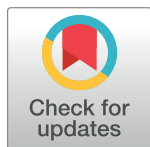# Bidirectional interactions facilitate the integration of a robot into a shoal of zebrafish *Danio rerio*

Vaios Papaspyros[1]*, **Frank Bonnet**[1], **Bertrand Collignon**[1,2], **Francesco Mondada**[1]

1 Biorobotics Laboratory, Institute of Bioengineering, École Polytechnique Fédérale de Lausanne (EPFL), Lausanne, Switzerland, **2** Unit of Social Ecology (USE), Université libre de Bruxelles (ULB), Bruxelles, Belgium

* vaios.papaspyros@epfl.ch

Abstract

Many studies on collective animal behavior seek to identify the individual rules that underlie collective patterns. However, it was not until the recent advancements of micro-electronic and embedded systems that scientists were able to create mixed groups of sensor-rich robots and animals and study collective interactions from the within a bio-hybrid group. In recent work, scientists showed that a robot-controlled lure is capable of influencing the collective decisions of zebrafish *Danio rerio* shoals moving in a ring and a two-room setup. Here, we study a closely related topic, that is, the collective behavior patterns that emerge when different behavioral models are reproduced through the use of a robotic lure. We design a behavioral model that alternates between obeying and disobeying the collective motion decisions in order to become socially accepted by the shoal members. Subsequently, we compare it against two extreme cases: a reactive and an imposing decision model. For this, we use spatial, directional and information theoretic metrics to measure the degree of integration of the robotic agent. We show that our model leads to similar information flow as in freely roaming shoals of zebrafish and exhibits leadership skills more often than the open-loop models. Thus, in order for the robot to achieve higher degrees of integration in the zebrafish shoal, it must, like any other shoal member, be bidirectionally involved in the decision making process. These findings provide insight on the ability to form mixed societies of animals and robots and yield promising results on the degree to which a robot can influence the collective decision making.

Introduction

From the middle of the twentieth century, modelling collective behavior has been in the center of attention of many biological, physical, and computational studies. The main motivation behind such studies lies in the rather complex behaviors that can arise from simple rules on the individual level [1]. Such phenomena are quite common in nature and range from the molecular level, to the subject of this study, collective animal behavior. Disentangling the

dynamics of such collective phenomena can reveal information concerning the underlying structure and organization of multiple physical phenomena observed in nature [2]. However, modelling these phenomena requires rigorous data analysis to extract formulation that can reproduce, at least in theory, the observed collective dynamics. In collective motion of fish, the focus of this paper, there exist various models describing their potential interactions rules [3–16]. More often than not, to compare the quality of different potential models, scientists have evaluated their likelihood to reproduce or predict experimental data [17, 18]. However, this approach does not guarantee any success since different sets of rules are able to reproduce the observed collective movements. One way to overcome these limitations is to test models in real conditions. However, their implementation and validation in physical, real-time and noisy systems is a technical challenge that was only recently solved thanks to the development of micro-electronic and embedded systems. Therefore, in the past years, robots and artificial lures allowed scientists to put these theoretical models to the test in real-life scenarios and with true feedback from the animals, in order to study their collective behavior. Thus, they have since been increasingly involved in inferring the rules of interaction among animals such as bees [19–21], fish [22–27], birds [28–30] and cockroaches [31].

Some of them [32–34] relied on the use of teleoperated devices that produce signals (e.g., visual, acoustic, electric) to attract or repel the animals, others rely on mobile robots that are not explicitly mimicking the animal under study (e.g., it could be a sheepdog among sheep [35]) and some relied on mimetic lures, that is, on lures that mimic the shape, size, and appearance or behavior [22, 23, 31, 36, 37]. These studies demonstrated that artificial agents able to perceive and emit pertinent and adapted signals can influence and control self-organized choices by mixed groups of animals and robots [38].

Although robots proved to be a valuable asset in this context, they also uncovered the need for behavioral models that are applicable to robotic platforms. As biological systems are characterized by great variability and diversity of interaction dynamics, it is very challenging to evaluate the degree to which a robot or robotic lure is integrated within it. A first strategy that addresses this issue relies on open-loop models [39–54] that do not actively react to or take into account the actions of the animals. In this case, it may be difficult to discern if the focal animal is reacting to external stimuli or to another agent considered as a shoal member. Therefore, they may often fail to provide a clear insight into the internal decision making process and the natural information flow between the shoal members. A second strategy is based on closed-loop models [23–26, 55–58] attempting to achieve a conspecific status among individuals by engaging in mutual information exchange and could reveal the intrinsic decision making mechanism of individuals. In [55], the authors compared the success of a robotic lure to integrate among a shoal of zebrafish *Danio rerio* by increasing the biomemitic characteristics of the artificial agent (biomimetic or non-biomimetic lure, movement pattern, and trajectory) and showed that the level of integration increases with the biomimetism of the robotic agent. While similar studies have been conducted with the use of computer animations [59], none of the former studies explicitly compared behavioral models of interaction to evaluate their relative performance when a physical device (i.e., a robot) is used for the interaction.

Therefore, the following research question arises: can we discriminate between different behavioral rules by implementing them in a robot interacting with a shoal of zebrafish? Here, we studied this research question as follows: (1) we make use of a circular corridor setup which serves as a baseline control experiment arena for numerous studies on collective behavior [27, 60–62]; (2) we designed three behavioral models exhibiting three different dynamics, namely, a purely reactive model that explicitly follows the fish, an imposing direction model constantly attempting to dictate the collective swimming direction decision and a biomimetic model mimicking the decisions of zebrafish in a circular corridor environment; (3) we

performed experiments with zebrafish-only and mixed groups of robots and zebrafish; (4) we analyzed the results using three inherently different approaches. We show that our robotic system was capable of participating in the collective decision making and blending in the shoal without perturbing its interaction dynamics. This, in turn, led to an improved integration state where the robot was not only accepted by the shoal, but it was also contributing to the collective decision and acting as a leader for the majority of the time.

Materials and methods

Animals

The authorization for the experiments conducted in this research work was approved by the state ethical board of the Department of Consumer and Veterinary Affairs of the Canton de Vaud (SCAV) of Switzerland (authorization № 2778).

For the experiments, 60 wild-type zebrafish *Danio rerio* with short fins were used (AB strain). The zebrafish were acquired from a pet shop and subsequently stored in a 60-litre aquarium. The average length of the zebrafish used was approximately 4 cm. Water in the housing aquarium was kept at a temperature of 26˚C. The fish were fed once per day with commercial food between 16:00 and 18:00. Furthermore, the aquarium was enriched with plastic plants, Cladophora, gravel, rocks, and aquatic snails.

Experimental setup

The experimental arena pictured in Fig 1 consists of a 10 cm wide circular corridor (from two circular walls: an outer of 58 cm diameter and (2) an inner of 38 cm diameter) placed in a $100 \times 100 \times 25$ cm^3 glass tank, as in [27]. This setup presented the zebrafish with a binary choice for movement: i.e they could either move clockwise (CW) or counter-clockwise (CCW). In fact, this is a common setup for behavioral studies [27, 61–63] because it allows for
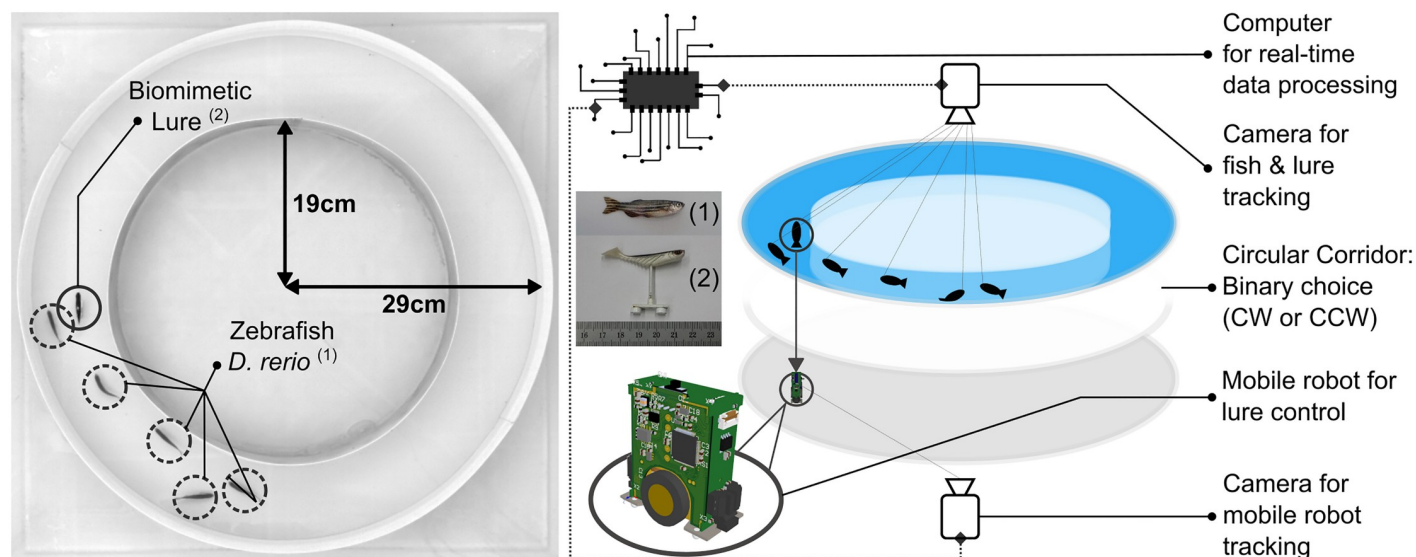
Fig 1. Experimental setup. (left) Top view depicting the setup's dimensions, i.e. inner ring radius of 19 cm and outer of 29 cm. The dotted circles indicate the positions of zebrafish while the full circle indicates the position of the biomimetic lure, and (right) breakdown view of the setup depicting individual components that are necessary for closed-loop interaction. The mobile robot (FishBot) is moving below the tank and drives a biomimetic lure inside the tank through a magnetic coupling. The top and bottom mounted cameras capture frames at a rate of 15 Hz and transmit the information to a computer. The computer will then fuse the information to determine the positions and heading of fish and robot(s) alike. In gray, we denote the conductive plates that are used to power the FishBot.

https://doi.org/10.1371/journal.pone.0220559.g001

setting aside spatial complexities and instead provides a symmetric arena that enables researchers to analyze multiple instances of the same behavioral traits (e.g., U-turns [60] where the fish will perform a direction change greater than or equal to 180˚) and quantify their consistency across different types of behavioral models.

The bottom part of the experimental tank was covered with a Teflon plate to allow for smoother motion of the robotic fish lure (see Robotic system for closed-loop zebrafish-robot interactions) and avoid any stimuli produced by reflections or by the mobile robot moving below the setup. Furthermore, the setup was confined behind white sheets to isolate the fish from the rest of the room, while also maintaining a consistent lighting environment. A uniform luminosity for the room was provided by four 110-W fluorescent lamps placed at each of the four sides of the tank.

Robotic system for closed-loop zebrafish-robot interactions

For the zebrafish-robot interaction experiments we used one miniature wheeled robot, the FishBot [22, 64] (see Fig 2B). The robot was placed between two conductive plates located below the experimental setup (see Fig 1) and was powered using brushes that were constantly in contact with them. This configuration allowed the robot to operate for long periods of time and powered the motors that were, in turn, capable of achieving the necessary speed and acceleration in order to quickly adapt to the rapid spatial displacements of zebrafish. The FishBot was additionally equipped with a Bluetooth chip that allowed it to wirelessly communicate with a computer that was providing the necessary motor commands.
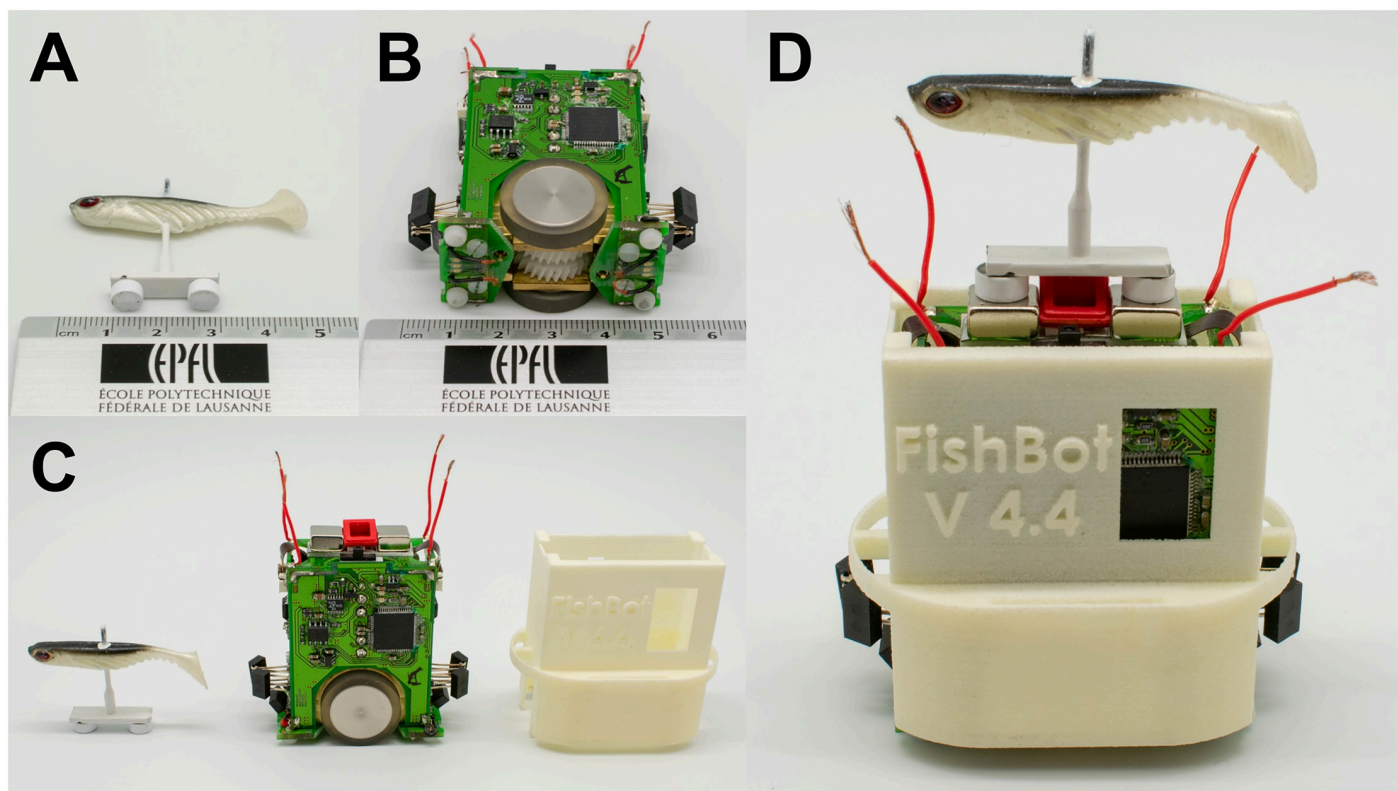
Fig 2. FishBot and biomimetic lure. (A) The biomimetic lure (approximately 4.5 cm long) fixed on a white carbon stick, (B) the FishBot on its side (approximately 5.5 cm long), (C) relative size of the lure, FishBot and FishBot cover (from left to right), and (D) assembled robotic system; the white cover protects critical parts of the FishBot, and the lure magnetically coupled to the FishBot.

A soft biomimetic lure of approximately 4.5 cm length (see Fig 2A) was selected to physically interact with the animals. This lure was designed to mimic the morphology of the zebrafish and passively beat its tail during its underwater motion. As described in [22] this specific lure achieved strong acceptance in groups of zebrafish. Subsequently, it was mounted on a carbon stick at a height of 3 cm to ensure that it was visible by neighboring fish. An iron plate located at the bottom of this stick carried two magnets that allowed for a magnetic coupling (similarly to [22–26], see Fig 2D) to the robot located below the setup.

Control and tracking software

In order to close the interaction loop between the fish and the robot we made use of the Control and tracking for multi-agent animal-robots groups (CATS) framework [65]. CATS continuously monitored the positions of the robot and the animals through image frames obtained by the two cameras located above and below the setup (see Fig 1). In particular, the overhead camera was set to simultaneously stream video in two resolutions; (1) a 1040 × 1040 stream that was recorded and used for the analysis of the experiment, and (2) a 512 × 512 stream that was used in CATS for the detection of the fish and/or lure in real-time. The image frames from the camera located at the bottom of the setup were directly fed to CATS for processing.

More specifically, the agents' positions were determined by feeding the image frames of the overhead camera to a corner detection method [66] implemented using the OpenCV [67] library; while the camera located below the setup was used to localize the mobile robot, which was equipped with 6 light-emitting diodes of blue color, using a blob detector. Subsequently, CATS fused information from both cameras to distinguish the artificial lure from the living individuals. We note that both cameras operated at a rate of 15 frames per second.

Once CATS had finished determining the positions, the resulting spatial information (2D position and heading direction) became available in the control layer of CATS. This layer is responsible for the higher level control procedure of the robot (i.e., deciding which is the next desired state for the robot). The behavioral models presented in the following section were implemented within the control layer of CATS and output higher level commands such as desired velocity, position, and orientation. Those commands were then fed to a micro-controller unit where a proportional-integral-derivative controller (PID) translated the higher level commands to motor commands, similarly to [64].

Additionally, to the online control procedure that was devised for the fish-robot interaction experiments, the videos of each experiment underwent post-processing using the idTracker software [68] to extract the trajectories of each agent for each experiment. This time-consuming and computationally expensive process is capable of recreating the trajectories of each identified agent (6 agents for 30 minute long experiments) with on average 95% accuracy, correcting any mistakes made in trajectories due to crossings that occurred naturally throughout the experiment.

Experimental procedure

For the duration of the experiments, we maintained a constant height of 6 cm of water in the setup. Under these conditions, the fish were not additionally stressed, and their movement was on average constrained to a specific height, thus, reducing spatial complexities on the z-axis. Prior to placing the zebrafish in the setup, the water temperature was brought to 26˚C. Thereafter, a shoal of zebrafish was randomly caught from the rearing tanks with a fishnet and placed in the experimental setup. After a 5 minute acclimatization period during which the FishBot remained stationary, we started and recorded the experiment for 30 minutes. No individual was used twice in the same day.

We conducted 10 experiments with shoals of six zebrafish and no FishBot to observe the baseline behavior of the fish (hereby referred to as "fish-only" experiments) when no artificial stimuli were provided. Then we conducted a total of 3×10 experiments with five zebrafish and one FishBot with three different behavioral models (described in the following section) for the FishBot. Each model was tested in random order to account for the fish getting accustomed to a specific behavior exhibited by the robot.

Behavioral models

Follower model (FM). We designed a closed-loop following model, the "Follower Model", where the robot was simply instructed to head towards the point in space that was on average most dense in terms of fish occupancy. FM is a purely reactive, passive model in the sense that it does not actively model or embed interaction in its design and instead reacts only to the position of the fish by always following it.

Despotic model (DM). We also designed an open-loop model, which is an adaptation of the approach described in [27] that uses only one robot which was instructed to perform a CW movement throughout the experiment. Contrary to FM, this model is always attempting to impose the collective movement direction decision, thus we call it "Despotic Model".

Feedback-Initiative model (FIM). Finally, we implemented a closed-loop parametric behavioral model similar to [61]. In [61], the authors described a model that operates in a one-dimensional decision space (i.e., CW or CCW movement). First, the circular corridor arena is divided into 40 equal cells of 9-degree arc length each. Then, the focal individual will take a directional decision according to the perceived heading directions of its neighbors and a probability to "disobey" this collective decision. More specifically, the next heading direction of an individual is given by the following expression:

$$h(\text{fish}_j, t+1) = \frac{h(\text{fish}_j, t) + \sum_{i=0, \ i \neq j}^{N_P} h(\text{fish}_i, t)}{\left| h(\text{fish}_j, t) + \sum_{i=0, \ i \neq j}^{N_P} h(\text{fish}_i, t) \right|}, \tag{1}$$

where $h(\text{fish}_j, t) \in \{-1, 1\}$ the heading direction at time t and N_P the subset of the set of all individuals, N_A, that are in the perceptual range of the focal individual. The perceptual range is defined as the set $F \subseteq N_A$ of the individuals that are within p_r cells in the forward direction of the focal individual. At every time-step and after computing the new heading direction, the focal individual is given a probability $1 - P_{\text{obey}}$ to choose the opposite direction to the one computed through the above interaction metric. We refer to this opposition to the collective decision of swimming direction as the initiative of the focal individual.

In order to investigate the degree to which our robot can influence the directional decisions taken by the group, we propose a variant of this model where the focal individual attempts to closely mimic zebrafish interaction in similar arenas [27, 61, 63] while at the same time it intelligently embeds initiative in its decision making according to the feedback perceived by the neighboring social companions. Therefore, we call this model "Feedback-Initiative Model". Similarly to [61], we discretized the circular corridor in cells and controlled the robot's direction of movement in the one-dimensional space of CW or CCW movement. More specifically, we separated the setup in 40 cells each of which corresponds to approximately 4 cm of arc length, the average length of a single zebrafish. Essentially, increasing the number of cells, i.e., decreasing the arc length per cell, would allow for more detailed separation of the fish in terms of cell occupancy, but would be subject to the noise produced by CATS (see Control and
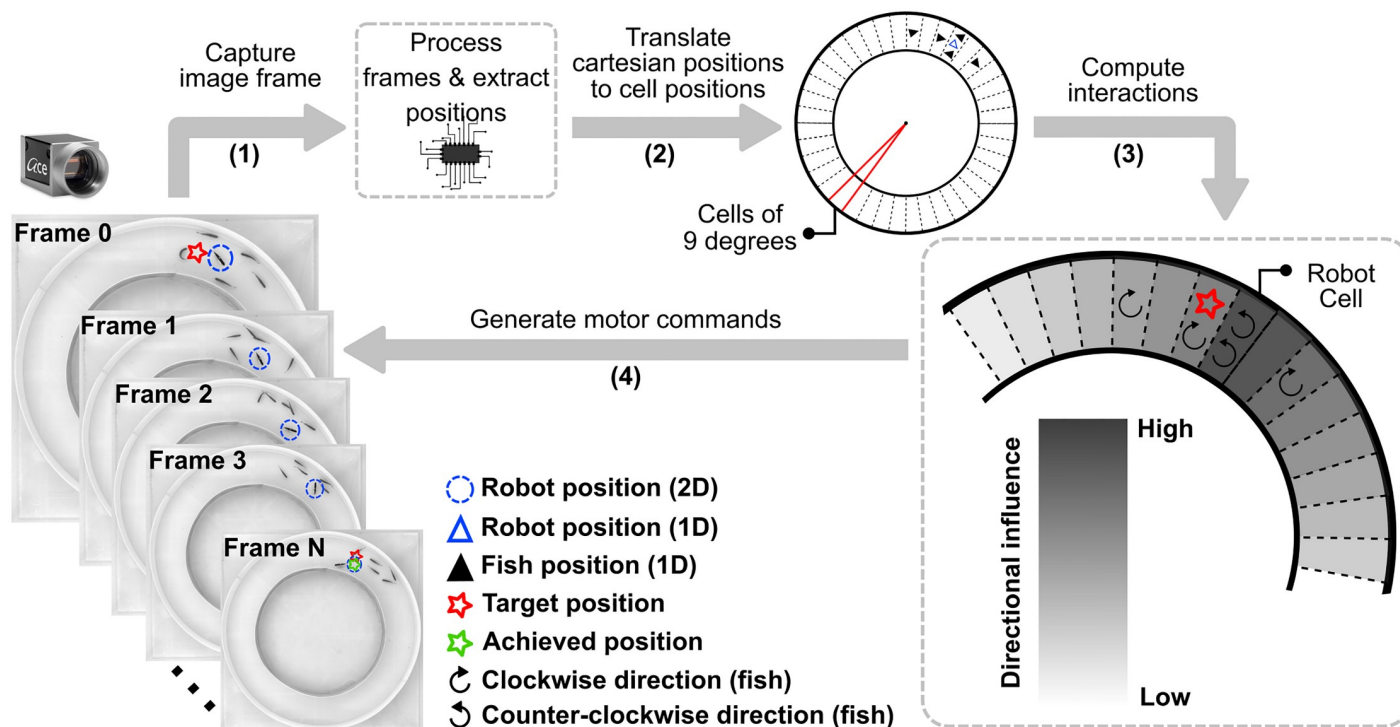
Fig 3. Closed-loop robot control. For each system cycle: (1) a high resolution image frame is captured by the overhead camera, (2) the frame sent to a high performance computer, where it is processed to determine the positions, velocities, and headings of each individual, (3) the extracted positions are discretized and each individual is placed in its corresponding cell and (4) the discretized positions and headings of each individual are forwarded to the FIM, which in turn, weighs the heading direction of the neighboring individuals and produces a desired position (red star) for the next timestep. After a few time-steps an approximation of the target position will be achieved (green star) and the process is repeated.

tracking software). This discretization process reduced the locomotion control dimensions to 1D space and the model need only to output a simple instruction at every time-step: move one cell CW or CCW. The model made a prediction about the best candidate direction at time $t + 1$, every 0.25 seconds (i.e., the controller time-step is equal to 0.25 seconds). In the context of this study, the best candidate swimming direction was considered to be the one that has the highest probability to elicit a collective U-turn (i.e., a switch of the swimming direction for the majority of individuals). A complete system cycle of the model is depicted in Fig 3.

Our source of inspiration for those interactions was derived by the innate behavior of zebrafish in similar setups and can be summarized in the following two key features for a focal individual: (1) an innate tendency to align with the shoal and (2) a tendency to perform a U-turn when few or no agents are in the field of view of the focal, initiating a direction change that might propagate throughout the shoal [3, 62, 63, 69]. The former is formalized as a weighted sum of the direction of a focal individual's neighbors and is defined as the dot product of the heading (-1 or 1) and an exponential function that can rapidly increase or decrease the impact of a neighbor according to its position, as follows:

$$h(\text{fish}_j, \ t + 1) = \sum_{i=0, \ i \neq j}^{N_A} h(\text{fish}_i, \ t) \cdot e^{\alpha \cdot p(\text{fish}_i, \ t)} \tag{2}$$

where $p(\text{fish}_j, \ t) \in [0, |N_C|) \subseteq \mathbb{Z}$ the position of the fish j at time t, N_C the set of cells that correspond to the discretization process and $a \in \mathbb{R}$ a regulatory parameter to set the slope.

To allow for more flexibility in FIM we used two separate parameters α_f and α_b for fish that are in leading and following positions respectively (see Algo. 1) and defined the sets F and B of fish in the forward and backward positions respectively. Intuitively, the difference between the parameters a_f and a_b can express biologically observed behaviors such as that fish within the immediate field of view of a focal individual have more influence on it [3, 60, 62, 69], while the followers can still be perceived due to the water flow [70] and might have less, but significant influence. The parameters α_f and α_b were manually tuned for this specific configuration and are given the values -0.2 and -0.5, respectively. Despite the sum's ability (see Eq 2) to phase out the influence of perceived agents over distance, we explicitly limited the robot's knowledge within 15 cells (p_f = 7 cells forward, p_b = 7 cells backwards and current cell occupied; or 135 degrees of perceptual range for the choice of 40 cells). Indeed, for very low values of α (see Eq 2), a conspecific could have been perceived in front and behind the focal fish due to the circular design of the setup. Finally, the focal individual's tendency to perform a U-turn and even disobey the collective decision of the shoal concerning the direction of movement was modeled as a probability P_{obey}. The probability P_{obey} is dependent on the amount of fish in the forward direction to account for an individual's intuition to not wander too far from the fish school or to simply initiate a random direction change. More specifically, this probability is regulated by two constant parameters (see Algo. 1): (1) $a_{influence}$ = 4, which allows one to increase or decrease the amount of influence of forward individuals concerning the obedience (2) a constant upper bound τ = 0.95, value which we estimated through past fish-only experiments. A subset of the parameter space that could be used in different scenarios include:

- $\alpha_f = \alpha_b = -inf$, $\tau = 1 \Rightarrow P_{obey} = 1$, will produce purely following behavior.

- $\alpha_f = \alpha_b = \tau = 0 \Rightarrow P_{obey} = 0$, will produce a behavior that always contradicts the collective.

- $a_{influence} = 0$, will retain a constant and neighbor-independent P_{obey}.

- Setting the perceptual range to zero (i.e., $Fish' = \emptyset$) and $\tau = 1$, will produce an imposing direction behavior (the initial direction will be followed throughout the experiment).

We designed this model to be parametric and include stochastic elements of decision making. The parametric design allows for modification of the model to comply with different scenarios of interaction or species of fish (e.g., the robot could be instructed to emphasize on following by changing a few parameters), either prior to deployment or during an experiment, while the stochasticity serves as a way to promote initiative in the model. In [61] the authors described a model where the focal individual's next direction will be with high certainty decided by the average swimming direction of the neighboring individuals. Conversely, here the goal was to elicit a different effect from the fish and influence them to change their swimming direction.

Algorithm 1
```
procedure STIMULATE(Fish′)
    Split Fish′ to:
        C = {fish in focal cell} ⊆ Fish′
        F = {fish in p_f forward cells} ⊆ Fish′
        F = {fish in p_b forward cells} ⊆ Fish′
```

$$s = \sum_{i=0,\ \text{fish}_i \in F}^{|F|} h(\text{fish}_i,\ t) \cdot e^{\alpha_f p(\text{fish}_i,\ t)} + \sum_{j=0,\ \text{fish}_j \in C \cup B}^{|C \cup B|} h(\text{fish}_i,\ t) \cdot e^{\alpha_b \cdot p(\text{fish}_j,\ t)}$$

$$h\prime = \begin{cases} s & ,\quad s \neq 0 \\ h(\text{focal},\ t) & ,\quad otherwise \end{cases}$$

$$P_{\text{obey}} = \begin{cases} \tau * \left(1 - (|F \cup C| + 1)^{-\alpha_{\text{influence}}}\right) & , \quad \text{Fish}' \neq \emptyset \\ \tau & , \quad otherwise \end{cases}$$

```
     with probability 1 - P_obey reverse h'
     return h'
   procedure MOVE(h(focal, t + 1)
     p(focal, t + 1) = p(focal, t) + (h(focal, t + 1))
   procedure FEEDBACKINITIATIVEMODEL
     ∀fish ∈ Fish, where |Fish| = |N_A|:
       Initialize position ∈ [0, |N_C|)
       Initialize heading ∈ {CW = -1, CCW = 1}
      while stopping criteria not met do
       ∀fish ∈ Fish:
         Fish' = {fish within the perceptual range of the focal fish} ⊆
         Fish ∪ ∅
         h(focal, t + 1) = STIMULATE(Fish')
         MOVE (h(focal, t + 1))
```

Data filtering

The 15 frames per second capture rate made it possible to detect even minor fluctuations in the displacement of an individual. On one hand, this rate is useful for tracking fast moving objects or animals, but, in the case of zebrafish that move with an average speed below 20-25 cm/second in this setup, it might induce noise due to temporary loss of the position or the image processing algorithm reporting minor differences in the position of an individual at every time-step. Therefore, throughout the following experiments we filtered the data reported in two ways: (1) 3 frames (0.2 seconds of interaction) were averaged to calculate the centroidal position, heading or velocity of each agent and (2) the behavioral models presented in following sections discretized the positions in bins, the number of which was selected to further filter the measurements where necessary.

Data analysis

In this section, we introduce a set of metrics based on spatial, directional and information theoretic measures, as well as the statistical methods followed to evaluate and compare the behavioral models. We note that all the raw trajectory data are available at https://github.com/epfl-mobots/plos_one_experiments.

Average angular distance. In collective behavior, denser groups often suggest a more cohesive, aligned and organized movement [63, 71], thus, density-based measures have been widely used for inferring the interaction rules within a group of animals [72, 73]. Here, we computed a similar measure, by calculating the average angular distance between all pairs of agents. The angular distance between two fish is defined by the angle $\theta_{ij}(t)$ at time t, where i, j are two individuals and $\theta_{ij}(t) \in [0, \pi]$ is the angle between i and j with respect to the origin point of the setup (center of both rings). We note that $\theta_{ij}(t)$ refers to the acute angle between the two individuals (i.e., we only evaluate the angular proximity). The average angular distance was computed as the average of all the pairwise angular distances and is summarized in the following expression:

$$\text{averageAngularDistance}(t) = \frac{1}{N_A(N_A - 1)} \sum_{i=1}^{N_A} \sum_{j=1, \ i \neq j}^{N_A} \theta_{ij}(t) \tag{3}$$

Collective U-turns.   Although the average angular distance provides useful topological information concerning the closeness of the group and thus its cohesive and synchronized movement, it would be incomplete without a complementary metric concerning the interactions within it. Here, we captured these interactions in the number of collective direction changes performed (e.g., from CW to CCW or vice versa), which in this binary choice scenario we defined as collective U-turns. The U-turn in schools of fish has attracted attention [60, 62, 63] as it provides insight on how information is propagated among individuals. Consequently, the effect that each behavioral model has on the occurrence of collective U-turns is representative of its ability to mimic, modulate or perturb the collective decision making.

To calculate the number of collective U-turn events we first defined the polarization of a zebrafish shoal in this context, as follows:

$$\text{pol}(t) = \frac{1}{N_A} \sum_{i=0}^{N_A} h(\text{fish}_i, \ t) \tag{4}$$

A collective U-turn occurs when the polarization of the shoal switches from one direction to another ($\text{pol}(t) \cdot \text{pol}(t-1) < 0$; i.e., we did not take into account transitions from CW/CCW to 0). Complementary figures concerning the duration of consecutive movement before a collective U-turn occurs are available in the Supporting information. section (S1, S2, S3, S4, S5, S6, S7 and S8 Figs).

Transfer entropy.   To complete the aforementioned metrics we also employed an information theoretic measure based on the Shannon entropy [74], called transfer entropy (TE) [75, 76]. Recent studies in collective behavior [43, 60, 77–80] have been increasingly using TE to provide insight on the mutual interactions of individuals over time or with time delays [81]. Here, we adopted the notation of TE with embedded delay, as defined in [82]. More specifically, given two time-series X and Y, TE measures the amount of information provided by the source X about the target Y and is defined as follows:

$$T_{X \rightarrow Y} = \sum p(y_{n+1}, y_n^{(k)}, x_n^{(l)}) \log \frac{p(y_{n+1} | y_n^{(k)}, x_n^{(l)})}{p(y_{n+1} | x_n^{(k)})} \tag{5}$$

where l and k are the history lengths for the two time-series:

$$x_n^{(l)} = \{x_n, \ x_{n-\tau_k}, \ x_{n-2\tau_k}, \dots, x_{n-(k-1) \cdot \tau_k}\} \tag{6}$$

$$y_n^{(l)} = \{y_n, \ y_{n-\tau_l}, \ y_{n-2\tau_l}, \dots, y_{n-(l-1) \cdot \tau_l}\} \tag{7}$$

and τ_k, τ_l the time delay for the source and destination signal respectively.

In the context of this research work, we defined the time-series X, Y to be the direction of two separate individuals over time. More specifically, we represented the direction of each agent in a discrete signal with values -1 (CW) or 1 (CCW), by sampling the direction of each individual every 0.2 seconds. For this computation, a more detailed trajectory is required and thus we used a discretization with a cell count of $N_C = 160$ (i.e., 1 cm per cell; for this procedure, we used the positions extracted from idTracker). In case an individual had not moved during this period, we assumed that its heading has remained the same as in the previous time-step. Furthermore, considering that the influence of one individual to another will be delayed in time, we shifted the source and target time-series by a factor τ_k and τ_l, respectively. The intuition behind this measure is that given the direction Y_n of an individual, we gain information about the next direction X_n of another individual. It is rather obvious that the direction

change of single individual will not propagate instantly and, thus, there exist the parameters τ_k and τ_l that express this delay.

However, choosing the latter parameters is a non-trivial task, as the values need be meaningful with respect to the experiment in question and at the same time expressive enough to allow for observing potential differences in the fish-robot experiments. To that end, we adopted the same technique of the authors in [60], that is, we run a simple search algorithm to find the parameters that maximize the average TE for the fish-only experiments. To reduce the size of the search space, we only considered values of $k \in [1, 15]$ and $\tau_k \in [1, 15]$ (i.e., up to three seconds of signal length and delay). We explicitly set the target delay to $\tau_l = 1$ and the length to $l = 1$ (i.e., 0.2 seconds) as we are interested in the effects of the source signal. We note that the robot-fish experiments were considered for this optimization step to account for the bias that was introduced due to the use of the lure and the models.

Subsequently, we calculated all the pairwise TE values and computed the mean TE across all individuals during one experiment. To do so, we used the JIDT [83] framework to calculate the TE with the optimized parameters $k = 4$ and $\tau_k = 1$ (i.e., $k = 4$ corresponds to 0.8 seconds of history and $\tau_k = 1$ corresponds to a delay of 0.2 seconds). The optimized parameters are empirically found to correspond the time that is necessary for an individual to perform a U-turn and fully propagate it to the shoal.

We computed two separate mean TE values: one for the outgoing and one for the incoming amount of information exchanged. For each case we sum the resulting TE for the trajectories of all the fish-fish or fish-robot pairs. Intuitively, the metric expresses the average information flow direction (incoming or outgoing) when the robot is used. For each behavioral model and each of the outgoing and incoming cases we computed three different quantities: (1) the overall TE for all individuals (2) the TE related to the robot alone (outgoing and incoming) and (3) the average TE of all fish (i.e., excluding the robot's contribution).

Statistical tests. To further validate the interpretation of the resulting data, we performed a Kruskal–Wallis (KW) test followed by a post hoc analysis using Tukey's honest significant difference (T-HSD), for each measure presented in the following section. The Kruskal–Wallis is chosen due to the fact that the variance of the data-sets in question differs depending on the experiment type.

Results

Average angular distance

First, we compiled the average angular distance between the shoal members across the entire observation period for each experiment of the four tested conditions (Fig 4). For the fish-only condition (i.e., where no artificial stimuli are present in the setup), we observed an average angular distance of 35.04 ± 12.62 degrees. Compared to this baseline measure, all models showed a higher inter-individual angular distance (FM: 64.72 ± 15.95 degrees, DM: 59.31 ± 11.87 degrees, FIM: 51.04 ± 10.68 degrees), with these distributions differing significantly from each other (KW test, $p < 0.0001$, $\chi^2 = 22.45$). A more detailed comparison revealed that both the FM and DM significantly differ from the fish-only (T-HSD post hoc test, $p < 0.0001$ and $p < 0.001$, respectively) while the FIM differs significantly from the FM and DM (T-HSD post hoc test, $p \approx 0.22$ and $p \approx 0.53$, respectively) but not from the fish-only distribution (T-HSD post hoc test, $p \approx 0.072$). Complementary statistics are available in S1 Table of the Supporting information. Thus, while FIM still did not perform as well as the control experiments (16 degrees $\approx$ 8 cm arc distance), its ability to mimic the collective decision making allowed the robot to maintain the cohesion of the shoal with on average 14 degrees better than FM and 9 degrees better than DM. Moreover, the results observed are consistent
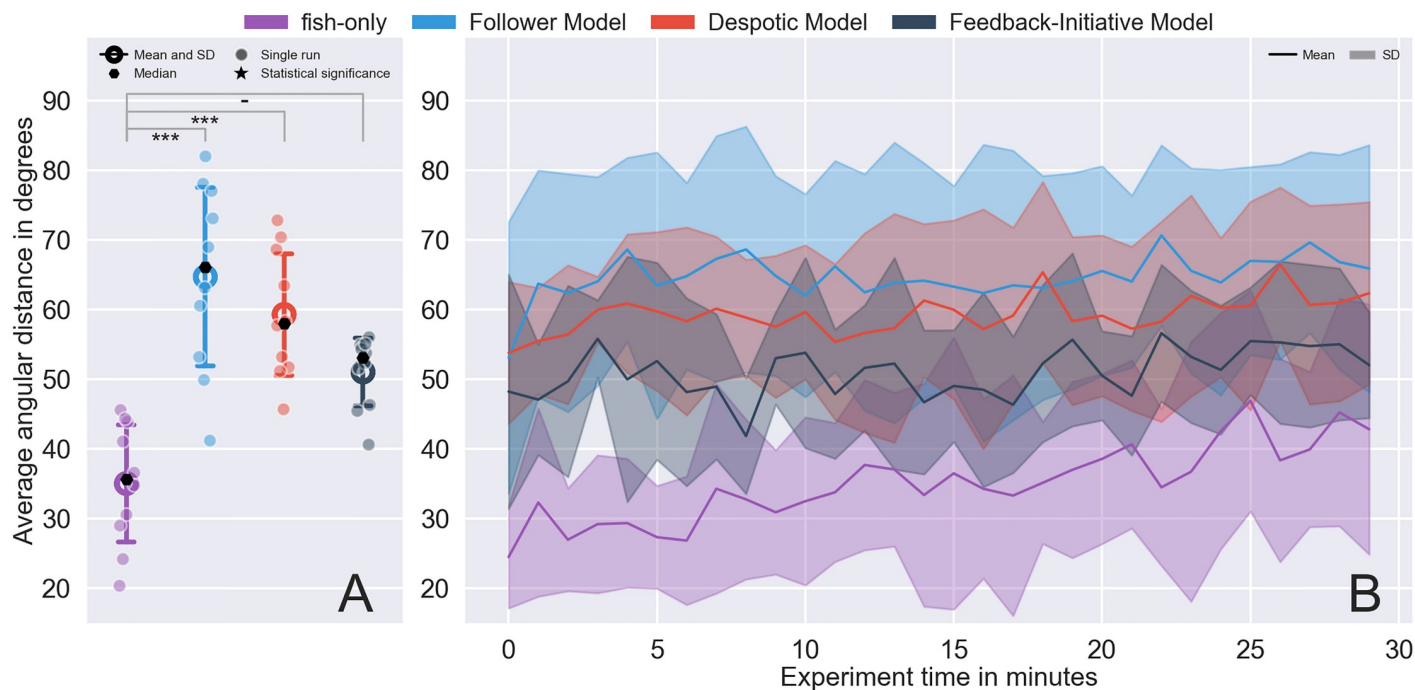
Fig 4. (A) Average angular distance in degrees over ten runs and (B) average angular distance in 1-minute time-steps over all replicates. Annotations of the statistical significance (Kruskal–Wallis test folowed by Tukey's honest significant difference criterion post hoc analysis) are marked with a dash or stars. The dash corresponds $p > 0.05$, a single star to $p < 0.05$, two stars to $p < 0.01$, three stars to $p < 0.001$. and four stars to $p < 0.0001$.

over time as shown in Fig 4B depicting the average angular distance for every minute of the experiment.

The mean performance and amount of variance in the FM model is indicative of its deficiency when it with regard to its ability to be accepted and integrated with the shoal. Furthermore, such result is in direct contradiction with FM's explicit goal, which was to head towards the densest point of the shoal and thus promote a more cohesive behavior, and could suggest that its movement patterns were too aggressive to be accepted by the shoal and contribute to its operation. DM, on the other hand, performed on average worse than the fish-only but exhibits similar variance and seemed to perturb the shoaling behavior less. Finally, FIM was the most consistent over time which could be indicative of an overall better acceptance by the shoal.

Collective U-turns

The collective U-turns performed per minute (see Fig 5). (1) fish-only had median of 13.05 turns and a mean of 12.38 ± 3.71, (2) FM had median of 7.25 turns and a mean of 8.27 ± 2.30, (3) DM had median of 6.97 turns and a mean of 7.91 ± 1.60 and (4) FIM had median of 12.38 turns and a mean of 11.91 ± 2.77. These distribution differ significantly from each other (KW test, $p < 0.001$ and $\chi^2 = 16.63$). An additional post hoc T-HSD analysis showed that: FM versus fish-only had significantly different mean rank ($p < 0.05$); DM versus fish-only also showed significantly different mean rank ($p < 0.001$); FIM versus fish-only showed no significant difference ($p > 0.99$) but FIM versus FM ($p < 0.05$) and DM ($p < 0.01$) showed a significant difference (a detailed table of the post hoc analysis is available in S2 Table of the Supporting information).
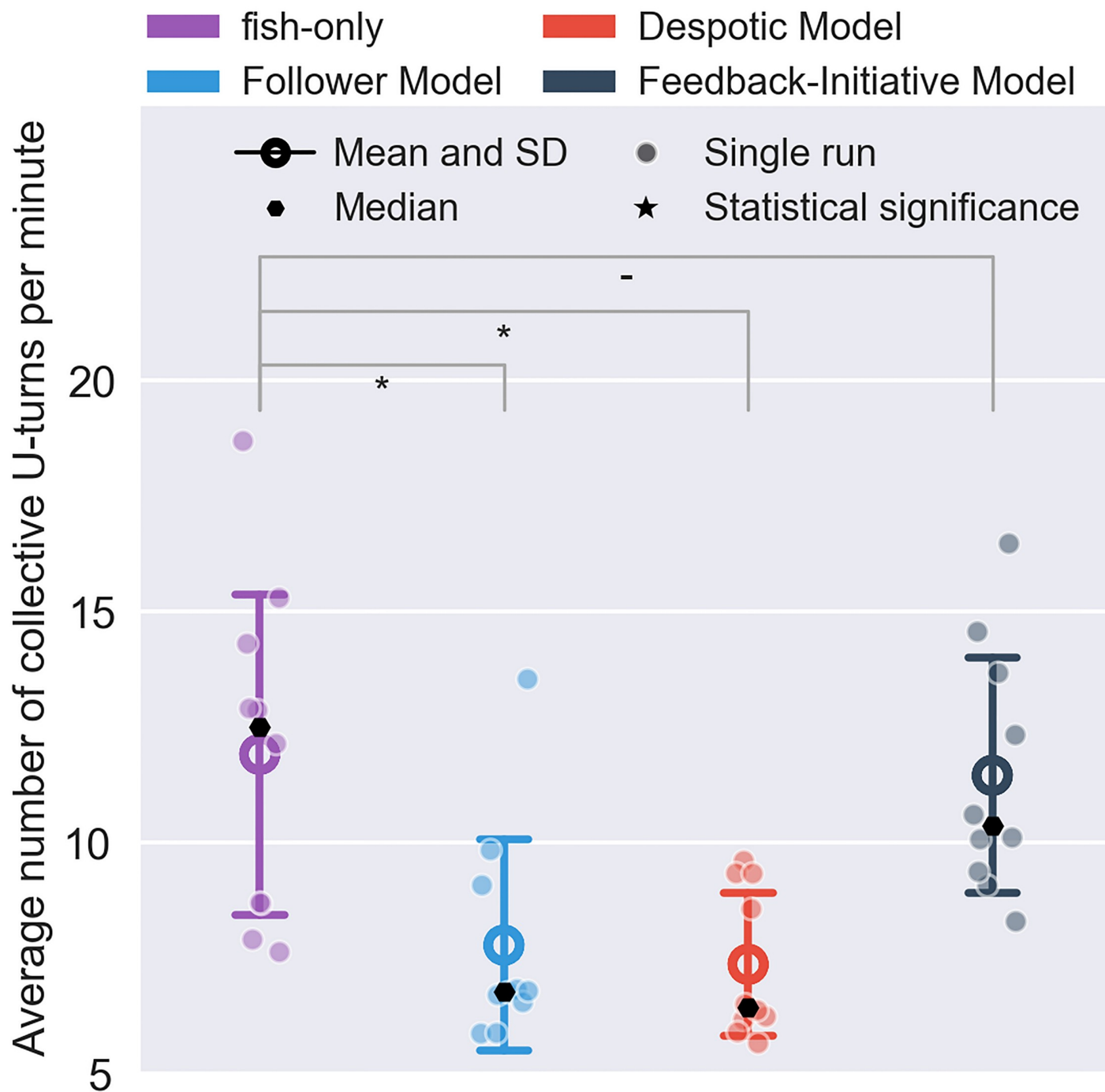
Fig 5. Average number of collective U-turns per minute over all replicates. Annotations of the statistical significance (Kruskal–Wallis test folowed by Tukey's honest significant difference criterion post hoc analysis) are marked with a dash or stars. The dash corresponds $p > 0.05$, a single star to $p < 0.05$, two stars to $p < 0.01$, three stars to $p < 0.001$. and four stars to $p < 0.0001$.

These results showed that FM's and DM's poor performance in terms of angular distance translated in poor performance in terms of collective U-turns. While this was to be expected for the DM that was instructing the robot to move CW, FM once again appeared to disrupt the collective dynamics of the shoal. More specifically, DM's low number of U-turns demonstrated
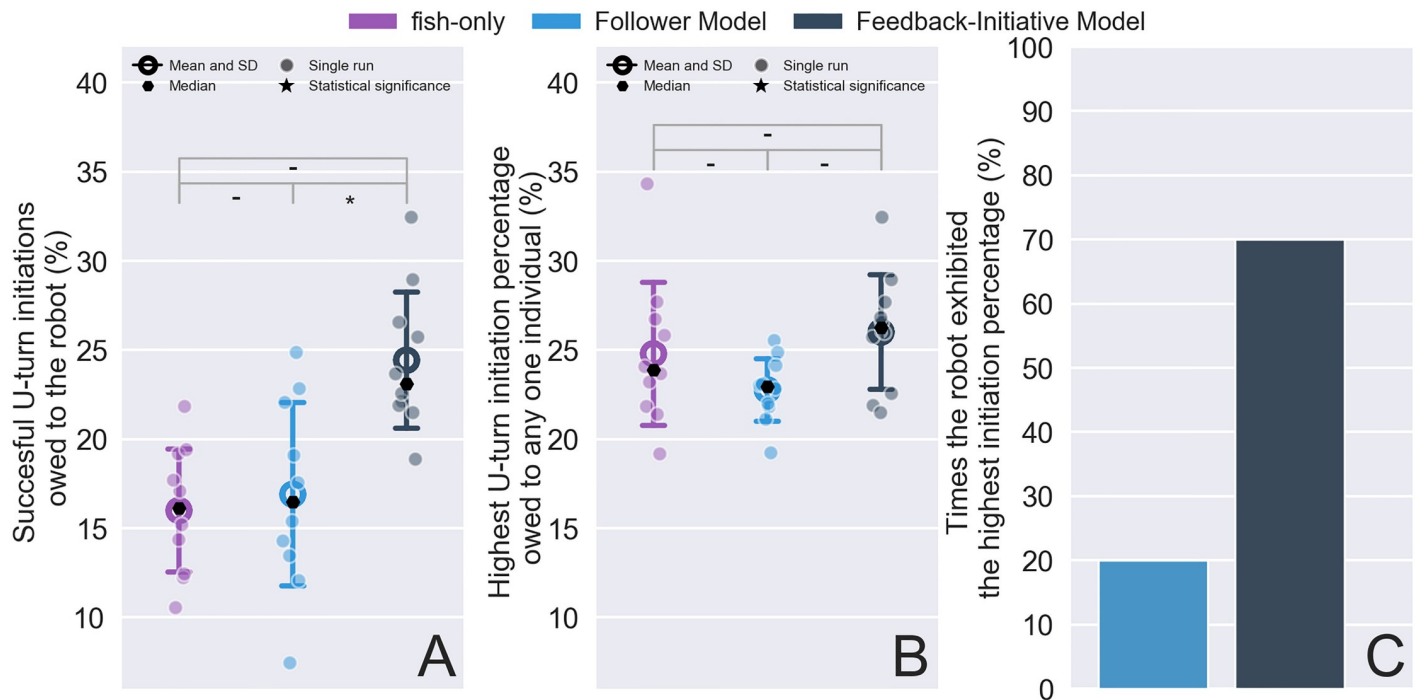
Fig 6. U-turn initiation success rates. (A) successful U-turns that were initiated by the robot (or a random individual in the case of fish-only), (B) highest success rate in an experiment owed to any one individual, and (C) percentage of experiments in which the robot had the highest success rate. Annotations of the statistical significance (Kruskal–Wallis test folowed by Tukey's honest significant difference criterion post hoc analysis) are marked with a dash or stars. The dash corresponds $p > 0.05$, a single star to $p < 0.05$, two stars to $p < 0.01$, three stars to $p < 0.001$. and four stars to $p < 0.0001$.

its ability to influence the collective decision making rather than to participate in it (the shoal moves CW ≈ 65% of the time similar to [27]). Conversely, there was no significant difference between FIM and the fish-only experiments regarding the collective U-turns. This indicates that FIM had strong biomimetic capabilities due to its design, that explicitly embedded the ability to follow but also initiate direction changes.

In addition to the collective U-turns, we also investigated the success rate of the robot to initiate a collective U-turn. In Fig 6A we depict the percentage of successful U-turns that were owed to the robot's motion, in Fig 6B we depict the highest percentage of successful U-turns exhibited by any one individual taking part in the experiment, and in Fig 6C we depict the percentage of the robot that was the most influential individual. We note that in the case of the fish-only experiments we chose one random individual and we excluded the DM experiments since the robot would never perform a U-turn. Intuitively, the above measurements can provide an estimate of the leadership characteristics of each model compared to the innate behavior of the zebrafish. Moreover, the distributions depicted in Fig 6A and 6B provide, once again, insight on the degree to which the robot might have been perturbing or naturally interacting with the living individuals.

We performed KW test for the successful U-turn initiation distribution and obtained the values of $p < 0.05$ and $\chi^2 = 7.65$. A follow-up T-HSD post hoc test revealed that fish-only does not differ significantly from the FM ($p > 0.5$), while it did indeed differ significantly from FIM ($p < 0.5$). FM also differs significantly from FIM ($p < 0.05$). While Fig 6B alone does not provide a lot of additional information (the distributions d not differ significantly KW test, $p > 0.05$), in combination with Fig 6A we notice that the robot's U-turn initiation success rate was very similar to the distribution for the fish-only individuals with the highest success rate.
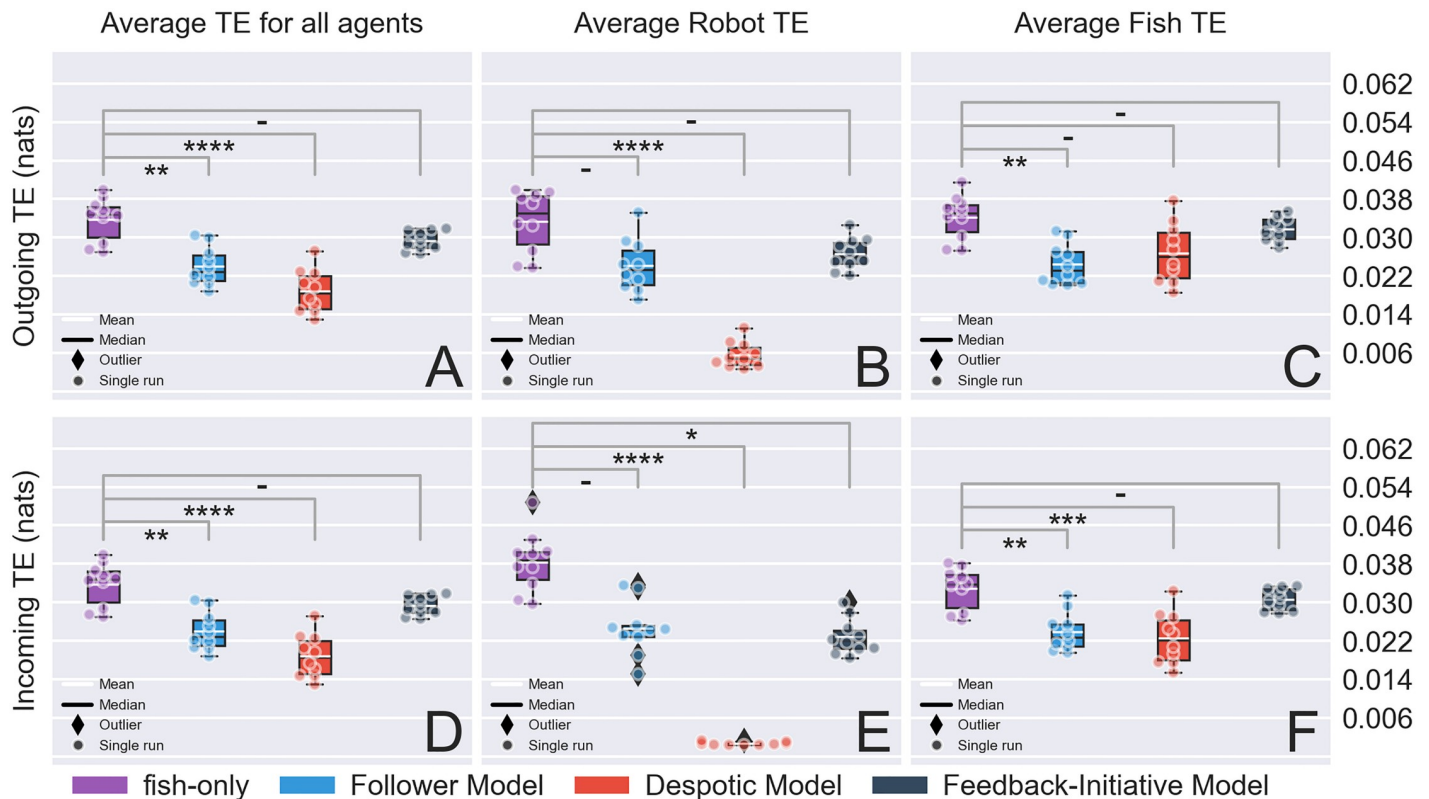
Fig 7. Average transfer entropy (TE) for direction time-series. The first row (i.e., Fig A, B, C) corresponds to the average outgoing TE of the shoal (i.e., amount of TE from focal towards other individuals) and the second (i.e., D, E, F) corresponds to the average incoming TE (i.e., amount of TE from other individuals towards the focal). A & D: TE for the mixed group (all individuals are considered), B & E: TE only for the robotic agent (for fish-only experiments a random fish replaces the robot) and C & F: TE only for living individuals (i.e. the robot is excluded in the computation and for fish-only experiments a random fish and excluded from the analysis). Annotations of the statistical significance (Kruskal–Wallis test followed by Tukey's honest significant difference criterion post hoc analysis) are marked with a dash or stars. The dash corresponds $p > 0.05$, a single star to $p < 0.05$, two stars to $p < 0.01$, three stars to $p < 0.001$. and four stars to $p < 0.0001$. The complete pairwise comparisons can be found in S5, S6, S7 and S8 Tables, name, S10 Table.

In Fig 6C we quantified the latter in terms of the percentage that the robot acted as the leading individual and found than in FM experiments this corresponds to 20% and in FIM experiments to 70%. While FIM was clearly more successful in initiating a U-turn, FM's success rate was greater than what could be expected by a following model. This is, was fact, directly linked to the densest point alternation (see Follower Model (FM)), which could very well have triggered a U-turn for the robot if the densest centroid appeared in the reverse direction. Overall, this provides evidence that the FIM was capable of producing patterns that did not perturb the collective and at the same time allowed it to have a leadership role with higher, compared to a random individual, percentage.

Transfer entropy

We complete this section by evaluating the information propagation capabilities of each model by resorting to information theory and more specifically to the use of TE (see section Transfer entropy). In Fig 7, we measured the TE for all shoal members (fish and robots) quantifying the average influence that individuals exerted on (outgoing TE) or received from (incoming TE) the others.

First, we analyzed the average outgoing entropy for all agents (Fig 7A). Again, the distributions significantly differ from each other across the different treatments (KW test, $p < 0.00001$,

and $\chi^2 = 26.78$). The fish-only condition, that showed the highest TE values, differs significantly from the FM (T-HSD post hoc test, $p < 0.01$) and DM (T-HSD post hoc test, $p < 0.0001$) but not from the FIM (T-HSD post hoc test, $p \approx 0.59$).

The lower performance observed in the mixed groups could be partly attributed to the robot's slow response to stimuli or the models' lack of locomotive aspects that might play an important role in good integration (e.g., biomimetic locomotion patterns). On the other hand, it is important to note that the amount of outgoing directional information exchanged did not significantly differ in the case of FIM versus fish-only which in turn implies that the robotic lure had a considerable impact on the shoal. In fact, FIM stood out compared to the rest of the models in terms of distribution similarity, therefore, we conclude that its biomimetic decision making was indeed important when it came to propagating information within the shoal.

To highlight the role played by the robot in the shoal dynamics, we separated the average outgoing TE of the robot (Fig 7B) and the average outgoing TE of the fish (Fig 7C). The intuition behind this threefold separation (average TE for all agents, average robot TE, average fish TE) is summarized as follows: (1) an overall estimate of how each model affected directional information transfer in the shoal, (2) a quantification of the robot's interaction with the fish and (3) an evaluation of the perturbations in information transfer between the fish due to the presence of the robot. We note that for the fish-only box-plots of Fig 7B and 7C, we chose one fish at random since no robot is used.

For the robot TE case in Fig 7B, the distributions differ significantly (KW test, $p < 0.00001$ and $\chi^2 = 27.19$). A multiple comparison of the distributions showed that: fish-only did not differ significantly from FM (T-HSD post hoc test, $p \approx 0.11$) and FIM (T-HSD post hoc test, $p \approx 0.54$) but differed significantly from DM (T-HSD post hoc test, $p < 0.0001$).

Then, we computed the average outgoing TE exchange only among the fish for the different conditions (Fig 7C). The KW test showed that the distributions differ significantly with each other (KW test, $p < 0.05$). The multiple comparisons post hoc test revealed that: fish-only differ significantly from FM (T-HSD post hoc test, $p < 0.01$), but not from DM (T-HSD post hoc test, $p > 0.05$) and FIM (T-HSD post hoc test, $p \approx 0.93$). In these cases, FIM performed closely to the control experiments. We believe that this could be linked to the degree of acceptance of the robot by the society. More specifically, if the robot does not perturb the directional information propagation, or ideally contributes to it, it might have higher chances to be accepted as a conspecific. In that respect, FIM seems to be the better model out of the ones we tested.

Similarly, we computed the average incoming TE for all individuals Fig 7D, only the robot Fig 7E, and only the fish for all conditions Fig 7F. For the incoming TE for all agents, the results were identical to the one obtained for the average outgoing TE shown in Fig 7A as the amount of information exchanged is preserved within the system but distributed differently among the fish and the robot.

Concerning the average incoming TE of the robot (see Fig 7E), the distributions were significantly different, as observed for the average outgoing TE (KW test, $p < 0.0001, \chi^2 = 32.21$). From the complementary T-HSD post hoc analysis we obtained the values $p \approx 0.07$, $p < 0.0001$, $p < 0.05$ for fish-only versus FM, fish-only versus DM and fish-only versus FIM, respectively. In this case, the FIM under-performed compared to FM. However, this could be expected as the FM, a purely reactive model, was constantly instructing the robot to follow the fish while the FIM could lead the robot to take an initiative that contradicted the behavior of the fish.

For the average incoming TE of only the fish (Fig 7F), we also observed a significant effect of the conditions on the average TE (KW test, $p < 0.001$). However, contrarily to the average outgoing TE, the multiple comparisons showed that fish-only significantly differs from the FM (T-HSD post hoc test, $p < 0.01$) and DM (T-HSD post hoc test, $p < 0.001$) but not from the

FIM (T-HSD post hoc test, $p \approx 0.89$). These results confirmed that the robot controlled by the FIM did not impede the transfer of information between the fish.

Discussion

Testing theoretical hypotheses in realistic conditions is an imperative step towards understanding the collective dynamics of natural systems. However, generating specific patterns that are valuable to validate those hypotheses requires sophisticated physical systems. In the case of animal studies, and specifically the study of zebrafish's group interactions, such systems must blend well enough in the shoal as to allow for natural and life-like interaction dynamics to emerge. Thus, apart from visual biomimetic cues, a robotic device ought to behave as close to the living creature as possible. In turn, this raises questions on the necessity of complex behavioral models in order for an artificial agent to socially interact with a high degree of integration in the group. Here, we showed that a model, that has been simplified to be implemented on a physical system, allowed the robot to establish life-like interactions with a shoal of fish through a bidirectional communication scheme.

While one could assume that a simple following model that instructs a robotic agent to move towards the fish would quite naturally succeed in "infiltrating" the shoal, we showed that our own Follower model (FM), failed to do so. Although further experiments need to be conducted to understand the underlying mechanisms that FM failed to capture, intuitively, a lure that is not attempting to interact with the living agents is less appealing to them and at times is completely disregarded (in the collective decision-making sense). Such hypotheses were already studied in related work (e.g. [26]), but, here, we were able to do so in a reproducible way, in experiments of 30 minutes, with a closed-loop interaction between the animals and the robot that was permanent during the entire experiment. The results of the previous section also showed that this passive control scheme perturbed the behavior of the fish as we observed greater mean inter-individual distance and at the same time fewer U-turns performed on the global scale. This was also validated by visual inspection of the corresponding experiment recordings. Moreover, we attempted to trace the source of this failure by using a TE metric to estimate the amount of information that is exchanged when this model is active on the robot. While such a metric can not be used to safely draw causal conclusions, the results implied that there was a significantly different trend in the information flow for this model (see Fig 7) that could explain the lack of similarity on the global scale (i.e., inter-individual distance, U-turns and successful U-turn initiation).

Another baseline, yet very informative, experiment we conducted is related to the Despotic model (DM). In contrary to the FM, we set out to test the response of the fish when the scenario is inverted, that is when the robot is disregarding their decisions concerning the direction of movement. More specifically, we aimed to test two extreme cases and observed the responses for each one. As shown in section Feedback-Initiative Model (FIM)., DM also fails to capture the interest of the living individuals for long periods of time. Conversely, the Feedback-Initiative model (FIM) managed to exhibit patterns that proved to lead to similar dynamics on a global scale. Especially the results depicted in Fig 5 concerning the collective U-turns, showed that the living individuals interact with the robot and between them in a similar manner as groups of only fish. Similarly, the TE measurements implied that the robot managed to establish stronger communication channels with the fish that could, in turn, explain the similarity of the U-turn distributions. Even more interestingly, we noticed that the robot had a leading role (i.e., a direction change of the robot was likely to propagate to the remaining group members) for the majority of the experiments and had a very similar influence to the most influential fish individual of the fish-only experiments.

In this study, we implemented different behavioural models on a robot-fish to test their ability to interact with the bio-hybrid group members. In particular, we segmented the model space in three categories spanning it; (1) passive models (FM), (2) reactive models (DM), and (3) intermediate models, like the FIM where the robot alternates between (1) and (2). The contribution of this work lies in the following two; (1) demonstrating that the robotic framework we developed is able to exhibit encoded models to study the collective behavior of small fish species, such as the one of [61] from which the FIM was inspired, and (2) the comparative study of those three fundamentally different approaches to the robot control problem and it constitutes a first towards building more complex models that can elicit even more complex responses by leveraging the findings of this study.

Limitations

Embedded systems have come a long way in the past century and have allowed for the miniaturization of robotic systems to the point where they can interact with small animals like fish [22–24, 26]. While most of those systems can achieve similar acceleration and velocity profiles to the zebrafish themselves, they are still constrained by the physical laws. That is, such systems need to account for friction, are often bound to the trade-off of size versus controllability, typically rely on additional parts that affect water flow and in the case of multi-robot systems they need to maintain a safe distance (e.g., fish tend to cross on the y-axis but to this day there is no physical robot that can achieve this). Here, we relied on a two-part system, a miniature robot and a lure, that has already proved to be significantly biomimetic to "blend" in a shoal of fish [22]. In some cases, we found that each of the models presented required a considerable amount of manual tuning of their parameters to reach realistic motion profiles. On that end, our control procedures could benefit from the advancements in machine learning that have been shown to be able to tune such models in real-time [84]. Moreover, while the models used did not require more computational power than what was available, we expect that future experiments with multiple robots and more complex models might and, therefore, the future work needs to focus on both data-efficient algorithms for the adaptation step of the robot and to making use of modern computational systems that can deal with this intensive load.

An alternative approach that was proposed in order to address some of the physical limitations of such a robotic system was the use of animated images (e.g., a fish image) projected into the experimental arena [59]. This would in principle allow for very rapid movements without the complexity owed to the use of physical systems that can not, for example, operate in very small distances. However, while such systems offer a great alternative, in the case of zebrafish and animals in general there might be interactions that are exclusively elicited through physical contact or properties (e.g., the zebrafish will perceive changes in water flow and adapt accordingly).

In addition to the physical limitations of such a complex robotic system, we were also faced with the task of ensuring that the individuals used in this study were not acclimatized to the effects of the robot. For this purpose, we used the same protocol established and validated in our prior studies, that is, the models were encoded in a random way for sequential experiments and no fish was used twice each day. The policy enforced by the host country (Switzerland) concerning the refined and reduced usage of animals for experimentation allowed for the of only two aquariums and a small number of fish (see section Experimental procedure). Since the fish had to be returned in the two aquariums available at the time, separating the groups tested was not feasible. To address this, we randomly caught fish from the aquariums to create mixed populations with different prior experiences and account for the same robot effect.

Future work

The evaluation of the extent to which the robot can "blend" in a group of living animals and socially interact with them is imperative in order to establish useful and informative baselines about the design, implementation and more importantly the prospects of mixed groups of animals and robots. A robotic tool capable of "convincing" its social companions that it is part of their group is likely the first step towards an ethological tool that allows for probing very specific responses from the bio-hybrid group. For this to happen, there is the need for realistic decision making models that can lead meaningful interactions, for robots agile enough to adapt to the quick responses of the fish and for robust and data-efficient algorithms that can potentially tune the models online to allow the robot to integrate in the group for long periods of time. The results conducted within this research work, imply that a highly integrated agent might be more influential and well integrated with a group of 5 zebrafish. While such a claim needs to be studied further, it becomes more obvious that to validate collective behavior models, data simulations are not sufficient as they can not model or predict its true impact on the individuals themselves, that is, there is a reality gap between simulation and physical world. Thus, the use of robotic systems becomes increasingly more important in view of more complex studies on the underlying mechanisms of natural systems. Thus far, our platform has allowed us to test scenarios, with various fish species interacting with one or more robots that could embed similar models to the one developed in this work, in order to study complex collective behavior with groups of multiple fish and robots.

Supporting information

S1 Text. Additional details on experimentation schedule and duration.
(PDF)

S1 Table. Post hoc analysis for the average angular distance distributions preceded by a Kruskal–Wallis test and using Tukey's honest significant difference.
(PDF)

S2 Table. Post hoc analysis for the average number of collective U-turns distributions preceded by a Kruskal–Wallis test and using Tukey's honest significant difference criterion.
(PDF)

S3 Table. Post hoc analysis for the U-turn initiation success rate distributions preceded by a Kruskal–Wallis test and using Tukey's honest significant difference criterion.
(PDF)

S4 Table. Post hoc analysis for the highest U-turn initiation success rate by any one individual distributions preceded by a Kruskal–Wallis test and using Tukey's honest significant difference criterion.
(PDF)

S5 Table. Post hoc analysis for the average outgoing transfer entropy distributions preceded by a Kruskal–Wallis test and using Tukey's honest significant difference criterion.
Average outgoing TE including all agents.
(PDF)

S6 Table. Post hoc analysis for the average outgoing transfer entropy distributions preceded by a Kruskal–Wallis test and using Tukey's honest significant difference criterion.
Average outgoing TE only for the robot's contribution.
(PDF)

S7 Table. Post hoc analysis for the average outgoing transfer entropy distributions preceded by a Kruskal–Wallis test and using Tukey's honest significant difference criterion. Average outgoing TE only for the fish contribution.
(PDF)

S8 Table. Post hoc analysis for the average incoming transfer entropy distributions preceded by a Kruskal–Wallis test and using Tukey's honest significant difference criterion. Average incoming TE including all agents (identical to outgoing case for all agents).
(PDF)

S9 Table. Post hoc analysis for the average incoming transfer entropy distributions preceded by a Kruskal–Wallis test and using Tukey's honest significant difference criterion. Average incoming TE only for the robot's contribution.
(PDF)

S10 Table. Post hoc analysis for the average incoming transfer entropy distributions preceded by a Kruskal–Wallis test and using Tukey's honest significant difference criterion. Average outgoing TE only for the fish contribution.
(PDF)

S1 Fig. Duration of consecutive movement towards a direction (across all 10 replicates). Fish-only case.
(TIF)

S2 Fig. Duration of consecutive movement towards a direction (across all 10 replicates). Fish-only case where a random individual is excluded.
(TIF)

S3 Fig. Duration of consecutive movement towards a direction (across all 10 replicates). Follower model.
(TIF)

S4 Fig. Duration of consecutive movement towards a direction (across all 10 replicates). Follower model with the robot excluded from the analysis.
(TIF)

S5 Fig. Duration of consecutive movement towards a direction (across all 10 replicates). Despotic model.
(TIF)

S6 Fig. Duration of consecutive movement towards a direction (across all 10 replicates). Despotic model with the robot excluded from the analysis.
(TIF)

S7 Fig. Duration of consecutive movement towards a direction (across all 10 replicates). Feedback-Initiative model.
(TIF)

S8 Fig. Duration of consecutive movement towards a direction (across all 10 replicates). Feedback-Initiative model with the robot excluded from the analysis.
(TIF)

S1 Video. Recordings (segments) of the models in action. A short video depicting each model's resulting dynamics for a duration of 60 seconds. The full videos are available upon request

due to the large file size.
(MP4)

Acknowledgments

Author Contributions

Conceptualization: Frank Bonnet, Bertrand Collignon.

Data curation: Vaios Papaspyros.

Formal analysis: Vaios Papaspyros.

Funding acquisition: Frank Bonnet, Francesco Mondada.

Investigation: Vaios Papaspyros, Frank Bonnet.

Methodology: Vaios Papaspyros, Frank Bonnet, Bertrand Collignon.

Project administration: Frank Bonnet, Francesco Mondada.

Software: Vaios Papaspyros.

Supervision: Francesco Mondada.

Visualization: Vaios Papaspyros.

Writing – original draft: Vaios Papaspyros.

Writing – review & editing: Vaios Papaspyros, Frank Bonnet, Bertrand Collignon.

References

1. Camazine S, Deneubourg JL, Franks NR, Sneyd J, Bonabeau E, Theraula G. Self-organization in biological systems. Princeton university press; 2003.

2. Sumpter DJ. Collective animal behavior. Princeton University Press; 2010.

3. Collignon B, Séguret A, Chemtob Y, Cazenille L, Halloy J. Collective departures in zebrafish: profiling the initiators. arXiv preprint arXiv:170103611. 2017;.

4. Aoki I. A simulation study on the schooling mechanism in fish. Bulletin of the Japanese Society of Scientific Fisheries (Japan). 1982;. https://doi.org/10.2331/suisan.48.1081

5. Bode NW, Franks DW, Wood AJ. Limited interactions in flocks: relating model simulations to empirical data. Journal of The Royal Society Interface. 2010; 8(55):301–304. https://doi.org/10.1098/rsif.2010.0397

6. Reynolds CW. Flocks, Herds, and Schools: A Distributed Behavioral Model. Computer Graphics. 1987; 21(4):25–34. https://doi.org/10.1145/37402.37406

7. Couzin ID, Krause J, James R, Ruxton GD, Franks NR. Collective memory and spatial sorting in animal groups. Journal of theoretical biology. 2002; 218(1):1–11. https://doi.org/10.1006/jtbi.2002.3065 PMID: 12297066

8. Lopez U, Gautrais J, Couzin ID, Theraulaz G. From behavioural analyses to models of collective motion in fish schools. Interface focus. 2012; 2(6):693–707. https://doi.org/10.1098/rsfs.2012.0033 PMID: 24312723

9. Niwa HS. Newtonian Dynamical Approach to Fish Schooling. Journal of Theoretical Biology. 1996; 1 (181):47–63. https://doi.org/10.1006/jtbi.1996.0114

10. Vicsek T, Czirók A, Ben-Jacob E, Cohen I, Shochet O. Novel type of phase transition in a system of self-driven particles. Physical review letters. 1995; 75(6):1226. https://doi.org/10.1103/PhysRevLett.75.1226 PMID: 10060237

11. Bertin E, Droz M, Grégoire G. Boltzmann and hydrodynamic description for self-propelled particles. Physical Review E. 2006; 74(2):022101. https://doi.org/10.1103/PhysRevE.74.022101

12. Chaté H, Ginelli F, Grégoire G, Peruani F, Raynaud F. Modeling collective motion: variations on the Vicsek model. The European Physical Journal B. 2008; 64(3-4):451–456. https://doi.org/10.1140/epjb/e2008-00275-9

13. Nagai KH, Sumino Y, Montagne R, Aranson IS, Chaté H. Collective motion of self-propelled particles with memory. Physical review letters. 2015; 114(16):168001. https://doi.org/10.1103/PhysRevLett.114.168001 PMID: 25955073

14. Gautrais J, Ginelli F, Fournier R, Blanco S, Soria M, Chaté H, et al. Deciphering interactions in moving animal groups. Plos computational biology. 2012; 8(9):e1002678. https://doi.org/10.1371/journal.pcbi.1002678 PMID: 23028277

15. Zienkiewicz A, Barton DA, Porfiri M, Di Bernardo M. Data-driven stochastic modelling of zebrafish locomotion. Journal of mathematical biology. 2015; 71(5):1081–1105. https://doi.org/10.1007/s00285-014-0843-2 PMID: 25358499

16. Mwaffo V, Anderson RP, Butail S, Porfiri M. A jump persistent turning walker to model zebrafish locomotion. Journal of The Royal Society Interface. 2015; 12(102):20140884. https://doi.org/10.1098/rsif.2014.0884

17. Strandburg-Peshkin A, Twomey CR, Bode NW, Kao AB, Katz Y, Ioannou CC, et al. Visual sensory networks and effective information transfer in animal groups. Current Biology. 2013; 23(17):R709–R711. https://doi.org/10.1016/j.cub.2013.07.059 PMID: 24028946

18. Mann R, Herbert-Read JE, Ma Q, Jordan LA, Sumpter DJ, Ward A. A model comparison reveals dynamic social information drives the movements of humbug damselfish (Dascyllus aruanus). Journal of the Royal Society Interface. 2014; 11(90):20130794. https://doi.org/10.1098/rsif.2013.0794

19. Landgraf T, Oertel M, Rhiel D, Rojas R. A biomimetic honeybee robot for the analysis of the honeybee dance communication system. In: Intelligent Robots and Systems (IROS), 2010 IEEE/RSJ International Conference on. IEEE; 2010. p. 3097–3102.

20. Stefanec M, Szopek M, Schmickl T, Mills R. Governing the swarm: Controlling a bio-hybrid society of bees & robots with computational feedback loops. In: Computational Intelligence (SSCI), 2017 IEEE Symposium Series on. IEEE; 2017. p. 1–8.

21. Griparić K, Haus T, Miklić D, Polić M, Bogdan S. A robotic system for researching social integration in honeybees. PLoS one. 2017; 12(8):e0181977. https://doi.org/10.1371/journal.pone.0181977 PMID: 28792955

22. Bonnet F, Binder S, de Oliveria ME, Halloy J, Mondada F. A miniature mobile robot developed to be socially integrated with species of small fish. In: 2014 IEEE International Conference on Robotics and Biomimetics (ROBIO 2014); 2014. p. 747–752.

23. Faria JJ, Dyer JR, Clément RO, Couzin ID, Holt N, Ward AJ, et al. A novel method for investigating the collective behaviour of fish: introducing 'Robofish'. Behavioral Ecology and Sociobiology. 2010; 64(8):1211–1218. https://doi.org/10.1007/s00265-010-0988-y

24. Swain DT, Couzin ID, Leonard NE. Real-time feedback-controlled robotic fish for behavioral experiments with fish schools. Proceedings of the IEEE. 2012; 100(1):150–163. https://doi.org/10.1109/JPROC.2011.2165449

25. Landgraf T, Nguyen H, Forgo S, Schneider J, Schröer J, Krüger C, et al. Interactive robotic fish for the analysis of swarm behavior. In: International conference in swarm intelligence. Springer; 2013. p. 1–10.

26. Landgraf T, Bierbach D, Nguyen H, Muggelberg N, Romanczuk P, Krause J. RoboFish: increased acceptance of interactive robotic fish with realistic eyes and natural motion patterns by live Trinidadian guppies. Bioinspiration & biomimetics. 2016; 11(1):015001. https://doi.org/10.1088/1748-3190/11/1/015001

27. Bonnet F, Gribovskiy A, Halloy J, Mondada F. Closed-loop interactions between a shoal of zebrafish and a group of robotic fish in a circular corridor. Swarm Intelligence. 2018;. https://doi.org/10.1007/s11721-017-0153-6

28. Vaughan R, Sumpter N, Henderson J, Frost A, Cameron S. Experiments in automatic flock control. Robotics and autonomous systems. 2000; 31(1-2):109–117. https://doi.org/10.1016/S0921-8890(99)00084-6

29. Jolly L, Pittet F, Caudal JP, Mouret JB, Houdelier C, Lumineau S, et al. Animal-to-robot social attachment: initial requisites in a gallinaceous bird. Bioinspiration & biomimetics. 2016; 11(1):016007. https://doi.org/10.1088/1748-3190/11/1/016007

30. Gribovskiy A, Halloy J, Deneubourg J, Mondada F. Designing a socially integrated mobile robot for ethological research. Robotics and Autonomous Systems. 2018; 103:42–55. https://doi.org/10.1016/j.robot.2018.02.003

31. Halloy J, Sempo G, Caprari G, Rivault C, Asadpour M, Tâche F, et al. Social integration of robots into groups of cockroaches to control self-organized choices. Science. 2007; 318(5853):1155–1158. https://doi.org/10.1126/science.1144259 PMID: 18006751

32. Pillot MH, Gautrais J, Gouello J, Michelena P, Bon R, et al. Moving together: Incidental leaders and naïve followers. Behavioural Processes. 2010; 83(3):235–241. https://doi.org/10.1016/j.beproc.2009.11.006

33. Correll N, Schwager M, Rus D. Social control of herd animals by integration of artificially controlled congeners. In: International Conference on Simulation of Adaptive Behavior. Springer; 2008. p. 437–446.

34. Worm M, Landgraf T, Prume J, Nguyen H, Kirschbaum F, von der Emde G. Evidence for mutual allocation of social attention through interactive signaling in a mormyrid weakly electric fish. Proceedings of the National Academy of Sciences. 2018; p. 201801283. https://doi.org/10.1073/pnas.1801283115

35. Strömbom D, Mann RP, Wilson AM, Hailes S, Morton AJ, Sumpter DJ, et al. Solving the shepherding problem: heuristics for herding autonomous, interacting agents. Journal of the royal society interface. 2014; 11(100):20140719. https://doi.org/10.1098/rsif.2014.0719

36. Kopman V, Porfiri M. Design, modeling, and characterization of a miniature robotic fish for research and education in biomimetics and bioinspiration. IEEE/ASME Transactions on mechatronics. 2013; 18(2):471–483. https://doi.org/10.1109/TMECH.2012.2222431

37. Marras S, Porfiri M. Fish and robots swimming together: attraction towards the robot demands biomimetic locomotion. Journal of The Royal Society Interface. 2012; p. rsif20120084. https://doi.org/10.1098/rsif.2012.0084

38. Romano D, Donati E, Benelli G, Stefanini C. A review on animal–robot interaction: from bio-hybrid organisms to mixed societies. Biological cybernetics. 2019; 113(3):201–225. https://doi.org/10.1007/s00422-018-0787-5

39. Phamduy P, Polverino G, Fuller R, Porfiri M. Fish and robot dancing together: bluefin killifish females respond differently to the courtship of a robot with varying color morphs. Bioinspiration & biomimetics. 2014; 9(3):036021. https://doi.org/10.1088/1748-3182/9/3/036021

40. Polverino G, Phamduy P, Porfiri M. Fish and robots swimming together in a water tunnel: robot color and tail-beat frequency influence fish behavior. PloS one. 2013; 8(10):e77589. https://doi.org/10.1371/journal.pone.0077589 PMID: 24204882

41. Polverino G, Porfiri M. Zebrafish (Danio rerio) behavioural response to bioinspired robotic fish and mosquitofish (Gambusia affinis). Bioinspiration & biomimetics. 2013; 8(4):044001. https://doi.org/10.1088/1748-3182/8/4/044001

42. Abaid N, Bartolini T, Macrì S, Porfiri M. Zebrafish responds differentially to a robotic fish of varying aspect ratio, tail beat frequency, noise, and color. Behavioural brain research. 2012; 233(2):545–553. https://doi.org/10.1016/j.bbr.2012.05.047 PMID: 22677270

43. Butail S, Ladu F, Spinello D, Porfiri M. Information flow in animal-robot interactions. Entropy. 2014; 16(3):1315–1330. https://doi.org/10.3390/e16031315

44. Ladu F, Bartolini T, Panitz SG, Chiarotti F, Butail S, Macrì S, et al. Live predators, robots, and computer-animated images elicit differential avoidance responses in zebrafish. Zebrafish. 2015; 12(3):205–214. https://doi.org/10.1089/zeb.2014.1041 PMID: 25734228

45. Ladu F, Mwaffo V, Li J, Macrì S, Porfiri M. Acute caffeine administration affects zebrafish response to a robotic stimulus. Behavioural brain research. 2015; 289:48–54. https://doi.org/10.1016/j.bbr.2015.04.020 PMID: 25907748

46. Polverino G, Abaid N, Kopman V, Macrì S, Porfiri M. Zebrafish response to robotic fish: preference experiments on isolated individuals and small shoals. Bioinspiration & biomimetics. 2012; 7(3):036019. https://doi.org/10.1088/1748-3182/7/3/036019

47. Spinello C, Macrì S, Porfiri M. Acute ethanol administration affects zebrafish preference for a biologically inspired robot. Alcohol. 2013; 47(5):391–398. https://doi.org/10.1016/j.alcohol.2013.04.003 PMID: 23725654

48. Polverino G, Porfiri M. Mosquitofish (Gambusia affinis) responds differentially to a robotic fish of varying swimming depth and aspect ratio. Behavioural brain research. 2013; 250:133–138. https://doi.org/10.1016/j.bbr.2013.05.008 PMID: 23684918

49. Ruberto T, Mwaffo V, Singh S, Neri D, Porfiri M. Zebrafish response to a robotic replica in three dimensions. Royal Society open science. 2016; 3(10):160505. https://doi.org/10.1098/rsos.160505 PMID: 27853566

50. Bartolini T, Mwaffo V, Showler A, Macrì S, Butail S, Porfiri M. Zebrafish response to 3D printed shoals of conspecifics: the effect of body size. Bioinspiration & biomimetics. 2016; 11(2):026003. https://doi.org/10.1088/1748-3190/11/2/026003

51. Kruusmaa M, Rieucau G, Montoya JCC, Markna R, Handegard NO. Collective responses of a large mackerel school depend on the size and speed of a robotic fish but not on tail motion. Bioinspiration & biomimetics. 2016; 11(5):056020. https://doi.org/10.1088/1748-3190/11/5/056020

52. Abaid N, Marras S, Fitzgibbons C, Porfiri M. Modulation of risk-taking behaviour in golden shiners (Notemigonus crysoleucas) using robotic fish. Behavioural processes. 2013;(100):9–12. https://doi.org/10.1016/j.beproc.2013.07.010 PMID: 23876393

53. Butail S, Bartolini T, Porfiri M. Collective response of zebrafish shoals to a free-swimming robotic fish. PLoS One. 2013; 8(10):e76123. https://doi.org/10.1371/journal.pone.0076123 PMID: 24146825

54. Butail S, Polverino G, Phamduy P, Del Sette F, Porfiri M. Influence of robotic shoal size, configuration, and activity on zebrafish behavior in a free-swimming environment. Behavioural brain research. 2014; 275:269–280. https://doi.org/10.1016/j.bbr.2014.09.015 PMID: 25239605

55. Cazenille L, Collignon B, Chemtob Y, Bonnet F, Gribovskiy A, Mondada F, et al. How mimetic should a robotic fish be to socially integrate into zebrafish groups? Bioinspiration & biomimetics. 2018; 13 (2):025001. https://doi.org/10.1088/1748-3190/aa8f6a

56. Kopman V, Laut J, Polverino G, Porfiri M. Closed-loop control of zebrafish response using a bioinspired robotic-fish in a preference test. Journal of the Royal Society Interface. 2013; 10(78):20120540. https://doi.org/10.1098/rsif.2012.0540

57. Porfiri M, Spinello C, Yang Y, Macrì S. Zebrafish adjust their behavior in response to an interactive robotic predator. Frontiers in Robotics and AI. 2019; 6:38.

58. Kim C, Ruberto T, Phamduy P, Porfiri M. Closed-loop control of zebrafish behaviour in three dimensions using a robotic stimulus. Scientific reports. 2018; 8(1):657. https://doi.org/10.1038/s41598-017-19083-2 PMID: 29330523

59. Abaid N, Spinello C, Laut J, Porfiri M. Zebrafish (Danio rerio) responds to images animated by mathematical models of animal grouping. Behavioural brain research. 2012; 232(2):406–410. https://doi.org/10.1016/j.bbr.2012.03.028 PMID: 22469628

60. Crosato E, Jiang L, Lecheval V, Lizier JT, Wang XR, Tichit P, et al. Informative and misinformative interactions in a school of fish. arXiv preprint arXiv:170501213. 2017;.

61. Abaid N, Porfiri M. Fish in a ring: spatio-temporal pattern formation in one-dimensional animal groups. Journal of The Royal Society Interface. 2010; p. rsif20100175. https://doi.org/10.1098/rsif.2010.0175

62. Lecheval V, Jiang L, Tichit P, Sire C, Hemelrijk CK, Theraulaz G. Social conformity and propagation of information in collective U-turns of fish schools. Proc R Soc B. 2018; 285(1877):20180251. https://doi.org/10.1098/rspb.2018.0251 PMID: 29695447

63. Jiang L, Giuggioli L, Perna A, Escobedo R, Lecheval V, Sire C, et al. Identifying influential neighbors in animal flocking. PLoS computational biology. 2017; 13(11):e1005822. https://doi.org/10.1371/journal.pcbi.1005822 PMID: 29161269

64. Bonnet F, Rétornaz P, Halloy J, Gribovskiy A, Mondada F. Development of a mobile robot to study the collective behavior of zebrafish. In: Biomedical Robotics and Biomechatronics (BioRob), 2012 4th IEEE RAS & EMBS International Conference on. Ieee; 2012. p. 437–442.

65. Bonnet F, Cazenille L, Gribovskiy A, Halloy J, Mondada F. Multi-robot control and tracking framework for bio-hybrid systems with closed-loop interaction. In: Robotics and Automation (ICRA), 2017 IEEE International Conference on. IEEE; 2017. p. 4449–4456.

66. Shi J, et al. Good features to track. In: Computer Vision and Pattern Recognition, 1994. Proceedings CVPR'94., 1994 IEEE Computer Society Conference on. IEEE; 1994. p. 593–600.

67. Bradski G. The opencv library (2000). Dr Dobb's Journal of Software Tools. 2000;.

68. Pérez-Escudero A, Vicente-Page J, Hinz RC, Arganda S, De Polavieja GG. idTracker: tracking individuals in a group by automatic identification of unmarked animals. Nature methods. 2014; 11(7):743. https://doi.org/10.1038/nmeth.2994 PMID: 24880877

69. Bumann D, Krause J. Front individuals lead in shoals of three-spined sticklebacks (Gasterosteus aculeatus) and juvenile roach (Rutilus rutilus). Behaviour. 1993; 125(3):189–198. https://doi.org/10.1163/156853993X00236

70. Dijkgraaf S. The functioning and significance of the lateral-line organs. Biological Reviews. 1963; 38 (1):51–105. https://doi.org/10.1111/j.1469-185X.1963.tb00654.x PMID: 14027866

71. Buhl J, Sumpter DJ, Couzin ID, Hale JJ, Despland E, Miller ER, et al. From disorder to order in marching locusts. Science. 2006; 312(5778):1402–1406. https://doi.org/10.1126/science.1125142 PMID: 16741126

72. Lukeman R, Li YX, Edelstein-Keshet L. Inferring individual rules from collective behavior. Proceedings of the National Academy of Sciences. 2010; p. 201001763. https://doi.org/10.1073/pnas.1001763107

73. Hemelrijk CK, Hildenbrandt H. Self-organized shape and frontal density of fish schools. Ethology. 2008; 114(3):245–254. https://doi.org/10.1111/j.1439-0310.2007.01459.x

74. Shannon CE, Weaver W. The Mathematical Theory of Information (Urbana, IL; 1949.

75. Schreiber T. Measuring information transfer. Physical review letters. 2000; 85(2):461. https://doi.org/10.1103/PhysRevLett.85.461 PMID: 10991308

76. Vicente R, Wibral M, Lindner M, Pipa G. Transfer entropy—a model-free measure of effective connectivity for the neurosciences. Journal of computational neuroscience. 2011; 30(1):45–67. https://doi.org/10.1007/s10827-010-0262-3

77. Porfiri M. Inferring causal relationships in zebrafish-robot interactions through transfer entropy: A small lure to catch a big fish. Animal Behavior and Cognition. 2018; 5(4):341–367. https://doi.org/10.26451/abc.05.04.03.2018

78. Mwaffo V, Butail S, Porfiri M. Analysis of pairwise interactions in a maximum likelihood sense to identify leaders in a group. Frontiers in Robotics and AI. 2017; 4:35. https://doi.org/10.3389/frobt.2017.00035

79. Porfiri M, Marín MR. Symbolic dynamics of animal interaction. Journal of theoretical biology. 2017; 435:145–156. https://doi.org/10.1016/j.jtbi.2017.09.005 PMID: 28916452

80. Ruberto T, Clément RJ, Spinello C, Neri D, Macri S, Porfiri M. The tagging procedure of visible implant elastomers influences zebrafish individual and social behavior. Zebrafish. 2018; 15(5):433–444. https://doi.org/10.1089/zeb.2018.1616 PMID: 30070967

81. Wibral M, Pampu N, Priesemann V, Siebenhühner F, Seiwert H, Lindner M, et al. Measuring information-transfer delays. PloS one. 2013; 8(2):e55809. https://doi.org/10.1371/journal.pone.0055809 PMID: 23468850

82. Takens F. Detecting strange attractors in turbulence. In: Dynamical systems and turbulence, Warwick 1980. Springer; 1981. p. 366–381.

83. Lizier JT. JIDT: An information-theoretic toolkit for studying the dynamics of complex systems. Frontiers in Robotics and AI. 2014; 1:11.

84. Cazenille L, Chemtob Y, Bonnet F, Gribovskiy A, Mondada F, Bredeche N, et al. How to Blend a Robot within a Group of Zebrafish: Achieving Social Acceptance through Real-time Calibration of a Multi-level Behavioural Model. arXiv preprint arXiv:180511371. 2018;.