Supplementary Information for

Predators attacking virtual prey reveal the costs and benefits of leadership

Christos C. Ioannou, Florence Rocque, James E. Herbert-Read, Callum Duffield, Josh A. Firth

Corresponding author: C.C. Ioannou
E-mail: c.c.ioannou@bristol.ac.uk

This PDF file includes:

Figs. S1 to S6
Tables S1 to S3

www.pnas.org/cgi/doi/10.1073/pnas.1816323116

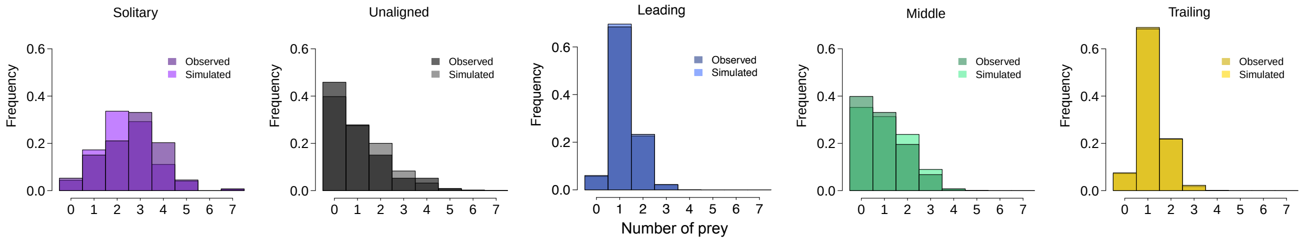

Fig. S1. Probability distributions of each prey position. The mean number of prey present in each of the different positions (solitary, grouped but unaligned, leader, middle, or trailing) was calculated across the 133 attacks (observed), and also separately for the 19,324 sampled instances in the simulation (i.e. every 100 time steps from the simulations presented to the fish; simulated). The difference in these mean values between the attacks and simulation samples was calculated for each prey type separately as the observed differences. To determine if these differences were to be expected at random, we randomly shuffled whether the number of each prey type was from the attacks or the simulation sample, maintaining the total number of attacks (133) and simulation samples (19,324). The difference in the mean values between the attacks and simulation samples for each prey position separately was recalculated, and the process iterated 10,000 times in total. A P value was then calculated for each prey position as the probability that the observed difference in mean values would be expected from a random process. It is rare that there are instances in the simulation where there are no solitary, leading or trailing prey.

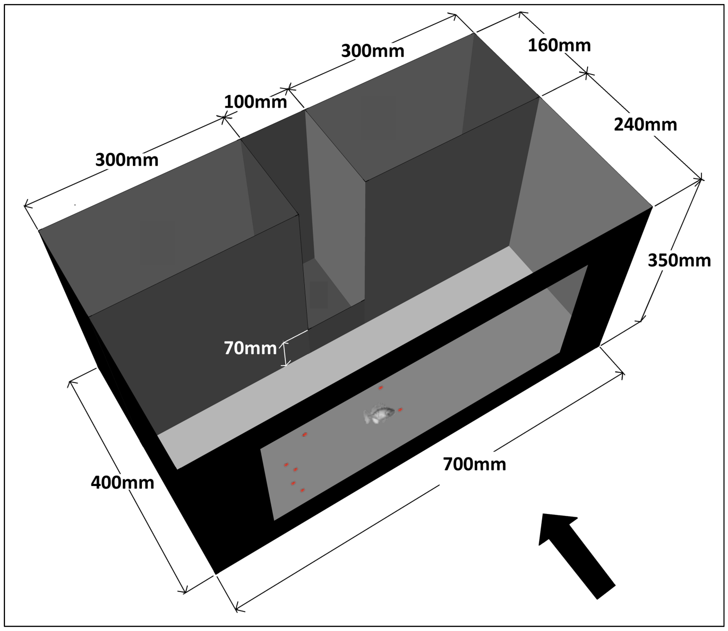

Fig. S2. The experimental tank. The arena used with virtual prey depicted (red dots) being projected (position of arrow) onto the screen in front of a stickleback (water depth 25 cm). The compartments on the back wall of the arena hold the companion fish. Both sides and back were covered in back plastic to reduce disturbance. A black sheet was also hung between the tank and projection and recording equipment to exclude disturbance and minimise reflections on the wall of the tank where the prey were projected. The tank was separated into the main arena and a refuge area at the back with two compartments either side. The refuge was created by shielding light from above with a piece of black plastic, and the addition of a small 7cm high vertical wall, behind which the test fish was placed at the start of the experiment. Inside each of the two compartments, five companion fish were held for each test day; these companion individuals were not used as companion fish on consecutive days or used as test fish. They were separated from the refuge area by 2mm black plastic mesh to ensure they had visual and olfactory contact to the fish being tested, facilitating habituation by the test fish and increasing their perception of food competition. The data projector (BenQ MW523, resolution 1280 x 800 pixels at 60Hz) was positioned below the level of the test tank to minimise reflected glare from the bulb on the wall of the tank. The screen consisted of a sheet of white translucent film (Rosco gel no. 252) mounted on the inside of the test tank, which was opaque enough to reflect the projected prey, and transparent enough to allow the test fish to be visible (Movie S1). The rectangular area within which projected prey could move measured 17.5 x 18.7 cm (height x width) and was positioned so that the top left corner was cm 1.1 left and 1.6 cm down from the top left corner of a black plastic border mounted around the screen. A Panasonic SD800 video recorder was positioned directly facing the screen, 40cm above the projector and 108cm from the front of the test tank, to record at 1920 x 1080 pixels and 25 frames per second. The behaviour of the fish was observed remotely on a computer monitor.

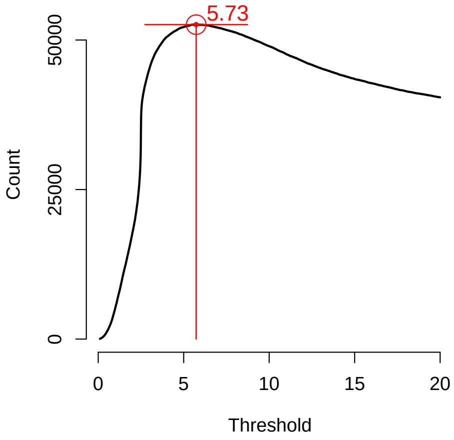

Fig. S3. Determining a threshold for whether two prey are near neighbours. Rather than select an arbitrary distance threshold to classify whether prey had a near neighbour (and hence whether they were in a group or solitary), we use the differences in nearest neighbour distance between the social and asocial prey types to determine a threshold as, due to the way they are programmed, we know follower types will behave socially and asocial prey will not. Across a range of potential thresholds from 0.1 to 20 patches (the unit of the coordinate system in Netlogo) in 0.01 increments (x axis), the number of the follower type prey classified as in a group minus the number of the asocial type classified as in a group from the 19,324 simulation samples (y axis) is plotted. This difference was maximised at 5.73 patches (equivalent to 2.9 cm as projected on the experimental tank), and this value was then used as the threshold for defining whether two prey were close enough to one another to be considered neighbours.

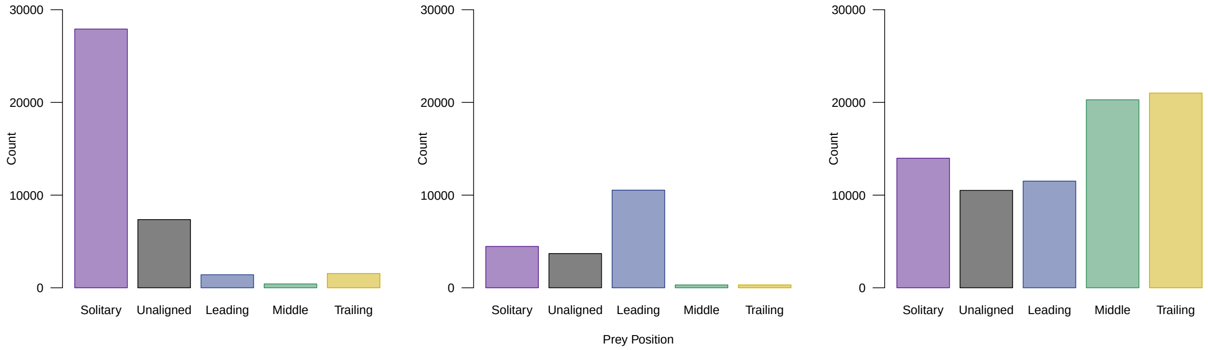

Fig. S4. The frequency of being in different relative social positions depending on prey type. Asocial types (left panel) are rarely found in aligned groups, and are mostly found alone. Leader types (middle panel) are most likely to be found leading other prey, but are also likely to be solitary or close to at least one other prey but not aligned. Follower types (right panel) are mostly likely found in the middle of a group or be the trailing individual in a group, although are also often found solitary, unaligned to others or leading. The common y axis reveals that a leading prey is just as likely to be a social prey type as it is a leading prey type, as there are four times as many social types as leader types. Middle and trailing prey are almost always the social prey type. Solitary individuals are most often asocial types, but often social types too (approximately a third of the time).

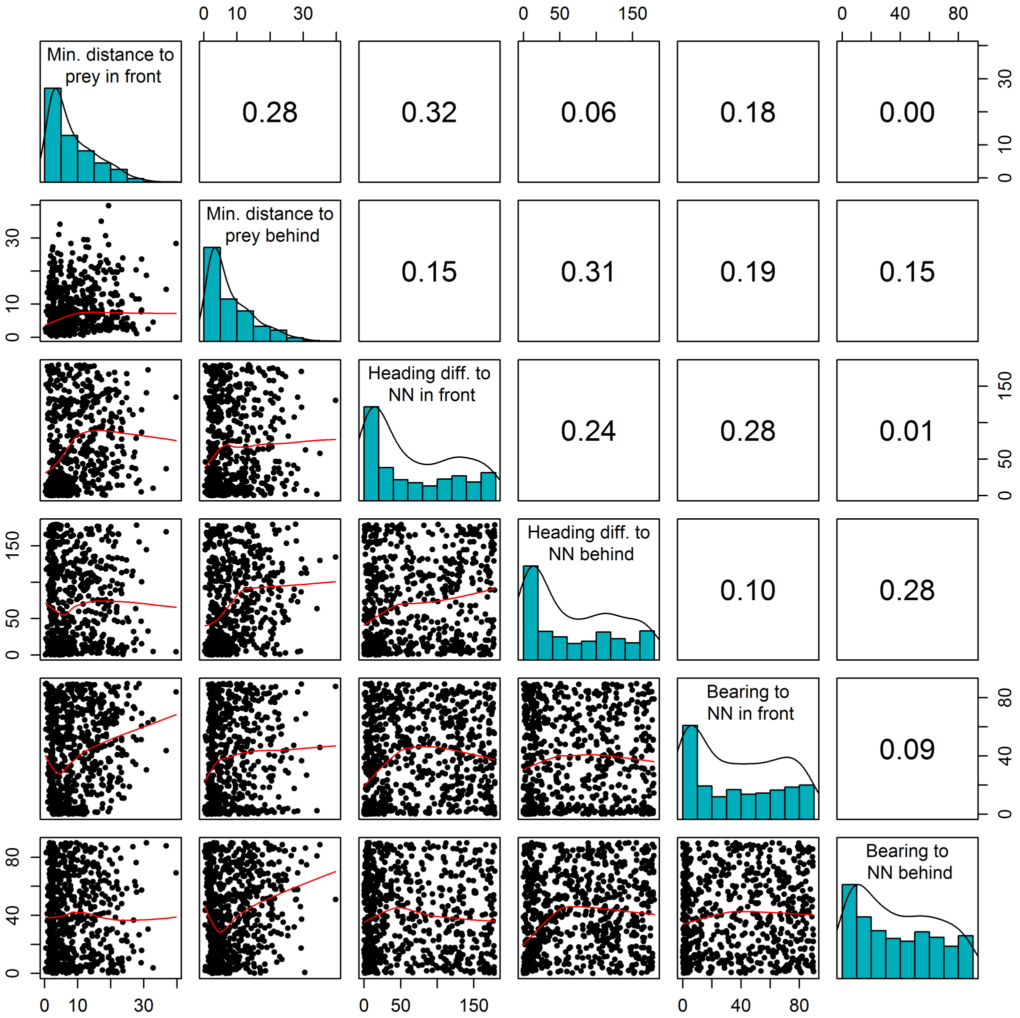

Fig. S5. Distributions of, and correlations between, the six continuous variables used to quantify each prey's social behaviour. Each variable (minimum distance, heading difference and bearing) are calculated separately for the nearest neighbour in front of the focal prey (determined from all prey with a bearing from the focal of $<90°$) and the nearest neighbour behind the focal (from all prey with a bearing from the focal of $>90°$). Correlation coefficients are Spearman's r_s. The red curve in the scatterplots are LOWESS smoothed curves. Note that the bearing for the nearest neighbour behind a prey individual is transformed so it ranges from zero (the nearest neighbour is directly behind) to 90 (the nearest neighbour is directly alongside). This matches the bearing for the nearest neighbour in front of the prey, where zero is the neighbour directly ahead, and 90° directly alongside the focal prey. The transformation is carried out so that these two bearing variables in the statistical models are not directly confounded to whether the bearing is for the prey in front (0-90°) or behind (90°-180° without the transformation).

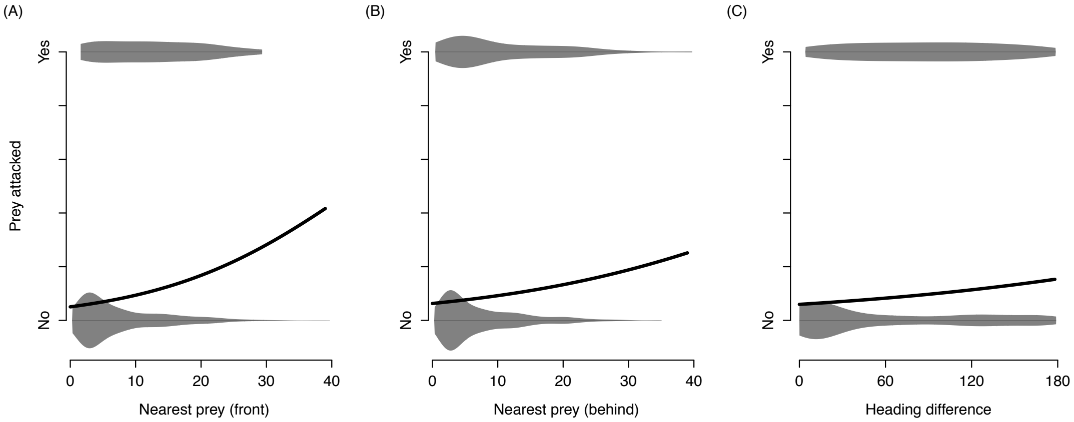

Fig. S6. The effect on per capita risk of distance to the nearest prey in front (A), nearest prey behind (B), and the heading difference to the nearest prey in front (C) of each prey. Fitted lines are calculated from the coefficients from the minimal adequate model which was the most likely model given the data from those included in the analysis (see SI Appendix, tables S2 and S3). In each panel, the other two variables in the model are fixed at their mean value. Also shown are the distributions of the data for prey that are attacked and not attacked.

Table S1. The difference in AICc between models explaining per capita risk for virtual prey including results from the randomisation tests.

Explanatory variable	Prey type binary: attraction or not	Prey type binary: can appear social or not	Prey type	Solitary or Grouping	Number of near neighbours	Total Group Size	Position
na (null model)	20.16	12.55	18.53	36.28	43.15	31.55	49.73
	-2.01	-2.01	-3.95	-2.00	-2.01	-2.01	-7.57
	3.98	3.91	4.38	3.01	2.06	1.34	3.43
Prey type binary: attraction or not		-7.61	-1.63	16.11	22.99	11.39	29.56
		-3.55	-2.01	-4.60	-4.62	-5.02	-8.44
		3.38	3.71	3.85	2.90	2.43	4.23
Prey type binary: can appear social or not			5.98	23.73	30.60	19.0	37.16
			-2.01	-4.5	-4.61	-4.76	-8.36
			3.86	3.71	2.79	2.26	4.33
Prey type				17.75	24.62	13.02	31.20
				-5.34	-5.41	-5.56	-8.96
				4.95	4.10	3.64	5.13
Solitary or Grouping					6.87	-4.73	13.45
					-2.64	-2.68	-5.83
					1.78	0.98	3.62
Number of Near Neighbours						-11.6	6.58
						-1.93	-6.28
						1.14	4.1
Total Group Size							18.18
							-5.97
							4.23

To test how prey type (asocial, leader or follower) and prey behaviour affected the risk of being attacked, we used binomial Generalised Linear Models (GLMs; the assumption that the dispersion parameters were approximately equal to 1 was met, i.e. between 0.5 and 2). Although prey at the time of each attack could be grouped by trial as a random factor, mixed models were not used as there was no variability between trials in the response variable because there was always one attacked prey and six unattacked prey. Note that all eight models are considered using the same approach; for example, we do not find a best fitting model and then compare this to a null model without any explanatory terms. To control for the variable frequencies of different prey behaviours, the randomisation test involved randomly shuffling which prey was attacked in each trial (there was still only one attacked prey each trial), and 10,000 randomised data sets were created. Using each of these randomised datasets, the GLMs were repeated, and the AICc from each model applied to each dataset was saved, as well as the fitted (i.e. predicted) values of per capita risk for the different prey positions (solitary, grouped but unaligned, leader, middle, or trailing) and the number of attacks on each of the prey positions. To test whether a model was more likely given the data than another model of the observed data, the values for the difference in the AICc between the two models in the observed data was compared to the 95% range of the null distribution, i.e. the 2.5% and 97.5% quantiles of the equivalent statistic from the randomisation test (the statistic was the difference in the AICc between the two models run on the same randomised data set). In each pairwise comparison between models, the first row is the observed difference in the AICc between the two models (the AICc of the model in the row - the AICc of the model in the column). Positive values indicate the model in the row is less likely given the data than the model in the column. The second and third rows in each comparison are the 2.5% and 97.5% percentiles (respectively) of the expected difference if targeting by the predators was random (but attacking at the same times in the simulation; see Materials and Methods). If the observed difference is outside of the 95% range, it is suggestive that one model is more likely than another, even after accounting for the non-independence in the data. Note that the 95% ranges always include 0, suggesting that any non-independence effect does not introduce a directional bias for any one model to be more likely than another. Similarly, to test whether differences in the predicted risk between the different prey positions in the models were likely to have occurred by chance, the differences in the predicted risk using the observed data were compared to the 95% range of the equivalent statistic from the randomisation tests (horizontal bars in Figure 3).

Table S2. The $\triangle$ AICc for models with continuous spatial and movement variables explaining per capita risk for virtual prey.

$\triangle$AICc	d.f	Model name	Explanatory variable(s)
0	4	MAM	Min. distance to prey in front, Min. distance to prey behind, Heading diff. to NN in front
4.8	7	Main effects	All main effects
8.9	2	Distance infront	Min. distance to prey in front
22.2	2	Heading diff. in front	Heading diff. to NN in front
22.9	2	Distance behind	Min. distance to prey behind
32.9	1	Null	na (null model)

The combination of binomial GLMs and the randomisation test was then used to assess the effects of, as continuous variables, the distance, heading difference and bearing of the nearest neighbour in front of, and separately behind, each prey (i.e. 6 explanatory variables). These variables were uncorrelated to one another (fig. S5). Cases where there was not a prey in front or behind a prey had to be excluded from the analysis due to missing data (261 out of 931 cases). The MAM (minimal adequate model) includes the three main effects identified as statistically significant ($P < 0.05$) in the main effects model based on Likelihood Ratio Tests. Models with only one of each of these three main effects were also included in the analysis to test which had the greatest effect on per capita risk, and whether any of these models were more likely than the MAM which included all three of these variables.

Table S3. The difference in AICc between models with continuous spatial and movement variables explaining per capita risk for virtual prey including results from the randomisation tests.

	Main effects	MAM	Distance infront	Distance behind	Heading diff. in front
Null model	28.14	32.93	24.06	10.00	10.74
	-85.11	-81.84	-79.52	-79.52	-79.44
	-38.78	-36.07	-34.34	-34.46	-34.42
Main effects		4.78	-4.08	-18.15	-17.40
		-2.92	-2.4	-2.31	-2.56
		5.9	9.36	9.33	9.35
MAM			-8.86	-22.93	-22.19
			-3.14	-3.26	-3.2
			3.99	3.9	3.99
Distance infront				-14.07	-13.32
				-4.33	-4.36
				4.24	4.08
Distance behind					0.74
					-4.25
					4.32

Layout as in SI Appendix, table S1

Movie S1. Clips from 4 trials where prey are attacked in different positions. The ‘Seed’ is the random seed set by the user to initiate the simulation. The ‘tick’ is the Netlogo term for the time step in the simulation.

Dataset S1. The Netlogo file for the virtual prey simulation used in the study. The Netlogo file’s code has been saved in .txt format.

Dataset S2. The data at the moment of each attack. The seed and time step (tick) are given for each attack, as well as the behavioural parameters and position (Saved X, Saved Y) for each prey and whether they were the attacked prey or not. NN and NND is the nearest neighbour and nearest neighbour distance, respectively. N near neighbours is the number of neighbours within the threshold distance around that prey.

Dataset S3. The data from the simulation sampled every 100 time steps. The seed and time step (tick) are given for each sample, as well as the behavioural parameters and position (Saved X, Saved Y) for each prey. NN and NND is the nearest neighbour and nearest neighbour distance, respectively. N near neighbours is the number of neighbours within the threshold distance around that prey.