Disentangling and modeling interactions in fish with burst-and-coast swimming reveal distinct alignment and attraction behaviors

Daniel S. Calovi[1], Alexandra Litchinko[1], Valentin Lecheval[1,2], Ugo Lopez[1], Alfonso Pérez Escudero[1,3], Hugues Chaté[4,5], Clément Sire[6], Guy Theraulaz[1] *

1 Centre de Recherches sur la Cognition Animale, Centre de Biologie Intégrative (CBI), Centre National de la Recherche Scientifique (CNRS) & Université de Toulouse – Paul Sabatier, 31062 Toulouse, France, **2** Groningen Institute for Evolutionary Life Sciences, University of Groningen, Centre for Life Sciences, Nijenborgh 7, 9747AG Groningen, The Netherlands, **3** Department of Physics, Massachusetts Institute of Technology, Cambridge, Massachusetts, United States of America, **4** Service de Physique de l'État Condensé, CEA – Saclay, 91191 Gif-sur-Yvette, France, **5** Computational Science Research Center, Beijing 100094, China, **6** Laboratoire de Physique Théorique, CNRS & Université de Toulouse – Paul Sabatier, 31062 Toulouse, France

* guy.theraulaz@univ-tlse3.fr

Abstract

The development of tracking methods for automatically quantifying individual behavior and social interactions in animal groups has open up new perspectives for building quantitative and predictive models of collective behavior. In this work, we combine extensive data analyses with a modeling approach to measure, disentangle, and reconstruct the actual functional form of interactions involved in the coordination of swimming in Rummy-nose tetra (*Hemigrammus rhodostomus*). This species of fish performs burst-and-coast swimming behavior that consists of sudden heading changes combined with brief accelerations followed by quasi-passive, straight decelerations. We quantify the spontaneous stochastic behavior of a fish and the interactions that govern wall avoidance and the reaction to a neighboring fish, the latter by exploiting general symmetry constraints for the interactions. In contrast with previous experimental works, we find that both attraction and alignment behaviors control the reaction of fish to a neighbor. We then exploit these results to build a model of spontaneous burst-and-coast swimming and interactions of fish, with all parameters being estimated or directly measured from experiments. This model quantitatively reproduces the key features of the motion and spatial distributions observed in experiments with a single fish and with two fish. This demonstrates the power of our method that exploits large amounts of data for disentangling and fully characterizing the interactions that govern collective behaviors in animals groups.

Author summary

We disentangle, quantify, and model the interactions involved in the control and coordination of burst-and-coast swimming in the fish *Hemigrammus rhodostomus*. We find that the interactions of fish with the arena wall result in avoidance behavior and those with a

neighbor result in a combination of attraction and alignment behaviors whose effects depend on the distance and the relative position and orientation to that neighbor. Then we show that a model entirely based on experimental data confirms that the combination of these individual level interactions quantitatively reproduces the dynamics of swimming of a single fish, the coordinated motion in groups of two fish, and the consequences of interactions on their spatial distribution.

Introduction

The study of physical or living self-propelled particles—active matter—has certainly become a booming field, notably involving biologists and physicists, often working together. Physical examples of active matter include self-propelled Janus colloids [1–8], vibrated granular matter [9–11], or self-propulsion mediated by hydrodynamical effects [12, 13], whereas biological examples are obviously ubiquitous: bacteria, cells, fish, humans, and simply speaking, most animals. In both physical and biological contexts, active matter can organize into rich collective phases. For instance, fish schools can be observed in a disordered swarming phase, or ordered schooling and vortex/milling phases [14, 15].

Yet, there are important differences between most physical and living self-propelled particles (SPP) [16]. For most physical SPP, interactions with other particles or obstacles do not modify the intrinsic or "desired" velocity of the particles but exert forces whose effect *adds up to their intrinsic velocity*. Most living SPP, like fish, birds, or humans, can also interact through physical forces (a fish or a pedestrian physically pushing another one or bumping into a wall) but mostly interact through "social forces" [17, 18]. For instance, a fish or a pedestrian wishing to avoid a physical obstacle or another individual will *modify its intrinsic velocity* in order to never actually touch it. Moreover, a physical force applied to an individual, in addition to its direct impact, can elicit a response in the form of a change in its intrinsic velocity. For instance, a fish can react to the physical force exerted by the fluid [19], resulting in a tendency for the fish to go along or against the flow. Moreover, social forces strongly break the Newtonian law of action and reaction [15, 16, 19, 20] since a fish or a pedestrian avoiding a physical obstacle obviously does not exert a contrary force on the obstacle. In addition, even between two individuals 1 and 2, the force exerted by 1 on 2 is most often not the opposite of the force exerted by 2 on 1, since social forces commonly depend on physical stimuli (light, sounds, temperature, pressure. . .) associated to an *anisotropic perception*: a pedestrian will most often react more to another pedestrian ahead than behind her/him. Similarly, social forces between two fish or two pedestrians also depend on their relative velocities or orientations: the need to avoid another individual will be in general greater when a collision is imminent than if it is unlikely, due to the velocity directions. Moreover, in some species of fish, individuals tend to explicitly align their heading to that of their neighbors.

Hence, if the understanding of the social interactions that underlie the collective behavior of animal groups is a central question in social ethology and behavioral ecology [21, 22], it has also a clear conceptual interest for physicists, since social and physical forces play very different roles in the dynamics of an active matter particle [16].

These social interactions play a key role in the ability of group members to coordinate their actions and collectively solve a wide range of problems, thus increasing their fitness [23–25]. In the past few years, the development of new methods based on machine learning algorithms for automating tracking and behavior analyses of animals in groups has improved to unprecedented levels the precision of available data on social interactions [26–28]. A wide variety of

biological systems have been investigated using such methods, from swarms of insects [29–31] to schools of fish [32–36], flocks of birds [37–40], groups of mice [41, 42], herds of ungulates [43, 44], groups of primates [45, 46], and human crowds [47, 48], bringing new insights on behavioral interactions and their consequences on individual and collective behavior.

The fine-scale analysis of individual-level interactions opens up new perspectives to develop quantitative and predictive models of collective behavior. One major challenge is to accurately identify the contributions and combination of each interaction involved at individual-level and then to validate with a model their role in the emergent properties at the collective level [49, 50]. Several studies on fish schools have explored ways to infer individual-level interactions directly from experimental data. The force-map technique [32] and the non-parametric inference technique [33] have been used to estimate from experiments involving groups of two fish the effective turning and speeding forces experienced by an individual. In the force-map approach, the implicit assumption considers that fish are particles on which the presence of neighboring fish and physical obstacles exert "forces". Visualizing these effective forces that capture the coarse-grained regularities of actual interactions has been a first step to characterize the local individual-level interactions [32, 33, 40]. However, none of these works incorporate or characterize the intrinsic stochasticity of individual behavior, and nor do they attempt to validate their findings by building trajectories from a model.

On the other hand, only a few models have been developed to connect a detailed quantitative description of individual-level interactions with the emergent dynamics observed at a group level [33–36]. The main difficulty to build such models comes from the entanglement of interactions between an individual and its physical and social environment. To overcome this problem, Gautrais et al. [34] have introduced an incremental approach that consists in first building from the experiments a model for the spontaneous motion of an isolated fish [51]. This model is then used as a dynamical framework to include the effects of interactions of that fish with the physical environment and with a neighboring fish. The validation of the model is then based on the agreement of its predictions with experiments on several observables in different conditions and group sizes. However, in these works, the interactions of a fish with the tank wall or another fish were only assumed to take reasonable functional forms (as a function of the fish distance and relative direction to the wall or another fish) and were depending on a few parameters which were determined by optimizing the agreement between the model predictions and the experimental results. In other words, the interactions functional forms were not truly extracted from the experimental data and a natural question which arises is whether the fair quantitative agreement of the model with experiments actually constitutes an implicit validation of the assumed forms of the interactions.

In the present work, we extend this approach to investigate the swimming behavior and interactions in the red nose fish *Hemigrammus rhodostomus*. This species performs a burst-and-coast type of swimming that makes it possible to analyze a trajectory as a series of discrete behavioral decisions in time and space. This discreteness of trajectories [36] is exploited to characterize the spontaneous motion of a fish, to identify the candidate stimuli (*e.g.* the distance, the orientation and velocity of a neighboring fish, or the distance and orientation of the tank wall), and to measure their effects on the behavioral response of a fish. We start from very general forms for the expected repulsive effect of the tank wall and for the repulsive/attractive and alignment interactions between two fish. These forms take into account the fish anisotropic perception of its physical and social environment and must satisfy some specific symmetry constraints which help us to differentiate these interactions and to disentangle and measure their relative contributions. The amount and precision of data accumulated and this modeling approach allow us to reconstruct the actual functional form of the response functions of fish governing their heading changes as a function of the distance, orientation, and

angular position relative to an obstacle or a neighbor. We show that the implementation of these interactions in a stochastic model of spontaneous burst-and-coast swimming quantitatively reproduces the motion and spatial distributions observed in experiments with a single fish and with two fish. Hence, our methodology and approach are the first which combine the precise determination of the spontaneous motion and the functional forms of the interactions *and* their implementation in a synthetic and faithful model.

Results

Characterization of individual swimming behavior

Hemigrammus rhodostomus fish have been monitored swimming alone and freely in shallow water in three different circular tanks of radius R = 176, 250, 353 mm (see the section Materials and methods for details, and Table 1). This species performs a burst-and-coast type of swimming characterized by sequences of sudden increase in speed followed by a mostly passive gliding period (see S1 Video). This allows the analysis of a trajectory as a series of discrete decisions in time. One can then identify the candidate stimuli (*e.g.* the distance, the orientation and velocity of a neighboring fish, or the distance and orientation of an obstacle) that have elicited a fish response and reconstruct the associated stimulus-response function. Most changes in fish heading occur exactly at the onset of the acceleration phase. We label each of these increases as a "kick".

Fig 1A and 1B show typical trajectories of *H. rhodostomus* swimming alone or in groups of two fish. After the data treatment (see the section Material and methods and Figs 7 and 8 there), it is possible to identify each kick (delimited by vertical lines in Fig 1C and 1D), which we use to describe fish trajectories as a group of straight lines between each of these events. While the average duration between kicks is close to 0.5 s for experiments with one or two fish (Fig 1G), the mean length covered between two successive kicks is slightly lower for two fish (Fig 1H). The typical velocity of the fish in their active periods (see Materials and methods) is of order 140 mm/s.

Quantifying the effect of the interaction of a single fish with the wall

Fig 2A shows the experimental probability density function (PDF) of the distance to the wall r_w after each kick, illustrating that the fish remains most of the time very close to the wall. We will see that the combination of the burst-and-coast nature of the trajectories (segments of average length $\sim$ 70 mm, but smaller when the fish is very close to the wall) and of the narrow distribution of angle changes between kicks (see Fig 2D) prevent a fish from efficiently escaping the curved wall of the tank (see S2 Video). Fig 2C shows the PDF of the relative angle of the fish to the wall θ_w, centered near, but clearly below 90°, as the fish remains almost parallel to the wall and most often goes toward it.

In order to characterize the behavior with respect to the walls, we define the signed angle variation $\delta\phi_+ = \delta\phi \times \mathrm{Sign}(\theta_w)$ after each kick, where $\delta\phi$ is the measured angle variation. $\delta\phi_+$ is hence positive when the fish goes away from the wall and negative when the fish is heading

Table 1. Experimental conditions.

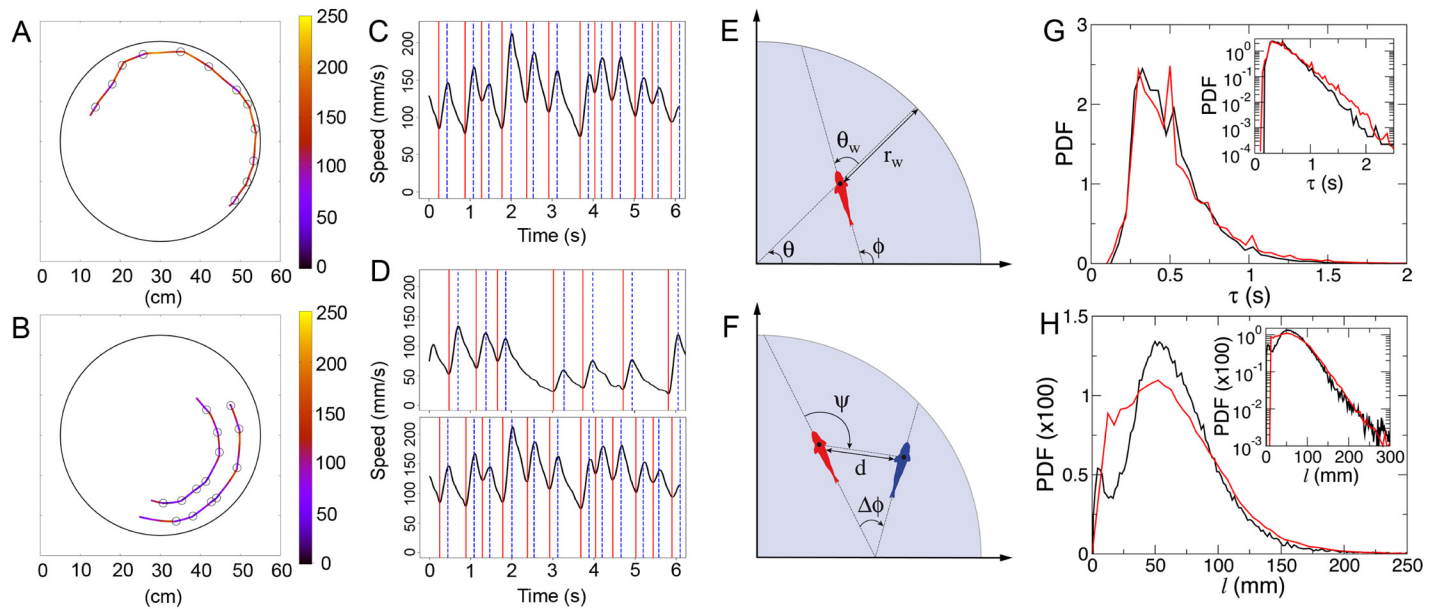| Arena radius (mm)	1 Fish			2 Fish
	176	**250**	**353**	**250**
Number of experiments	11	12	8	16
Temperature (˚C, mean ± se)	27.7 ± 0.2	26.8 ± 0.3	27.4 ± 0.3	26.1 ± 0.3
Body length (mm, mean ± se)	35.0 ± 1.1	28.3 ± 1.3	30.9 ± 1.5	30.3 ± 1.0

Fig 1. Burst-and-coast motion. Trajectories along with the bursts (circles) of a fish swimming alone (A) and a group of 2 fish (B). The color of trajectories indicates instantaneous speed. The corresponding speed time series are shown in C and D, along with the acceleration/burst phase delimited by red and blue vertical lines. E defines the variables r_w and θ_w (distance and relative orientation to the wall) in order to describe the fish interaction with the wall. F defines the relevant variables d, ψ, and $\Delta\phi$ (distance, viewing angle, relative orientation of the focal fish with respect to the other fish) in order to describe the influence of the blue fish on the red one. G and H show respectively the probability distribution function (PDF) of the duration and distance traveled between two kicks as measured in the one (black) and two (red) fish experiments (tank of radius $R = 250$ mm). Insets show the corresponding graphs in semi-log scale.

https://doi.org/10.1371/journal.pcbi.1005933.g001

towards it. The PDF of $\delta\phi_+$ is wider than a Gaussian and is clearly centered at a positive $\delta\phi_+ \approx 15°$ (tank of radius $R = 353$ mm), illustrating that the fish works at avoiding the wall (Fig 2D). When one restricts the data to instances where the fish is at a distance ($r_w > 60$ mm from the wall, for which its influence becomes negligible (see the section hereafter, addressing the interaction with the wall), the PDF of $\delta\phi_+$ indeed becomes symmetric, independent of the tank in which the fish swims, and takes a quasi Gaussian form of width of order $20°$ (inset of Fig 2D). The various quantities displayed in Fig 2 will ultimately be used to calibrate and test the predictions of our model.

Modeling and direct measurement of fish interaction with the wall

We first define a simple model for the spontaneous burst-and-coast motion of a single fish without any wall boundaries, and then introduce the fish-wall interaction, before considering the interaction between two fish in the next subsection. The large amount of data accumulated (more than 300000 recorded kicks – $\sim$ 100000 in each tank – for 1 fish, and 200000 for 2 fish in the $R = 250$ mm tank; see Materials and methods) permits us to not only precisely characterize the interactions, but also to provide a stringent test of the model by comparing its results to various experimental quantities (e.g. the full fish-wall and fish-fish distance and angle distributions instead of simply their mean, as only considered in many studies [34, 51]) which would be very sensitive to a change in the model structure and/or parameters.

Swimming dynamics without any interaction. We model the burst-and-coast motion by a series of instantaneous kicks each followed by a gliding period where fish travel in straight lines with a decaying velocity. At the n-th kick, the fish located at $\vec{x}_n$ at time t_n with angular direction ϕ_n randomly selects a new heading angle ϕ_{n+1}, a start or peak speed v_n, a kick duration τ_n, and a kick length l_n. During the gliding phase, the speed is empirically found to
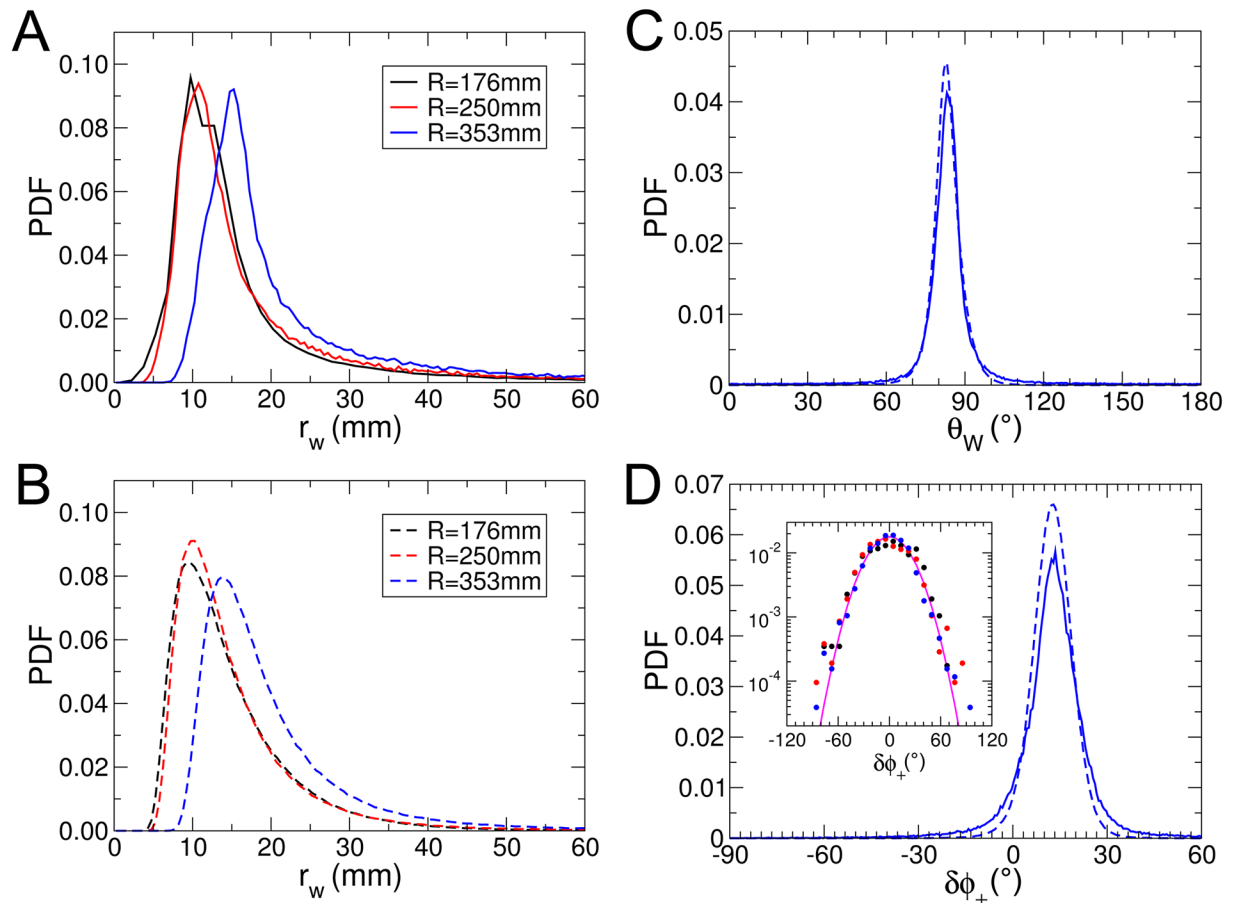
Fig 2. Quantification of the spatial distribution and motion of a fish swimming alone. Experimental (A; full lines) and theoretical (B; dashed lines) PDF of the distance to the wall r_w after a kick in the three arenas of radius $R = 176, 250, 353$ mm. C: experimental (full line) and theoretical (dashed line) PDF of the relative angle of the fish with the wall θ_w ($R = 353$ mm). D: PDF of the signed angle variation $\delta\phi_+ = \delta\phi \times \mathrm{Sign}(\theta_w)$ after each kick ($R = 353$ mm). The inset shows the distribution of $\delta\phi_+$ when the fish is near the center of the tank ($r_w > 60$ mm), for $R = 176, 250, 353$ mm (colored dots), which becomes centered at $\delta\phi_+ = 0°$ and Gaussian of width $\approx 20°$ (full line).

decrease quasi exponentially to a good approximation, as shown in Fig 3, with a decay or dissipation time $\tau_0 \approx 0.80$s, so that knowing v_n and τ_n or v_n and l_n, the third quantity is given by $l_n = v_n \tau_0 (1 - \exp[-\frac{\tau_n}{\tau_0}])$. At the end of the kick, the position and time are updated to

$$\vec{x}_{n+1} = \vec{x}_n + l_n \vec{e}(\phi_{n+1}), \quad t_{n+1} = t_n + \tau_n, \tag{1}$$

where $\vec{e}(\phi_{n+1})$ is the unit vector along the new angular direction ϕ_{n+1} of the fish. In practice, we generate v_n and l_n, and hence τ_n from simple model bell-shaped probability density functions (PDF) consistent with the experimental ones shown in Fig 1G and 1H. In addition, the distribution of $\delta\phi_R = \phi_{n+1} - \phi_n$ (the R subscript stands for "random") is experimentally found to be very close to a Gaussian distribution when the fish is located close to the center of the tank, i.e. when the interaction with the wall is negligible (see the inset of Fig 2D). $\delta\phi_R$ describes the spontaneous decisions of the fish to change its heading:

$$\phi_{n+1} = \phi_n + \delta\phi_R = \phi_n + \gamma_R g, \tag{2}$$

where g is a Gaussian random variable with zero average and unit variance, and γ_R is the
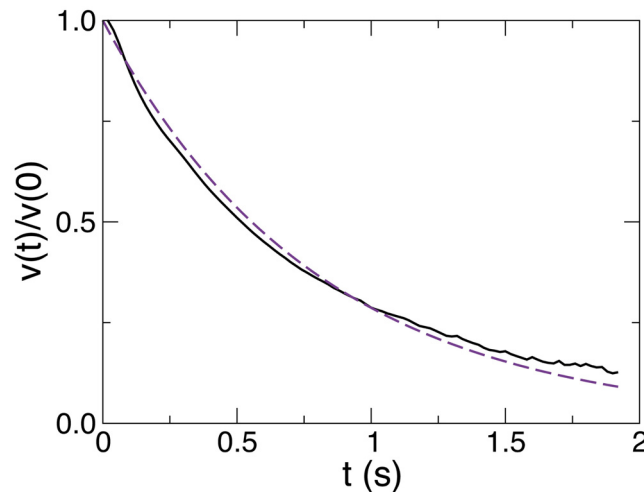
Fig 3. Average decay of the fish speed right after a kick. This decay can be reasonably described by an exponential decay with a relaxation time $\tau_0 \approx 0.80$s (violet dashed line).

intensity of the heading direction fluctuation, which is found to be of order 0.35 radian ($\approx 20°$) in the three tanks.

By exploiting the burst-and-coast dynamics of *H. rhodostomus*, we have defined an effective kick dynamics, of length and duration l_n and τ_n. However, it can be useful to generate the full continuous time dynamics from this discrete dynamics. For instance, such a procedure is necessary to produce "real-time" movies of fish trajectories obtained from the model (see S2, S4, and S5 Videos). As already mentioned, during a kick, the speed is empirically found to decrease exponentially to a good approximation (see Fig 3), with a decay or dissipation time $\tau_0 \approx 0.80$s. Between the time t_n and $t_{n+1} = t_n + \tau_n$, the viscous dynamics due to the water drag for $0 \leq t \leq \tau_n$ leads to

$$\vec{x}(t_n + t) = \vec{x}_n + l_n \frac{1 - \exp\left[-\frac{t}{\tau_0}\right]}{1 - \exp\left[-\frac{\tau_n}{\tau_0}\right]} \vec{e}(\phi_{n+1}), \tag{3}$$

so that one recovers $\vec{x}(t_n + \tau_n) = \vec{x}(t_{n+1}) = \vec{x}_n + l_n \vec{e}(\phi_{n+1}) = \vec{x}_{n+1}$.

Fish interaction with the wall. In order to include the interaction of the fish with the wall, we introduce an extra contribution $\delta\phi_W$

$$\delta\phi = \delta\phi_R(r_w) + \delta\phi_W(r_w, \theta_w), \tag{4}$$

where, due to symmetry constraints in a circular tank, $\delta\phi_W$ can only depend on the distance to the wall r_w, and on the angle θ_w between the fish angular direction ϕ and the normal to the wall (pointing from the tank center to the wall; see Fig 1E). We did not observe any statistically relevant left/right asymmetry, which imposes the symmetry condition

$$\delta\phi_W(r_w, -\theta_w) = -\delta\phi_W(r_w, \theta_w). \tag{5}$$

The random fluctuations of the fish direction are expected to be reduced when it stands near the wall, as the fish has less room for large angles variations (compare the main plot and the inset of Fig 2D), and we now define

$$\delta\phi_R(r_w) = \gamma_R[1 - \alpha f_w(r_w)]g. \tag{6}$$

$f_{\mathrm{w}}(r_{\mathrm{w}}) \to 0$, when $r_{\mathrm{w}} \gg l_{\mathrm{w}}$ (where l_{w} sets the range of the wall interaction), recovering the free spontaneous motion in this limit. In addition, we define $f_{\mathrm{w}}(0) = 1$ so that the fluctuations near the wall are reduced by a factor $1 - \alpha$, which is found experimentally to be close to $1/3$, so that $\alpha \approx 2/3$.

If the effective "repulsive force" exerted by the wall on the fish (first considered as a physical particle) tends to make it go toward the center of the tank, it must take the form $\delta\phi_{\mathrm{W}}(r_{\mathrm{w}}, \theta_{\mathrm{w}}) = \gamma_{\mathrm{W}} \sin(\theta_{\mathrm{w}}) f_{\mathrm{w}}(r_{\mathrm{w}})$, where the term $\sin(\theta_{\mathrm{w}})$ is simply the projection of the normal to the wall (*i.e.* the direction of the repulsion "force" due to the wall) on the angular acceleration of the fish (of direction $\phi + 90°$). For the sake of simplicity, $f_{\mathrm{w}}(r_{\mathrm{w}})$ is taken as the same function as the one introduced in Eq (6), as it satisfies the same limit behaviors. In fact, a fish does not have an isotropic perception of its environment. In order to take into account this important effect in a phenomenological way [15], we introduce $\epsilon_{\mathrm{w}}(\theta_{\mathrm{w}}) = \epsilon_{w,1} \cos(\theta_{\mathrm{w}}) + \epsilon_{w,2} \cos(2\theta_{\mathrm{w}}) + \ldots$, an even function (by symmetry) of θ_{w}, which, we assume, does not depend on r_{w}. Finally, we define

$$\delta\phi_{\mathrm{W}}(r_{\mathrm{w}}, \theta_{\mathrm{w}}) = \gamma_{\mathrm{W}} \sin(\theta_{\mathrm{w}})[1 + \epsilon_{\mathrm{w}}(\theta_{\mathrm{w}})] f_{\mathrm{w}}(r_{\mathrm{w}}), \qquad (7)$$

where γ_{W} is the intensity of the wall repulsion.

Once the displacement l and the total angle change $\delta\phi$ have been generated as explained above, we have to eliminate the instances where the new position of the fish would be outside the tank. More precisely, and since $\vec{x}$ refers to the position of the center of mass of the fish (and not of its head) before the kick, we introduce a "comfort length" l_c, which must be of the order of one body length (BL; 1 BL $\sim$ 30 mm; see Materials and methods), and we reject the move if the point $\vec{x} + (l + l_c)\vec{e}(\phi + \delta\phi)$ is outside the tank. When this happens, we regenerate l and $\delta\phi$ (and in particular, its random contribution $\delta\phi_{\mathrm{R}}$), until the new fish position is inside the tank. Note that in the rare cases where such a valid couple is not found after a large number of iterations (say, 1000), we generate a new value of $\delta\phi_{\mathrm{R}}$ uniformly drawn in $[-\pi, \pi]$ until a valid solution is obtained. Such a large angle is for instance necessary (and observed experimentally), when the fish happens to approach the wall almost perpendicularly to it ($\delta\phi \sim 90°$ or more; S5 Video at 20 s, where the red fish performs such a large angle change).

In order to measure experimentally $\epsilon_{\mathrm{w}}(\theta_{\mathrm{w}})$ and $f_{\mathrm{w}}(r_{\mathrm{w}})$, and confirm the functional form of Eq (7), we define a fitting procedure which is explicitly described in Materials and Methods, by minimizing the error between the experimental $\delta\phi$ and a general product functional form $\delta\phi_{\mathrm{W}}(r_{\mathrm{w}}, \theta_{\mathrm{w}}) = f_{\mathrm{w}}(r_{\mathrm{w}}) O_{\mathrm{w}}(\theta_{\mathrm{w}})$, where the only constraint is that $O_{\mathrm{w}}(\theta_{\mathrm{w}})$ is an *odd* function of θ_{w} (hence the name O), in order to satisfy the symmetry condition of Eq (5). Since multiplying O_{w} by an arbitrary constant and dividing f_{w} by the same constant leaves the product unchanged, we normalize O_{w} (and all angular functions appearing below) such that its average square is unity: $\frac{1}{2\pi}\int_{-\pi}^{+\pi} O_{\mathrm{w}}^2(\theta_{\mathrm{w}})\, d\theta_{\mathrm{w}} = 1$.

For each of the three tanks, the result of this procedure is presented as a scatter plot in Fig 4A and 4B respectively, along with the simple following functional forms (solid lines)

$$O_{\mathrm{w}}(\theta_{\mathrm{w}}) \propto \sin(\theta_{\mathrm{w}})[1 + 0.7\cos(2\theta_{\mathrm{w}})], \qquad (8)$$

$$f_{\mathrm{w}}(r_{\mathrm{w}}) = \exp\left[-(r_{\mathrm{w}}/l_{\mathrm{w}})^2\right], \text{ with } l_{\mathrm{w}} \approx 2\text{BL}. \qquad (9)$$

Hence, we find that the range of the wall interaction is of order $l_{\mathrm{w}} \approx 2$ BL $\sim$ 60 mm, and is strongly reduced when the fish is parallel to the wall (corresponding to a "comfort" situation), illustrated by the deep (*i.e.* lower response) observed for $\theta_{\mathrm{w}} \approx 90°$ in Fig 4B ($\cos(2\theta_{\mathrm{w}}) \approx -1$). Moreover, we do not find any significative dependence of these functional forms with the radius of the tank, although the interaction strength γ_W is found to decrease as the radius of
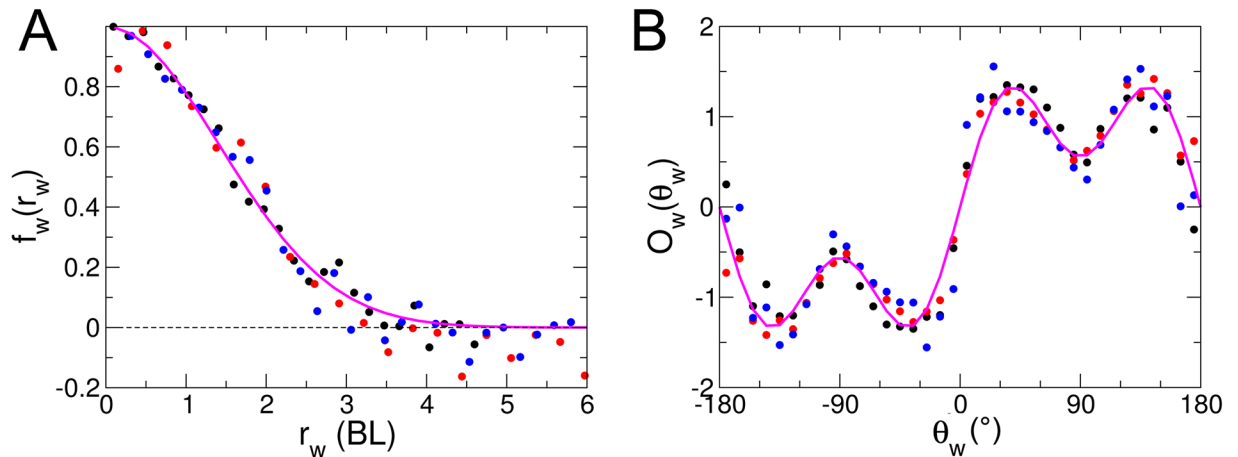
Fig 4. Interaction of a fish with the tank wall. Its intensity is shown as a function of the fish distance r_w (A) and relative orientation to the wall θ_w (B) as measured experimentally in the three tanks of radius $R = 176$ mm (black), $R = 250$ mm (blue), $R = 353$ mm (red). The full lines correspond to the analytic forms of $f_w(r_w)$ and $O_w(\theta_w)$ given in the text. In particular, $f_w(r_w)$ is well approximated by a Gaussian of width $l_w \approx 2$ BL ~ 60 mm.

https://doi.org/10.1371/journal.pcbi.1005933.g004

the wall increases (see Table 2). The smaller the tank radius (of curvature), the more effort is needed by the fish to avoid the wall.

Note that the fitting procedure used to produce the results of Fig 4 (described in detail in Materials and methods) does not involve any regularization scheme imposing the scatter plots to fall on actual continuous curves. The fact that they actually do describe such fairly smooth curves (as we will also find for the interaction functions between two fish) is an implicit validation of our procedure.

In Fig 2, and for the three tank radii considered, we compare the distribution of distance to the wall r_w, relative angle to the wall θ_w, and angle change $\delta\phi$ after each kick, as obtained experimentally and in extensive numerical simulations of the model, finding an overall satisfactory agreement. The numerical values of the parameters of the model are listed in Table 2. On a more qualitative note, the model fish dynamics mimics fairly well the behavior and motion of a real fish (see S2 Video). The fish making only small angle changes between each kick/burst cannot escape efficiently from the curved wall, especially in the 2 smallest circular tanks.

Quantifying the effect of interactions between two fish

Experiments with two fish were performed using the tank of radius $R = 250$ mm and a total of around 200000 kicks were recorded (see Materials and methods for details).

In Fig 5, we present various experimental PDF which characterize the swimming behavior of two fish resulting from their interaction, and which will permit to calibrate and test our

Table 2. Parameters used in the simulations.

# fish; R (mm)	l_c (mm)	γ_R (rad)	γ_R^0 (rad)	γ_W (rad)
1; 176	21	0.35	0.12	0.40
1; 250	30	0.35	0.14	0.12
1; 353	45	0.35	0.11	0.10
2; 250	30	0.45	0.15	0.15

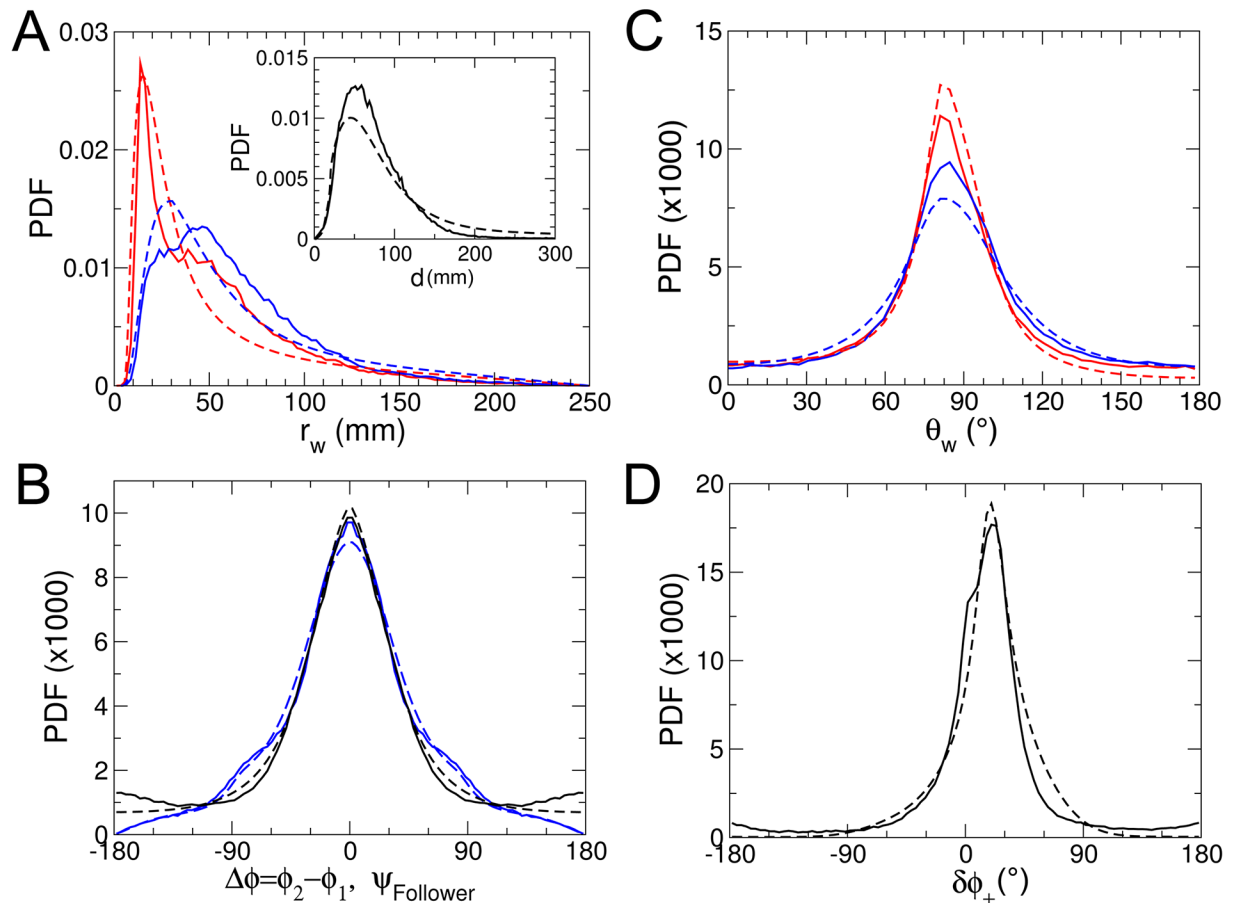https://doi.org/10.1371/journal.pcbi.1005933.t002

Fig 5. Quantification of the spatial distribution and motion in groups of two fish. In all graphs, full lines correspond to experimental results and dashed lines to numerical simulations of the model. A: PDF of the distance to the wall, for the geometrical leader (red) and follower (blue) fish; the inset displays the PDF of the distance d between the two fish. B: PDF of the relative orientation $\Delta\phi = \phi_2 - \phi_1$ between the two fish (black) and PDF of the viewing angle ψ of the follower (blue). C: PDF of the relative angle to the wall θ_w for the leader (red) and follower fish (blue). D: PDF (averaged over both fish) of the signed angle variation $\delta\phi_+ = \delta\phi \times \text{Sign}(\theta_w)$ after each kick.

model. Fig 5A shows the PDF of the distance to the wall, for the geometrical "leader" and "follower" fish. The geometrical leader is defined as the fish with the largest viewing angle $|\psi| \in [0, 180°]$ (see Fig 1F where the leader is the red fish), that is, the fish which needs to turn the most to directly face the other fish. Note that the geometrical leader is not always the same fish, as they can exchange role (see S5 Video). We find that the geometrical leader is much closer to the wall than the follower, as the follower tries to catch up and hence hugs the bend. Still, both fish are farther from the wall than an isolated fish is (see Fig 2A; also compare S2 and S4 Videos). The inset of Fig 5A shows the PDF of the distance d between the two fish, illustrating the strong attractive interaction between them.

Fig 5C shows the PDF of θ_w for the leader and follower fish, which are again much wider than for an isolated fish (see Fig 2C). The leader, being closer and hence more parallel to the wall, displays a sharper distribution than the follower. Fig 5B shows the PDF of the relative orientation $\Delta\phi = \phi_2 - \phi_1$ between the two fish, illustrating their tendency to align, along with the PDF of the viewing angle ψ of the follower. Both PDF are found to be very similar and peaked at 0°. Finally, Fig 5D shows the PDF (averaged over both fish) of the signed angle variation $\delta\phi_+ = \delta\phi \times \text{Sign}(\theta_w)$ after each kick, which is again much wider than for an isolated fish

([Fig 2D](#)). Due to their mutual influence, the fish swim farther from the wall than an isolated fish, and the wall constrains less their angular fluctuations.

Modeling and direct measurement of interactions between two fish

In the presence of another fish, the total heading angle change now reads

$$
\begin{aligned}
\delta\phi \quad = \quad & \delta\phi_{\mathrm{R}}(r_{\mathrm{w}}) + \delta\phi_{\mathrm{W}}(r_{\mathrm{w}}, \theta_{\mathrm{w}}) + \\
& \delta\phi_{\mathrm{Att}}(d, \psi, \Delta\phi) + \delta\phi_{\mathrm{Ali}}(d, \psi, \Delta\phi),
\end{aligned}
\tag{10}
$$

where the random and wall contributions are given by Eqs ([6](#), [7](#), [8](#), [9](#)), and the two new contributions result from the expected attraction ("Att" subscript) and alignment ("Ali" subscript) interactions between fish. The distance between fish d, the relative position or viewing angle ψ, and the relative orientation angle $\Delta\phi$ are all defined in [Fig 1F](#). By mirror symmetry already discussed in the context of the interaction with the wall, one has the exact constraint

$$
\delta\phi_{\mathrm{Att,\ Ali}}(d, -\psi, -\Delta\phi) = -\delta\phi_{\mathrm{Att,\ Ali}}(d, \psi, \Delta\phi),
\tag{11}
$$

meaning that a trajectory of the two fish observed from above the tank has the same probability of occurrence as the same trajectory as it appears when viewing it from the bottom of the tank. We hence propose the following product expressions

$$
\delta\phi_{\mathrm{Att}}(d, \psi, \Delta\phi) = F_{\mathrm{Att}}(d) O_{\mathrm{Att}}(\psi) E_{\mathrm{Att}}(\Delta\phi),
\tag{12}
$$

$$
\delta\phi_{\mathrm{Ali}}(d, \psi, \Delta\phi) = F_{\mathrm{Ali}}(d) O_{\mathrm{Ali}}(\Delta\phi) E_{\mathrm{Ali}}(\psi),
\tag{13}
$$

where the functions O are odd, and the functions E are even. For instance, O_{Att} must be odd as the focal fish should turn by the same angle (but of opposite sign) whether the other fish is at the same angle $|\psi|$ to its left or right. Like in the case of the wall interaction, we normalize the four angular functions appearing in Eqs ([12](#), [13](#)) such that their average square is unity. Both attraction and alignment interactions clearly break the law of action and reaction, as briefly mentioned in the introduction. Although the heading angle difference perceived by the other fish is simply $\Delta\phi' = -\Delta\phi$, its viewing angle ψ' is in general not equal to $-\psi$ (see [Fig 1F](#)).

As already discussed in the context of the wall interaction, an isotropic radial attraction force between the two fish independent of the relative orientation, would lead exactly to [Eq (12)](#), with $O_{\mathrm{Att}}(\psi) \sim \sin(\psi)$ and $E_{\mathrm{Att}}(\Delta\phi) = 1$. Moreover, an alignment force tending to maximize the scalar product, *i.e.* the alignment, between the two fish headings takes the natural form $O_{\mathrm{Ali}}(\Delta\phi) \sim \sin(\Delta\phi)$, similar to the one between two magnetic spins, for which one has $E_{\mathrm{Ali}}(\psi) = 1$. However, we allow here for more general forms satisfying the required parity properties, due to the fish anisotropic perception of its environment, and to the fact that its behavior may also be affected by its relative orientation with the other fish. For instance, we anticipate that $E_{\mathrm{Ali}}(\psi)$ should be smaller when the other fish is behind the focal fish ($\psi = 180°$; bad perception of the other fish direction) than when it is ahead ($\psi = 0°$).

As for the dependence of F_{Att} with the distance between fish d, we expect F_{Att} to be negative (repulsive interaction) at short distance $d \leq d_0 \sim 1\mathrm{BL}$, and then to grow up to a typical distance l_{Att}, before ultimately decaying above l_{Att}. Note that if the attraction force is mostly mediated by vision at large distance, it should be proportional to the 2D solid angle produced by the other fish, which decays like $1/d$, for large d. These considerations motivate us to introduce an
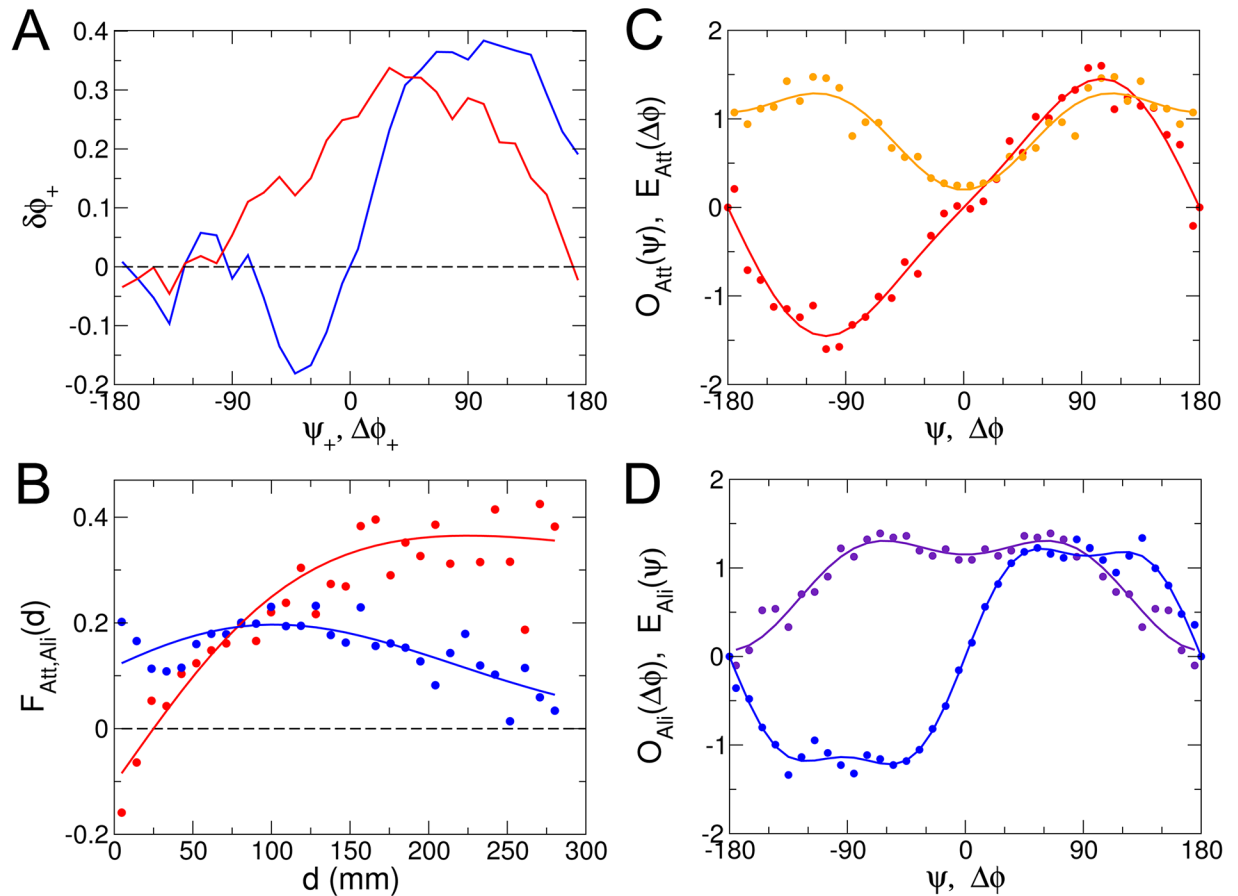
Fig 6. Quantification and modeling of interactions between pairs of fish. A: we plot the average signed angle change after a kick $\delta\phi_+ = \delta\phi \times \text{Sign}(\psi)$ vs $\Delta\phi \times \text{Sign}(\psi)$ (red) and $\delta\phi_+ = \delta\phi \times \text{Sign}(\Delta\phi)$ vs $\psi \times \text{Sign}(\Delta\phi)$ (blue) (see text). B: dependence of the attraction ($F_{\text{Att}}(d)$ in red) and alignment ($F_{\text{Ali}}(d)$ in blue) interactions with the distance d between fish. The full lines correspond to the physically motivated form of Eq (14) (red), and the fit proposed in the text for $F_{\text{Ali}}(d)$ (blue). C: $O_{\text{Att}}(\psi)$ (odd function in red) and $E_{\text{Att}}(\Delta\phi)$ (even function in orange) characterize the angular dependence of the attraction interaction, and are defined in Eq (12). D: $O_{\text{Ali}}(\Delta\phi)$ (odd function in blue) and $E_{\text{Ali}}(\psi)$ (even function in violet), defined in Eq (13), characterize the angular dependence of the alignment interaction. Dots in B, C, and D correspond to the results of applying the procedure explained in Materials and Methods to extract the interaction functions from experimental data.

https://doi.org/10.1371/journal.pcbi.1005933.g006

explicit functional form satisfying all these requirements:

$$F_{\text{Att}}(d) \propto \frac{d - d_0}{1 + (d/l_{\text{Att}})^2}. \tag{14}$$

F_{Ali} should be dominant at short distance, before decaying for d greater than some l_{Ali} defining the range of the alignment interaction. For large distance d, the alignment interaction should be smaller than the attraction force, as it becomes more difficult for the focal fish to estimate the precise relative orientation of the other fish than to simply identify its presence.

Fig 6A shows strong evidence for the existence of an alignment interaction. Indeed, we plot the average signed angle change after a kick $\delta\phi_+ = \delta\phi \times \text{Sign}(\psi)$ vs $\Delta\phi \times \text{Sign}(\psi)$ and $\delta\phi_+ = \delta\phi \times \text{Sign}(\Delta\phi)$ vs $\psi \times \text{Sign}(\delta\phi)$. In accordance with Eqs (12, 13), a strong positive $\delta\phi_+$ when the corresponding variable is positive indicates that the fish changes more its heading if it favors mutual alignment (reducing $\Delta\phi$), for the same viewing angle ψ. In other words, Fig 6A shows that the turning angle change is *amplified* when this angle change would tend to align both fish, for a given viewing angle of the focal fish and relative position of both fish.

As precisely explained in Materials and Methods, we have determined the six functions appearing in Eqs (12, 13) by minimizing the error with the measured $\delta\phi$, only considering kicks for which the focal fish was at a distance $r_w > 2$ BL from the wall, in order to eliminate its effect (see Fig 4A). This procedure leads to smooth and well behaved measured functions displayed in Fig 6. As shown in Fig 6B, the functional form of Eq (14) adequately describes $F_{Att}(d)$, with $l_{Att} \approx 200$ mm, and with an apparent repulsive regime at very short range, with $d_0 \approx 30$ mm ~ 1 BL. The crossover between a dominant alignment interaction and a dominant attraction interaction is also clear. The blue full line in Fig 6B, a guide to the eye reproducing appropriately $F_{Ali}(d)$, corresponds to the phenomenological functional form

$$F_{Ali}(d) \propto (d + d_0')\exp[-(d/l_{Ali})^2], \tag{15}$$

with $l_{Ali} \approx 200$ mm. Note that $F_{Att}(d)$ and $F_{Ali}(d)$ cannot be properly measured for $d > 280$ mm due to the lack of statistics, the two fish remaining most of the time close to each other (see the inset of Fig 6A; the typical distance between fish is $d \sim 75$ mm).

Fig 6C shows $O_{Att}(\psi) \propto \sin(\psi)[1 + \epsilon_{Att,1} \cos(\psi) + \ldots]$ (odd function) and $E_{Att}(\Delta\phi) \propto 1 + \eta_{Att,1} \cos(\Delta\phi) + \ldots$ (even function) along with fits involving no more than 2 non zero Fourier coefficients (and often only one; see Materials and methods for their actual values). $E_{Att}(\Delta\phi)$ has a minimum for $\Delta\phi = 0$ indicating that the attraction interaction is reduced when both fish are aligned. Similarly, Fig 6D shows $O_{Ali}(\Delta\phi)$ and $E_{Ali}(\psi)$ and the corresponding fits. As anticipated, the alignment interaction is stronger when the influencing fish is ahead of the focal fish ($|\psi| < 90°$), and almost vanishes when it is behind ($\psi = \pm180°$).

In Fig 5, we compare the results of extensive numerical simulations of the model including the interactions between fish to experimental data, finding an overall qualitative (S4 and S5 Videos) and quantitative agreement.

As a conclusion of this section, we would like to discuss the generality of the product functional forms of Eqs (12, 13) for the interaction between fish, or of Eq (7) in the context of the wall interaction. As already briefly mentioned, for a physical point particle interacting through a physical force like gravity, the angle change $\delta\phi_{Att}(d, \psi)$ would be the projection of the radial force onto the angular acceleration (normal to the velocity of angular direction ψ relative to the vector between the two particles) and would then exactly take the form $F_{Att}(d) \times \sin(\psi)$. Hence, Eq (12) (resp. Eq (7), for the wall interaction) is the simplest generalization accounting for the fish anisotropic perception of its environment, while keeping a product form and still obeying the left/right symmetry condition of Eq (11) (resp. of Eq (5)). In principle, $\delta\phi_{Att}(d, \psi, \Delta\phi)$ should be written most generally as an expansion $\Sigma_i F_{Att,i}(d)O_{Att,i}(\psi)E_{Att,i}(\Delta\phi)$. However, as the number of terms of this expansion increases, we run the risk of overfitting the experimental data by the procedure detailed in Materials and Methods. In addition, the leading term of this expansion would still capture the main behavioral effects of the interaction and should be very similar to the results of Fig 6, while the weaker remaining terms would anyway be difficult to interpret. Note that the same argument applies to the alignment interaction, when exploiting the analogy with the magnetic alignment force between two spins. Eq (13) is the simplest generalization of the interaction $\delta\phi_{Ali}(d, \Delta\phi) = F_{Ali}(d)\sin(\Delta\phi)$ obtained in this case, while preserving the left/right symmetry and product form. Considering the fact that no regularization or smoothing procedure was used in our data analysis (see Materials and methods), the quality (low noise, especially for angular functions) of the results presented in Figs 4 and 6 strongly suggests that the generalized product forms used here capture most of the features of the actual experimental angle change.

Discussion

Characterizing the social interactions between individuals as well as their behavioral reactions to the physical environment is a crucial step in our understanding of complex collective dynamics observed in many group-living species and their impact on individual fitness [21, 24]. In the present work, we have analyzed the behavioral responses of a fish to the presence in its neighborhood of an obstacle and to a conspecific fish. In particular, we used the discrete decisions (kicks) of *H. rhodostomus* to control its heading during burst-and-coast swimming as a proxy to measure and model individual-level interactions. The large amount of data accumulated allowed us to disentangle and quantify the effects of these interactions on fish behavior with a high level of accuracy.

We have quantified the spontaneous swimming behavior of a fish and modeled it by a kick dynamics with Gaussian distributed angle changes. We found that the interactions of fish with an obstacle and a neighboring fish result from the combination of four behavioral modes:

1. wall avoidance, whose effect starts to be effective when the fish is less than 2 BL from a wall;

2. short-range repulsion between fish, when inter-individual distance is less than 30 mm (~ 1 BL);

3. attraction to the neighboring fish, which reaches a maximum value around 200 mm (~ 6 to 7 BL) in our experimental conditions;

4. alignment to the neighbor, which saturates around 100 mm (~ 3 BL).

In contrast to previous phenomenological models, these behavioral modes are not fixed to discrete and somewhat arbitrary zones of distances in which the neighboring fish are found [52–54]. Instead, there is a continuous combination of attraction and alignment as a function of the distance between fish. Alignment dominates attraction up to ~ 75 mm (~ 2.5 BL) while attraction becomes dominant for larger distances. As distance increases even more, attraction must decrease as well. However, the limited size of the experimental tanks and the lack of sufficient data for large d prevented us from measuring this effect, suggesting the long-range nature of the attraction interaction mediated by vision. Note that a cluster of fish can elicit a higher level of attraction, proportional to the 3D solid angle of the fish group as seen by the focal fish, as suggested by models based on visual perception [55, 56], and as captured by the power-law decay proposed in Eq (14). Designing experiments to test and quantify the long-range nature of the attraction interaction between fish would be of clear interest.

Moreover, the behavioral responses are strongly modulated by the anisotropic perception of fish. The wall repulsion effect is maximum when the orientation of the fish with regards to the wall is close to 45° and minimum when the fish is parallel to the wall. Likewise, the maximum amplitude alignment occurs when a neighboring fish is located on the front left or right and vanishes as its position around the focal fish moves towards the back.

To quantify separately the effects of attraction and alignment, we exploited physical analogies and symmetry considerations to extract the interactions between a focal fish and the wall and with another fish. Previous studies have shown that in the Golden shiners [32] and the Mosquito fish [33], there was no clear evidence for an explicit matching of body orientation. In these species, the alignment between fish was supposed to results from a combination of attraction and repulsion. However, at least in the Mosquito fish, it is likely that the strength of alignment could have been underestimated because the symmetry constraints on alignment and attraction were not taken into consideration. In the Rummy-nose tetra, we find strong evidence for the existence of an explicit alignment.

The characterization and the measurement of burst-and-coast swimming and individual interactions were then used to build and calibrate a model that quantitatively reproduces the dynamics of swimming of fish alone and in groups of two and the consequences of interactions on their spatial and angular distributions. The model shows that the wall avoidance behavior coupled with the burst-and-coast motion results in an unexpected concentration of fish trajectories close to the wall, as observed in our experiments. In fact, this phenomenon is well referenced experimentally for run-and-tumble swimming (for instance, in sperm cells [57] or bacteria [58]). It can be explained theoretically and reproduced in simple models [59, 60], as the effective discreteness of the trajectories separated in bursts or tumbles prevents the individuals from escaping the wall. Our model also reproduces the alternation of temporary leaders and followers in groups of two fish, the behavior of the temporary leader being mostly governed by its interactions with the wall, while the temporary follower is mostly influenced by the behavior of the temporary leader.

This validated model can serve as a basis for testing hypotheses on the combination of influence exerted by multiples neighbors on a focal fish in tanks of arbitrary shape. Moreover, it would certainly be interesting to study theoretically the dynamics of many fish swimming without any boundary and according to the found interactions. The study of the phase diagram as a function of the strength of the attraction and alignment interactions (and possibly their range) should show the emergence of various collective phases (schooling phase, vortex phase...) [14, 15].

Finally, our method has proved successful in disentangling and fully characterizing the interactions that govern the behavior of pairs of animals when large amounts of data are available. Hence, it could be successfully applied to collective motion phenomena occurring in various biological systems at different scales of organization.

Materials and methods

Experimental procedures and data collection

Ethics statement. Our experiments have been approved by the Ethics Committee for Animal Experimentation of the Toulouse Research Federation in Biology N˚1 and comply with the European legislation for animal welfare. During the experiments, no mortality occurred.

Study species. *Hemigrammus rhodostomus* (rummy-nose tetras, Fig 7) were purchased from Amazonie Labège (http://www.amazonie.com) in Toulouse, France. This species was chosen because it exhibits a strong schooling behavior and it is very easy to handle in controlled conditions. Fish were kept in 150 L aquariums on a 12:12 hour, dark:light photoperiod, at 26.8˚C (±1.6˚C) and were fed *ad libitum* with fish flakes. Body lengths (BL) of the fish used in these experiments were on average 31 mm (see Table 1).

The experimental tank (120×120 cm) was made of glass and was set on top of a box to isolate fish from vibrations. The setup, placed in a chamber made by four opaque white curtains, was surrounded by four LED light panels giving an isotropic lighting. Circular tanks (of radius R = 176, 250, and 353 mm) were set inside the experimental tank filled with 7 cm of water of controlled quality (50% of water purified by reverse osmosis and 50% of water treated by activated carbon) heated at 26.69˚C (±1.19˚C) (details in Table 1). Reflections of light due to the bottom of the experimental tank are avoided thanks to a white PVC layer. Each trial started by setting one or two fish randomly sampled from their breeding tank into a circular tank. Fish were let for 10 minutes to habituate before the start of the trial. A trial consisted in one or three hours of fish freely swimming (*i.e.* without any external perturbation) in a circular tank (Table 1 and S1 Table). Fish trajectories were recorded by a Sony HandyCam HD camera filming from above the set-up at 50 Hz (50 frames per second) in HDTV resolution (1920×1080p).

Fig 7. A group of two fish from the study species *Hemigrammus rhodostomus*. Credits to David Villa ScienceImage/CBI/ CNRS, Toulouse, 2015.

https://doi.org/10.1371/journal.pcbi.1005933.g007

Two main sources of uncertainty in the measures from video recorded from above with only one camera occur:

1. by not knowing the water depth at which a fish swims, between 0 to 7 cm from the bottom of the tank;

2. because of parallax issues (the bigger the angle between a swimming fish and the camera axis, the bigger the error made estimating the position of the fish).

The contribution of each source to the uncertainty of our measures has been estimated by computing the lengths of the cells of a chessboard set at the bottom of the tank ($Z = 0$ cm) and at the top of the water level (here $Z = 6$ cm), coming from photographs shot at two zoom levels, the one used to record experiments in the tank of radius $R = 250$ mm and the one used to record experiments in the tank of radius $R = 353$ mm. As a result, the uncertainty due to the unknown position of the fish in the water column is higher than the uncertainty due to parallax (3.5% vs 0.5%).

Data extraction and pre-processing

Positions of fish on each frame have been tracked with the tracking software idTracker 2.1 [5]. The idTracker output format gives fish identity and barycenter positions of individuals in the image (in pixels), where the latter needs to be converted into position in the experimental frame of reference (in millimeters). The intermediary Matlab files issued by the tracker store the background image, which is the information used to calculate wall positions and thus colliding distances more accurately). Also, the intermediary files give the area of the detected fish, which can be used to determine fish heading from shape detection, independently of the trajectories. The processing of the output and intermediary files is processed with a custom-built Matlab script, which is structured into several procedures:

1. Detection of tank walls

2. Conversion to metric frame of reference

3. Fish shape detection and body length/width measurements

4. Fish activity selection and sampling

5. Segmentation

6. Segmented variables estimation

This section aims to document each of these procedures.

Detection of tank walls. From the tracking software idTracker, several files associated with one video are produced. In particular, a matrix is generated, with as many entries as the pixel resolution of the videos (1920×1080 pixels in our case), containing light intensity of each pixel, that is coded in a grayscale image of the video (Fig 8A). The intersection of the bottom of the tank with the tank floor is shadowed (Fig 8A). This shadow is manually enhanced to improve the detection of the tank walls.

Before running the tracking on idTracker, the user has to define a mask in order to exclude areas where individuals (here fish) cannot be tracked (e.g. outside the tank) (Fig 8B). This mask gives a raw circular estimation of the contour of the tank with radius R (*outer circle*) and a second raw circular estimation with arbitrary radius 0.85×R (*inner circle*) is derived. For N_θ radii with angle $\theta_i \in [0, 2\pi)$, $i \in [1, 2, \ldots, N_\theta]$, light intensity is measured from grayscale image every pixel from inner circle to outer circle. To measure the light intensity of a pixel, the average over the focal pixel and its 8 nearest neighbors in the image matrix is considered. We take $N_\theta = 5 \times 360 = 1800$ to oversample these measures. Mean position (x_i, y_i) of the three smallest values of light intensity (*i.e.* darker values) associated with each θ_i is computed, yielding a noisy and discrete estimate of lower image positions of the tank walls. A first smoothing procedure is run on the positions (x_i, y_i) to exclude bad walls estimates (e.g. detection of rust in the upper edge of the wall). A raw center (x_0, y_0) is defined as the mean position of all the (x_i, y_i) and estimate N_θ radii r_i (in pixels) for each θ_i. The following criterion is used: two consecutive radii cannot differ by more than 5 pixels. If it is the case for the ith radius, it is replaced by the previous one and associated (x_{i-1}, y_{i-1}) are recomputed given θ_i and the new radius r_i (in pixels). This procedure gives a new series of radii r_i where the previous procedure excludes outliers (Fig 8D). Assuming that fish swim at constant height h from the bottom, we analytically calculate from the detected positions of tank walls at the bottom the positions of the tank walls at height h using the formula

$$\vec{r}_h = \left(1 + \frac{h}{h + D}\right)(\vec{r} - \vec{r}_{CCD}) + \vec{r}_{CCD}, \qquad (16)$$
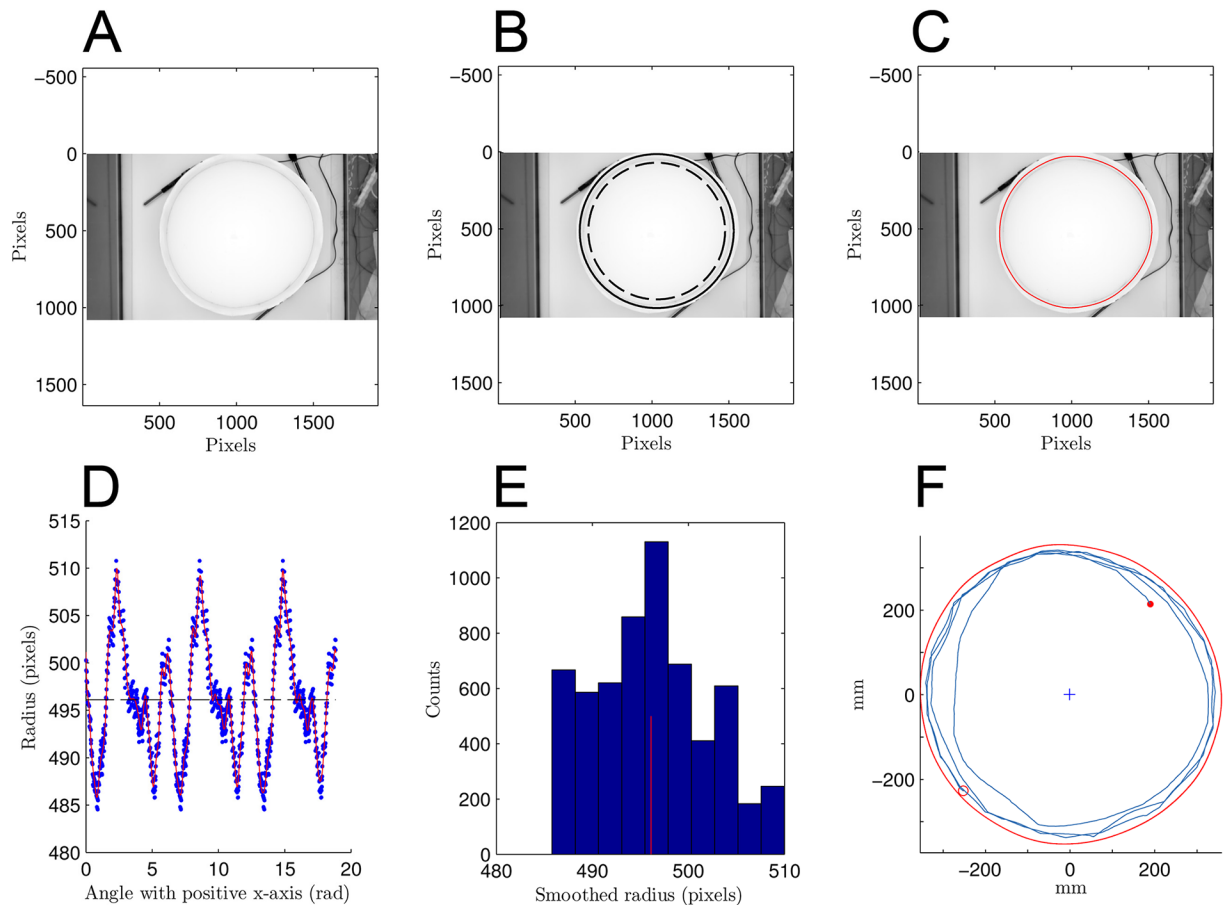
Fig 8. Tracking. A: Background image in grayscale extracted from a video file of an experiment with the biggest tank (radius R = 353 mm). B: Arena estimated from user-defined mask. The outer bold circle of radius R is derived from the mask drawn by the user of the tracking software and defining the area where tracking occurs. Inner dashed circle has arbitrary radius $0.85 \times R$. C: Estimated walls of the tank. D: Estimation of the radius along the circle. Red line stands for cubic spline smoothing and extrapolation over 2000 angular points. The signal is repeated 3 times to improve estimation on limits (at 0 and 2π). The second period is kept to compute wall distances. Estimation of local polynomials is done on 30 equally spaced ranges over one period. Dashed line shows the average radius. E: Distribution of estimated radius in pixels. Red line stands for estimation of the average, used as radius approximation to compute the ratio of pixel to millimeters (PixelsToMm ratio is equal to 0.71 for this video). F: Trajectory of a fish during 40 seconds (2000 points) reported inside the estimated walls. Filled and empty circle respectively stand for start and end points.

with D the metric distance between the optical center of the camera and the bottom of the tank alongside the optical axis, and r_{CCD} the position of the center of the image.

Conversion to metric frame of reference. A cubic spline estimation is computed to smooth the noisy signal of the radius of the detected tank (red line on Fig 8D). The radius signal is repeated over 3 periods to avoid border effects, that is to make estimations at 0 and 2π connected. The Matlab package `immoptibox` (http://www2.imm.dtu.dk/projects/immoptibox/) is used to estimate the cubic splines with function `splinefit`. Splines are piecewise-defined polynomial functions. Knots of the spline are chosen equally spaced. We take 30 knots on each period: the `splinefit` function estimates a local polynomial on each interval defined by two knots. The spline estimated over the second period (2π to 4π) is used for any subsequent calculations to ensure border continuity. The `splineval` function is used to find the radius of the tank wall corresponding to any angular position. Fig 8C shows positions of estimated walls for 2000 θ (*i.e.* of the same magnitude as the number of pixels

describing the tank contours) which allow to compute the center of the tank (x_0, y_0) which will be used as the center of the coordinate system of fish positions. Given the mean radius derived from these estimates and the *a priori* known radius in millimeters, the number of pixels per millimeters is computed (PixelsToMm ratio) (Fig 8E). This value is the conversion ratio to translate image coordinates into the experimental metric frame of reference, which origin is taken to be the centroid of the estimated tank positions. Fig 8F exemplifies the metric fish positions and velocities. The tank coordinates are also converted to metric coordinates through the same translation and a new metric spline is evaluated to obtain the tank metric coordinates for any angular position.

Fish shape detection and body length/width measurement. By removing the image background information from every frame, idTracker is able to detect an approximate fish shape, *i.e.* a set of pixels' coordinates with their corresponding light intensity, which is stored as intermediary files. These files are read to find the main axis of the fish through a principal component analysis, to estimate the typical body length (BL) and body width (BW) for every fish in all frames. BL and BW are calculated as the difference of the maximum and minimum value constituted by the projection of fish points along respectively main axis and secondary axis, then converted to metric values through the above-mentioned conversion ratio. From the main axis, a heading can be derived using the lower light intensity of fish image due to the black eye of the fish. Detection of the head direction along the main axis is done by evaluating the position relative to the barycenter of the blackest five percent of fish image points. To avoid inaccurate shape detection due to the identification of fish shade as fish shape when the fish is stopped against the wall, we approximate BL and BW as their mean value when the fish is moving faster than 15 mm.s^{-1}.

Fish activity selection and sampling. Inter-fish variability in terms of activity is reduced by selecting the phase where a sustained swimming is observed. Considering that the observation period is much longer than the typical activity phase, these detected activity phases are sampled into sections of two minutes, allowing us to grasp the intra-fish variability occurring along an activity section. This procedure is based on the evaluation of fish speed relative to its mean body-length $u = \frac{v}{\text{BL}}$, evaluated through a centered difference scheme with 0.16 s amplitude. First, the program detects whether the fish are *swimming*, *pausing* or *stopping*. *Swimming* is defined as swimming at a speed greater than a threshold velocity $u_{\min}$. *Pausing* is defined as the fastest fish of the group swimming at a speed smaller than $u_{\min}$ for a period of time smaller or equal to $\tau_s = 4$ s. *Stopping* is defined as the fastest fish of the group swimming at a speed smaller than $u_{\min}$ during more than $\tau_s = 4$ s. The program extracts sequences of frames where the fish is either *swimming* or *pausing*, removing *stopping* behavior. From these sequences where the fish are active, *i.e.* not *stopping*, the program cuts series of the same length τ_l. For each experiment, the program will give discontinued series where the fish are swimming. The number of series for each experiment can be used to estimate how much an experiment will participate in the statistics. The values $\tau_l = 120$ s and $\tau_s = 4$ s are chosen as a compromise between the amount of data available and the insensitivity of the results of activity selection to the mere parameters. The value of $u_{\min} = 0.5$ BL.s^{-1} is a reasonably low threshold that allows to exclude the low activity phases where the fish uses pectoral fin swimming, of no interest for our study describing regular fish motion using body and caudal fin swimming. The proportion of time where individuals are detected active over the whole experiment is listed in S1 Table (column *Proportion of active swimming*).

Segmentation. *H. rhodostomus* swims in a burst-and-coast (or burst-and-glide) style. There is a succession of short acceleration phases during which the fish may also change its heading and each acceleration phase is followed by a gliding phase during which the velocity

decreases and then the cycle starts again (Fig 1C, main text). The points of acceleration exhibited by fish when "bursting" is used to detect these decisions. Most heading changes occur at these decision points also called "kicks". Our assumption is that these kicks are sufficient to describe fish swimming behavior, the passive phases containing only a consequence of the previous action, being entirely determined by physical forces. Thus, we can minimize the amount of noise given by barycenter estimation and minor trajectory deviations by describing the trajectory as segments between kicks. In order to properly identify the acceleration events, we have to smooth the raw speed time series obtained by taking the modulus of the velocity vector through a centered difference scheme over a moving time window of bandwidth 0.08 s (4 frames). We use a Savitsky-Golay [61] filter of degree three over a 0.36 s time window (18 frames) to smooth the raw time series, allowing us to classify the time series into accelerating and decelerating state [38]. To limit remaining noise, we merge any consecutive pair of accelerations separated by a deceleration lasting less than 0.08 s. We then discard any acceleration lasting less than 0.08 s as it is a too short period of time regarding the typical duration of a body motion. We assume that the times of the kicks coincide with the starting of the acceleration periods.

Segmented variable estimation. Assuming a fish instantaneously takes a new direction and velocity every time its body motion produces an acceleration, we reduce the full time-sampled trajectories of every two minutes activity sample b of a fish Id to a set of positions and times of interest $\{\vec{x}_i, t_i\}_b^{Id}$ corresponding to a set of decision events. From this point of view, the statistics of interest used in both data and simulation are:

- The length and time intervals between two decision events

- The absolute and relative to rotation direction change in orientation due to a decision event

- The distance between the fish centroid and the closest point on the wall (*wall distance*)

- The top speed between decision events

and are discussed below.

Length and time intervals
Length between two decision events is defined as the Euclidean distance between both decision points $l_i = \parallel \vec{x}_{i+1} - \vec{x}_i \parallel$. The duration $\tau_i = t_{i+1} - t_i$ of the kick initiated at t_i is calculated from decision times.

Heading
Heading of the fish ϕ_i during the kick initiated at t_i is identified to the direction of the vector between two decision points.

Computation of wall distances
θ_i, the angle in radians between the positive x-axis of the frame and (x_i, y_i) is computed from the current position of the fish (x_i, y_i). The radius r_{θ_i} for θ_i is computed based on the spline estimation described in the previous section. The wall distance is the Euclidean distance between fish position and estimated wall position as in $r_{w,i} = \parallel r_{\theta_i} \vec{e}_{\theta_i} - \vec{x}_i \parallel$.

Top speed between kicks
The top speed v_i between kicks is determined from the smoothed speed time series used by the segmentation procedure, taking the maximum value reached between kicks at time t_i and t_{i+1}.

Symmetrization of the data. We did not observe any statistically relevant left/right asymmetry in the distribution of angles θ_w (1 and 2 fish; see Fig 1E), or ψ and $\Delta\phi$ (2 fish; see Fig 1F).

Assuming perfect left/right symmetry amounts to saying that a trajectory as observed from the top of the tank (as we did) has exactly the same probability to occur as the very same trajectory but as seen from under the tank ("mirror trajectory"). For the mirror trajectory, all angles θ_w, ψ, and $\Delta\phi$ have the opposite sign compared to the original trajectory. Hence, the systematic angle change $\delta\phi$ of a fish due to the interaction with the wall (1 or 2 fish experiments) and with another fish (2 fish experiments) must exactly satisfy the symmetry condition

$$\delta\phi(r_w, -\theta_w) = -\delta\phi(r_w, \theta_w), \tag{17}$$

for 1 fish experiments, and

$$\delta\phi(r_w, -\theta_w, d, -\psi, -\Delta\phi) = -\delta\phi(r_w, \theta_w, d, \psi, \Delta\phi), \tag{18}$$

for 2 fish experiments. In order to analyze and disentangle the interactions, notably the attraction and alignment interactions between fish, we have imposed general functional forms (see Eqs (7, 12, 13) in the main text) obeying these conditions. Accordingly, exploiting this assumed but reasonable left/right symmetry, we have effectively doubled our data set by adding the mirror trajectory associated to each observed trajectory. This procedure not only reduces the statistical uncertainty on quantities depending on angles (by a factor $\sqrt{2}$, by the law of large numbers), but it also helps stabilizing the optimization procedure used to extract the various components of the interactions from $\delta\phi$, which is detailed in the next section.

Analysis of the interactions

Interaction with the wall of a single fish. The position and orientation of a fish relative to the wall is fully determined by r_w and θ_w (see the main Fig 1E). As explained below Eq (4), in addition to the random component of the angle change $\delta\phi$ between kicks, which accounts for the fish spontaneous motion, we look for a systematic angle change due to the presence of the wall of the form $\delta\phi_W(r_w, \theta_w) = f_w(r_w)O_w(\theta_w)$, where $O_w(\theta_w)$ is an odd function of θ_w. If the wall interaction tends to push back the fish toward the center of the tank and is isotropic, one has exactly that $O_w(\theta_w) \propto \sin(\theta_w)$ (the projection of a radial force on the angular acceleration, which is perpendicular to the velocity). Hence, the ansatz $\delta\phi_W(r_w, \theta_w) = f_w(r_w)O_w(\theta_w)$ is the simplest generalization accounting for the fish anisotropic perception of its environment, while keeping a product form and still obeying left/right symmetry. Note that in the Gautrais *et al.* model [34, 51], the sign function was phenomenologically used instead of the sin function. Despite having a qualitatively similar shape, and being both odd functions of θ_w as requested by symmetry (see above), the sign function has the unphysical/unbiological drawback of attributing a sharp discontinuous response to a fish when it approaches the wall from an arbitrary small angle from the left or the right (an angle sign that the fish could not measure with such a perfect precision).

In order to measure the actual $f_w(r_w)$ and $O_w(\theta_w)$, we first define a discrete mesh of the two-dimensional space (r_w, θ_w), with each direction $r_w \in [0, R]$ and $\theta_w \in [-\pi, \pi]$ partitioned respectively in I and J boxes (typical values are $I = 40$ and $J = 30$). We tabulate the unknown functions $f_w(r_w)$ and $O_w(\theta_w)$ by defining f_i as the (mean) value of $f_w(r_w)$ when r_w falls in box i, and O_j as the (mean) value of $O_w(\theta_w)$ when θ_w falls in box j. We finally define ϵ_{ij} as the number of data points falling in the squared box of index i and j, and $\delta\phi_{ij}$ as the averaged experimental angle change for data points in this box.

f_i and O_j are then determined by minimizing the error

$$\Delta = \sum_{i=1}^{I}\sum_{j=1}^{J}\epsilon_{ij}(\delta\phi_{ij} - f_iO_j)^2. \tag{19}$$

This minimization is achieved by writing the equation $\partial\Delta/\partial f_i = 0$ and $\partial\Delta/\partial O_j = 0$ under the form

$$f_i = \frac{\sum_{j=1}^{J}\epsilon_{ij}\delta\phi_{ij}O_j}{\sum_{j=1}^{J}\epsilon_{ij}O_j^2},\tag{20}$$

$$O_j = \frac{\sum_{i=1}^{I}\epsilon_{ij}\delta\phi_{ij}f_i}{\sum_{i=1}^{I}\epsilon_{ij}f_i^2}.\tag{21}$$

It is straightforward to realize that if the experimental $\delta\phi_{ij}$ were exactly of the form $f_i\times O_j$, the right-hand side of Eqs (20, 21) would indeed exactly recover f_i and O_j.

This system is solved iteratively starting from reasonable initial conditions, but most importantly with O_j being an odd function of θ_w. We checked that this procedure leading to the results of Fig 4 does not depend on the initial conditions.

In practice, knowing the f_i's and O_j's at a given iteration, we generate their values at the next iteration by computing

$$f_i' = (1-p)f_i + p\hat{f}_i,\tag{22}$$

$$O_j' = (1-p)O_j + p\hat{O}_j,\tag{23}$$

where $\hat{f}_i$ and $\hat{O}_j$ are given by the right-hand side of Eqs (20, 21), and p is a damping parameter that we took equal to 0.25. Obviously, the fixed point solution of Eqs (22, 23) ultimately coincides with the wanted solution of Eqs (20, 21).

Since multiplying O_j by an arbitrary constant and dividing f_i by the same constant leaves the product $f_i O_j$ unchanged, we choose to normalize O_j (and all angular functions appearing in Figs 4 and 6) after each iteration such that its average square is unity:

$$\frac{1}{J}\sum_{j=1}^{J}O_j^2 = \frac{1}{2\pi}\int_{-\pi}^{+\pi}O_w^2(\theta_w)\,d\theta_w = 1\tag{24}$$

The procedure described here converges to a relative accuracy of 10^{-6} in typically 100 iterations, leading to the result of Fig 4.

Attraction and alignment interaction between two fish. We define a procedure identical in spirit as above, but involving more unknown interaction functions now depending on the 3 parameters d, ψ, and $\Delta\phi$ defined in Fig 1F. We choose to restrict our analysis to data points for which the focal fish was at a distance greater than 2 BL ~ 60 mm, for which the wall interaction is found to be negligible by the previous analysis.

Again, we partition the three-dimensional space (d, ψ, $\Delta\phi$) in a mesh of $K\times L\times M$ boxes (with typically $K = 40$, $L = M = 30$). As previously, we define ϵ_{klm} as the number of data points falling in the cubic box of index k, l and m, and $\delta\phi_{klm}$ as the averaged experimental angle change of the focal fish for data points in this box.

As explained below Eq (11), the systematic angle change $\delta\phi = \delta\phi_{\text{Att}}(d, \psi, \Delta\phi) + \delta\phi_{\text{Ali}}(d, \psi, \Delta\phi)$ due to the attraction and alignment forces is parameterized by 6 unknown functions

$$\delta\phi_{\text{Att}}(d,\psi,\Delta\phi) = F_{\text{Att}}(d)O_{\text{Att}}(\psi)E_{\text{Att}}(\Delta\phi),\tag{25}$$

$$\delta\phi_{\text{Ali}}(d,\psi,\Delta\phi) = F_{\text{Ali}}(d)O_{\text{Ali}}(\Delta\phi)E_{\text{Ali}}(\psi),\tag{26}$$

with parity constraints (O functions are odd, E functions are even). These rather general

functional forms translate into mathematical forms the notion of attraction and alignment which can in fact have cumulative or contrary effects depending on the relative position or orientation of the two fish (see Fig 6A). Intuitively, attraction means that if the other fish is on the right, the focal fish should turn to the right, and should turn by the same amount to the left in the "mirror situation" (see the "Symmetrization of the data" section above in Materials and methods).

Again we tabulate these 6 functions, depending each on only one variable, as $F_{\mathrm{Att},k}$, $O_{\mathrm{Att},l}$, $E_{\mathrm{Att},m}$, $F_{\mathrm{Ali},k}$, $O_{\mathrm{Ali},m}$, $E_{\mathrm{Ali},l}$. These functions are determined by minimizing the error

$$\Delta = \sum_{k=1}^{K}\sum_{l=1}^{L}\sum_{m=1}^{M}\epsilon_{klm} \times (\delta\phi_{klm} - F_{\mathrm{Att},k}O_{\mathrm{Att},l}E_{\mathrm{Att},m} \\ - F_{\mathrm{Ali},k}O_{\mathrm{Ali},m}E_{\mathrm{Ali},l})^2. \tag{27}$$

Expressing that the derivative of Δ with respect to the 6 tabulated function is zero at the minimum, we obtain 6 equations similar to Eqs (20, 21), although a bit more complicated. For instance, $F_{\mathrm{Att},k}$ satisfies the fixed point equation

$$F_{\mathrm{Att},k} = \frac{\sum_{l,m}\epsilon_{klm}O_{\mathrm{Att},l}E_{\mathrm{Att},m}(\delta\phi_{klm} - F_{\mathrm{Ali},k}O_{\mathrm{Ali},m}E_{\mathrm{Ali},l})}{\sum_{l,m}\epsilon_{klm}O_{\mathrm{Att},l}^2 E_{\mathrm{Att},m}^2}. \tag{28}$$

Note the presence of the counter term $F_{\mathrm{Ali},k} O_{\mathrm{Ali},m} E_{\mathrm{Ali},l}$ between the parentheses of Eq (28), meaning that the attraction force is evaluated by subtracting the estimated alignment interaction to the actual experimental angle change. Again, it is straightforward to check that if the experimental $\delta\phi_{klm}$ takes exactly the form $F_{\mathrm{Att},k} O_{\mathrm{Att},l} E_{\mathrm{Att},m} + F_{\mathrm{Ali},k} O_{\mathrm{Ali},m} E_{\mathrm{Ali},l}$, Eq (28) exactly recovers the correct d dependence of the attractive force $F_{\mathrm{Att},k}$.

Finally, the resulting system of 6 equations is solved by the same iterative procedure as for the wall interaction (including the normalization of the average square of angular functions), leading to the results of Fig 6B, 6C, and 6D. In particular, the angular functions obtained are plotted with simple analytical forms given by

$$O_{\mathrm{Att}}(\psi) \quad \propto \quad \sin(\psi)[1 - 0.33\cos(\psi)], \tag{29}$$

$$E_{\mathrm{Att}}(\Delta\phi) \quad \propto \quad 1 + 0.48\cos(\Delta\phi) - 0.31\cos(2\Delta\phi), \tag{30}$$

$$O_{\mathrm{Ali}}(\Delta\phi) \quad \propto \quad \sin(\Delta\phi)[1 + 0.30\cos(2\Delta\phi)], \tag{31}$$

$$E_{\mathrm{Ali}}(\psi) \quad \propto \quad 1 + 0.60\cos(\psi) - 0.32\cos(2\psi), \tag{32}$$

the overall multiplicative constant being fixed by the normalization of the average square of these functions to unity. The analytical forms for $F_{\mathrm{Att}}(d)$ and $F_{\mathrm{Ali}}(d)$ are respectively discussed above Eqs (14) and (15).

One of the main interests of assuming these reasonable product forms is to drastically limit the number of fitting parameters, but also to derive from these forms an *explicit* model. Indeed, if we were to produce a three-dimensional map of $\delta\phi$ on the $K{\times}L{\times}M$ mesh made of 36000 boxes ($K = 40$, $L = M = 30$), we would need 36000 fitting parameters at this resolution, barely smaller than the number of kicks available from experiment ($\sim$ 200000 for 2 fish). The present procedure only requires $2{\times}(K + L + M) = 200$ fitting parameters, which allows us to extract the main feature of the interactions with a high resolution and yet a rather small noise (see Figs 4 and 6, and in particular, the various angular functions).

Parameter estimation and simulations

As shown in the inset of Fig 2D, the spontaneous angle change has a nearly Gaussian distribution of zero mean (left/right symmetry) and variance $\gamma_R \approx 0.35$ rad, as measured when the fish is far from the wall ($r_w > 2$ BL). This value of γ_R was used in all simulations of the one fish dynamics, in the three circular tanks of radius $R = 176, 250, 353$ mm. For two fish, γ_R is found experimentally to be slightly larger, probably because each fish activity is stimulated by the presence of the other fish.

The total angle change PDF shown in the main graph of Fig 2D (one fish) and in Fig 5D (two fish) is dominated by the majority of kicks where the fish are very close to the wall, and has a width reduced by a factor nearly 3, so that a fair estimate of the spontaneous angle change intensity near the wall is (see Eq (6))

$$\gamma_R^0 = (1 - \alpha)\gamma_R \sim \gamma_R/3. \tag{33}$$

The exact value of γ_R^0 (or α), of the comfort distance to the wall l_c introduced below Eq (7), which should naturally be of order of 1 BL ~ 30 mm, and of the intensity of the interaction with the wall γ_W are tuned near their experimental expected value, resulting in the very satisfactory agreement with experiments obtained in Figs 2 and 5. The values of the parameters used in our simulations are reported in Table 2.

For each of the 4 conditions (the three tank sizes for one fish and one tank size with two fish; see S1 Table), the graphs of Fig 2 (one fish) and in Fig 5 (two fish) are obtained by typically simulating 100 runs with 10^6 kicks each, compared to the 100000-200000 kicks recorded experimentally for each condition/tank.

Supporting information

S1 Table. List of experiments.
(PDF)

S1 Video. One fish trajectories. Left: A typical experiment with 1 fish swimming in a circular tank of radius 250 mm. Right: 3D representation of the tracking and analysis output. The successive positions of the fish are represented by spheres whose color depends on its current speed (maximum and minimum speeds are shown in red and light blue respectively). The xy plane represents the position of fish as seen in the top down recording, and the z axis corresponds to the time evolution in the last 5 s. Newer positions appear at the top, while positions older than the aforementioned threshold are erased from the animation.
(MOV)

S2 Video. Experiment vs model—one fish. Comparison of the dynamics of a single fish in an experiment (left) and in a typical run of the model (right; starting from approximately the same location as in the experiment), in a circular tank of radius 250 mm. The animated fish size is not strictly proportional to its actual dimensions, but used to facilitate visualization. The kick-segmented dynamics obtained in the model was exploited to produce a continuous trajectory by the procedure detailed in the main text.
(MOV)

S3 Video. Two fish trajectories. Left: A typical experiment with 2 fish swimming in a circular tank of radius 250 mm. Right: 3D representation of the tracking and analysis output. The successive positions of fish are represented by spheres whose color depends on their current speed (maximum and minimum speeds are shown in red and light blue respectively). The xy plane represents the position of fish as seen in the top down recording, and the z axis corresponds to

the time evolution in the last 5 s. Newer positions appear at the top, while positions older than the aforementioned threshold are erased from the animation.
(MOV)

S4 Video. Experiment vs model—two fish. Comparison of the dynamics of groups of two fish in an experiment (left) and in a typical run of the model (right), in a circular tank of radius 250 mm (see also the legend of S1 Video for details). On the short 40 s time scale of both videos, fish display a greater variety of behaviors than for one fish experiments/simulations: wide excursions and crossing of the tank, change from clockwise to anticlockwise motion along the wall. Both videos also illustrate the fact that the geometrical leader (the red fish; see main text) can persist for a long time, although leader changes are observed experimentally and in the model on longer time scales (see S5 Video).
(MOV)

S5 Video. Leader change. This video illustrates an exchange (at 20 s) of geometrical leader and follower in a typical run of the model with two fish in a circular tank of radius 250 mm. Note that the large angle change of the red fish at 20 s is probably due to the rejection procedure detailed in the main text, sometimes leading to such a large spontaneous random angle change $\delta\phi_R$ when a fish arrives almost perpendicularly to the wall.
(MOV)

Author Contributions

Conceptualization: Clément Sire, Guy Theraulaz.

Data curation: Daniel S. Calovi, Valentin Lecheval, Ugo Lopez.

Formal analysis: Daniel S. Calovi, Alexandra Litchinko, Ugo Lopez, Clément Sire.

Funding acquisition: Guy Theraulaz.

Investigation: Daniel S. Calovi, Valentin Lecheval, Ugo Lopez, Hugues Chaté, Clément Sire, Guy Theraulaz.

Methodology: Alexandra Litchinko, Clément Sire, Guy Theraulaz.

Project administration: Guy Theraulaz.

Software: Daniel S. Calovi, Valentin Lecheval, Alfonso Pérez Escudero.

Supervision: Clément Sire, Guy Theraulaz.

Validation: Clément Sire, Guy Theraulaz.

Visualization: Daniel S. Calovi, Valentin Lecheval, Alfonso Pérez Escudero.

Writing – original draft: Valentin Lecheval, Clément Sire, Guy Theraulaz.

Writing – review & editing: Valentin Lecheval, Clément Sire, Guy Theraulaz.

References

1. Brown A. and Poon W. (2014) Ionic effects in self-propelled Pt-coated Janus swimmers. *Soft matter* 10.22:4016–4027. https://doi.org/10.1039/C4SM00340C PMID: 24759904

2. Walther A. and Müller A. H. E. (2008) Janus particles. *Soft Matter* 4.4:663. https://doi.org/10.1039/b718131k

3. Howse J. R., Jones R. A., Ryan A. J., Gough T., Vafabakhsh R. and Golestanian R. (2007), Self-motile colloidal particles: from directed propulsion to random walk. *Phys. Rev. Lett.* 99.4:048102. https://doi.org/10.1103/PhysRevLett.99.048102 PMID: 17678409

4. Theurkauff I., Cottin-Bizonne C., Palacci J., Ybert C., and Bocquet L. (2012) Dynamic clustering in active colloidal suspensions with chemical signaling. *Phys. Rev. Lett.* 108.26:268303. https://doi.org/10.1103/PhysRevLett.108.268303 PMID: 23005020

5. Palacci J., Cottin-Bizonne C., Ybert C., and Bocquet L. (2010) Sedimentation and effective temperature of active colloidal suspensions. *Phys. Rev. Lett.* 105.8:088304. https://doi.org/10.1103/PhysRevLett.105.088304 PMID: 20868136

6. Palacci J., Sacanna S., Steinberg A. P., Pine D. J., and Chaikin P. M. (2013) Living crystals of light-activated colloidal surfers. *Science* 339.6122:936–940. https://doi.org/10.1126/science.1230020 PMID: 23371555

7. Buttinoni I., Volpe G., Kümmel F., Volpe G., and Bechinger C. (2012) Active Brownian motion tunable by light. *J. Phys. C* 24.28:284129.

8. Ginot F., Theurkauff I., Levis D., Ybert C., Bocquet L., Berthier L., and Cottin-Bizonne C. (2015) Non-equilibrium equation of state in suspensions of active colloids. *Phys. Rev. X* 5.1:011004.

9. Narayan V., Ramaswamy S., and Menon N. (2007) Long-lived giant number fluctuations in a swarming granular nematic. *Science* 317.5834:105–108. https://doi.org/10.1126/science.1140414 PMID: 17615353

10. Kudrolli A., Lumay G., Volfson D., and Tsimring L. S. (2008) Swarming and swirling in self-propelled polar granular rods. *Phys. Rev. Lett.* 100.5:058001. https://doi.org/10.1103/PhysRevLett.100.058001 PMID: 18352433

11. Deseigne J., Dauchot O., and Chaté H. (2010) Collective motion of vibrated polar disks. *Phys. Rev. Lett.* 105.9:098001. https://doi.org/10.1103/PhysRevLett.105.098001

12. Thutupalli S., Seemann R., and Herminghaus S. (2011) Swarming behavior of simple model squirmers. *New J. Phys.* 13.7:073021. https://doi.org/10.1088/1367-2630/13/7/073021

13. Bricard A., Caussin J.-B., Desreumaux N., Dauchot O., and Bartolo D. (2013) Emergence of macroscopic directed motion in populations of motile colloids. *Nature* 503.7474:95–98. https://doi.org/10.1038/nature12673 PMID: 24201282

14. Tunstrøm K., Katz Y., Ioannou C. C., Huepe C., Lutz M. J., and Couzin I. D. (2013) Collective states, multistability and transitional behavior in schooling fish. *PLoS Comput. Biol.* 9(2):e1002915. https://doi.org/10.1371/journal.pcbi.1002915 PMID: 23468605

15. Calovi D. S., Lopez U., Ngo S., Sire C., Chaté H., and Theraulaz G. (2014) Swarming, Schooling, Milling: Phase diagram of a data-driven fish school model. *New J. Phys.*, 16:015026.

16. Sire C. (2017) Active and passive self-propelled particles. Preprint.

17. Helbing D. and Molnar P. (1995) Social force model for pedestrian dynamics. *Phys. Rev. E* 51:4282.

18. Warburton K. (1997) Social forces in animal congregations: interactive, motivational, and sensory aspects. In: Parrish J.K. & Hamner W.M., *Animal groups in three dimensions*, p. 313–335. Cambridge: Cambridge University Press.

19. Filella A., Nadal F., Sire C., Kanso E., and Eloy C. (2017) Model of collective fish behavior with hydrodynamic interactions. *arXiv:1705.07821*.

20. Barberis L., Peruani F. (2016) Large-scale patterns in a minimal cognitive flocking model: incidental leaders, nematic patterns, and aggregates. *Phys. Rev. Lett.*, 117:248001. https://doi.org/10.1103/PhysRevLett.117.248001 PMID: 28009185

21. Camazine S., Deneubourg J. L., Franks N., Sneyd J., Theraulaz G., and Bonabeau E. (2001) Self-Organization in Biological Systems. Princeton, NJ: Princeton University Press.

22. Giardina I. (2008) Collective behavior in animal groups: Theoretical models and empirical studies. *HFSP Journal* 2:205–219. https://doi.org/10.2976/1.2961038 PMID: 19404431

23. Sumpter D. J. T. (2010) Collective Animal Behavior. Princeton, NJ: Princeton University Press.

24. Krause J. and Ruxton G. D. (2002) Living in groups. Oxford University Press, Oxford.

25. Parrish J. K. and Edelstein-Keshet L. (1999) Complexity, pattern, and evolutionary trade-offs in animal aggregation. *Science* 284(5411), 99–101. https://doi.org/10.1126/science.284.5411.99 PMID: 10102827

26. Branson K., Robie A. A., Bender J., Perona P., and Dickinson M. H. (2009) High-throughput ethomics in large groups of Drosophila. *Nat. Methods* 6:451–457. https://doi.org/10.1038/nmeth.1328 PMID: 19412169

27. Pérez-Escudero A., Vicente-Page J., Hinz R. C., Arganda S., and de Polavieja G. G. (2014) idTracker: tracking individuals in a group by automatic identification of unmarked animals. *Nature Methods*, 11:743–748. https://doi.org/10.1038/nmeth.2994 PMID: 24880877

28. Dell A. I. et al. (2014) Automated image-based tracking and its application in ecology. *Trends Ecol. Evol.* 29(7):417–428. https://doi.org/10.1016/j.tree.2014.05.004 PMID: 24908439

29. Buhl J.et al. (2006) From disorder to order in marching locusts. *Science* 312(5778):1402–1406. https://doi.org/10.1126/science.1125142 PMID: 16741126

30. Attanasi A.et al. (2014) Collective behaviour without collective order in wild swarms of midges. *PLoS Comput Biol* 10(7):e1003697. https://doi.org/10.1371/journal.pcbi.1003697 PMID: 25057853

31. Schneider J. and Levine J. D. (2014) Automated identification of social interaction criteria in Drosophila melanogaster. *Biol. Lett.* 10:20140749. https://doi.org/10.1098/rsbl.2014.0749 PMID: 25354920

32. Katz Y., Tunstrøm K., Ioannou C., Huepe C., and Couzin I. D. (2011) Inferring the structure and dynamics of interactions in schooling fish. *Proc. Natl. Acad. Sci. USA* 108(46):18720–18725. https://doi.org/10.1073/pnas.1107583108 PMID: 21795604

33. Herbert-Read J.et al. (2011) Inferring the rules of interaction of shoaling fish. *Proc. Natl. Acad. Sci. USA* 108(46):18726–18731. https://doi.org/10.1073/pnas.1109355108 PMID: 22065759

34. Gautrais J.et al. (2012) Deciphering interactions in moving animal groups. *PLoS Comp. Biol.* 8(9): e1002678. https://doi.org/10.1371/journal.pcbi.1002678

35. Mwaff V., Anderson R. P., Butail S., and Porfiri M. (2014) A jump persistent turning walker to model zebrafish locomotion. *J. R. Soc. Interface* 12:20140884. https://doi.org/10.1098/rsif.2014.0884

36. Herbert-Read J. E., Rosén E., Szorkovszky A., Ioannou C., Rogell B., Perna A., Ramnarine I. W., Kotrschal A., Kolm N., Krause J., Sumpter D. J. T. (2017) How predation shapes the social interaction rules of shoaling fish. *Proceedings of the Royal Society B*, 284:20171126. https://doi.org/10.1098/rspb.2017.1126 PMID: 28855361

37. Ballerini M.et al. (2008) Interaction Ruling Animal Collective Behavior Depends on Topological Rather than Metric Distance: Evidence from a Field Study. *Proc. Natl. Acad. Sci.* USA 105:1232–1237. https://doi.org/10.1073/pnas.0711437105 PMID: 18227508

38. Nagy M., Akos Z., Biro D., and Vicsek T. (2010) Hierarchical group dynamics in pigeon flocks. *Nature* 464(7290):890–U99. https://doi.org/10.1038/nature08891 PMID: 20376149

39. Bialek W.et al. (2014) Social interactions dominate speed control in poising natural flocks near criticality. *Proc. Natl. Acad. Sci.* 111:7212–7217. https://doi.org/10.1073/pnas.1324045111 PMID: 24785504

40. Lukeman R., Li Y. X., Edelstein-Keshet L. (2010) Inferring individual rules from collective behavior. *Proc. Natl. Acad. Sci.* 107 (28), 12576–12580. https://doi.org/10.1073/pnas.1001763107 PMID: 20616032

41. de Chaumont F.et al. (2012) Computerized video analysis of social interactions in mice. *Nat. Methods* 9: 410–U134. https://doi.org/10.1038/nmeth.1924 PMID: 22388289

42. Shemesh Y.et al. (2013) High-order social interactions in groups of mice. *eLife* 2:e00759. https://doi.org/10.7554/eLife.00759 PMID: 24015357

43. Ginelli F.et al. (2015) Intermittent collective dynamics emerge from conflicting imperatives in sheep herds. *Proc. Natl. Acad. Sci. USA*, 112:12729–12734. https://doi.org/10.1073/pnas.1503749112 PMID: 26417082

44. King A. J.et al. (2012) Selfish-herd behaviour of sheep under threat. *Curr. Biol.* 22(14):R561–R562. https://doi.org/10.1016/j.cub.2012.05.008 PMID: 22835787

45. Strandburg-Peshkin A., Farine D. R., Couzin I. D., and Crofoot M. C. (2015) Shared decision-making drives collective movement in wild baboons. *Science* 348:1358–1361. https://doi.org/10.1126/science.aaa5099 PMID: 26089514

46. Ballesta S., Reymond G., Pozzobon M., and Duhamel J. R. (2014) A real-time 3D video tracking system for monitoring primate groups. *J. Neurosci. Methods* 234:147–152. https://doi.org/10.1016/j.jneumeth.2014.05.022 PMID: 24875622

47. Moussaïd M., Helbing D., and Theraulaz G. (2011) How simple rules determine pedestrian behavior and crowd disasters. *Proc. Natl. Acad. Sci. USA* 108:6884–6888. https://doi.org/10.1073/pnas.1016507108 PMID: 21502518

48. Gallupa A. C.et al. (2012) Visual attention and the acquisition of information in human crowds. *Acad Proc. Natl. Sci. USA* 109:7245–7250. https://doi.org/10.1073/pnas.1116141109

49. Lopez U., Gautrais J., Couzin I. D., and Theraulaz G. (2012) From behavioural analyses to models of collective motion in fish schools. *Interface Focus* 2:693–707. https://doi.org/10.1098/rsfs.2012.0033 PMID: 24312723

50. Herbert-Read J. E. (2016) Understanding how animal groups achieve coordinated movement. *J. Exp. Biol.* 219:2971–2983. https://doi.org/10.1242/jeb.129411 PMID: 27707862

51. Gautrais J.et al. (2009) Analyzing fish movement as a persistent turning walker. *J. Math. Biol.* 58:429–445. https://doi.org/10.1007/s00285-008-0198-7 PMID: 18587541

52. Aoki I. (1982) A simulation study on the schooling mechanism in fish. *Bull. J. Soc. Sci. Fish* 48 (8):1081–1088. https://doi.org/10.2331/suisan.48.1081

53. Huth A. and Wissel C. (1992) The simulation of the movement of fish schools. *J. Theor. Biol.* 156 (3):365–385. https://doi.org/10.1016/S0022-5193(05)80681-2

54. Couzin I. D., Krause J., James R., Ruxton G., and Franks N. (2002) Collective memory and spatial sorting in animal groups. *J. Theor. Biol.* 218(5):1–11. https://doi.org/10.1006/jtbi.2002.3065 PMID: 12297066

55. Pita D., Moore B. A., Tyrrell L. P., and Fernandez-Juricic E. (2015) Vision in two cyprinid fish: implications for collective behavior. *PeerJ* 3:e1113. https://doi.org/10.7717/peerj.1113 PMID: 26290783

56. 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. https://doi.org/10.1098/rsos.150473 PMID: 26909173

57. Elgeti J., Kaupp U. B., and Gompper G. (2010) Hydrodynamics of sperm cells near surfaces. *Biophys. J.* 99:1018–26. https://doi.org/10.1016/j.bpj.2010.05.015 PMID: 20712984

58. Vladescuat I. D.et al. (2014) Filling an emulsion drop with motile bacteria. *Phys. Rev. Lett.* 113:268101. https://doi.org/10.1103/PhysRevLett.113.268101

59. Tailleur J. and Cates M. E. (2009) Sedimentation, trapping, and rectification of dilute bacteria. *EPL* 86:60002. https://doi.org/10.1209/0295-5075/86/60002

60. Elgeti J. and Gompper G. (2015) Run-and-tumble dynamics of self-propelled particles in confinement. *EPL* 109:58003. https://doi.org/10.1209/0295-5075/109/58003

61. Savitzky A. and Golay M. J. E. (1964) Smoothing and Differentiation of Data by Simplified Least Squares Procedures. *Anal. Chem.* 36(8):1627–39. https://doi.org/10.1021/ac60214a047