

PAPER • OPEN ACCESS

Quantifying the biomimicry gap in biohybrid robot-fish pairs

To cite this article: Vaios Papaspyros *et al* 2024 *Bioinspir. Biomim.* **19** 046020

View the [article online](#) for updates and enhancements.

You may also like

- [Design methodologies and engineering applications for ecosystem biomimicry: an interdisciplinary review spanning cyber, physical, and cyber-physical systems](#)
Kathryn Hinkelman, Yizhi Yang and Wangda Zuo
- [Bibliometric analysis of global research trends on biomimetics, biomimicry, bionics, and bio-inspired concepts in civil engineering using the Scopus database](#)
Naim Sedira, Jorge Pinto, Isabel Bentes et al.
- [Biomimicry in agrotechnology: Future solution of water problem for the agriculture industry?](#)
N I Othmani, N M Sahak and M Y M Yunos

Bioinspiration & Biomimetics



PAPER

Quantifying the biomimicry gap in biohybrid robot-fish pairs

OPEN ACCESS

RECEIVED
17 March 2024

REVISED
23 May 2024




ACCEPTED FOR PUBLICATION
12 June 2024

PUBLISHED
28 June 2024

Original Content from this work may be used under the terms of the [Creative Commons Attribution 4.0 licence](https://creativecommons.org/licenses/by/4.0/).

Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.



Vaios Papaspyros^{1,*} , Guy Theraulaz² , Clément Sire^{3,*}  and Francesco Mondada¹ 

¹ Mobile Robotic Systems (MOBOTS) group, School of Computer Science, École Polytechnique Fédérale de Lausanne (EPFL), CH-1015 Lausanne, Switzerland

² Centre de Recherches sur la Cognition Animale, Centre de Biologie Intégrative, CNRS, Université de Toulouse III—Paul Sabatier, 31062 Toulouse, France

³ Laboratoire de Physique Théorique, CNRS, Université de Toulouse III—Paul Sabatier, 31062 Toulouse, France

* Authors to whom any correspondence should be addressed.

E-mail: vaios.papaspyros@gmail.com, clement.sire@univ-tlse3.fr, guy.theraulaz@univ-tlse3.fr and francesco.mondada@epfl.ch

Keywords: animal–robot interaction, ethorobotics, collective behavior, biomimicry, deep learning, reality gap.

Supplementary material for this article is available [online](#)

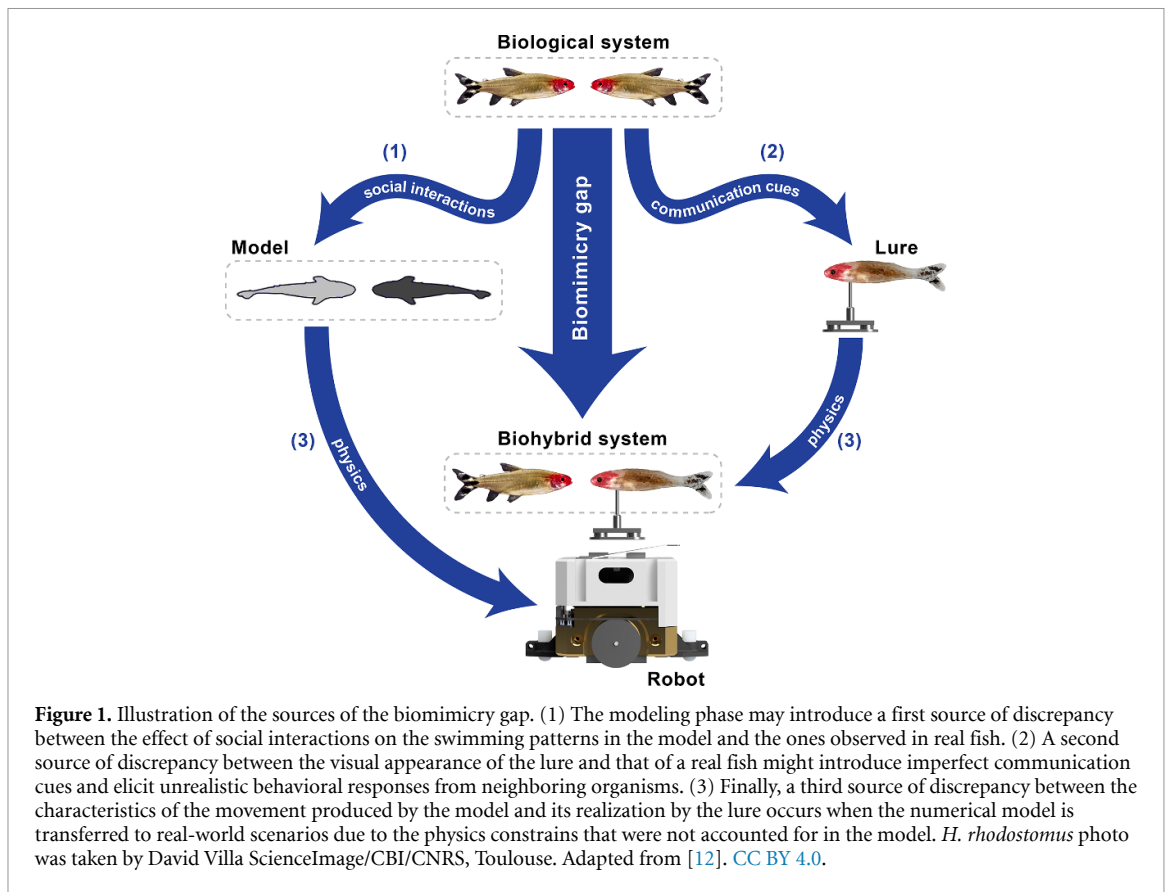
Abstract

Biohybrid systems in which robotic lures interact with animals have become compelling tools for probing and identifying the mechanisms underlying collective animal behavior. One key challenge lies in the transfer of social interaction models from simulations to reality, using robotics to validate the modeling hypotheses. This challenge arises in bridging what we term the ‘biomimicry gap’, which is caused by imperfect robotic replicas, communication cues and physics constraints not incorporated in the simulations, that may elicit unrealistic behavioral responses in animals. In this work, we used a biomimetic lure of a rummy-nose tetra fish (*Hemigrammus rhodostomus*) and a neural network (NN) model for generating biomimetic social interactions. Through experiments with a biohybrid pair comprising a fish and the robotic lure, a pair of real fish, and simulations of pairs of fish, we demonstrate that our biohybrid system generates social interactions mirroring those of genuine fish pairs. Our analyses highlight that: 1) the lure and NN maintain minimal deviation in real-world interactions compared to simulations and fish-only experiments, 2) our NN controls the robot efficiently in real-time, and 3) a comprehensive validation is crucial to bridge the biomimicry gap, ensuring realistic biohybrid systems.

Robot–animal interactions have been increasingly gaining momentum as means to study collective behavior. Biohybrid systems, composed of living organisms and artificial agents, are particularly compelling as they enable researchers to investigate the way animals respond to controlled interactions. This is typically achieved through autonomous robotic devices equipped with species-specific communication channels, which can be employed to evoke responses in a biomimetic or non-biomimetic manner [41]. Robots offer the advantage of conducting repetitive and repeatable experiments, even when driven by complex behavioral models. This is particularly important in the context of social interactions, which encompass considerable complexity when scaling from short-term interactions at the individual level to long-term emergent collective patterns.

Social fish species, such as the rummy-nose tetra (*Hemigrammus rhodostomus*) and zebrafish

(*Danio rerio*), are frequently selected for these studies due to the intricacy of their short- and long-term interactions and their suitability for laboratory environments [10, 41], as well as the abundance of general knowledge about their behavior, genetics, and housing conditions [19, 22, 31, 33, 48]. As a matter of fact, many fish-robot systems have been proposed to investigate various aspects of fish behavior, employing behavioral models with diverse degrees of detail and realistic features, and typically relying on analytical modeling approaches based on observation of fish interaction [6, 7, 15–17, 19, 21, 29, 30, 34, 39, 42, 44, 47]. Concurrently, machine learning-based modeling approaches have gained a growing interest [13, 14, 20, 23, 36], but only a handful have been tested in real-time with a robotic device [13]. These machine learning approaches are usually intended to study collective behavior by predicting motion in simulations alone [13, 20, 23], while the studies that exploit



robotic systems typically evaluate only instantaneous group-level quantities in the short-term timescale [14].

A few flocking models for fish behavior, analytical or machine learning, have been evaluated in extended simulations to study long-term emergent collective behavior [12, 19, 36]. However, these models have not been tested and validated in biohybrid groups. Conversely, numerous models have been implemented on robotic devices without being tested in simulations [2, 8, 14–17, 34]. Furthermore, the majority of these studies involve robot experiments lasting no more than 30 min, with the resulting interaction patterns being rarely or only superficially quantified. Consequently, none of these models have been stringently benchmarked on both short- and long-term timescales within both simulation and fish-robot biohybrid experiments. In fact, previous research indicates that certain models may yield satisfactory biomimetic outcomes in the short term while failing to reproduce emergent dynamics accurately on longer time scales [36].

Moreover, the transfer of computer models of social interactions into robot controllers that operate in real situations involving animals is not straightforward and can generate a discrepancy with the corresponding numerical simulations, akin to the reality gap observed when transferring simulated robot controllers to real-world applications [32]. As depicted in

figure 1, several sources of discrepancy can combine and feed this gap:

- (1) subtle behavioral patterns of the actual animals that social interaction models may fail to capture;
- (2) non-trivial physics related to the operation of the robot in real life that were not accounted for, and which would alter the rendering of the social model;
- (3) the extent of biomimicry exhibited by artificial lures and behaviors [35, 43], resulting in an animal not interacting with the robot as it would do with a conspecific.

We refer to the cumulative effect of these discrepancies with the term ‘biomimicry gap’. Therefore, the biomimicry gap is an inherent aspect of the multifaceted, cross-domain process of creating biohybrid groups composed of animals and robots. To the best of our knowledge, the feasibility of bridging this biomimicry gap—achieved by conducting extended experiments in both simulated and real-world environments, and comparing their results—has yet to be conclusively and rigorously validated across all these levels in a single approach.

In this study, we investigate this notion by employing the (pretrained) machine learning model presented in [36]. We implement this model on a

robotic system which is shown to achieve unprecedented levels of biomimicry, the LureBot [35]. The LureBot consists of an agile mobile robot capable of generating accelerations and velocities that closely mimic those of *H. rhodostomus*. The robot moves between two plates under the tank where the real fish swim. Additionally, a highly biomimetic artificial lure is magnetically attached to the robot and moves in the same tank as the fish. This lure is meticulously designed to faithfully resemble a real *H. rhodostomus*. We exploit the robotic system to execute approximately 11 h of multiple pair experiments wherein the biomimetic lure interacts with a single *H. rhodostomus*. This allows us to measure the behavioral differences between actual and simulated pairs *H. rhodostomus*, as well as, pairs of 1 biomimetic lure and 1 *H. rhodostomus*. In turn, this yields the first end-to-end approach aimed at minimizing the biomimicry gap, and presented in the following sections.

1. Methods

1.1. Experimental procedure

All experiments were conducted with approximately 60 *Hemigrammus rhodostomus* (rummy-nose tetra) fish that were purchased from Amazonie Labège in Toulouse, France. They were housed at the animal facilities of the Centre de Biologie Intégrative (Université de Toulouse—Paul Sabatier) in two 16 L aquariums set to a 12/12 h, dark/light photoperiod, with a 30 min dimming period between light and dark phases. The water temperature in the housing tanks and the experimental setup was maintained at 27°C. Trained technicians regularly measured the fish length (found to be on average 35 mm long), verified the housing conditions and fed the fish daily between 8:30 am and 9:30 am.

The experiments were conducted in a circular arena of radius $R = 25$ cm, where the water was maintained at a height of approximately 5 cm. We used the same source of water used in the housing tanks to assure the same water quality and salinity. All experiments were strictly conducted after feeding and within the fully lit period (between 9:00 am and 20:30 pm). For our experiments, we randomly selected fish from one aquarium. Then, we allowed the fish a 15 min acclimation period. In the case of biohybrid experiments, the robot would not operate during these 15 min. At the end of each experiment, the fish were returned to a temporary tank to ensure that they would not be part of the experiment twice on the same day. In fact, on the next day, we repeated this process with fish from the second tank, thus, ensuring that no single fish is used two days in a row.

1.2. Real-time tracking and robot control

We made use of the LureBot [35], an agile robotic system, along with a high-fidelity biomimetic lure modeled after the *H. rhodostomus* and shown in [35] to be of significant importance to elicit social responses from the fish. The LureBot can achieve linear speeds up to 100 cm s^{-1} with a maximum acceleration measured at 175 cm s^{-2} . This typically corresponds to three times the maximum speed and acceleration achieved by the considered fish species, showing that the LureBot can comfortably match the physical capabilities of the fish.

To close the interaction loop, the experiments were performed with the behavioral observation and biohybrid interaction (BOBI) framework [35], which includes the LureBot, the biomimetic lure, ancillary robot software, and hardware mounted around the aquarium. A 30 Hz camera mounted at the top of the setup keeps track of fish and the artificial lure swimming inside a circular water tank of radius $R = 25$ cm. In addition, a second 30 Hz camera on the bottom of the setup tracks the LureBot. The information is combined to distinguish which of the individuals seen by the top camera is the lure. Furthermore, BOBI is able to track multiple agents (here, only 2 are used) in real time, while maintaining unique IDs for each agent's trajectory. These agent-specific sequences of spatial movement can be exploited by a behavioral model (see section 1.3) to compute real-time individual and collective quantities concerning the biohybrid group, and close the loop of interaction by adapting the robot's behavior with instructions on future movements.

In BOBI (see [35] for more details), the output of such a behavioral model is communicated to a motion controller and converted to motor commands for the differential drive of the LureBot. Here, we use the proportional–integral–derivative (PID) controller as defined in BOBI, that incorporates *a priori* velocity information provided by the behavioral model. The PID combines the linear and angular errors between the LureBot's current and desired position, as well as the model's predicted velocity profile, to smoothly displace the robot.

1.3. Deep Learning Interaction model

We use a pretrained version of the Deep Learning Interaction (DLI) model [36], to generate real-time goal positions for the LureBot [35]. The DLI consists of 7 layers (see figure 2(b): 1st and 4th are LSTM layers [24]; the remaining are densely connected layers; ReLU activations are used for all layers except for the last one which is linear. For a single agent i , the state at time t is defined as a 1×5 vector $\mathbf{s}_i(t)$:

$$\mathbf{s}_i(t) = (\vec{u}_i(t), \vec{v}_i(t), r_w^i(t)) \in \mathbb{R}^5, \quad (1)$$

where $\vec{u}_i(t)$, $\vec{v}_i(t)$ are the 2D position and velocity, respectively, and where $r_w^i(t)$ is the distance of the

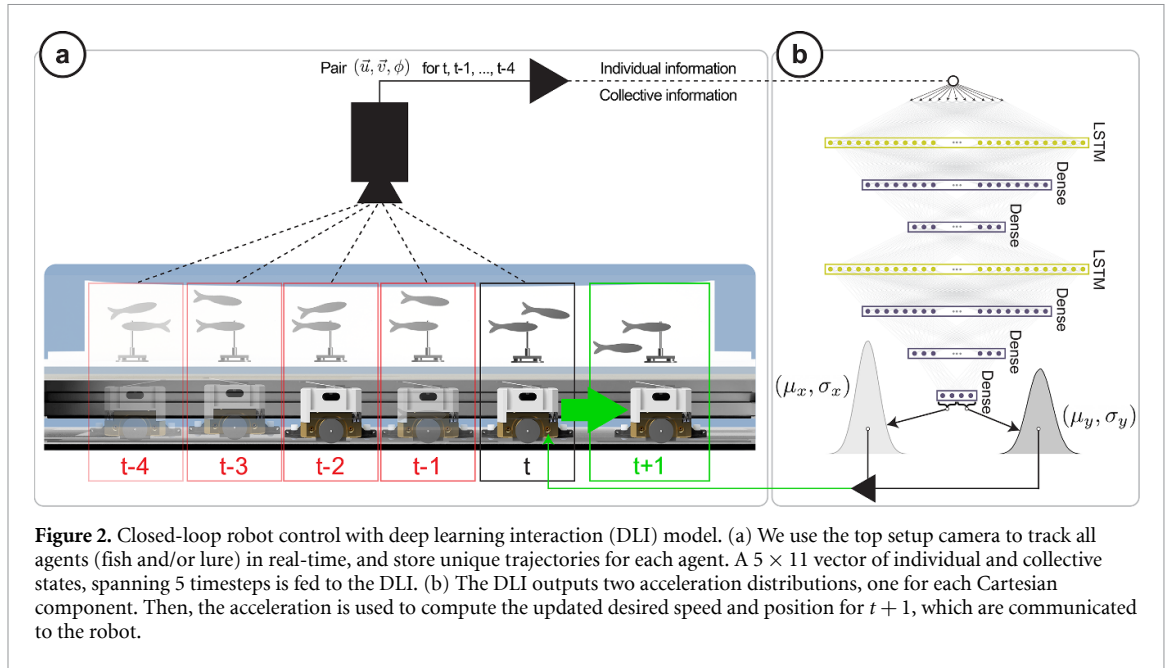


Figure 2. Closed-loop robot control with deep learning interaction (DLI) model. (a) We use the top setup camera to track all agents (fish and/or lure) in real-time, and store unique trajectories for each agent. A 5×11 vector of individual and collective states, spanning 5 timesteps is fed to the DLI. (b) The DLI outputs two acceleration distributions, one for each Cartesian component. Then, the acceleration is used to compute the updated desired speed and position for $t + 1$, which are communicated to the robot.

individual i from the wall at time t . Then, the pairwise state at time t is summarized in the following 1×11 vector:

$$\mathbf{S}_{ij}(t) = \left(\underbrace{\mathbf{s}_i(t)}_{\text{individual (focal) information}}, \underbrace{\mathbf{s}_j(t)}_{\text{individual (neighbor) information}}, \underbrace{d_{ij}(t)}_{\text{collective information}} \right) \in \mathbb{R}^{11}, \quad (2)$$

with i the focal individual for which we generate trajectory predictions, j its neighbor, and d_{ij} their inter-individual distance. In real-time, we feed the DLI with a 5×11 sequence $(\mathbf{S}(t-4), \dots, \mathbf{S}(t))$ of the pair-wise states (see figure 2(a), where we make sure that i (focal individual) corresponds to the LureBot.

Subsequently, the DLI model outputs the expected acceleration mean and standard deviation value, (μ_x, σ_x) and (μ_y, σ_y) , of the Cartesian components x and y . Assuming a Gaussian distribution for the acceleration [18], we sample this distribution to produce acceleration predictions $\vec{a} = (a_x, a_y)$ and use the following motion equations to generate velocity commands and the goal position of the LureBot at time $t + 1$:

$$\vec{v}_i(t+1) = \vec{v}_i(t) + \Delta t \vec{a}, \quad (3)$$

$$\vec{u}_i(t+1) = \vec{u}_i(t) + \Delta t \vec{v}_i(t+1), \quad (4)$$

where $\Delta t = 0.12$ s, a choice made with respect to the data filtering procedure applied on the raw data to generate an intermediate training dataset for the DLI [36]. The 2D velocity commands, defined in equation (3), and goal position, defined in equation (4), are given to the BOBI's PID [35], and eventually translated to motor commands (see section 1.2; see figures 2(a) and (b).

In [36], this approach was validated in long simulations and was shown to be capable of reproducing

the social dynamics of *H. rhodostomus* pairs faithfully compared to experiments. In the following sections, we test the extent to which the DLI can produce faithful interactions when deployed on a physical robot-fish group instead of a simulated group.

1.4. Evaluating the outcome of short- and long-term interactions between fish and the LureBot

Evaluating the extent to which models can reproduce the social dynamics of animal groups, here *H. rhodostomus* fish, is a non-trivial task. As explored in [36], such models may succeed in reproducing quantities in the short-term timescale, but may also fail to reproduce the emergent dynamics in the long term. Here, we opt to benchmark our results by exploiting the 9 observables considered in [25, 36].

The first 3 observables correspond to instantaneous quantities at the *individual* level, for which we measure their probability density function (PDF): the speed of an individual, V ; its distance to the wall, r_w ; and its heading angle relative to the normal to the wall, θ_w . 3 additional observables probe the instantaneous *collective* dynamics: the distance d_{ij} between the pair of individuals; the difference $|\phi_{ij}|$ between the heading directions of the two individuals; and the viewing angle ψ_{ij} at which an individual perceives its neighbor (see figure 3). Finally, we consider 3 *temporal correlation functions* that probe the social dynamics at a very fine level [25], and which are generally particularly difficult to reproduce:

$$C_X(t) = \left\langle [\vec{u}_i(t+t') - \vec{u}_i(t')]^2 \right\rangle, \quad (5)$$

$$C_V(t) = \langle \vec{v}_i(t+t') \cdot \vec{v}_i(t') \rangle, \quad (6)$$

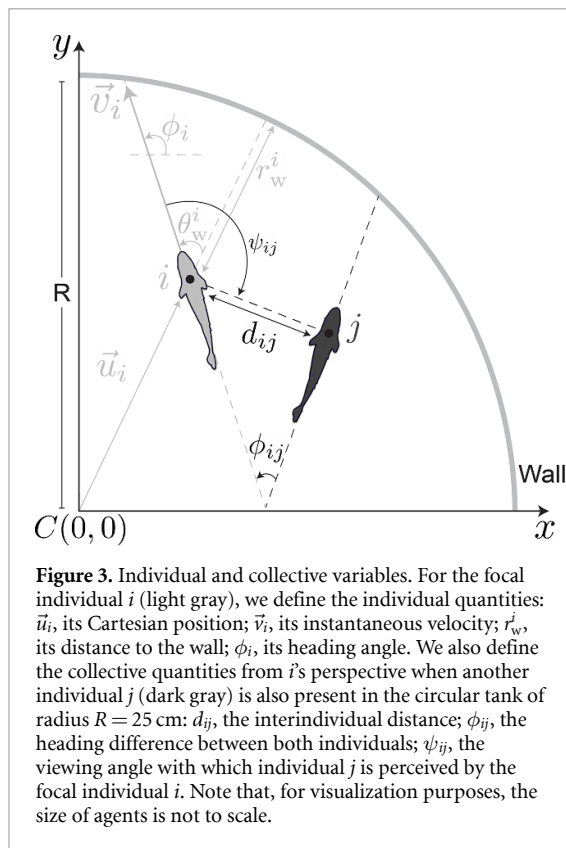


Figure 3. Individual and collective variables. For the focal individual i (light gray), we define the individual quantities: \vec{u}_i , its Cartesian position; \vec{v}_i , its instantaneous velocity; r_w^i , its distance to the wall; ϕ_i , its heading angle. We also define the collective quantities from i 's perspective when another individual j (dark gray) is also present in the circular tank of radius $R = 25$ cm: d_{ij} , the interindividual distance; ϕ_{ij} , the heading difference between both individuals; ψ_{ij} , the viewing angle with which individual j is perceived by the focal individual i . Note that, for visualization purposes, the size of agents is not to scale.

$$C_{\theta_w}(t) = \langle \cos [\theta_w^i(t+t') - \theta_w^i(t')] \rangle. \quad (7)$$

C_X is the mean-squared displacement, C_V the velocity autocorrelation, and C_{θ_w} the autocorrelation of the angle of incidence to the wall. In general, we denote $C_q(t) = \langle q(t+t')q(t') \rangle$ as the average of the quantity $q(t')q(t+t')$ over the reference times t' , over individuals, and over different experiments. Assuming the stationarity of the system, the temporal correlation function $C_q(t)$ only depends on the time difference between observations, and is often noted $C_q(t) = \langle q(t)q(0) \rangle$ (implicitly implying an average over the reference time $t' = 0$).

1.5. Quantifying the (dis)similarity between two PDF: the Hellinger distance

The comparison between the different test cases exploits the 9 observables introduced above and supplementary videos for fish-only experiments, DLI simulated pairs (DLI-SP), and biohybrid pairs (DLIv2-SP)⁴.

For all quantities (PDF and correlation functions), we have computed the statistical and sample to sample standard error by using a bootstrap method. In addition, for each PDF, we report the mean and standard deviation (SD) in table 1, as well as their standard error that we will omit to mention in the

hereafter analysis of the results, for readability (except when their value is relevant to the discussion).

Moreover, in order to compare the PDF for a given quantity between two given test cases, we compute the Hellinger distance between these distributions in table 2. For two PDF F and G for the same quantity x , the Hellinger distance $H(F|G)$ quantifies their (dis)similarity [4, 5]:

$$H(F|G) = \frac{1}{2} \int (\sqrt{F(x)} - \sqrt{G(x)})^2 dx, \quad (8)$$

$$= 1 - \int \sqrt{F(x)}\sqrt{G(x)} dx, \quad (9)$$

where we have used the normalization of the PDE, $\int F(x) dx = \int G(x) dx = 1$, to obtain the last equality. The first definition of $H(F|G)$ clarifies that it measures the overall difference between $F(x)$ and $G(x)$. Meanwhile, the second equivalent definition provides a complementary interpretation in terms of the *overlap* of both PDF. Indeed, the second definition measures the distance from unity of the scalar product of $\sqrt{F(x)}$ and $\sqrt{G(x)}$ seen as vectors of unit Euclidean norm (a consequence of the normalization, $\int \sqrt{F(x)}^2 dx = 1$). The Hellinger distance is 0 if both PDFs are identical, and is bounded by 1, a limit reached if the distributions have a non-overlapping support. In general, a Hellinger distance $H(F|G) \lesssim 0.1$ points to a good agreement between both PDF, $0.1 \lesssim H(F|G) \lesssim 0.2$ points to a fair similarity between them, while $H(F|G) \gtrsim 0.2$ indicates that the two distributions are significantly dissimilar.

1.6. Data for the dynamics of pairs of agents

In this work, we focus on the social dynamics that arise from pairwise interactions in three different conditions. First, we consider ≈ 11 h of experiments involving pairs of *H. rhodostomus*, to characterize and quantify the spontaneous social interactions when no artificial devices are present in the tank.

Secondly, we consider ≈ 16 h of effective trajectories for DLI-SP [36], as a baseline to the robot's underlying model in ideal conditions. This DLI model was originally trained in [36] on a different series of experimental data obtained in [12] for the same species (*H. rhodostomus*), but in different conditions. More specifically, we used a different tank of the same radius $R = 25$ cm, but made of a higher-quality material that is compatible with our robotic system [35]. In addition, our lighting conditions (which greatly impact the fish behavior) are also slightly different and adapted to the constraints for the real-time fish tracking algorithm. We will also mention the results obtained after retraining the DLI model with the fish data considered in the present work, which we will refer to as the DLIv2 simulated pairs (DLIv2-SP) (see table 1 and figures 7–9).

⁴ The videos are also available at <https://doi.org/10.5281/zenodo.8253256>.

Finally, we have conducted ≈ 11 h of experiments where the LureBot propels a biomimetic lure moving inside the circular arena, which is interacting in closed-loop with an actual *H. rhodostomus*. For brevity, in the following analysis of the results, we will simply refer to the LureBot and the lure attached to it as the LureBot. The LureBot is given a pre-trained copy of the DLI model of [36], which is queried in real time to generate biomimetic trajectories (see section 1.3). We refer to these data as DLI biohybrid pairs (DLI-BP). We did not perform experiments with the LureBot trained with the DLiv2 model, since our experimental campaign obviously predated the training of the DLiv2 model, which required these new experimental results.

In all experiments involving fish pairs or LureBot-fish pairs, we have explicitly designed a protocol which did not allow the use of the same fish in an experiment for at least 48 h after their first test, to avoid potential learning effects when the fish interact with the lure (see section 1.1). The fish housing conditions and experiments have been approved by the local ethical committee (see Ethic Statement) and are described in detail in [35].

2. Results

This section reports the detailed comparison between the three test cases: (fish-only) experiments with pairs of *H. rhodostomus*; DLI-SP; DLI-BP, that consist of the LureBot interacting in closed loop with a *H. rhodostomus*. Our results are also qualitatively illustrated by a supplementary video (see <https://zenodo.org/doi/10.5281/zenodo.8253256>) displaying side-by-side trajectories for the three test cases. In addition, at the end of this section, we will briefly present results for DLiv2 simulated pairs (DLiv2-SP; trained on the fish-only experimental data of the present work).

As mentioned above, this section will also exploit the results of table 1 (means and SD and their standard error) and table 2 (Hellinger distances between PDF corresponding to two different conditions). At the end of table 2, we also report the Hellinger distances resulting from the inherent variability observed in fish-only experiments. They were obtained by a bootstrap method by randomly splitting the 14 fish-only experiments in two sets. Then, for each pair of sets, we compute the corresponding Hellinger distances between their associated PDF and average the results over the random draws. We find a mean Hellinger distance (averaged over the 6 observables) of $\bar{H} = 0.10$, which constitutes a baseline for comparing the results of fish-only experiments to other conditions involving biohybrid or simulated pairs.

2.1. Instantaneous individual observables

Figure 4(a) shows the speed PDF for the three cases we considered. Fish pairs swim at a mean speed of 10.5 cm s^{-1} , associated to a standard deviation (SD) of 5.7 cm s^{-1} (see table 1). DLI-SP produce a rather similar speed PDF (Hellinger distance $H = 0.09$; see table 2), albeit slightly wider (SD of 7.0 cm s^{-1}), with a nearly identical mean of 11.1 cm s^{-1} . For biohybrid pairs, the fish and the LureBot have a very similar mean speed (identical within error bars; see table 1), but which is 20% smaller than in the fish-only experiments, although the SD is similar to that of the fish experiments, resulting in a Hellinger distance of $H = 0.18$ (see section 1.5).

In figure 4(b), we plot the PDF of the distance to the wall, r_w , for each case. Fish pairs swim very close to the wall, with a mean distance of 4.4 cm and a SD of 3.9 cm, both comparable to the typical fish body length (~ 3.5 cm). This is a consequence of the burst-and-coast swimming mode exhibited by *H. rhodostomus*, as shown in [12]. Indeed, the motion of this species is characterized by a succession of sudden acceleration periods ('kicks' or bursts of typical duration 0.1 s), each followed by a longer gliding period of typical duration 0.5 s, during which the fish moves in a quasi straight line. Because of the rather narrow distribution of heading changes between kicks, even observed when a fish is far from the wall [12], the fish is unable to escape the concave boundaries of the wall, except when rare large heading changes occur. The mean distance to the wall is 5.7 cm for the DLI-SP, and the associated PDF compared to fish experiments has a Hellinger distance of $H = 0.13$, showing that the DLI model captures reasonably well the tendency of the fish to move close to the wall. For biohybrid pairs, we found that the fish swims farther from the wall than in fish-only experiments, with a mean distance of 5.5 cm. In this case, the LureBot is even farther to the wall, at a mean distance of 6.6 cm, which likely also causes the fish to swim farther to the wall than in fish-only experiments.

Finally, in figure 4(c), we plot the PDF of the absolute value of the heading angle relative to the normal to the wall, $|\theta_w|$. As a consequence of the agents (fish, DLI model, or LureBot) moving close to the wall, we naturally find that the mean of $|\theta_w|$ is very close to, but slightly below 90° (see table 1), with a difference which is statistically significant. Indeed, as already reported in the experiments of [12], the agents spend slightly more time heading toward the wall ($|\theta_w| < 90^\circ$) than moving away from it ($|\theta_w| > 90^\circ$). The PDF for the three considered cases are symmetric around their mean, but we find that the fish experiments lead to the narrowest distribution, with a SD of 22° , compared to a SD of 35° for the DLI-SP, and a SD of 33° and 43° for the fish and the LureBot in a biohybrid pair. The values of these SD are naturally correlated with the mean distance of the agent to the

Table 1. Means and standard deviations. For fish-only experiments, DLI simulated pairs (DLI-SP), and biohybrid pairs (DLI-SP), we report the mean and the standard deviation (SD) of the 6 observables introduced in section 1.4, along with their respective standard error. The speed V is given in cm s^{-1} , the distances r_w and d_{ij} are given in cm, and the angles $|\theta_w|$, $|\phi_{ij}|$, and ψ_{ij} are in degrees. Note the small standard error in the case of the (DLI-SP) resulting from extensive simulations (16.6 h long, almost twice the amount of data collected for other cases) and the fact that the 2 agents are statistically identical. For the biohybrid experiments, we report the mean and SD for V , r_w , and $|\theta_w|$, averaged over the fish and the LureBot, as well as for each of them. Finally, we present the corresponding results for a DLI model retrained on the present fish experiments (DLIv2-SP).

Pair	Quantity	Mean	Standard deviation
Fish-only	V	10.50 ± 0.60	5.73 ± 0.36
	r_w	4.39 ± 0.43	3.86 ± 0.22
	$ \theta_w $	87.42 ± 0.39	21.91 ± 1.46
	d_{ij}	8.05 ± 0.71	5.11 ± 0.43
	$ \phi_{ij} $	26.72 ± 1.91	29.81 ± 1.24
	ψ_{ij}	7.96 ± 4.73	108.98 ± 1.19
DLI-SP	V	11.06 ± 0.04	7.04 ± 0.02
	r_w	5.66 ± 0.03	4.42 ± 0.03
	$ \theta_w $	88.07 ± 0.06	34.55 ± 0.16
	d_{ij}	7.43 ± 0.03	4.38 ± 0.04
	$ \phi_{ij} $	38.06 ± 0.19	38.63 ± 0.17
	ψ_{ij}	-4.11 ± 0.33	107.13 ± 0.06
DLI-BP	V	8.60 ± 0.22	5.93 ± 0.12
	r_w	6.05 ± 0.25	4.76 ± 0.06
	$ \theta_w $	86.44 ± 0.17	38.07 ± 0.73
	d_{ij}	9.96 ± 0.48	6.27 ± 0.33
	$ \phi_{ij} $	58.60 ± 0.91	48.38 ± 0.24
	ψ_{ij}	-7.42 ± 4.16	110.41 ± 0.51
DLI-BP (fish)	V	8.44 ± 0.26	5.13 ± 0.21
	r_w	5.54 ± 0.35	4.54 ± 0.09
	$ \theta_w $	87.46 ± 0.19	32.76 ± 1.25
DLI-BP (robot)	V	8.74 ± 0.16	6.62 ± 0.12
	r_w	6.59 ± 0.15	4.91 ± 0.05
	$ \theta_w $	85.42 ± 0.24	42.78 ± 0.79
DLIv2-SP	V	10.53 ± 0.48	6.18 ± 0.28
	r_w	4.64 ± 0.23	4.37 ± 0.05
	$ \theta_w $	87.56 ± 0.11	26.47 ± 0.47
	d_{ij}	8.39 ± 0.07	6.15 ± 0.11
	$ \phi_{ij} $	30.54 ± 0.30	33.11 ± 0.29
	ψ_{ij}	11.72 ± 0.87	109.08 ± 0.19

wall: the farther the agent, the larger are the fluctuations (SD) of its heading angle relative to the wall.

In summary, the DLI-BP system presents a fair agreement with the experimental results for all quantities. Concurrently, DLI-BP and DLI-SP show smaller dissimilarity, indicating that the transposition of the simulated model into the robot was successful (see table 2). We also observe that, in some cases (e.g. as reflected in the PDF of $|\theta_w|$), the fish's behavior guides the DLI-powered robot, the latter behaving more closely as a fish than the virtual/simulated DLI agents. Nonetheless, the observables test the DLI's performance at a very fine level, especially in the case of DLI-BP, where the physical aspect is also impeding the precise reproduction of the social dynamics, either due to the imperfect (with respect to fish) motion of the robot or the varying degree the robotic system's acceptance by the fish.

2.2. Instantaneous collective observables

H. rhodostomus have a natural tendency to swim in close proximity to each other. In our experiments, fish pairs typically maintain a median interindividual distance d_{ij} of less than two body lengths (see figure 5(a), with a mean distance of 8.05 ± 0.7 cm and a SD of 5.1 cm (see table 1). The dynamics of DLI-SP results in a very similar PDF ($H = 0.16$), with a mean of 7.43 ± 0.03 cm, which is within one standard error (for the fish experiments) from the mean obtained for fish. As for the biohybrid pair, it is less bound than pairs of fish or DLI, with a mean distance between the fish and the LureBot of 9.96 ± 0.5 cm. The distribution is also slightly wider, with a SD of 6.3 cm. In fact, although the peak of the interindividual distance PDF is located at a similar value as for fish or DLI pairs (5–6 cm in the three cases), the biohybrid pairs are more often separated by a distance larger than 15 cm.

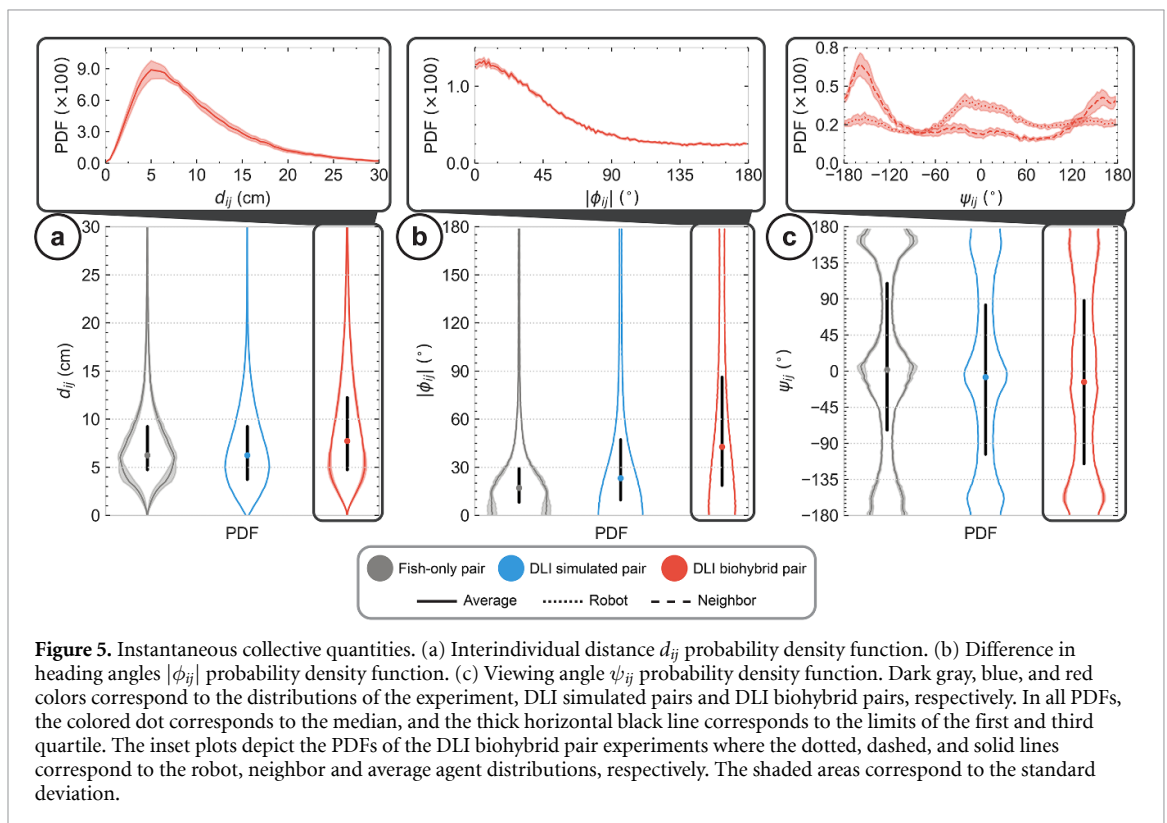
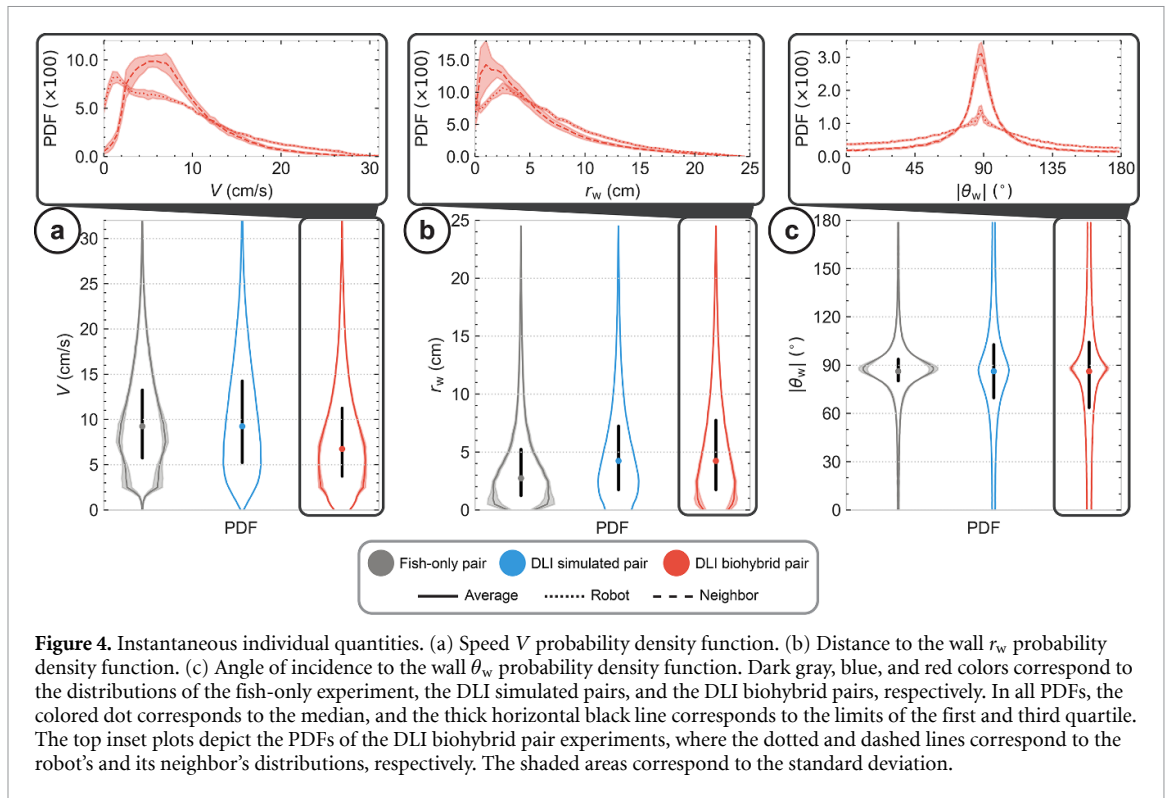
Table 2. Hellinger distances. We exploit the Hellinger distance between two PDF (see section 1.5) to compare the PDF of the 6 observables introduced in section 1.4, for fish-only experiments, DLI simulated pairs (DLI-SP and DLiv2-SP), and biohybrid pairs (DLI-BP). The last condition describes the inherent variability between the 14 fish experiments and is obtained by a bootstrap method by randomly splitting these 14 experiments in two sets, and computing the Hellinger distance between their 6 corresponding PDF. We also report the average Hellinger distance \bar{H} for each condition.

Pair	Quantity	Hellinger distance H
Fish-only vs DLI-SP	V	0.09
	r_w	0.13
	$ \theta_w $	0.23
	d_{ij}	0.12
	$ \phi_{ij} $	0.14
	ψ_{ij}	0.09
	Average \bar{H}	0.13
Fish-only vs DLI-BP	V	0.18
	r_w	0.15
	$ \theta_w $	0.25
	d_{ij}	0.16
	$ \phi_{ij} $	0.30
	ψ_{ij}	0.15
	Average \bar{H}	0.20
DLI-SP vs DLI-BP	V	0.14
	r_w	0.04
	$ \theta_w $	0.04
	d_{ij}	0.18
	$ \phi_{ij} $	0.17
	ψ_{ij}	0.07
	Average \bar{H}	0.11
Fish-only vs DLiv2-SP	V	0.05
	r_w	0.08
	$ \theta_w $	0.08
	d_{ij}	0.14
	$ \phi_{ij} $	0.06
	ψ_{ij}	0.04
	Average \bar{H}	0.08
Fish-only vs Fish-only (Bootstrap)	V	0.07
	r_w	0.09
	$ \theta_w $	0.04
	d_{ij}	0.16
	$ \phi_{ij} $	0.15
	ψ_{ij}	0.09
	Average \bar{H}	0.10

H. rhodostomus is a social species, often found to form well aligned schools. In fact, their pairwise alignment interaction was quantitatively measured in [12], showing that this interaction remains strong up to three body lengths, well within the typical distance between fish. In figure 5(b), to quantify the alignment within pairs of agents, we plot the distribution of the absolute value of the difference between the heading angles of the two agents, $|\phi_{ij}|$ (see the graphical definition in figure 3). The mean heading difference observed in fish experiments is 27° , with a rather narrow PDF associated with a SD of 30° , confirming the good level of alignment between the two fish. The DLI-SP are not as aligned as fish pairs, with a larger mean and SD equal to 38° , although the Hellinger distance between the two PDF ($H = 0.14$) remains satisfactory. The corresponding

PDF for biohybrid pairs exhibits the largest disagreement with the fish experiments of all the PDF presented here ($H = 0.30$). Indeed, despite also being peaked at $|\phi_{ij}| = 0$, the PDF has a non-negligible weight for $|\phi_{ij}| > 90^\circ$, resulting in a much larger mean of 59° and a SD of 48° . This wider PDF is a consequence of the fact that the fish and the LureBot, despite remaining close to each other on average, have a much higher probability than fish pairs to be at a distance above the range of the alignment interaction. Moreover, when the fish and the LureBot are far apart and attempt to get closer, they have a high chance to be actually anti-aligned during this process, hence the significant weight of the PDF near $|\phi_{ij}| = 180^\circ$.

Finally, figure 5(c) shows the PDF of the angle of perception ψ_{ij} , defined in figure 3. For pairs of



fish, the PDF presents clear peaks at $\psi_{ij} = 0^\circ$ and near $|\psi_{ij}| = 180^\circ$. This indicates that the well aligned fish are following each other rather than swimming side by side. For DLI-SP, the same pattern is observed but with slightly less pronounced peaks, although the Hellinger distance of $H = 0.05$ confirms the excellent agreement between both PDF. As for the biohybrid

pair, the PDF averaged over the fish and the LureBot again presents the same peaks as before, but even less pronounced. Again, the less sharp peaks are a consequence of the fact that the biohybrid pairs stand farther from the wall than fish pairs, and above all, of the fact that their distance has a higher probability to be large enough so that their angle of perception

ψ_{ij} becomes uncorrelated. The lesser alignment of the biohybrid pairs (see above) originates from the same causes, and in turn also results in a more homogeneous distribution of the angle of perception. However, the apparent reasonable agreement with the PDF for the fish-only and DLI-SP pairs masks the difference between the PDF for the fish and for the LureBot shown in the top inset of figure 5(c). There, we observe that the peak near $\psi_{ij} = 0^\circ$ is dominated by the contribution of the fish, showing that the fish more often follows the LureBot than the converse. In addition, we find that the PDF for the fish is also peaked slightly above $\psi_{ij} = -180^\circ$, while the PDF for the LureBot has a corresponding peak slightly below $\psi_{ij} = +180^\circ$. By periodicity of 360° , these two peaks are obviously located at almost the same angle, but this slight angular shift translates to the fact that the fish is, on average, slightly closer to the wall than the LureBot, as noted in section 2.1.

The instantaneous collective quantities demonstrate that despite the dissimilarities measured in the individual behavior of both DLI-SP and DLI-BP with respect to the fish-only experiment, the collective dynamics are fairly reproduced. Furthermore, the DLI is transferred in a physical system with good agreement compared to its simulated version, and the living agent responds positively. However, the angular control of the robot is arguably less precise, which contributes to the general deviation from the experimental angle-related distributions.

2.3. Temporal correlation functions

In figure 6, we plot the three observables used to quantify the temporal correlations that emerge in the system during the long-term dynamics, which are defined in section 1.4.

Figure 6(a) shows the mean square displacement of the agents, $C_X(t)$, in the three considered cases. After a rapid growth, $C_X(t)$ presents a peak and an ultimate decay to a mean level equal to twice the mean square of the distance to the center of the tank. Indeed, for large time difference, the positions at time t' and $t + t'$ become uncorrelated, and we obtain

$$\begin{aligned} C_X(t) &= \left\langle [\vec{u}_i(t+t') - \vec{u}_i(t')]^2 \right\rangle \\ &\underset{t \rightarrow +\infty}{\approx} \left\langle \vec{u}_i^2(t+t') \right\rangle + \left\langle \vec{u}_i^2(t') \right\rangle \\ &= 2 \left\langle \vec{u}_i^2(t') \right\rangle, \end{aligned} \quad (10)$$

which becomes time-independent due to the stationarity of the dynamics. Although $C_X(t)$ has the same qualitative form in the three cases, one observes differences in the position and height of the peak and in the asymptotic value. The latter is explained by the fact that the closer the agents are to the wall, the larger is the mean square of their distance to the center of the tank, $\left\langle \vec{u}_i^2(t') \right\rangle$. Indeed, we have found, in section 2.1, that fish pairs swim closest to the wall, while biohybrid pairs are the farthest, which

is consistent with the asymptotic behavior of $C_X(t)$ observed in figure 6(a). Furthermore, the top inset of figure 6(a) for the biohybrid pairs shows that $C_X(t)$ for the fish is systematically larger than for the LureBot, which is also consistent with the fact that the fish swims slightly closer to the wall than the LureBot. As for the position of the peaks in figure 6(a), it roughly corresponds to the time for the corresponding agent to travel half of the tank perimeter. This time is directly correlated with the mean speed of the agent. In section 2.1, we found that the fish pairs and DLI-SP had essentially the same mean speed, which explains the agreement between the position of the corresponding peaks in $C_X(t)$. However, we also found that the biohybrid pairs were 20% slower, which explains the fact that the peak in their $C_X(t)$ is reached at a later time than for fish and DLI pairs.

Figure 6(b) shows the velocity autocorrelation, $C_V(t)$, in the three considered cases, which vanishes for t large enough, when the velocity at time $t + t'$ becomes uncorrelated with that at time t' . It can be formally shown that $C_V(t) = \frac{1}{2} \frac{d^2 C_X}{dt^2}(t)$ (although this relation is only approximate, when the 2 quantities are observed independently over a finite sampling time), so that the interpretation of the shape of $C_V(t)$ results from the analysis that we have presented above for $C_X(t)$. In particular, the peaks of the first two oscillations in $C_V(t)$ roughly correspond to the two inflection points just before and after the main peak in $C_X(t)$. In addition, $C_V(t=0)$ is the mean square velocity, and we indeed observe an agreement between its value for fish and DLI pairs, while the slower biohybrid pairs result in a lower initial value of $C_V(t=0)$ in this case.

Finally, the (most subtle) temporal correlation function of the heading of an agent relative to the wall, $C_{\theta_w}(t) = \left\langle \cos[\theta_w^i(t+t') - \theta_w^i(t')] \right\rangle$, is shown in figure 6(c). For very large time t , $C_{\theta_w}(t)$ must obviously decay, but we observe that for fish pairs, we still have $C_{\theta_w}(t=30\text{s}) \approx 0.35$, indicating strong correlations. For DLI-SP, we find that $C_{\theta_w}(t)$ vanishes very rapidly ($C_{\theta_w}(t=15\text{s}) \approx 0$). Finally, for biohybrid pairs, we still observe some weak remnant correlations at $t=30\text{s}$, with $C_{\theta_w}(t=30\text{s}) \approx 0.1$ (although the correlation is dominated by the contribution of the fish, as shown in the top inset of figure 6(c). Here, the decay rate of $C_{\theta_w}(t)$ is strongly related to the sharpness of the peak near $\theta_w = 90^\circ$ in the PDF of θ_w (see figure 4(c) and section 2.1). Indeed, a sharp peak suggests that it can take a long time to explore values of θ_w far from 90° , leading to a slower decay of $C_{\theta_w}(t)$. Accordingly, we indeed found that the least sharp peak in the PDF of θ_w is observed for DLI-SP, resulting in the fastest decay of $C_{\theta_w}(t)$ in this case.

Both the DLI-SP and DLI-BP fail to precisely reproduce the correlation function $C_{\theta_w}(t)$, producing a very similar sharp decay compared to the one of real fish. This is again due to the DLI's tendency

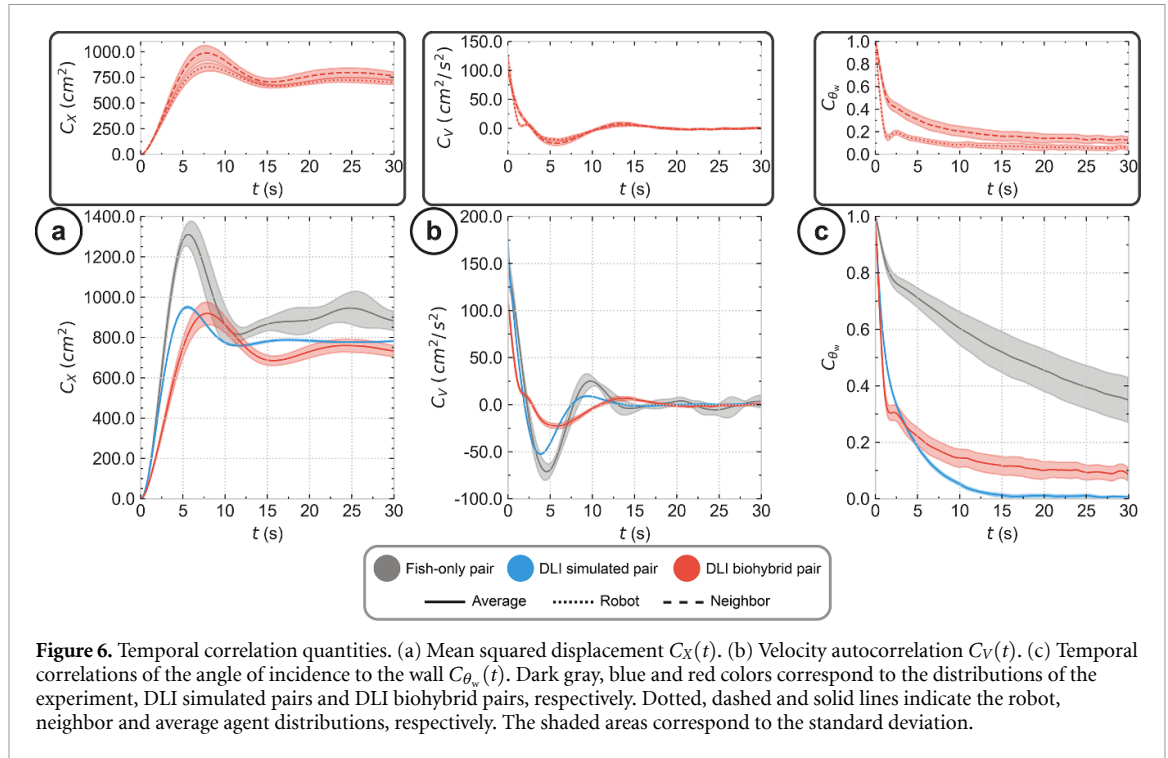


Figure 6. Temporal correlation quantities. (a) Mean squared displacement $C_X(t)$. (b) Velocity autocorrelation $C_V(t)$. (c) Temporal correlations of the angle of incidence to the wall $C_{\theta_w}(t)$. Dark gray, blue and red colors correspond to the distributions of the experiment, DLI simulated pairs and DLI biohybrid pairs, respectively. Dotted, dashed and solid lines indicate the robot, neighbor and average agent distributions, respectively. The shaded areas correspond to the standard deviation.

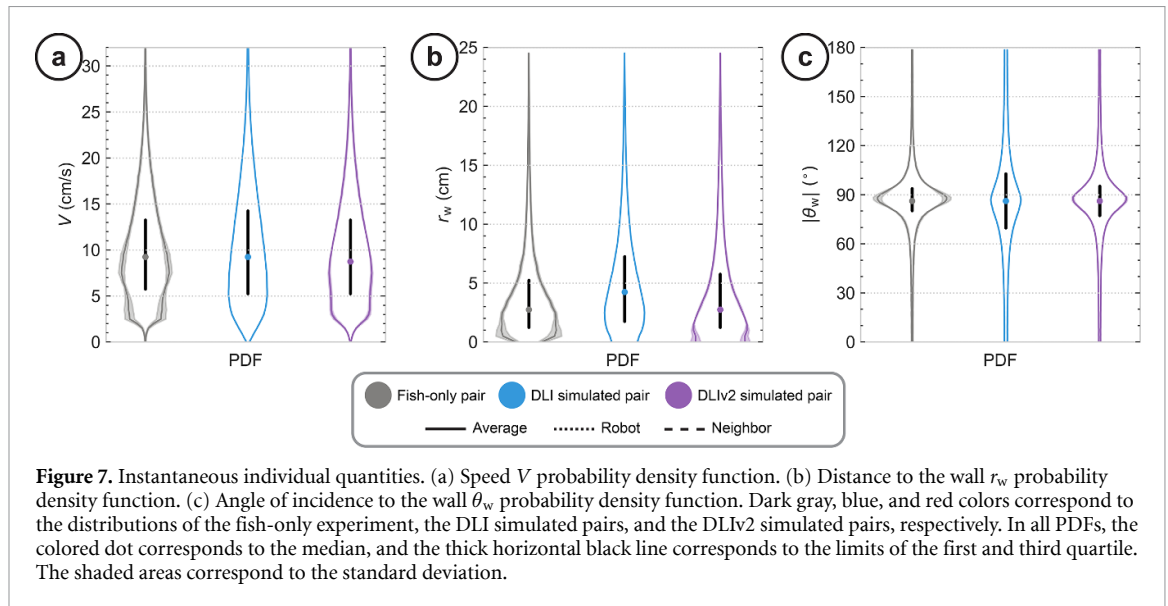


Figure 7. Instantaneous individual quantities. (a) Speed V probability density function. (b) Distance to the wall r_w probability density function. (c) Angle of incidence to the wall θ_w probability density function. Dark gray, blue, and red colors correspond to the distributions of the fish-only experiment, the DLI simulated pairs, and the DLIv2 simulated pairs, respectively. In all PDFs, the colored dot corresponds to the median, and the thick horizontal black line corresponds to the limits of the first and third quartile. The shaded areas correspond to the standard deviation.

to frequently produce trajectories farther from the wall than what observed in the experiment. Despite that, the DLI-BP remains fairly faithful to the DLI-SP, which indicates that the DLI is missing some aspects of the social dynamics before being implemented on the robot, but that the robot performs reasonably well in reproducing its underlying model.

2.4. Complementary results for DLIv2 simulated pairs

In addition to the DLI pretrained network utilized in the previous sections, we have also considered an updated version, the DLIv2. This version was retrained on data gathered from the present fish-only experiments under new lighting conditions,

concurrently to the robot experiments presented in this work, so that retraining was only feasible after their completion. However, it provided us with the opportunity to test the scalability and predictive performance of the pretrained DLI with new input samples, which, while not fundamentally different, originated from altered social dynamics. For this purpose, we conducted extensive simulations with the DLIv2, and found that their results are in excellent agreement with the present fish-only pair experiments (see tables 1 and 2 for further details) for the individual (see figure 7) and collective (see figure 8) observables, and for the temporal correlation functions (see figure 9). The performance of the simulated DLIv2 model present a significant

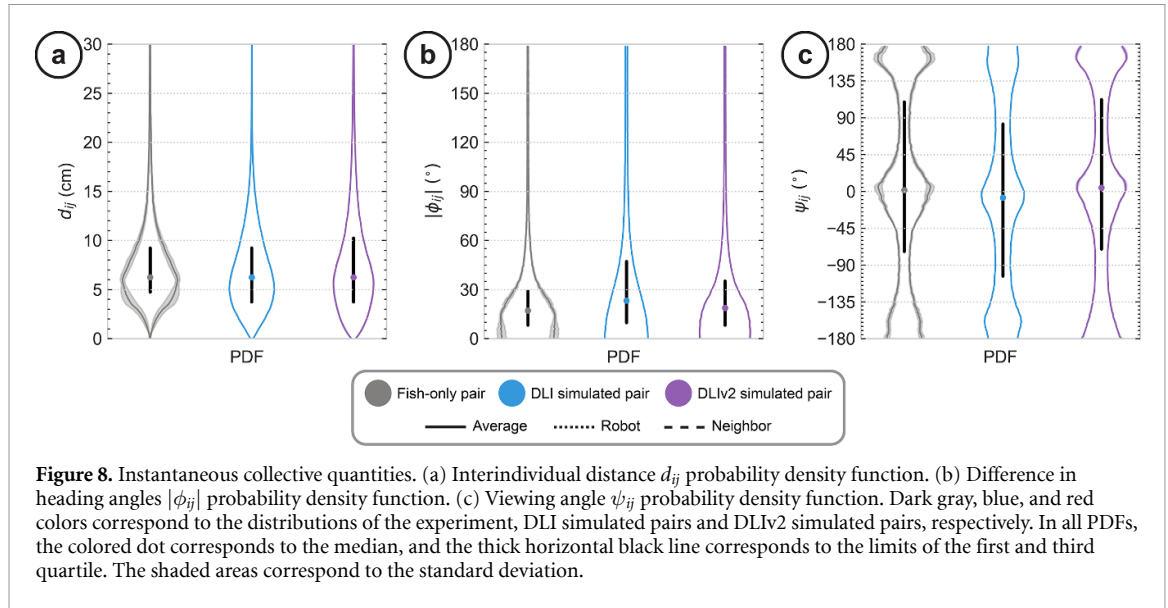


Figure 8. Instantaneous collective quantities. (a) Interindividual distance d_{ij} probability density function. (b) Difference in heading angles $|\phi_{ij}|$ probability density function. (c) Viewing angle ψ_{ij} probability density function. Dark gray, blue, and red colors correspond to the distributions of the experiment, DLI simulated pairs and DLiv2 simulated pairs, respectively. In all PDFs, the colored dot corresponds to the median, and the thick horizontal black line corresponds to the limits of the first and third quartile. The shaded areas correspond to the standard deviation.

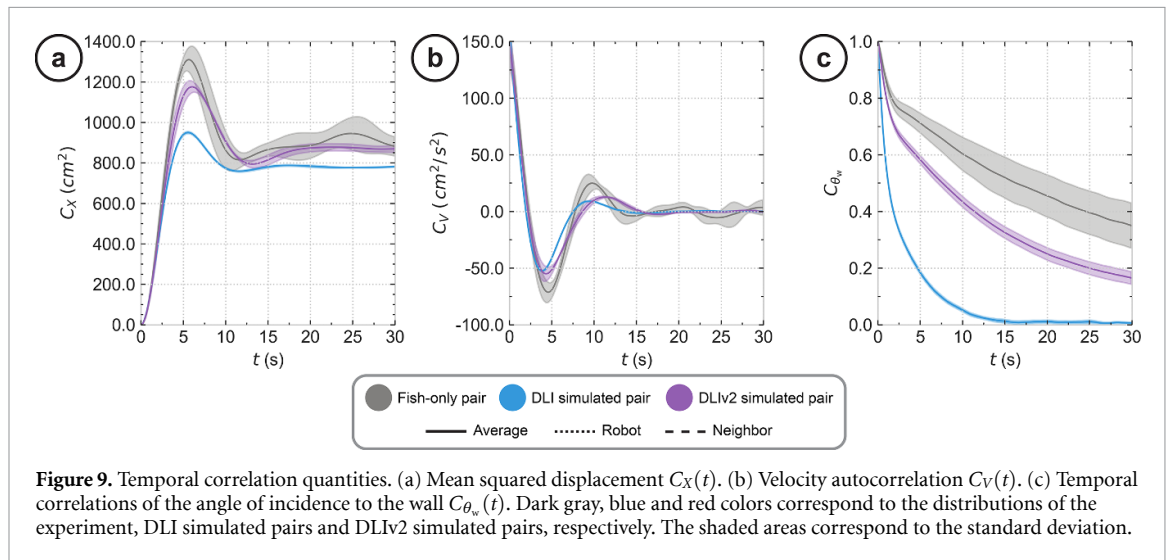


Figure 9. Temporal correlation quantities. (a) Mean squared displacement $C_X(t)$. (b) Velocity autocorrelation $C_V(t)$. (c) Temporal correlations of the angle of incidence to the wall $C_{\theta_w}(t)$. Dark gray, blue and red colors correspond to the distributions of the experiment, DLI simulated pairs and DLiv2 simulated pairs, respectively. The shaded areas correspond to the standard deviation.

improvement compared to that of the pretrained DLI model, and one could expect that the LureBot commanded by the DLiv2 model would lead to better results than for the LureBot commanded by the pretrained DLI model. Yet, our point here is that the pretrained DLI model, in different experimental conditions, can still interact with a fish in a similar way as a fish would do.

3. Discussion and Conclusion

Despite the wealth of studies on fish-robot interactions, to our knowledge, no prior research has drawn a systematic and quantitative comparison between the social interaction dynamics produced by fish-only, biohybrid, and simulated groups, with a robot commanded by a machine learning model. This comparison also raises an intriguing issue: while the reality gap in robotics [32] typically pertains to the transferability of robot controllers from simulation to real-life conditions, a parallel can be drawn for

biohybrid social interactions, termed the biomimicry gap. Addressing this gap is complicated by:

- (1) subtle behavioral patterns that behavioral mathematical models or machine learning models fail to capture;
- (2) the imperfect or absence of the rendition of the real physics in most models;
- (3) the inherently imperfect biomimetic properties of artificial lures and devices.

Constructing biohybrid systems with minimal or, ideally, no biomimicry gap, thus making them indistinguishable from pure animal groups, could open doors to groundbreaking research in the study of collective phenomena in animal groups. In particular, this would allow to accurately gauge the reactions of an animal or an animal group to a controlled perturbation (for instance, a robot changing its behavior by adopting a different mean speed or aiming at a

target location). Such endeavors require that any non-biomimetic effects of the robot be stringently assessed and resolved. Moreover, it is crucial to ensure that models do not simply overfit experimental data, but can be genuinely transferred to real-world scenarios, through robotic systems that *faithfully* execute the instructions of these models when interacting with animals.

Unfortunately, despite significant strides in the integration of behavioral modeling and robotics hardware, which has long been touted as crucial for deciphering and comprehending the mechanisms underlying collective behavior in animal groups, the closing of the biomimicry gap lacks convincing support in the literature. In particular, several biohybrid implementations tend to limit the set of decisions of the robot (spatial choices, like clockwise/anticlockwise motion) [9, 34]. In addition, many of these systems rely on simplified passive (open-loop) [1, 3, 11, 27, 28, 37, 38, 45, 46] or reactive (closed-loop) [13, 17, 21, 26, 29, 30, 40, 47] models, with only a handful utilizing biomimetic models. Even fewer biomimetic models have been successfully tested in biohybrid groups [13, 17] to emulate real-life dynamics of fish groups. Moreover, to our knowledge, no end-to-end machine learning (ML) model has been examined in this context, despite the booming field of ML. As developed in [36], assessing a model's fidelity is particularly challenging in the case of ML models, which are often black-box (i.e. not easily explainable). The methodology presented here to validate a model *and* the robotic platform in which it is implemented offers a preliminary solution to this conundrum. Finally, the importance of high-fidelity biomimetic lures and agile robotic devices capable of reproducing the typical motion patterns of the considered animal (speed, acceleration...) is usually underplayed [35].

In this work, through the precise and comparative quantification of collective behavior in pairs of agents (fish-only pairs, DLI-SP, and DLI-BP), we demonstrate that our biomimetic lure and robot system [35], combined with the DLI, are capable of bridging a substantial part of the biomimicry gap. More specifically, our study reveals that the overall gap between actual pairs of *H. rhodostomus* and DLI-SP is fairly small (mean Hellinger distance of $\bar{H} = 0.13$), while biohybrid pairs (DLI-BP) and fish-only pairs are more dissimilar, but in fair agreement ($\bar{H} = 0.20$). Despite this larger difference, the DLI-BP and DLI-SP remain in very good agreement ($\bar{H} = 0.11$). We also found that the inherent variability between fish experiments results in an average Hellinger distance of $\bar{H} = 0.10$ (see the end of table 2), which could be considered as the target performance for future studies following the benchmarking paradigm exploited in this study. In essence, our DLI model is successful in generating realistic social interactions [36], our

robotic system faithfully replicates its instructions, but the transferred model results in greater discrepancies and the gap widens compared to the simulation (see table 2). Nonetheless, the biohybrid pair is not fully aligned compared to fish groups: the Hellinger distance for the PDF of the angle of incidence to the wall is $H = 0.25$ and that of the heading difference is $H = 0.30$. These observables are the largest contributors to widening the social interaction discrepancies (i.e. the largest contributor, out of all observables, to increasing the mean Hellinger distance). These discrepancies are consequently observed for the correlation function of the angle of incidence to the wall. However, it is important to recognize that even when comparing two independent fish-only series of experiments, the mean Hellinger distance would not vanish ($\bar{H} = 0.10$, in our case), owing to the inherent variability in fish behavior across experiments. This implies that the main objective for robotic systems should be to notably narrow the gap between experimental results, rather than completely eradicating it.

Moreover, the present work, complementing [35, 36], also presents a systematic methodology for a comprehensive assessment of the extent of the biomimicry gap. This is accomplished by introducing nine observables (easily generalizable to larger groups or other species) that quantify the instantaneous individual and collective behavior, as well as the temporal correlations present in the system. In addition, this methodology is supplemented by the utilization of the Hellinger distance quantifier. We strongly encourage researchers in the field to explore a similar methodology to evaluate the biomimicry gap in their respective systems of study.

Despite the positive results highlighted in our study, we demonstrated that further closing the biomimicry gap necessitates efforts to minimize all three discrepancy sources depicted in figure 1. First, it would require that we refine our modeling approach, e.g. by repeating the biohybrid experiments with the DLIv2 model. Secondly, the physics-related discrepancies, primarily attributable to the transposition of the model into the robot, remains relatively small, but also requires measurable improvement in the robotic system's operation to fully bridge the gap. Finally, discrepancies in the communication cues pose a considerable challenge in terms of evaluation and could only be fully measured in the absence of the other two sources of discrepancies.

We believe that our study may mark the beginning of many endeavors that integrate animal experiments, biomimetic biohybrid experiments, and simulations of a model commanding the robot, all within a single end-to-end approach. As demonstrated, this approach, combined with a systematic methodology to quantify the biomimicry gap, offers a deeper understanding of the factors contributing

to behavioral inaccuracies in biohybrid experiments, thereby highlighting areas in need of improvement. This, in turn, contributes to two main objectives:

- (1) establishing a more robust experimentation pipeline to explore the diverse sources of the biomimicry gap (such as physical limitations of the robot, social interactions as depicted by the models, and potential discrepancies in the communication cues used to elicit responses);
- (2) drawing more definitive and insightful behavioral conclusions without the introduction of unrealistic effects inherent in robotic systems and social interaction models.

In future research, we aim to extend our experiments to involve multiple individuals and other species, thereby enhancing our understanding of how our robotic platform and DLI model scale to multi-agent interactions and whether large groups of living animals respond similarly to the DLI-driven artificial agent. Additionally, we intend to consistently report the biomimicry gap score, as presented here, with the hope that future studies may adopt a standardized methodology to evaluate the fidelity of biohybrid systems compared to natural ones.

Ethics statement

Experiments were approved by the local ethical committee for experimental animals and were performed at the Centre de Recherches sur la Cognition Animale, Centre de Biologie Intégrative, Toulouse, in an approved fish facility (A3155501) under permit APAFIS#27303-2020090219529069 v8 in agreement with the French legislation.

Funding

V P and F M were supported by the Swiss National Science Foundation Project ‘Self-Adaptive Mixed Societies of Animals and Robots’, Grant No. 175 731. G T and C S were supported by the French National Research Agency (ANR-20-CE45-0006-01).

Conflict of interest declaration


We declare no competing interests.

Data availability statement

The data for the fish-only pairs experiments, DLI pairs numerical simulations, and DLI biohybrid pairs experiments, i.e. the data that support the findings of this study, are openly available at the following URL/DOI: <https://doi.org/10.5281/zenodo.8253256>.

ORCID iDs

Vaios Papaspyros  <https://orcid.org/0000-0001-8257-1891>

Guy Theraulaz  <https://orcid.org/0000-0002-3062-4215>

Clément Sire  <https://orcid.org/0000-0003-4089-4013>

Francesco Mondada  <https://orcid.org/0000-0001-8641-8704>

References

- [1] Abaid N, Bartolini T, Macrì S and Porfiri M 2012 Zebrafish responds differentially to a robotic fish of varying aspect ratio, tail beat frequency, noise and color *Behav. Brain Res.* **233** 545–53
- [2] Abaid N and Porfiri M 2010 Fish in a ring: spatio-temporal pattern formation in one-dimensional animal groups *J. R. Soc. Interface* **7** 1441–53
- [3] Bartolini T, Mwaffo V, Showler A, Macrì S, Butail S and Porfiri M 2016 Zebrafish response to 3d printed shoals of conspecifics: the effect of body size *Bioinspir. Biomim.* **11** 026003
- [4] Basu A, Harris I R and Basu S 1997 2 minimum distance estimation: the approach using density-based distances *Handbook Stat.* **15** 21–48
- [5] Beran R 1977 Minimum hellinger distance estimates for parametric models *Ann. Stat.* **5** 445–63 (available at: www.jstor.org/stable/2958896)
- [6] Bonnet F, Binder S, Oliveria M E de, Halloy J and Mondada F 2014 A miniature mobile robot developed to be socially integrated with species of small fish 2014 *IEEE Int. Conf. on Robotics and Biomimetics (ROBIO 2014)* (IEEE) pp 747–52
- [7] Bonnet F, Gribovskiy A, Halloy J and Mondada F 2018 Closed-loop interactions between a shoal of zebrafish and a group of robotic fish in a circular corridor *Swarm Intell.* **12** 227–44
- [8] Bonnet F, Kato Y, Halloy J and Mondada F 2016 Infiltrating the zebrafish swarm: design, implementation and experimental tests of a miniature robotic fish lure for fish–robot interaction studies *Artif. Life Robot.* **21** 239–46
- [9] Bonnet F, Mills R, Szopek M, Schönwetter-Fuchs S, Halloy J, Bogdan S, Correia L, Mondada F and Schmickl T 2019 Robots mediating interactions between animals for interspecies collective behaviors *Sci. Robot.* **4** eaau7897
- [10] Butail S, Abaid N, Macrì S and Porfiri M 2015 Fish–robot interactions: robot fish in animal behavioral studies *Robot Fish: Bio-inspired Fishlike Underwater Robots* 359–77
- [11] Butail S, Polverino G, Phamduy P, Del Sette F and Porfiri M 2014 Influence of robotic shoal size, configuration and activity on zebrafish behavior in a free-swimming environment *Behav. Brain Res.* **275** 269–80
- [12] Calovi D S, Litchinko A, Lecheval V, Lopez U, Pérez Escudero A, 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
- [13] Cazenille L, Bredeche N and Halloy J 2018 Evolutionary optimisation of neural network models for fish collective behaviours in mixed groups of robots and zebrafish *Biomimetic and Biohybrid Systems: 7th Int. Conf., Living Machines 2018 (Paris, France, 17–20 July 2018) (Proc.)* vol 7 (Springer) pp 85–96
- [14] Cazenille L, Bredeche N and Halloy J 2019 Automatic calibration of artificial neural networks for zebrafish collective behaviours using a quality diversity algorithm 38–50

- [15] Cazenille L, Chemtob Y, Bonnet F, Gribovskiy A, Mondada F, Bredeche N and Halloy J 2017 Automated calibration of a biomimetic space-dependent model for zebrafish and robot collective behaviour in a structured environment 107–18
- [16] Cazenille L, Chemtob Y, Bonnet F, Gribovskiy A, Mondada F, Bredeche N and Halloy J 2018 How to blend a robot within a group of zebrafish: achieving social acceptance through real-time calibration of a multi-level behavioural model 73–84
- [17] Cazenille L, Collignon B, Chemtob Y, Bonnet F, Gribovskiy A, Mondada F, Bredeche N and Halloy J 2018 How mimetic should a robotic fish be to socially integrate into zebrafish groups? *Bioinspir. Biomim.* **13** 025001
- [18] Chua K, Calandra R, McAllister R and Levine S 2018 Deep reinforcement learning in a handful of trials using probabilistic dynamics models 4754–65
- [19] Collignon B, Séguret A and Halloy J 2016 A stochastic vision-based model inspired by zebrafish collective behaviour in heterogeneous environments *R. Soc. Open Sci.* **3** 150473
- [20] Costa T, Laan A, Heras F J H and de Polavieja G G 2020 Automated discovery of local rules for desired collective-level behavior through reinforcement learning *Front. Phys.* **8** 200
- [21] Faria J J, Dyer J R G, Clément R O, Couzin I D, Holt N, Ward A J W, Waters D and Krause J 2010 A novel method for investigating the collective behaviour of fish: introducing ‘robotfish’ *Behav. Ecol. Sociobiol.* **64** 1211–8
- [22] Harpaz R, Aspiras A C, Chambule S, Tseng S, Bind M-A’ele, Engert F, Fishman M C and Bahl A 2021 Collective behavior emerges from genetically controlled simple behavioral motifs in zebrafish *Sci. Adv.* **7** eabi7460
- [23] Heras F J H, Romero-Ferrero F, Hinz R C and de Polavieja G G 2019 Deep attention networks reveal the rules of collective motion in zebrafish *PLOS Comput. Biol.* **15** 1–23
- [24] Hochreiter S and Schmidhuber J 1997 Long short-term memory *Neural Comput.* **9** 1735–80
- [25] Jayles B, Escobedo R, Pasqua R, Zanon C, Blanchet A, Roy M, Trédan G, Theraulaz G and Sire C 2020 Collective information processing in human phase separation *Phil. Trans. R. Soc. B* **375** 20190801
- [26] Kim C, Ruberto T, Phamduy P and Porfiri M 2018 Closed-loop control of zebrafish behaviour in three dimensions using a robotic stimulus *Sci. Rep.* **8** 657
- [27] Kruusmaa M, Rieucan G, Carlos Castillo Montoya J, Markna R and Olav Handegard N 2016 Collective responses of a large mackerel school depend on the size and speed of a robotic fish but not on tail motion *Bioinspir. Biomim.* **11** 056020
- [28] Ladu F, Mwaffo V, Li J, Macrì S and Porfiri M 2015 Acute caffeine administration affects zebrafish response to a robotic stimulus *Behav. Brain Res.* **289** 48–54
- [29] Landgraf T, Bierbach D, Nguyen H, Muggelberg N, Romanczuk P and Krause J 2016 Robotfish: increased acceptance of interactive robotic fish with realistic eyes and natural motion patterns by live trinidadian guppies *Bioinspir. Biomim.* **11** 015001
- [30] Landgraf T, Nguyen H, Forgo S, Schneider J, Schröer J, Krüger C, Matzke H, Clément R O, Krause J and Rojas R 2013 Interactive robotic fish for the analysis of swarm behavior *Advances in Swarm Intelligence: 4th Int. Conf., ICSI 2013 (Harbin, China, 12–15 June 2013) (Proc.)* vol 4 (Springer) pp 1–10
- [31] Miller N and Gerlai R 2012 From schooling to shoaling: patterns of collective motion in zebrafish (*danio rerio*) *PLoS One* **7** e488865
- [32] Mouret J-B and Chatzilygeroudis K 2017 20 years of reality gap: a few thoughts about simulators in evolutionary robotics 1121–4
- [33] Orger M B and de Polavieja G G 2017 Zebrafish behavior: opportunities and challenges *Annu. Rev. Neurosci.* **40** 125–47
- [34] Papaspyros V, Bonnet F, Collignon B and Mondada F 2019 Bidirectional interactions facilitate the integration of a robot into a shoal of zebrafish *danio rerio* *PLoS One* **14** e0220559
- [35] Papaspyros V, Burnier D, Cherfan R, Theraulaz G, Sire C and Mondada F 2023 A biohybrid interaction framework for the integration of robots in animal societies *IEEE Access* **11** 67640–59
- [36] Papaspyros V, Escobedo R, Alahi A, Theraulaz G, Sire C and Mondada F 2024 Predicting the long-term collective behaviour of fish pairs with deep learning *J. R. Soc. Interface* **21** 20230630
- [37] Phamduy P, Polverino G, Fuller R C and Porfiri M 2014 Fish and robot dancing together: bluefin killifish females respond differently to the courtship of a robot with varying color morphs *Bioinspir. Biomim.* **9** 036021
- [38] Polverino G and Porfiri M 2013 Zebrafish (*danio rerio*) behavioural response to bioinspired robotic fish and mosquitofish (*gambusia affinis*) *Bioinspir. Biomim.* **8** 044001
- [39] Porfiri M 2018 Inferring causal relationships in zebrafish-robot interactions through transfer entropy: a small lure to catch a big fish *Animal Behav. Cogn.* **5** 341–67
- [40] Porfiri M, Spinello C, Yang Y and Macrì S 2019 Zebrafish adjust their behavior in response to an interactive robotic predator *Front. Robot. AI* **6** 38
- [41] Romano D, Donati E, Benelli G and Stefanini C 2019 A review on animal–robot interaction: from bio-hybrid organisms to mixed societies *Biol. Cybern.* **113** 201–25
- [42] Romano D and Stefanini C 2021 Unveiling social distancing mechanisms via a fish-robot hybrid interaction *Biol. Cybern.* **115** 565–73
- [43] Romano D and Stefanini C 2022 Any colour you like: fish interacting with bioinspired robots unravel mechanisms promoting mixed phenotype aggregations *Bioinspir. Biomim.* **17** 045004
- [44] Romano D and Stefanini C 2022 Robot-fish interaction helps to trigger social buffering in neon tetras: the potential role of social robotics in treating anxiety *Int. J. Soc. Robot.* **14** 963–72
- [45] Ruberto T, Mwaffo V, Singh S, Neri D and Porfiri M 2016 Zebrafish response to a robotic replica in three dimensions *R. Soc. Open Sci.* **3** 160505
- [46] Spinello C, Macrì S and Porfiri M 2013 Acute ethanol administration affects zebrafish preference for a biologically inspired robot, *Alcohol* **47** 391–8
- [47] Swain D T, Couzin I D and Ehrlich Leonard N 2011 Real-time feedback-controlled robotic fish for behavioral experiments with fish schools *Proc. IEEE* **100** 150–63
- [48] Zienkiewicz A K, Ladu F, Barton D A W, Porfiri M and Di Bernardo M D 2018 Data-driven modelling of social forces and collective behaviour in zebrafish *J. Theor. Biol.* **443** 39–51