fish. Nippon Suisan Gakkaishi *48*, 1081–1088. https://doi.org/10.2331/suisan.48.1081.
3. Huth, A., and Wissel, C. (1994). The simulation of fish schools in comparison with experimental data. Ecol. Modell. *75–76*, 135–146. https://doi.org/10.1016/0304-3800(94)90013-2.
4. Huth, A., and Wissel, C. (1992). The simulation of the movement of fish schools. J. Theor. Biol. *156*, 365–385. https://doi.org/10.1016/S0022-5193(05)80681-2.
5. Vicsek, T., Czirók, A., Ben-Jacob, E., Cohen, I., and Shochet, O. (1995). Novel type of phase transition in a system of self-driven particles. Phys. Rev. Lett. *75*, 1226–1229. https://doi.org/10.1103/PhysRevLett.75.1226.
6. Jhawar, J., Morris, R.G., Amith-Kumar, U.R., Danny Raj, M., Rogers, T., Rajendran, H., and Guttal, V. (2020). Noise-induced schooling of fish. Nat. Phys. *16*, 488–493. https://doi.org/10.1038/s41567-020-0787-y.
7. Katz, Y., Tunstrøm, K., Ioannou, C.C., Huepe, C., and Couzin, I.D. (2011). Inferring the structure and dynamics of interactions in schooling fish. Proc.

Natl. Acad. Sci. USA *108*, 18720–18725. https://doi.org/10.1073/pnas.1107583108.

8. Hinz, R.C., and de Polavieja, G.G. de (2017). Ontogeny of collective behavior reveals a simple attraction rule. Proc. Natl. Acad. Sci. USA *114*, 2295–2300. https://doi.org/10.1073/pnas.1616926114.
9. Calovi, D.S., Litchinko, A., Lecheval, V., Lopez, U., Pérez Escudero, A.P., Chaté, H., Sire, C., and Theraulaz, G. (2018). Disentangling and modeling interactions in fish with burst-and-coast swimming reveal distinct alignment and attraction behaviors. PLoS Comput. Biol. *14*, e1005933. https://doi.org/10.1371/journal.pcbi.1005933.
10. Cavagna, A., and Giardina, I. (2010). Large-scale behaviour in animal groups. Behav. Processes *84*, 653–656. https://doi.org/10.1016/j.beproc.2010.02.026.
11. Sumpter, D.J.T. (2010). Collective Animal Behavior (Princeton University Press). https://doi.org/10.1515/9781400837106.
12. Harpaz, R., Nguyen, M.N., Bahl, A., and Engert, F. (2021). Precise visuomotor transformations underlying collective behavior in larval zebrafish. Nat. Commun. *12*, 6578. https://doi.org/10.1038/s41467-021-26748-0.
13. Vicsek, T., and Zafeiris, A. (2012). Collective motion. Phys. Rep. *517*, 71–140. https://doi.org/10.1016/j.physrep.2012.03.004.
14. Sridhar, V.H., Li, L., Gorbonos, D., Nagy, M., Schell, B.R., Sorochkin, T., Gov, N.S., and Couzin, I.D. (2021). The geometry of decision-making in individuals and collectives. Proc. Natl. Acad. Sci. USA *118*, e2102157118. https://doi.org/10.1073/pnas.2102157118.
15. Harpaz, R., Tkačik, G., and Schneidman, E. (2017). Discrete modes of social information processing predict individual behavior of fish in a group. Proc. Natl. Acad. Sci. USA *114*, 10149–10154. https://doi.org/10.1073/pnas.1703817114.
16. Harpaz, R., Phillips, M., Goel, R., Fishman, M.C., and Engert, F. (2024). Experience-dependent modulation of collective behavior in larval zebrafish. Preprint at bioRxiv. https://doi.org/10.1101/2024.08.02.606403.
17. Stednitz, S.J., Lesak, A., Fecker, A.L., Painter, P., Washbourne, P., Mazzucato, L., and Scott, E.K. (2025). Coordinated social interaction states revealed by probabilistic modeling of zebrafish behavior. Curr. Biol. *35*, 2903–2915.e6. https://doi.org/10.1016/j.cub.2025.05.031.
18. Bod'ová, K., Mitchell, G.J., Harpaz, R., Schneidman, E., and Tkačik, G. (2018). Probabilistic models of individual and collective animal behavior. PLoS One *13*, e0193049. https://doi.org/10.1371/journal.pone.0193049.
19. Johnson, R.E., Linderman, S., Panier, T., Wee, C.L., Song, E., Herrera, K.J., Miller, A., and Engert, F. (2020). Probabilistic models of larval zebrafish behavior reveal structure on many scales. Curr. Biol. *30*, 70–82.e4. https://doi.org/10.1016/j.cub.2019.11.026.
20. Wiltschko, A.B., Johnson, M.J., Iurilli, G., Peterson, R.E., Katon, J.M., Pashkovski, S.L., Abraira, V.E., Adams, R.P., and Datta, S.R. (2015). Mapping sub-second structure in mouse behavior. Neuron *88*, 1121–1135. https://doi.org/10.1016/j.neuron.2015.11.031.
21. Berman, G.J., Choi, D.M., Bialek, W., and Shaevitz, J.W. (2014). Mapping the stereotyped behaviour of freely moving fruit flies. J. R. Soc. Interface *11*, 20140672. https://doi.org/10.1098/rsif.2014.0672.
22. Stephens, G.J., Johnson-Kerner, B., Bialek, W., and Ryu, W.S. (2008). Dimensionality and dynamics in the behavior of C. elegans. PLoS Comput. Biol. *4*, e1000028. https://doi.org/10.1371/journal.pcbi.1000028.
23. Leem, J.-B., Jeon, W., Yun, C.-Y., and Lee, S.-H. (2012). Quantitative analysis of fish schooling behavior with different numbers of medaka (Oryzias latipes) and goldfish (Carassius auratus). Ocean Sci. J. *47*, 445–451. https://doi.org/10.1007/s12601-012-0040-4.
24. Nakamura, Y. (1952). Some experiments on the shoaling reaction in Oryzias latipes (TEMMINCK et SCHLEGEL). Nippon Suisan Gakkaishi *18*, 93–101. https://doi.org/10.2331/suisan.18.93.
25. Isoe, Y., Konagaya, Y., Yokoi, S., Kubo, T., and Takeuchi, H. (2016). Ontogeny and sexual differences in swimming proximity to conspecifics in response to visual cues in medaka fish. Zool. Sci. *33*, 246–254. https://doi.org/10.2108/zs150213.
26. Marques, J.C., Lackner, S., Félix, R., and Orger, M.B. (2018). Structure of the zebrafish locomotor repertoire revealed with unsupervised behavioral clustering. Curr. Biol. *28*, 181–195.e5. https://doi.org/10.1016/j.cub.2017.12.002.
27. Lauder, G.V. (2015). Fish locomotion: recent advances and new directions. Annu. Rev. Mar. Sci. *7*, 521–545. https://doi.org/10.1146/annurev-marine-010814-015614.
28. Videler, J.J. (1993). Fish Swimming (Springer). https://doi.org/10.1007/978-94-011-1580-3.
29. Videler, J.J., and Weihs, D. (1982). Energetic advantages of burst-and-coast swimming of fish at high speeds. J. Exp. Biol. *97*, 169–178. https://doi.org/10.1242/jeb.97.1.169.
30. Rajan, G., Lafaye, J., Faini, G., Carbo-Tano, M., Duroure, K., Tanese, D., Panier, T., Candelier, R., Henninger, J., Britz, R., et al. (2022). Evolutionary divergence of locomotion in two related vertebrate species. Cell Rep. *38*, 110585. https://doi.org/10.1016/j.celrep.2022.110585.
31. Mearns, D.S., Hunt, S.A., Schneider, M.W., Parker, A.V., Stemmer, M., and Baier, H. (2025). Diverse prey capture strategies in teleost larvae. eLife *13*, RP98347. https://doi.org/10.7554/eLife.98347.
32. Li, G., Kolomenskiy, D., Liu, H., Godoy-Diana, R., and Thiria, B. (2023). Intermittent versus continuous swimming: An optimization tale. Phys. Rev. Fluids *8*, 013101. https://doi.org/10.1103/PhysRevFluids.8.013101.
33. Weihs, D. (1974). Energetic advantages of burst swimming of fish. J. Theor. Biol. *48*, 215–229. https://doi.org/10.1016/0022-5193(74)90192-1.
34. Ashraf, I., Van Wassenbergh, S., and Verma, S. (2020). Burst-and-coast swimming is not always energetically beneficial in fish (Hemigrammus bleheri). Bioinspir. Biomim. *16*, 016002. https://doi.org/10.1088/1748-3190/abb521.
35. Amichay, G., Li, L., Nagy, M., and Couzin, I.D. (2024). Revealing the mechanism and function underlying pairwise temporal coupling in collective motion. Nat. Commun. *15*, 4356. https://doi.org/10.1038/s41467-024-48458-z.
36. Peravali, R. (2018). Morphometric and quantitative behavioral analysis of inbred medaka lines. PhD thesis (Ruperto-Carola University of Heidelberg). https://doi.org/10.11588/heidok.00024134.
37. Miller, N., and Gerlai, R. (2007). Quantification of shoaling behaviour in zebrafish (Danio rerio). Behav. Brain Res. *184*, 157–166. https://doi.org/10.1016/j.bbr.2007.0h557.007.
38. Dreosti, E., Lopes, G., Kampff, A.R., and Wilson, S.W. (2015). Development of social behavior in young zebrafish. Front. Neural Circuits *9*, 39. https://doi.org/10.3389/fncir.2015.00039.
39. Miller, N., and Gerlai, R. (2012). From schooling to shoaling: patterns of collective motion in zebrafish (Danio rerio). PLoS One *7*, e48865. https://doi.org/10.1371/journal.pone.0048865.
40. Gautrais, J., Ginelli, F., Fournier, R., Blanco, S., Soria, M., Chaté, H., and Theraulaz, G. (2012). Deciphering interactions in moving animal groups. PLoS Comput. Biol. *8*, e1002678. https://doi.org/10.1371/journal.pcbi.1002678.
41. Pitcher, T.J. (1986). Functions of shoaling behaviour in teleosts. In The Behaviour of Teleost Fishes, T.J. Pitcher, ed. (Springer), pp. 294–337. https://doi.org/10.1007/978-1-4684-8261-4_12.
42. Radakov, D.V. (1973). Schooling in the Ecology of Fish (John Wiley & Sons Inc).
43. Bolton, A.D., Haesemeyer, M., Jordi, J., Schaechtle, U., Saad, F.A., Mansinghka, V.K., Tenenbaum, J.B., and Engert, F. (2019). Elements of a stochastic 3D prediction engine in larval zebrafish prey capture. eLife *8*, e51975. https://doi.org/10.7554/eLife.51975.
44. Oteiza, P., Odstrcil, I., Lauder, G., Portugues, R., and Engert, F. (2017). A novel mechanism for mechanosensory-based rheotaxis in larval zebrafish. Nature *547*, 445–448. https://doi.org/10.1038/nature23014.
45. Naumann, E.A., Fitzgerald, J.E., Dunn, T.W., Rihel, J., Sompolinsky, H., and Engert, F. (2016). From whole-brain data to functional circuit models: the zebrafish optomotor response. Cell *167*, 947–960.e20. https://doi.org/10.1016/j.cell.2016.10.019.

46. Muto, A., and Kawakami, K. (2013). Prey capture in zebrafish larvae serves as a model to study cognitive functions. Front. Neural Circuits *7*, 110. https://doi.org/10.3389/fncir.2013.00110.

47. Krishnan, K., Muthukumar, A., Sterrett, S., Pflitsch, P., Fairhall, A.L., Fishman, M., Bahl, A., Zwaka, H., and Engert, F. (2025). Attentional switching in larval zebrafish. Sci. Adv. *11*, eads4994. https://doi.org/10.1126/sciadv.ads4994.

48. Sridhar, G., Vergassola, M., Marques, J.C., Orger, M.B., Costa, A.C., and Wyart, C. (2024). Uncovering multiscale structure in the variability of larval zebrafish navigation. Proc. Natl. Acad. Sci. USA *121*, e2410254121. https://doi.org/10.1073/pnas.2410254121.

49. Harpaz, R., and Schneidman, E. (2020). Social interactions drive efficient foraging and income equality in groups of fish. eLife *9*, e56196. https://doi.org/10.7554/eLife.56196.

50. Herbert-Read, J.E., Rosén, E., Szorkovszky, A., Ioannou, C.C., Rogell, B., Perna, A., Ramnarine, I.W., Kotrschal, A., Kolm, N., Krause, J., et al. (2017). How predation shapes the social interaction rules of shoaling fish. Proc. Biol. Sci. *284*, 20171126. https://doi.org/10.1098/rspb.2017.1126.

51. Groneberg, A.H., Marques, J.C., Martins, A.L., Diez del Corral, R., de Polavieja, G.G., and Orger, M.B. (2020). Early-life social experience shapes social avoidance reactions in larval zebrafish. Curr. Biol. *30*, 4009–4021.e4. https://doi.org/10.1016/j.cub.2020.07.088.

52. Ko, H., Girma, A., Zhang, Y., Pan, Y., Lauder, G., and Nagpal, R. (2025). Beyond planar: fish schools adopt ladder formations in 3D. Sci. Rep. *15*, 20249. https://doi.org/10.1038/s41598-025-06150-2.

53. Macrì, S., Neri, D., Ruberto, T., Mwaffo, V., Butail, S., and Porfiri, M. (2017). Three-dimensional scoring of zebrafish behavior unveils biological phenomena hidden by two-dimensional analyses. Sci. Rep. *7*, 1962. https://doi.org/10.1038/s41598-017-01990-z.

54. Larsch, J., and Baier, H. (2018). Biological motion as an innate perceptual mechanism driving social affiliation. Curr. Biol. *28*, 3523–3532.e4. https://doi.org/10.1016/j.cub.2018.09.014.

55. Harpaz, R., Aspiras, A.C., Chambule, S., Tseng, S., Bind, M.-A., Engert, F., Fishman, M.C., and Bahl, A. (2021). Collective behavior emerges from genetically controlled simple behavioral motifs in zebrafish. Sci. Adv. *7*, eabi7460. https://doi.org/10.1126/sciadv.abi7460.

56. Heras, F.J.H., Romero-Ferrero, F., Hinz, R.C., and de Polavieja, G.G. de (2019). Deep attention networks reveal the rules of collective motion in zebrafish. PLoS Comput. Biol. *15*, e1007354. https://doi.org/10.1371/journal.pcbi.1007354.

57. Yu, J.-H., Napoli, J.L., and Lovett-Barron, M. (2024). Understanding collective behavior through neurobiology. Curr. Opin. Neurobiol. *86*, 102866. https://doi.org/10.1016/j.conb.2024.102866.

58. Ahrens, M.B., and Engert, F. (2015). Large-scale imaging in small brains. Curr. Opin. Neurobiol. *32*, 78–86. https://doi.org/10.1016/j.conb.2015.01.007.

59. Kappel, J.M., Förster, D., Slangewal, K., Shainer, I., Svara, F., Donovan, J.C., Sherman, S., Januszewski, M., Baier, H., and Larsch, J. (2022). Visual recognition of social signals by a tectothalamic neural circuit. Nature *608*, 146–152. https://doi.org/10.1038/s41586-022-04925-5.

60. Zada, D., Schulze, L., Yu, J.-H., Tarabishi, P., Napoli, J.L., Milan, J., and Lovett-Barron, M. (2024). Development of neural circuits for social motion perception in schooling fish. Curr. Biol. *34*, 3380–3391.e5. https://doi.org/10.1016/j.cub.2024.06.049.

61. Vohra, S., Herrera, K., Tavhelidse-Suck, T., Wittbrodt, J., Knoblich, S., Seleit, A., Aulehla, A., Boulanger-Weill, J., Chambule, S., Aspiras, A., et al. (2024). Multi-species community platform for comparative neuroscience in teleost fish. Preprint at bioRxiv. https://doi.org/10.1101/2024.02.14.580400.

62. Fouke, K.E., He, Z., Loring, M.D., and Naumann, E.A. (2025). Neural circuits underlying divergent visuomotor strategies of zebrafish and Danionella cerebrum. Curr. Biol. *35*, 2457–2466.e4. https://doi.org/10.1016/j.cub.2025.04.027.

63. Puckett, K.J., and Dill, L.M. (1984). Cost of sustained and burst swimming to juvenile coho salmon (Oncorhynchus kisutch). Can. J. Fish. Aquat. Sci. *41*, 1546–1551. https://doi.org/10.1139/f84-192.

64. Windsor, S.P., Tan, D., and Montgomery, J.C. (2008). Swimming kinematics and hydrodynamic imaging in the blind Mexican cave fish (Astyanax fasciatus). J. Exp. Biol. *211*, 2950–2959. https://doi.org/10.1242/jeb.020453.

65. Teyke, T. (1985). Collision with and avoidance of obstacles by blind cave fish Anoptichthys jordani (Characidae). J. Comp. Physiol. A *157*, 837–843. https://doi.org/10.1007/BF01350081.

66. Wang, W., Escobedo, R., Sanchez, S., Han, Z., Sire, C., and Theraulaz, G. (2025). Collective phases and long-term dynamics in a fish school model with burst-and-coast swimming. R. Soc. Open Sci. *12*, 240885. https://doi.org/10.1098/rsos.240885.

67. Yang, Y., Turci, F., Kague, E., Hammond, C.L., Russo, J., and Royall, C.P. (2022). Dominating lengthscales of zebrafish collective behaviour. PLoS Comput. Biol. *18*, e1009394. https://doi.org/10.1371/journal.pcbi.1009394.

68. Wang, W., Escobedo, R., Sanchez, S., Sire, C., Han, Z., and Theraulaz, G. (2022). The impact of individual perceptual and cognitive factors on collective states in a data-driven fish school model. PLoS Comput. Biol. *18*, e1009437. https://doi.org/10.1371/journal.pcbi.1009437.

69. Hilgers, L., and Schwarzer, J. (2019). The untapped potential of medaka and its wild relatives. eLife *8*, e46994. https://doi.org/10.7554/eLife.46994.

70. Ansai, S., Hiraki-Kajiyama, T., Ueda, R., Seki, T., Yokoi, S., Katsumura, T., and Takeuchi, H. (2025). The Medaka approach to evolutionary social neuroscience. Neurosci. Res. *214*, 32–41. https://doi.org/10.1016/j.neures.2024.10.005.

71. Okuyama, T., Yokoi, S., Abe, H., Isoe, Y., Suehiro, Y., Imada, H., Tanaka, M., Kawasaki, T., Yuba, S., Taniguchi, Y., et al. (2014). A neural mechanism underlying mating preferences for familiar individuals in medaka fish. Science *343*, 91–94. https://doi.org/10.1126/science.1244724.

72. Nakayasu, T., and Watanabe, E. (2014). Biological motion stimuli are attractive to medaka fish. Anim. Cogn. *17*, 559–575. https://doi.org/10.1007/s10071-013-0687-y.

73. Savitzky, A., and Golay, M.J.E. (1964). Smoothing and differentiation of data by simplified least squares procedures. Anal. Chem. *36*, 1627–1639. https://doi.org/10.1021/ac60214a047.

74. Huang, C., Ling, F., and Kanso, E. (2024). Collective phase transitions in confined fish schools. Proc. Natl. Acad. Sci. USA *121*, e2406293121. https://doi.org/10.1073/pnas.2406293121.

75. Zaki, H., Lushi, E., and Severi, K.E. (2021). Larval zebrafish exhibit collective circulation in confined spaces. Front. Phys. *9*, 545. https://doi.org/10.3389/fphy.2021.678600.

76. Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. J. R. Stat. Soc. B *58*, 267–288. https://doi.org/10.1111/j.2517-6161.1996.tb02080.x.

Please cite this article in press as: Harpaz et al., Social interactions in medaka fish depend on discrete kinematic states of swimming behavior, Current Biology (2026), https://doi.org/10.1016/j.cub.2026.01.039

CellPress
OPEN ACCESS

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Deposited data		
All raw data	This paper	https://dataverse.harvard.edu/dataset.xhtml?persistentId=doi:10.7910/DVN/QTNPJL
Experimental models: Organisms/strains		
wild-type zebrafish, *Danio rerio* (AB strain)	Engert lab fish facility	N/A
Wild-type medaka, *Oryzia latipes* (drR)	Engert lab fish facility	N/A
Software and algorithms		
Matlab 2023a; used for tracking and analysis	Written specifically for this study, based on Harpaz et al.[12,15]	N/A

EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS

Fish Husbandry

Medaka

All fish used in the experiments were obtained from crosses of adult wild-type medaka (drR). Larvae were raised in low densities of approximately 10 - 20 fish in medium petri dishes (D = 9 cm). Dishes were filled with filtered fish facility water and were kept at 28 °C, on a 14–10 h light-dark cycle. After hatching (8 dpf), fish were fed paramecia once a day. On day 2 after hatching (10 dpf), fish that were not tested in behavioral experiments, were returned to the fish facility where they were raised in 2L tanks filled with 1.5 inch nursery water (2.5 ppt; prepared by dissolving 4.75g Instant Ocean sea salt mix per liter of RO water), with ∼15 fish in each tank and no water flow. On days 10–12 post fertilization, water flow was turned on, and fish were fed artemia 3 times a day until tested at subsequent ages.

Zebrafish

Experimental fish were obtained by crossing adult wild-type zebrafish (AB). Larvae were raised in conditions similar to medaka. Fish were fed from 5 dpf onwards.

For both adult zebrafish and adult medaka, male-to-female ratios were maintained at approximately 1:1 and individuals were randomly assigned to experimental groups.

All experiments followed institution IACUC protocols as determined by the Harvard University Faculty of Arts and Sciences standing committee on the use of animals in research and teaching.

METHOD DETAILS

Behavioral experiments

Juvenile and adult fish

Juvenile medaka (34 dpf), adult medaka (>60 dpf), and adult zebrafish (>75 dpf) were transferred from their holding tanks to 3 arenas comprised of an opaque Polypropylene plastic pan (McMaster-Carr)(D = 26cm) filled with ∼ 2-3 cm of water (depending on the size and age of the fish). Experimental arenas were filmed from above using an overhead Bassler acA2040-90umNIR camera equipped with Cinegon 1.9/10(0901) lens. Each arena was lit from below using a specialized infrared 24x24 inch IR light panel (850nm, Knema) and from above by indirect light coming from 4 32 W fluorescent lights. Recorded images were collected at ∼50fps and tracked offline using software written in house using Matlab (see Harpaz et al.[15] for details).

Larval fish

All behavioral experiments involving larval fish (10-24 dpf) followed the experimental protocols described previously.[12] Briefly, fish were transferred from their holding tanks to custom-designed experimental arenas of sizes d = 6.5, 9.2, 12.6 cm (for 10, 16 and 24 dpf fish respectively), filled with filtered fish facility water up to a height of ∼0.8 cm. Experimental arenas were made inhouse from sandblasted 1/ 16" clear PETG plastic and had a flat bottom and curved walls (half a sphere of radius 0.5 cm) to encourage fish to swim away from the walls. Every experimental arena was filmed using an overhead camera (Grasshopper3-NIR, FLIR System, Zoom 7000, 18–108 mm lens, Navitar) and a long pass filter (R72, Hoya). Arenas were lit from below using 4 infrared LED panels (940 nm panels, Cop Security) and from above by indirect light coming from 4 32 W fluorescent lights. Recorded images were immediately background subtracted and thresholded to segment fish bodies. The segmented binary images were saved for subsequent offline tracking and further processing. All acquisition and online segmentation were performed using custom-designed software written in Matlab.[12]

Please cite this article in press as: Harpaz et al., Social interactions in medaka fish depend on discrete kinematic states of swimming behavior, Current Biology (2026), https://doi.org/10.1016/j.cub.2026.01.039

Each group was imaged for 20 minutes, following 5–10 minutes of acclimation to the arena. Groups were eliminated from subsequent analysis in the case that one or more of the fish were immobile for more than 60% of the experiment.[12] Only 3 groups of 10dpf larvae were eliminated from the experiment. Choosing a more stringent, or a less stringent criteria for elimination did not change the qualitative nature of the results.

Tracking

Segmented fish images were subsequently analyzed offline to extract the center of mass position $\vec{x}_i(t)$of each animal i at time t. Subsequent positions were then concatenated to obtain continuous individual fish trajectories.[12,15] Fish trajectories were smoothed using a Savitzky–Golay filter spanning ~1/6 of a second.[7,15,73] Fish images were additionally analyzed to obtain the body orientation $\vec{d}_i(t)$ of the animals using second-order image moments, and to extract fish midline for tracking of body and tail undulations (see below).

Midline extraction and tail tracking

Binary fish images were analyzed to extract the midline of the body from head to tail. Briefly, we 'skeletonize' the binary image to get a single pixel width skeleton of the image. We detect branch points and endpoints along the skeleton and choose the two endpoints that correspond to the edges of the image. We then used geodesic distances inside the skeleton to compute a shortest path connecting those endpoints. The resulting path was then smoothed and resampled to obtain the final midline of the fish (Figure 1B; Videos S2 and S3).

Individual and group level swim properties

Fish velocity was calculated from the center of mass positions of the fish such that $\vec{v}_i(t) = \left[\vec{x}_i(t+dt) - \vec{x}_i(t-dt)\right]/2dt$, where dt is one frame (~1/50 s). Speed is then defined as the magnitude of the velocity $s_i(t) = |\vec{v}_i(t)|$. We measure group attraction as $Attraction(t) = -log\left(NN(t)/NN^{Shuffled}\right)$, where $NN(t)$ is the average nearest neighbor distance of the fish in a group and $NN^{Shuffled}$ is the average nearest neighbor distance obtained from shuffling the identity and time signature of fish across groups (i.e. all fish in the shuffled groups are taken from different real groups)(Figure 1C). Group alignment is defined as $Alignment(t) = |\sum_i \vec{d}_i|/n$, where $\vec{d}_i$ is the unit direction vector of fish i, and n is the number of fish in the group. Alignment is therefore bounded between 0 (if all fish point maximally away from each other, and 1, if all fish point in the same direction). For reference we also calculate the average alignment of shuffled fish groups (Figure 1d). The alignment of a fish i with the direction of the group (Figure 1F) is defined as the projection of $\vec{d}_i$ onto $\sum_i \vec{d}_i/n$, which is the average direction of the group, and this value is bounded between -1 and 1.

Neighbor Maps

Density maps (Figure 1G) were obtained by calculating a 2d histogram of relative positions of all neighbors with respect to a focal individual situated at (0,0) pointing north. Maps were smoothed using a 2d rotationally symmetric gaussian filter with $\sigma^2 = 0.15$cm. Scale bar represents percent of the most visited bin (which is set to equal 1). Angular deviation maps were calculated by collecting the heading direction relative to a focal fish situated at (0,0) pointing north, of all neighbors that occupy a specific spatial bin around a focal fish. We then calculate an average direction vector for that bin and the angular deviation of that vector from north. Histograms were smoothed using a similar gaussian filter. For both maps, we discarded all data points in which fish were at a distance <0.3 Arena radii to the walls of the arena (3.9 cm for adult medaka) to reduce the effects of arena boundaries on the structure of the maps[15,74,75] (Figure S1B).

Segmenting fish speed profiles into discrete states

We used the speed profile $s_i(t)$ of each fish to segment fish swimming into acceleration, deceleration and constant speed states. We define a 'constant speed' epoch, as the longest segment we can detect that has similar start and end speeds, and low speed variability within the segment. We used the following algorithm: start from time t - and look for the farthest point $t+t_c$ such that $|S(t) - S(t_c)| < \delta_1$ and $variance[S(t..t_c)] < \delta_2$, where $\delta_1 = 0.1\ (cm/s)$ and $\delta_2 = 0.2\ (cm/s)^2$. Segments had a minimal length of 150ms. Accelerations were then taken as all times that are not detected as constant speed epochs and had a positive acceleration sign, while deceleration epochs had a negative acceleration sign. Very low speeds (<0.6cm/s) were unassigned and were excluded from further analysis. Low speeds accounted for ~8% of all data. These choices of $\delta_1$and δ_2 resulted in 3 well separated clusters, and slight variations in these thresholds did not change the detected clusters (Figure S2B).

For comparison, we also segmented fish speed profiles to only two states - accelerations and decelerations (Figures S3C and S3D) according to the sign of the acceleration, similarly removing very low speeds from further analysis.

Please cite this article in press as: Harpaz et al., Social interactions in medaka fish depend on discrete kinematic states of swimming behavior, Current Biology (2026), https://doi.org/10.1016/j.cub.2026.01.039

Function fitting to kinematic states

Every identified speed epoch $s_i(t_1 \ldots t_{n_i})$ was assigned one of 3 labels - 'acceleration', 'deceleration' or 'constant speed' according to the algorithms described above. We then use the following function families to describe these epochs:

$$\text{Sigmoidal function}: S(t) = \frac{max[S(t)]}{1+e^{-(t-t_0)/\tau_{sig}}}$$

where $max[S(t)]$ is the maximum speed in that epoch, t_0 is the midpoint of the sigmoidal function and τ_{sig} is the time constant of the segment. We optimize t_0, and τ_{sig} for each epoch.

$$\text{Exponential function}: S(t) = S_0 \cdot e^{-t/\tau_{exp}}$$

where S_0 is the starting speed and τ_{exp} is the decay time constant. We optimize S_0 and τ_{exp} for each epoch.

$$\text{Sine wave function}: S(t) = A \cdot sin(\omega \cdot t + \phi) + \bar{s}$$

Where A is the amplitude, ω is the angular velocity, ϕ is a phase shift and $\bar{s}$ is the average speed of the epoch. We optimize A, ω and ϕ for each epoch.

For comparison, we fitted all function families to all epoch types and chose the functions that gave the best overall fit to a specific state. Sigmodal and Exponential families both gave a good fit to acceleration and deceleration epochs (but not to constant speed epochs). However, the sigmoidal family gave a consistently better fit to both acceleration and deceleration epochs, with a similar number of free parameters.

Operational definition of social responsiveness

We define 'social responsiveness' here, as the degree to which a focal fish's velocity changes (turning and speed changes) are predictable from the velocities of its neighbors, given the detailed spatiotemporal map that charts the influence that neighbors exert on the focal fish (Figures 3A–3C and 4A).

'Receptive field' models to predict active velocity changes of individuals

We predict the active velocity changes of fish i at time t, $\Delta\vec{v}_i^{\,Active}(t)$, from the velocities of its neighbors in the recent past $\vec{v}_j(t-k\Delta t)$ and their spatial configuration, as described in Equation 3. To test for varying spatiotemporal effects in the different kinematic states, we fit a separate model for each of the kinematic states for every group tested. Model parameter values, or β weights, were learned using least squares Lasso regression,[76] by minimizing the following cost function:

$$min_{\vec{\beta}}\left\{\frac{1}{2N}\sum_{t=1}^{N}\left[\Delta\vec{v}(t) - \vec{z}(t)^T\cdot\vec{\beta}\right]^2 + \lambda\sum_{p=1}^{P}|\beta_p|\right\} \quad \text{(Equation 4)}$$

Where $\Delta\vec{v}(t)$ is the change in velocity of the fish, $\vec{z}(t)$ are neighbor velocities in the spatiotemporal bins around the focal fish, $\vec{\beta}$ is the set of weights to be optimized, and N is the number of training observations. The second term in the parenthesis, penalizes the cost function for having many non-zeros parameters in the model, with hyper parameter λ controlling the magnitude of the penalty. Optimizing model parameters on held out test data usually results in a sparser model, setting noncontributing weights to 0. We standardized all predictors $\vec{z}(t)$ and target vectors $\Delta\vec{v}(t)$ within each kinematic state and inside every cross-validation training fold to zero mean and unit variance before learning model parameters $\vec{\beta}$, to ensure that differences in the scales or magnitude of the observations between the different states (Figures 2H and 2I), will not bias model fit accuracies or the scale of the learned weights. For each group, we discarded all data where the distance to the nearest wall was smaller than the median of the distribution of wall distances[15,74,75] (Figure S1B).

'Mean velocity' model to predict active velocity changes of individuals

To establish a baseline for model prediction accuracies, and to assess the unique contribution of spatiotemporal features in describing fish movement decisions, we also trained a simpler model that ignores the spatial properties of neighbors and only used their velocity information to predict the change in velocity of a focal fish - $\Delta\vec{v}_i^{\,Active}$:

$$\Delta\vec{v}_i^{\,Active}(t) = \beta\sum_j\vec{z}_j(t-\Delta t) \quad \text{(Equation 5)}$$

where $\vec{z}_j(t-\Delta t)$ is the velocities of all neighboring fish and the direction of the closest wall $\vec{d}_l(t-\Delta t)$ at time $t-\Delta t$.

Surrogate data of shuffled groups

To confirm that prediction accuracies indeed depend on unique social interaction computations in the different kinematic states, and not on inherent difference of the observations between the states, we created a 'shuffled identity' dataset that

Please cite this article in press as: Harpaz et al., Social interactions in medaka fish depend on discrete kinematic states of swimming behavior, Current Biology (2026), https://doi.org/10.1016/j.cub.2026.01.039

retains the individual swim kinematics of the fish, yet randomizes within group coupling, by mixing fish identities from different groups. Specifically, every new 'surrogate' group of fish is composed of the complete swimming trajectories of fish chosen from different real groups (such that no two fish in a surrogate group ever swam together). We created the same number of surrogate groups as real groups and we fit the RF models to these datasets in the same way we fit data from real groups (Figure S3D). Prediction accuracies from this surrogate dataset are similar to the ones obtained from other shuffling methods, such as keeping fish identities intact but shuffling the temporal contingency of observations and predictors and the spatial organization of neighbors' velocities in the context of the RF models (dashed black lines in Figures 3D, S3C, and S3D).

Complexity matched datasets

To specifically test whether the state-dependent differences in prediction accuracy (Figure 3D) were driven by differences in motion complexity (Figures 2H and 2I), we generated a complexity-matched dataset in which turning and speed were equalized across all three states. Specifically, for every fish in every real group, we subsampled the velocity changes $\Delta\overrightarrow{v_i}$ from the acceleration and deceleration epochs such that the joint distribution of turning and speeding will be best matched across the 3 states (Figure S3E). We first calculate a reference 2d histogram of the target vectors in the constant speed state. We then use the bin edges from the target histogram, to calculate a similar 2d histogram for the acceleration and deceleration states and subsample the observations in each bin to best match the number of observations in the target histogram. Consequently, the total number of samples in the subsampled histogram can only be smaller or equal to that of the reference histogram, and observations that are not represented in the constant speed state are excluded from the new subsampled datasets. Since the original distributions of motion vectors in the acceleration and deceleration states were more diverse than those of the constant speed state and the number of observations was larger, this subsampling procedure gave a very good match to the target histogram ($C_{constant\text{-}acceleration}$ =0.84±0.04, $C_{constant\text{-}deceleration}$=0.85±0.05; mean ± sd for all fish in all groups, Pearson's correlation coefficient). We then fit our RF models to these complexity matched datasets in the same way we fit the original datasets (Figure S3E).

Model parameters

The following parameters were used in training state dependent spatiotemporal RF models (Equation 3), based on values reported previously,[15] and the spatiotemporal properties of medaka (Figures 1G, 1H, and S3A).

Parameter name	Description	Values used
Distance bins	# of rings in each spatial filter	6 bins, 0-6 body length at 1 body length increments
Angular bins	# of sectors in each spatial filter	6 bins, covering 60 degrees each.
Time delays	# of spatial filters measured at different Δt times into the past.	$\Delta t_1 \approx$-160ms (8 frames), $\Delta t_2 \approx$-390ms (20 frames), $\Delta t_3 \approx$-630ms (32 frames). For the 'simple average' model we use $\Delta t \approx$-160ms only. *first Δt interval was chosen to be $\approx \tau_c$(Equation 2), which is the estimated decay time constant of the autocorrelation functions (Figure S3A).

Weight map plots

β weight maps (Figure 4A) show the average weights per spatiotemporal bin over all fitted groups. We perform a significance test for every spatiotemporal bin, using a one sample t-test, and only plot β weights with *p-value* < 0.05. Since our weight estimation procedure is already conservative (weights are learned using a Lasso regularized cross-validation procedure for each group, which tends to shrink parameters towards 0), we do not perform additional post-hoc corrections.

QUANTIFICATION AND STATISTICAL ANALYSIS

Linear models linking group properties to kinematic features

We used linear models to regress group alignment on mean speed, mean nearest neighbor distance and percent of time spent swimming at constant speed (Figures 2K and S2H). This combination of predictors gave the best fit to the data and allowed us to rule out other predictor combinations by comparing AICc values of the models. For example, adding mean turning rates or wall distance did not improve AICc values (ΔAICc = -0.65). To assess sensitivity of the results to any specific group, we performed a leave-one-out re-fitting of the models, and we recorded the model coefficients for every fit. This analysis confirmed that fitted model parameters are not biased by any single group.

Full model details:

$$Alignment = \beta_0 + \beta_1 \cdot NNd + \beta_2 \cdot Speed + \beta_3 \cdot \%T_{constant}$$

$F_{model}(11) = 47.7$, $p_{model} = 1.35 \cdot 10^{-6}$, $R^2 = 0.93$

Please cite this article in press as: Harpaz et al., Social interactions in medaka fish depend on discrete kinematic states of swimming behavior, Current Biology (2026), https://doi.org/10.1016/j.cub.2026.01.039

	β	SE	t-statistic	p-value
Intercept	0.37	0.11	3.23	0.008
NNd	-0.05	0.011	-4.81	0.0005
Speed	0.11	0.025	4.76	0.0006
$\%T_{constant}$	0.005	0.001	3.62	0.004

To quantify the unique contribution of each variable to group alignment, we computed the partial correlation between alignment and the variable of interest (e.g., % time at constant speed) after regressing out the variance explained by the other two predictors (e.g., nearest-neighbor distance and speed). This corresponds to the correlation between the residual alignment values and the residuals of the predictor after controlling for the remaining variables. High partial correlation values indicate a strong unique contribution of that variable, independent of the other explanatory variables (Figures 2K and S2H).

Sample sizes, trial numbers and power estimation

Sample sizes were chosen based on previous experiments studying groups of larvae, juvenile and adult zebrafish for reference.[12,15,55] These group sizes were expected to allow accurate estimation of individual and group level properties and to allow comparison between experimental conditions with sufficient statistical power.

Statistical testing

We used the non-parametric Wilcoxon's signed rank tests to assess differences between different models fitted to the same groups and report exact p-values and effects sizes using Cohen's d (Figure 4B). Confidence intervals of the mean are calculated using the parametric t-distribution with N-1 degrees of freedom, and $\alpha = 2.5\%$ at both ends of the interval. We used R^2 values to assess goodness of fit measures when fitting functions to different kinematic epochs (Figure S2C). We used Spearman's correlation coefficient to assess the monotonic relationship between the different parameters of the function fits and the amplitude of the speed profile (Figure 2C) and estimated their statistical significance using a permutation test with 10000 shuffled repetitions (therefore the limit of p-value accuracy is 1/10000). We used one sample t-test to detect average model weights that are statistically different from 0 (Figure 4A). Before performing a specific statistical procedure, we made sure all model assumptions are fulfilled.

Current Biology, Volume *36*

Supplemental Information

Social interactions in medaka fish depend on discrete kinematic states of swimming behavior

Roy Harpaz, Pedro Piquet, Yasuko Isoe, Mark C. Fishman, and Florian Engert

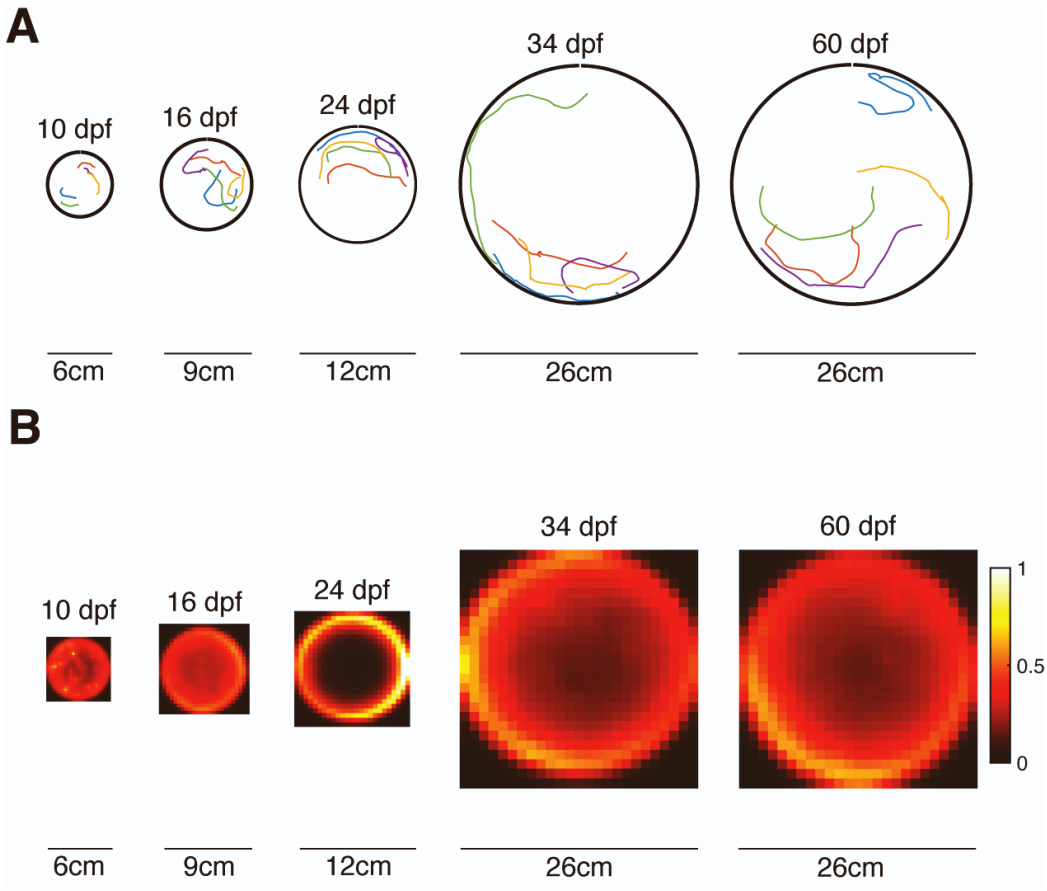


Figure S1 - Social behavior in medaka develops early and rapidly. Related to Figure 1. A. Example trajectories (5s long) of groups of 5 fish at different developmental stages. Colors represent different fish. **B.** Heatmaps show the aggregated spatial positions of all fish across all groups for each developmental stage. Colors represent a fraction of the most occupied spatial bin (which is set to 1).

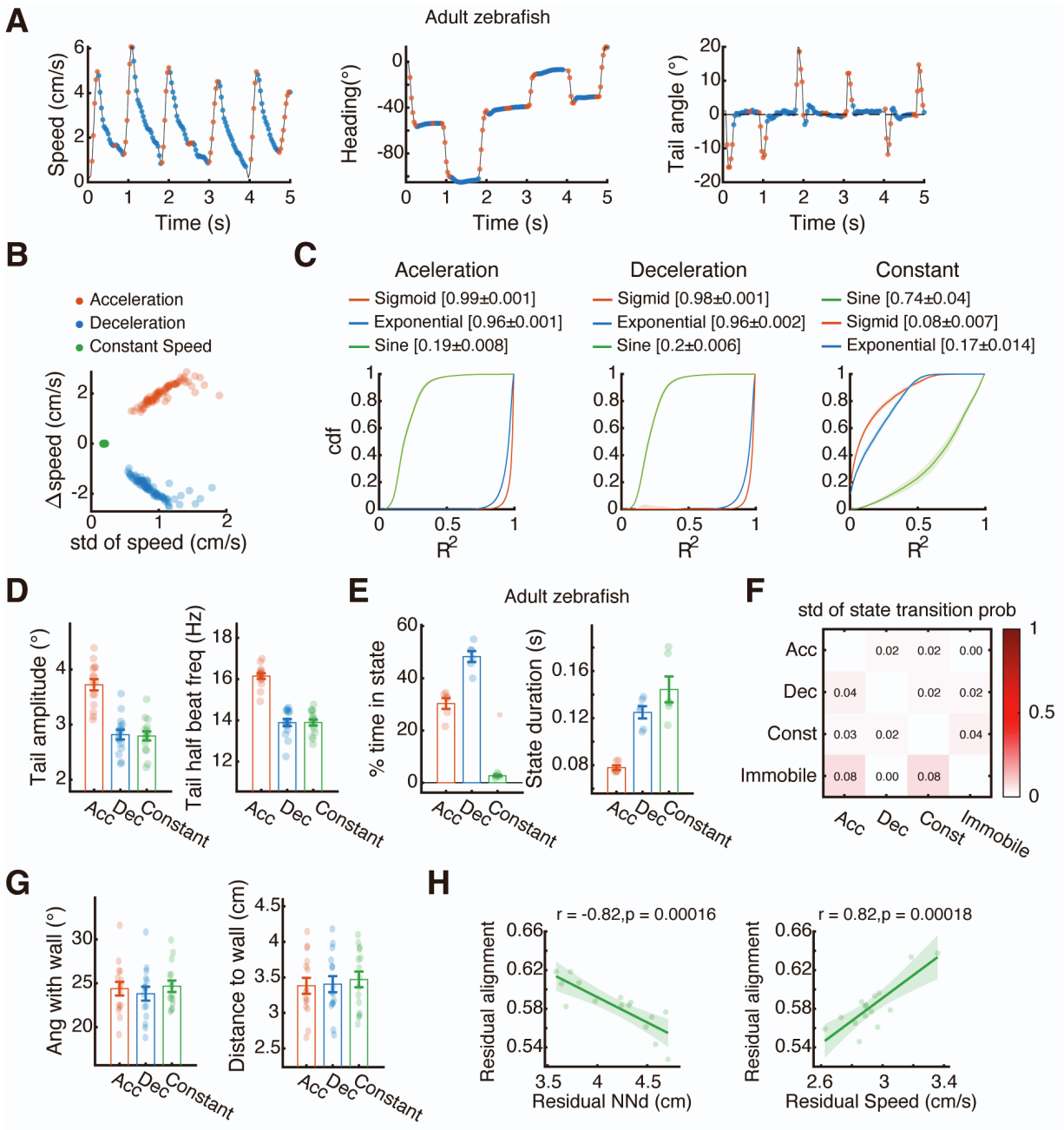

A
Adult zebrafish
Speed (cm/s)
Time (s)
Heading(°)
Time (s)
Tail angle (°)
Time (s)
B
Acceleration
Deceleration
Constant Speed
Δspeed (cm/s)
std of speed (cm/s)
C
Aceleration
Sigmoid [0.99±0.001]
Exponential [0.96±0.001]
Sine [0.19±0.008]
Deceleration
Sigmid [0.98±0.001]
Exponential [0.96±0.002]
Sine [0.2±0.006]
Constant
Sine [0.74±0.04]
Sigmid [0.08±0.007]
Exponential [0.17±0.014]
cdf
R2
D
Tail amplitude (°)
Tail half beat freq (Hz)
Acc
Dec
Constant
E
Adult zebrafish
% time in state
State duration (s)
F
std of state transition prob
Acc
Dec
Const
Immobile
G
Ang with wall (°)
Distance to wall (cm)
H
r = -0.82,p = 0.00016
Residual alignment
Residual NNd (cm)
r = 0.82,p = 0.00018
Residual Speed (cm/s)

Figure S2 - Distinct behavioral swim states of medaka and zebrafish. Related to Figure 2. A. Examples of the speed (left), heading angle (middle) and tail angle deviation from the body axis (right) of an adult zebrafish swimming in a group of 5 fish. Red and blue colors represent acceleration and deceleration epochs (constant speed epochs were not detected in this example). **B.** Average change in speed calculated as ΔSpeed = [S(end)-S(start)] plotted against the average within epochs standard deviation across the three states. Constant speed epochs were detected using δ_1=0.1 (cm/s) for maximum absolute change in speed, and δ_2=0.2 (cm^2/s^2) for maximum within epoch variance (Methods). **B.** Cumulative density functions of goodness of fit values (R^2) for acceleration (left), deceleration (middle) and constant speed epochs (right). Different colors represent the R^2 values obtained by fitting different functions to epochs of a specific state. Solid lines and shaded areas are mean±SEM of the calculated cdf's of different groups. Values in brackets are mean±sd. **D.** Mean tail amplitude (left) and half beat frequency (right) of adult medaka for the three kinematic states (Methods). **E.** Mean percent time spent in each of the kinematic states (left) and mean duration of the detected states (right) of adult zebrafish. **F.** Standard deviations of the mean state-transition probabilities shown in Figure 2g. **H.** Mean angle between fish heading direction and the direction tangent to the closest wall (left) and mean distance to the closest wall (right) for the three kinematic states for adult medaka. **H. Left:** Relationship between group alignment and nearest neighbor distances, after controlling for the effects of speed and %time spent at constant speed. Dots represent groups, solid line is the best fit linear model, shaded area is the model's 95% CI. **Right:** Same, but for the relationship between alignment and speed after controlling for nearest neighbor distances and %time spent at constant speed. In D, E and G, dots represent groups and error bars are mean±SEM.

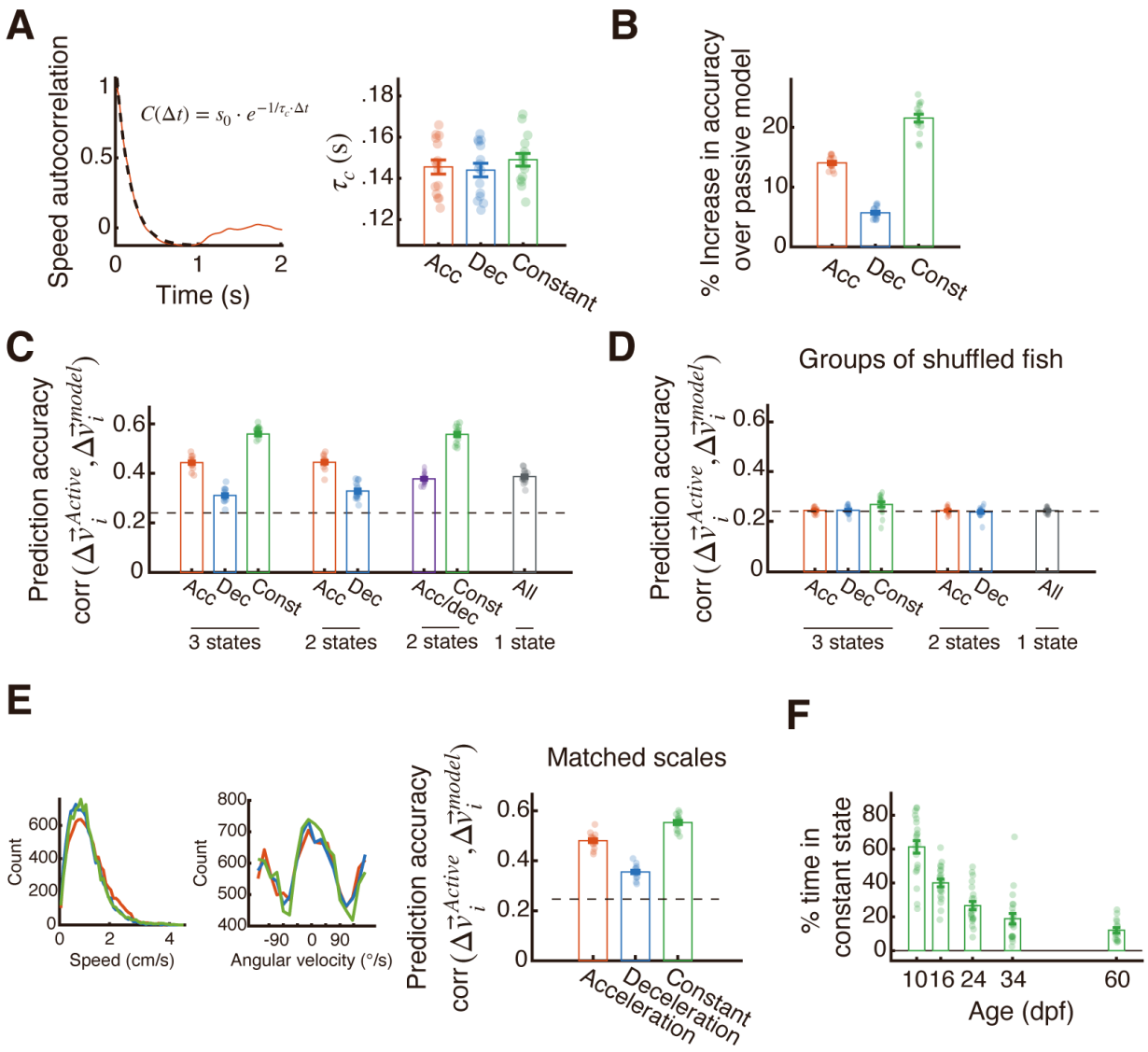
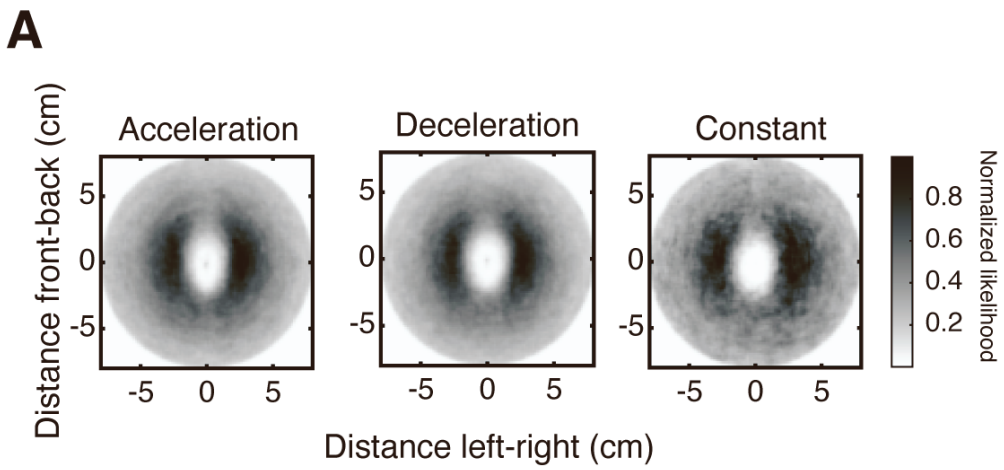
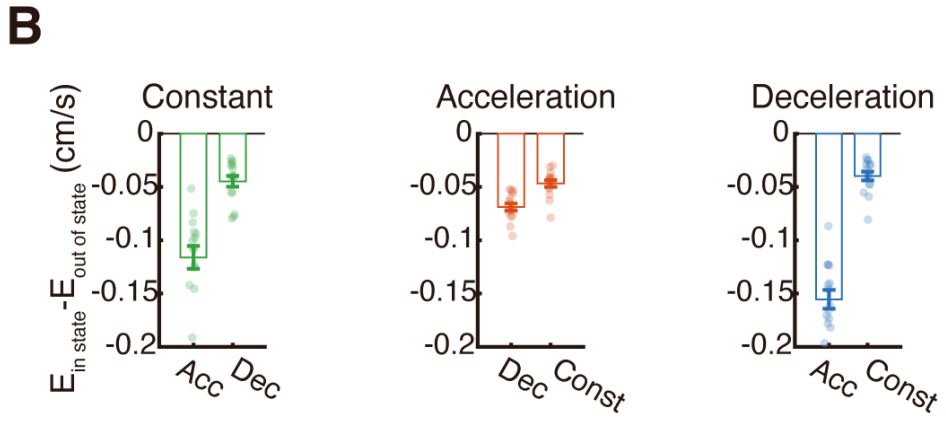
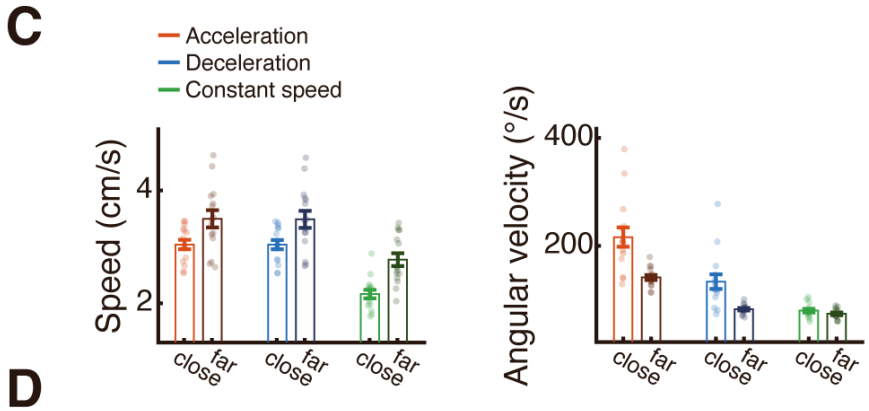

A
Speed autocorrelation
$C(\Delta t) = s_0 \cdot e^{-1/\tau_c \cdot \Delta t}$
Time (s)
τ_c (s)
Acc
Dec
Constant
B
% Increase in accuracy over passive model
Acc
Dec
Const
C
Prediction accuracy
corr $(\Delta \vec{v}_i^{Active}, \Delta \vec{v}_i^{model})$
Acc
Dec
Const
Acc
Dec
Acc/dec
Const
All
3 states
2 states
2 states
1 state
D
Groups of shuffled fish
Prediction accuracy
corr $(\Delta \vec{v}_i^{Active}, \Delta \vec{v}_i^{model})$
Acc
Dec
Const
Acc
Dec
All
3 states
2 states
1 state
E
Count
Speed (cm/s)
Count
Angular velocity (°/s)
Matched scales
Prediction accuracy
corr $(\Delta \vec{v}_i^{Active}, \Delta \vec{v}_i^{model})$
Constant
Deceleration
Acceleration
F
% time in constant state
Age (dpf)

Figure S3 - Distinct swim states imply varying levels of social responsiveness. Related to Figure 3. A. Left: Example autocorrelation function (red) and the corresponding function fit of the form $C(\Delta t) = s_0 \cdot e^{-1/\tau_c \cdot \Delta t}$ (black dashed line) calculated for acceleration events of an adult medaka swimming in a group of 5. **Right:** Decay time constant τ_c of the auto-correlation function fits for each of the kinematic states. **B.** Increase in accuracy of the full model (active and passive components) over the passive component alone, calculated as $(1 - Erorr_{full}/Erorr_{passive}) \cdot 100$, where $Erorr = |\Delta\vec{v}_i^{fish} - \Delta\vec{v}_i^{model}|$. The % increase in model accuracy is 14.04±1.06 [13.46, 14.63] for accelerations, 5.71±0.93 [5.2, 6.23] for decelerations, and 21.55±2.53 [20.15, 22.95] for constant speed epochs; mean±sd, 95% CI in brackets. **C.** Prediction accuracy of the spatiotemporal models (eq. 3), similar to Figure 3d, but for different segmentations of the data - 3 states, 2 states (no constant speed), 2 states (acceleration and deceleration merged into one state) or a single state. **D.** Prediction accuracy for spatiotemporal models fitted to surrogate groups created by mixing fish identities from different groups (Methods). All analysis on surrogate data was done in the same manner as for real fish data (Figure 3d). The dashed black line represents prediction accuracy when models are fitted to data shuffled over groups and time (Methods). **E. Left:** Subset matched speed and angular velocity distributions for the three kinematic states, for an example group of 5 fish. The original distributions of velocity vectors from acceleration (red) and deceleration (blue) epochs were subsampled to match the joint distribution of speed and angular velocity in constant speed epochs (green)(Correlation between matched and constant-speed distributions across all fish and groups: $C_{constant\text{-}acceleration}$=0.84±0.04, $C_{constant\text{-}deceleration}$=0.85±0.05; mean±sd, pearson correlation coefficient). **Right:** Prediction accuracy (similar to Figure 3d) for the complexity matched datasets: $Accuracy_{acc}$=0.48±0.02 [0.46, 0.49], $Accurcy_{dec}$=0.35±0.03 [0.33, 0.37], $Accuracy_{constant}$=0.56±0.03 [0.54, 0.58] (mean±sd; 95% CI in brackets). **F**. Percent of time spent in the constant speed state for medaka groups at different ages. In all panels, dots represent groups and error bars are mean±SEM**.**


A
Acceleration
Deceleration
Constant
Distance front-back (cm)
Distance left-right (cm)
Normalized likelihood
0.8
0.6
0.4
0.2
5
0
-5


B
Constant
Acceleration
Deceleration
$E_{in\ state}$-$E_{out\ of\ state}$ (cm/s)
0
-0.05
-0.1
-0.15
-0.2
Acc
Dec
Const


C
Acceleration
Deceleration
Constant speed
Speed (cm/s)
4
2
Angular velocity (°/s)
400
200
close
far

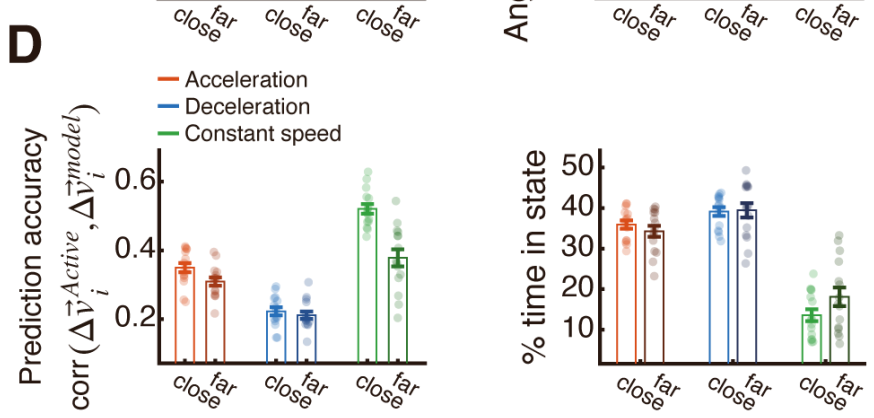

D
Acceleration
Deceleration
Constant speed
Prediction accuracy
$corr(\Delta\vec{v}_i^{Active}, \Delta\vec{v}_i^{model})$
0.6
0.4
0.2
% time in state
50
40
30
20
10
close
far

Figure S4 - Fish utilize different spatiotemporal filters in each of the kinematic states. Related to Figure 4. A. Distributions of nearest neighbor positions with respect to a focal fish, situated at [0,0] pointing north, for groups of adult medaka. Grayscale values represent a fraction of the most occupied spatial bin (which is set to 1). **B.** Difference in prediction error, defined as $E = |\vec{v}_i^{fish}(t) - \vec{v}_i^{Model}(t)|$, when using the model that was trained in-state - for example predicting held out velocities in the acceleration state using models trained on the acceleration state - and predictions of the same velocities using models that were trained on other kinematic states (out-of-state). For every group and state, prediction error is consistently higher when using out-of-state models. **C. Left:** Average speed when fish swim close or far away from their neighbors (lower 20% and upper 20% of the NNd distribution) for the three kinematic state (ΔS_{acc}=-0.45±0.4 [-0.68, -0.23], ΔS_{dec}=-0.45±0.4 [-0.67, -0.23], ΔS_{const}=-0.61±0.42 [-0.84, -0.38] cm/s, mean±sd; 95% CI in brackets). **Right:** same, but for angular velocity (ΔAV_{acc} = 75±53 [45, 104], ΔAV_{dec}=51±45 [26, 76], ΔAV_{const}=6±8 [1, 10] deg/s). **D. Left:** Prediction accuracy for fish that swim in close proximity to neighbors and for fish that are far away from their neighbors ($\Delta Accuracy_{acc}$=0.04±0.063 [0.006, 0.075], $\Delta Accuracy_{dec}$=0.01±0.071 [-0.03, 0.05], $\Delta Accuracy_{const}$=0.14±0.095 [0.09, 019]; mean±sd; 95% CI in brackets). **Right:** Time spent at different kinematic states at close and far neighbor distances ($\Delta\%T_{acc}$ = 1.6±2.8 [0.1, 3.17], $\Delta\%T_{dec}$ = -0.3±3.9 [-2.5, 1.8], $\Delta\%T_{const}$ = -4.5±4.3 [-2, -7]; mean±sd, 95% CI in brackets). Nearest neighbor distances for 'close neighbors' conditions are: NNd_{acc} = 1.77±0.3, NNd_{dec}=1.81±0.3, NNd_{const}=1.88±0.28 cm (mean± sd), and for the 'far away neighbors' condition: NNd_{acc} = 8.12±1.41, NNd_{dec}=8.08±1.44, NNd_{const}=8.17± 1.47 cm. In B-D dots represent different groups and errorbars represent mean±SEM.