

A common rule for decision making in animal collectives across species

Sara Arganda^{1,2}, Alfonso Pérez-Escudero², and Gonzalo G. de Polavieja³

Instituto Cajal, Consejo Superior de Investigaciones Científicas, 28002 Madrid, Spain

Edited by Simon A. Levin, Princeton University, Princeton, NJ, and approved November 2, 2012 (received for review June 21, 2012)

A diversity of decision-making systems has been observed in animal collectives. In some species, choices depend on the differences of the numbers of animals that have chosen each of the available options, whereas in other species on the relative differences (a behavior known as Weber's law), or follow more complex rules. We here show that this diversity of decision systems corresponds to a single rule of decision making in collectives. We first obtained a decision rule based on Bayesian estimation that uses the information provided by the behaviors of the other individuals to improve the estimation of the structure of the world. We then tested this rule in decision experiments using zebrafish (*Danio rerio*), and in existing rich datasets of argentine ants (*Linepithema humile*) and sticklebacks (*Gasterosteus aculeatus*), showing that a unified model across species can quantitatively explain the diversity of decision systems. Further, these results show that the different counting systems used by animals, including humans, can emerge from the common principle of using social information to make good decisions.

collective behavior | public information | probability matching

Sensory data always has some degree of ambiguity, so animals need to make decisions by estimating the properties of the environment from uncertain sensory data (1–5). This estimation has been shown to be close to optimal in many cases, making optimal Bayesian decision making a successful framework shared by behavioral, neurobiological, and psychological studies (1–7).

A richer scenario for decision making takes place when animals move in groups. In this case, the behaviors of other animals are an extra source of information (6–34). Animals of different species have been observed to incorporate this extra information in their decisions in different ways. Some species make decisions that can be explained using the differences of the numbers of animals taking each option (21, 22), others according to the relative differences (Weber's law) (23, 24) or using other rules (25–34). This diversity of decision schemes has translated into a diversity of models (21, 22, 24–34).

To search for a unified framework having the diversity of decision-making schemes as particular cases, we generalized Bayesian decision making to the case of animal collectives. Our previous attempt at building such a theory predicted that the only relevant social information is the difference of the numbers of individuals already choosing each available option, and not the numbers themselves or the relative differences (or Weber's law) (22). However, this theory was limited to the particular case in which only one of the options could be a good option (22). We have now generalized the theory, allowing all available options to be good or bad options. We found that this generalization explains the diversity of decision rules observed in collectives, maintaining the same conceptual and mathematical simplicity, and containing our previous theory as a particular case. We have tested the theory experimentally in decision experiments using zebrafish (*Danio rerio*), but to cover the diversity of decision systems, we have also tested it using rich datasets of decision making in argentine ants (*Linepithema humile*) (24) and three-spined sticklebacks (*Gasterosteus aculeatus*) (25, 26). We found a quantitative match between the theory and the different decision systems of these representative species.

Results

We studied how the behaviors of others should be taken into account to improve the estimations of the structure of the world and make decisions in animal collectives. For a situation with two identical options to choose from (Fig. 1A), we looked for the probability that one option, say x , is a good option given that n_x and n_y animals have already chosen options x and y , respectively. We used Bayesian theory to find an approximated analytic expression for this probability as (SI Text)

$$P(x \text{ is good}) = \frac{1}{1 + as^{-(n_x - kn_y)}}. \quad [1]$$

Parameter a measures the quality of nonsocial information available to the deciding individual, and s measures how reliably an individual that has chosen x indicates to the deciding individual that x is a good option. According to Eq. 1, the higher the number of individuals that chose option x , n_x , the higher the probability that option x is good for the deciding individual, and more so the higher the reliability s of the information from the individuals that already chose x . However, each individual that chooses y decreases the probability that x is a good option. Parameter k measures the relative impact of these two opposing effects. Individuals need to decide based on the estimated probabilities in Eq. 1. A common decision rule in animals, from insects to humans, is probability matching, according to which the probability of choosing a behavior is proportional to the estimated probability (35–44),

$$P_x = \frac{P(x \text{ is good})}{P(x \text{ is good}) + P(y \text{ is good})}. \quad [2]$$

This rule is known to be optimal when there is competition for resources (39, 40) and when the estimated probabilities change in time (41–44). Probability matching in Eq. 2, together with the estimation in Eq. 1, gives that the probability of choosing x is

$$P_x = \left(1 + \frac{1 + as^{-(n_x - kn_y)}}{1 + as^{-(n_y - kn_x)}} \right)^{-1}, \quad [3]$$

and $P_y = 1 - P_x$ is the probability of choosing y . The main implications of Eq. 3 are apparent in its plot (Fig. 1B). First, decision making in collectives is predicted to be different for low and high numbers of individuals. For low numbers, there is a fast

Author contributions: S.A., A.P.-E., and G.G.d.P. designed research, performed research, analyzed data, and wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

¹Present address: Centre de Recherches sur la Cognition Animale, Centre National de la Recherche Scientifique Unité Mixte de Recherche 5169, Université Paul Sabatier, 31062, Toulouse, France.

²S.A. and A.P.-E. contributed equally to this work.

³To whom correspondence should be addressed. E-mail: gonzalo.polavieja@cajal.csic.es.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1210664109/-DCSupplemental.

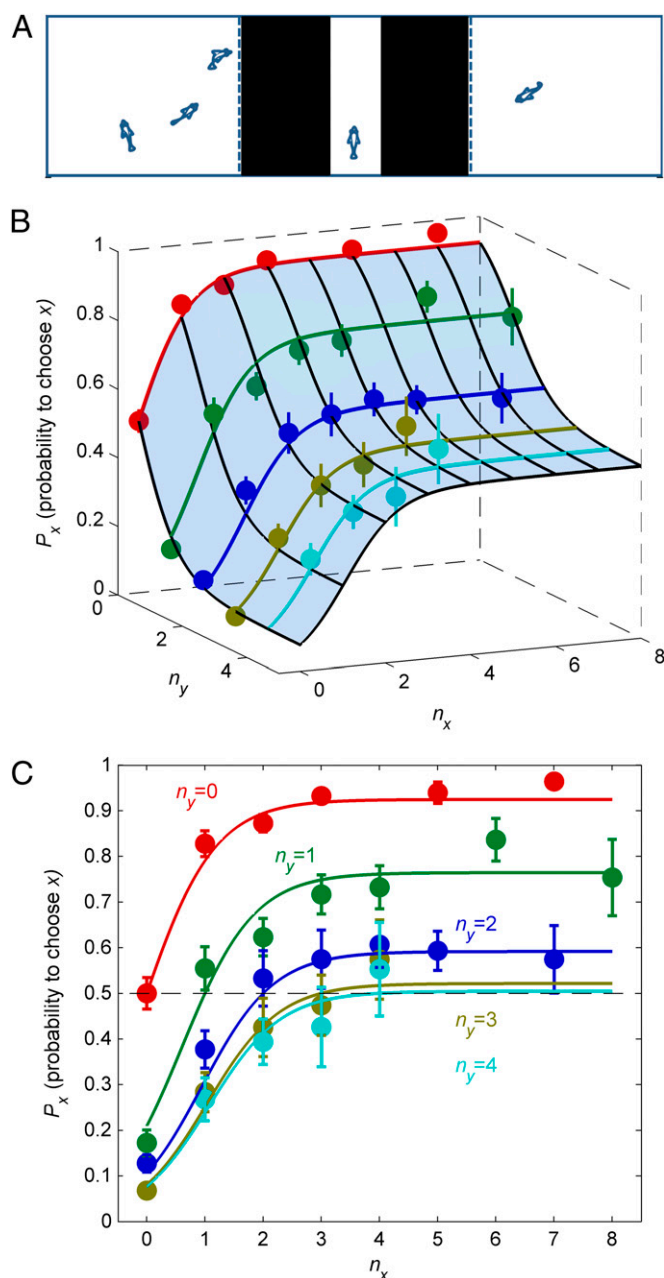


Fig. 2. Zebrafish choices correspond to the general rule of decisions in collectives. (A) Focal fish choosing between two sites with different number of zebrafish, separated from the focal fish by glass. (B) Probability of choosing option x for different numbers of zebrafish at sites x and y , n_x and n_y . Theoretical probabilities for $a = 11.2$ and $s = 5$, and $k = 0$ in Eq. 3 represented as a surface and experimental data represented as dots indicating the mean value of typically 15 animals at each configuration. Different dot colors correspond to different values of n_y and bars are SEM. (C) Same as B but plotted only as a function of n_x and different colors representing the value of n_y .

although at a slower rate, with how many have chosen option y . This situation might be common, for example, in food search. Animals choosing one option can indicate that there is a food source in that direction, but also that there might not be a food source at the other option. In this case, the probability of choosing x has a plateau in which both options are equally likely, but increasing the number of animals that have chosen x , n_x , reaches a transition region of rapid increase in probability (Fig. 1B). This transition region follows a straight line of slope k in the probability

plot (Fig. 1B). This line obeys for high number of animals that $n_y \approx k n_x$. This is a Weber law (23, 24), according to which the just-noticeable difference between two groups is proportional to the total number of individuals. Indeed, if we substitute $n_y \approx k n_x$ into $\Delta N/N \equiv (n_x - n_y)/(n_x + n_y)$ we obtain a constant of value $(k - 1)/(k + 1)$. A second prediction of the model is that decisions should deviate from Weber behavior at low numbers (below the transition point τ in Fig. 1B).

We have found that decisions made by the Argentine ant (*L. humile*) correspond to the case $0 < k < 1$. Ants' choices to turn left and right have been recorded by Perna et al. (24), and we found that they have choice probabilities well described by Eq. 3, except that experimental probabilities do not reach values as close to 0 or 1 as the theory. This difference might be due simply to the fact that ants are not always making turn decisions based on pheromones, but responding to other factors, such as roughness of terrain or collisions with other ants. We therefore considered that ants choose at random with a given probability and otherwise make a decision according to Eq. 3 (Eq. 4). This modification only introduces an overall rescaling in the probabilities, so all structural features described below are present in Eq. 3 (Fig. S5). We obtain a good correspondence with data for high (Fig. 3A) and low numbers of animals (Fig. 3B) with a fit that is robust (Fig. S6). The experimental data are smoother than the theory, without a central plateau, but still with a close correspondence, as also shown in the following analysis. According to Weber's law, isoproductivity curves should be horizontal lines in the $\Delta N/N \equiv (n_x - n_y)/(n_x + n_y)$ vs. $N \equiv n_x + n_y$ plane, and this is true both for the theory and experiments for high numbers of total animals N (Fig. 3C). The advantage of this plot is that it magnifies the region of low N , where the data deviate from Weber's law similarly to the theoretical prediction. A further quantitative analysis revealing the close correspondence between theory and data are shown in Fig. 3D. We performed a linear fit to the experimental probability along the lines of constant $n_x + n_y$ depicted in Fig. 3D Inset. The slope of each linear fit was then plotted against the total number of animals N (Fig. 3D, blue dots). The experimental data has a very close correspondence with the theoretical values in this plot (Fig. 3D, red line). For a high number of animals, both theory and data show Weber behavior, corresponding in this logarithmic plot to a straight line with slope -1 (Fig. 3D, black line) (24). Interestingly, for low numbers of animals, the theoretical prediction of a deviation from Weber behavior corresponds to the data.

The last case we consider has $k = 1$, for which Eq. 3 depends only on the variable $\Delta N \equiv n_x - n_y$. This situation could take place when there is a high probability that only one of the options is good, and those animals choosing x indicate that x may be the good one in a similar way that those choosing y may indicate that x might not be the good one. We have previously shown (22) that the simple decision rule $P_x = 1/(1 + a s^{-\Delta N})$ explains well a large dataset of collective decisions in sticklebacks, *G. aculeatus* (25, 26). In these experiments, animal groups were made to choose in two-choice setups with different combinations of social and nonsocial information (Fig. 4A, Far Left). Interestingly, Eq. 3 has the simple rule $P_x = 1/(1 + a s^{-\Delta N})$ as a particular case for $k = 1$ (SI Text). Indeed, all experimental results (blue histograms in Fig. 4A and Fig. S7) are fit using Eq. 3 with parameters $s = 2.5$, $k = 1$ (Fig. 4A, red lines). Additionally, for low numbers of animals (up to τ in Fig. 1B), an approximated ΔN rule can also be found for any value of k but with different values of the nonsocial reliability parameter a (SI Text). Therefore, the stickleback data can be fit with any value of k (green and blue lines in Fig. 4A and Fig. S7 for $k = 0.5$ and $k = 0$, respectively), with robust fits (Fig. S8). The reason why in this case k can have any value is that its main effect is to control the slope of the boundaries of the plateau of probability 0.5, which is not present in the experimentally explored region of the stickleback dataset (Fig. 4B, white triangle). Still, all these fits

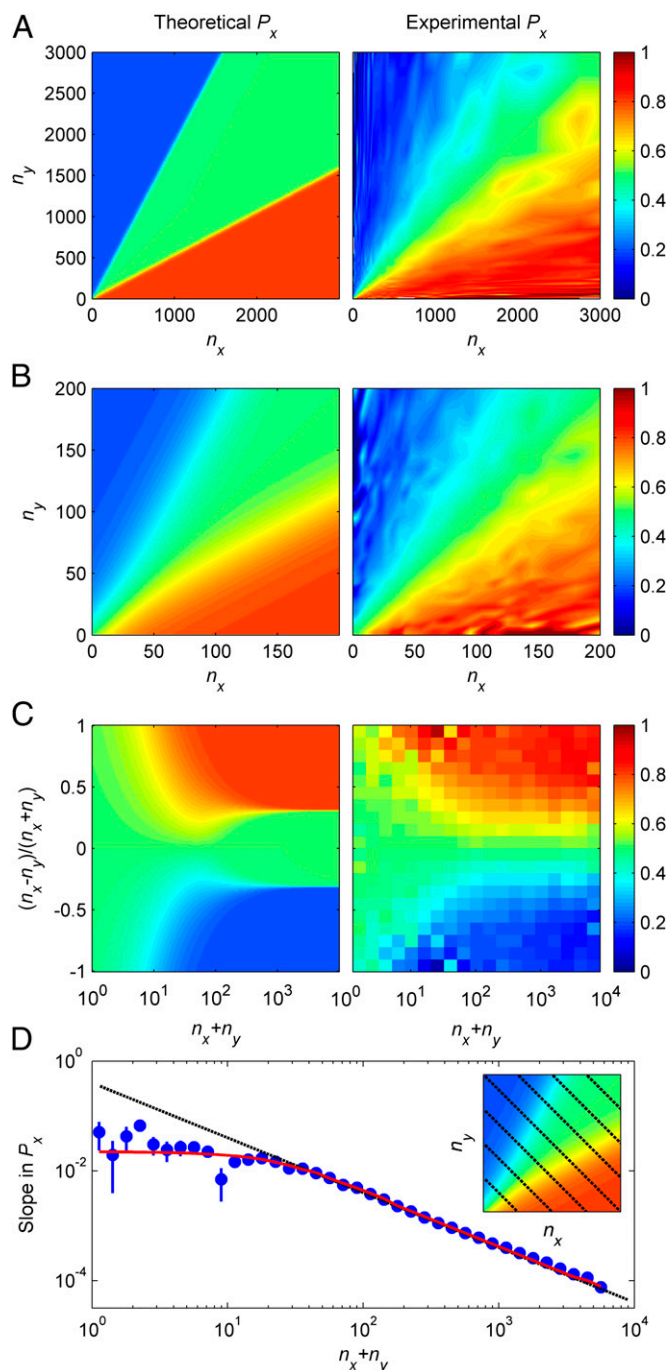


Fig. 3. Ant choices correspond to the general rule of decisions in collectives. (A) Probability of choosing option x as a function of how many ants have previously been at locations x and y , n_x and n_y , for theory (Left) using Eq. 4 with $a = 2.5$, $s = 1.07$, $k = 0.53$, $p_{\text{rand}} = 0.39$, and experiments (Right) from Perna et al. (24). (B) Detail of A. (C) Same as A but represented as a function of $\Delta N/N$ and N . (D) Slope of the probability of choosing x in A as obtained from a linear fit along the lines depicted in Inset. Experimental values (blue dots; error bars are 95% confidence interval), theory (red line), and Weber's law (black line).

have in common an effective ΔN rule for the experimental region (Fig. 4B), giving strong support to this rule in this dataset.

Discussion

Our results support that estimation by the brain using social information to counteract the ambiguity of sensory data is a fundamental principle in collective decision making. The theory explains

also the diversity in number discrimination schemes used in collective decisions, including counting up to a given number of animals, counting the difference of animals choosing among options, ΔN , or the relative difference, $\Delta N/N$, as well as observed deviations from these ideal cases and the existence of different counting regimes for high and low numbers as observed in many species, including humans (47, 48). A single mathematical rule contains all these cases and can be used as a first-principles approach to quantitatively study decisions in animal collectives.

One important ingredient of our theory is the use of probability matching (Eq. 2); for symmetric decisions, it implies a functional form of the type $P_x = f(x, y)/(f(x, y) + f(y, x))$. Our model in Eq. 3 is a particular case of this function, with $f(x, y)$ derived from an approximation to Bayesian estimation. Interestingly, many previous approaches derive from the form $P_x = f(x)/(f(x) + f(y))$ (21, 22, 27, 28), which is also a particular case of $P_x = f(x, y)/(f(x, y) + f(y, x))$, and therefore compatible with probability matching. In other cases, the basic form $P_x = f(x)/(f(x) + f(y))$ has been modified by adding constant terms (29, 30) or an extra function (25), as $P_x = f(x)/(f(x) + f(k))$, with k a constant when animals have access to a single choice (31, 32, 34). Weber behavior can also be seen as a particular case. It has been previously described using a function (24) that can be expressed as $f(x, y) = 1/2 + \delta(n_x - n_y)/(n_x + n_y)$, with δ between 0 and 1/2. This function obeys $f(x, y) + f(y, x) = 1$, so in this case $P_x = f(x, y)$, following Weber behavior.

These previous functions are very useful when applied to particular datasets because they may use few parameters in these conditions. In particular, our previous model (22), a particular case of Eq. 3 (SI Text), used only one parameter in the symmetric experiments with sticklebacks, and a model with two parameters described the ants dataset (24). However, these two models cannot fit the three datasets or even two of them (Fig. S9 A and D and S10). For the zebrafish data in Fig. 2, none of the previously proposed functions (21, 22, 24, 27–29) give a good fit of the plateau in the data (Fig. S9). Our approach has been developed to be applied in very different species and conditions, here tested for three large datasets in three different species. One important factor in this ability to describe different datasets is that our basic function $f(x, y)$ has a term $s^{-(n_x - k n_y)}$ that captures how the estimated quality of an option depends not only on the animals choosing that option but also on the animals choosing the other option. These two sources of information are balanced by parameter k , and different datasets are found to correspond to different balances k .

Previous functions describing ant foraging include a constant term that represents a threshold of pheromone concentration below which ants do not react (24, 27, 28). In this way, these functions can describe the deviation from Weber's law at low pheromone concentration (24). In our case, the theory naturally shows this behavior as one more particular case of the predicted difference between a low and high number of animals. Comparing the two approaches, it is interesting to consider that the behavior for low numbers that is predicted from estimation theory can be achieved in ants using a threshold of pheromone concentration.

An advantage of our approach is that the form of the function f is derived for any type of setup simply from estimation given non-social sensory data and the behaviors of others (SI Text). For example, we predict for a symmetric setup with N options a generalization of Eq. 1 of the form

$$P(x \text{ is good}) = \frac{1}{1 + a s^{-(n_x - k M)}},$$

with $M = \sum_{i \neq x}^N n_i$ the total number of animals choosing any option except x (see Eq. S10 for the more general case of asymmetric choices).

A further advantage is that the parameters a , s , and k are not only fitting parameters but have expressions, Eqs. S4, S9, and

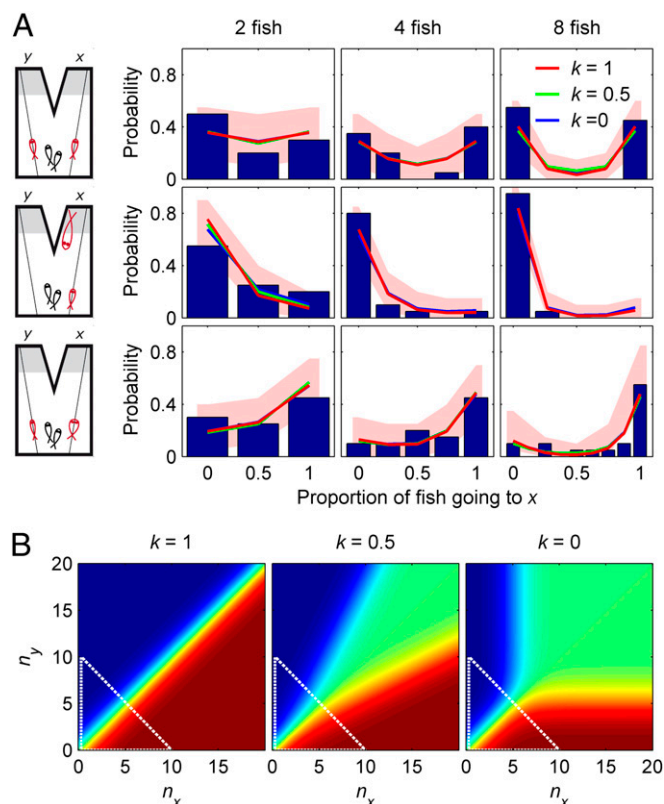


Fig. 4. Stickleback choices correspond to the general rule of decisions in collectives. (A) Probability of finding a final proportion of sticklebacks choosing option x (blue histograms are experimental results from refs. 25 and 26 and theoretical values as lines for $k = 1$, $k = 0.5$, and $k = 0$) for different group sizes (two, four, and eight fish) and for three types of setups: a symmetric setup with different numbers of replica fish going to x and y (Top), a setup with a replica predator at x and different replica fish going to x (Middle), and a symmetric setup with modified replica fish (Bottom). See model parameters and 68 additional experiments with fits in Fig. S7. (B) Theoretical P_x for $k = 1$, $a = 1$ (Left), $k = 0.5$, $a = 5$ (Center), and $k = 0$, $a = 224$ (Right), and $s = 2.5$ in the three cases. All models require an effective ΔN rule to compare with data for the number of animals used in experiments (triangle).

S16, respectively, which give additional predictions. For example, the social reliability parameter is given by

$$s = \frac{P(\beta|X, C)}{P(\beta|\bar{X}, C)},$$

with β a given animal behavior. This expression means that the social reliability parameter s is higher for a behavior β that is produced with high probability when x is a good option, and with very low probability when it is not a good option. Among all behaviors, those with higher s allow an individual to obtain a higher probability that option x is a good one (Eq. 1), so we expect them to have a larger effect on collective decision making.

Another advantage of an approach based on a theory of estimation is that generalizations of the theoretical expressions can be envisaged deriving models using fewer assumptions. For example, including dependencies in the behaviors of the other individuals and explicit space and time variables should be natural extensions of the theory.

Materials and Methods

Experimental Protocol for Zebrafish. All procedures met with European Economic Community (EEC) guidelines for animal experiments under Directive 86/609/EEC. Experimental procedures were approved by the Bioethics Sub-

committee of Consejo Superior de Investigaciones Científicas. We used wild-type adult zebrafish, *D. rerio*, of both sexes. Fish were acclimatized to the setup water for 1 d before the experiments (Fig. S11). At 1 h before the experiment, each fish was isolated and fed to ensure uniform nutritional status across individuals. A focal fish entered the setup and swam freely in a central chamber between two social chambers with different number of fish and separated from the choice chamber by glass. Once a fish had been recorded for 5 min it could be placed in one of the lateral chambers as a social stimulus for another fish. The fish in the lateral chambers were interchanged between trials to ensure uniformity, and sides were randomized. The central chamber of the setup was washed between trials to remove odor traces. We computed the probability P_x as the fraction of time the focal fish spent on the black region close to one of the social chambers, x . This fraction of time converges to P_x for a fish that makes repeated decisions choosing x (y) with probability P_x ($P_y = 1 - P_x$). A total of 238 fish were tested only once. To test the effect of previous experience, another 233 trials were performed with fish that were tested several times. We found no significant difference between the two groups in the mean times spent at each side (Fig. S12), so all data were pooled for Fig. 2.

Model with Noise Added to the Decision Rule. The model in Eq. 3 has a good agreement with data from experiments using the Argentine ant, *L. humile* (24), except that experimental probabilities do not reach values as close to 0 or 1 as the theory. To account for the experimental data, we made a simple modification of the model by assuming that the ant has some probability p_{rand} of making the decision at random motivated by unknown factors. Then, with probability $(1 - p_{\text{rand}})$, the ant makes the decision according to Eq. 3. Therefore, the probability of turning toward x is

$$P_x = \frac{p_{\text{rand}}}{2} + (1 - p_{\text{rand}}) \left(1 + \frac{1 + a s^{-(n_x - k n_y)}}{1 + a s^{-(n_y - k n_x)}} \right)^{-1}. \quad [4]$$

The parameters that best fit the ant data are $a = 2.5$, $s = 1.07$, $k = 0.53$, and $p_{\text{rand}} = 0.39$. This same model can be applied to the zebrafish and stickleback datasets, but in these cases the best fit is obtained for $p_{\text{rand}} \approx 0$, which corresponds to Eq. 3.

Analysis of the Ants Dataset. Both the raw dataset and preprocessing routines were provided by Perna et al. (24), and we used their data assuming no evaporation of pheromone (this assumption does not change the results significantly). We calculated from the data the probability of turning right or left, not a continuous angle, to compare directly to our predicted probabilities. To reduce the noise in the experimental maps of Fig. 3, we symmetrized the data so that the probability shown at point (n_x, n_y) is obtained as $(P_x(n_x, n_y) + (1 - P_x(n_y, n_x)))/2$.

Experimental data from Perna et al. (24) measures a quantity that is proportional to the number of ants previously at the left/right of the deciding ant, not directly the numbers, so the number of ants (n_x, n_y) used in the plots are related to the actual number of ants that count for the decision $(n_{x, \text{true}}, n_{y, \text{true}})$ by $n_x = \lambda n_{x, \text{true}}$, $n_y = \lambda n_{y, \text{true}}$, where λ is an unknown proportionality constant. This relation means that the model still applies but with $s = s_{\text{true}}/\lambda$, where s_{true} is the actual value of the reliability parameter.

Asymmetric models. For the case of sticklebacks deciding in the presence of a predator (Fig. 4A, Middle), we derived the more general asymmetric version of Eq. 1, Eq. S22 (Figs. S13 and S14). We also derived Eq. S23 for the case with two different types of replicas (Fig. 4A, Bottom).

Fitting Procedures. To fit the model's parameters to the data, we performed 2D exhaustive searches in the space of parameters. For functions with more than two parameters, we performed the search successively with all possible pairs of parameters. In these cases we repeated the fit several times starting from different initial conditions, always getting the same final result.

ACKNOWLEDGMENTS. We thank Andrea Perna for sharing the raw ant data before publication; David Sumpter, Jacques Gautrais, Raphael Jeanson, Matthieu Moreau, and members of G.G.d.P.'s laboratory for fruitful discussions; and two reviewers for constructive comments. Funding for this work was provided by the Ministerio de Ciencia e Innovación (MICINN, Spain) as Plan Nacional and as partners of the ERASysBio+ initiative (G.G.d.P.) supported under the European Union European Research Area Networks (ERA-NET) Plus scheme in Framework Program 7, including contracts to S.A. and A.P.E. A.P.E. is the recipient of a Formación de Profesorado Universitario (FPU) fellowship from MICINN (Spain).

1. Knill DC, Pouget A (2004) The Bayesian brain: The role of uncertainty in neural coding and computation. *Trends Neurosci* 27(12):712–719.
2. Gold JJ, Shadlen MN (2007) The neural basis of decision making. *Annu Rev Neurosci* 30:535–574.
3. Tenenbaum JB, Kemp C, Griffiths TL, Goodman ND (2011) How to grow a mind: Statistics, structure, and abstraction. *Science* 331(6022):1279–1285.
4. Biernaskie JM, Walker SC, Gegebar RJ (2009) Bumblebees learn to forage like Bayesians. *Am Nat* 174(3):413–423.
5. McNamara JM, Green RF, Olsson O (2006) Bayes theorem and its applications in animal behaviour. *Oikos* 112(2):243–251.
6. Banerjee AV (1992) A simple model of herd behavior. *Q J Econ* 107(3):797–817.
7. Bikhchandani S, Hirshleifer D, Welch I (1992) A theory of fads, fashion, custom, and cultural change as informational cascades. *J Polit Econ* 100(5):992–1026.
8. Valone TJ, Templeton JJ (2002) Public information for the assessment of quality: A widespread social phenomenon. *Philos Trans R Soc Lond B Biol Sci* 357(1427):1549–1557.
9. Giraldeau LA, Valone TJ, Templeton JJ (2002) Potential disadvantages of using socially acquired information. *Philos Trans R Soc Lond B Biol Sci* 357(1427):1559–1566.
10. Wagner RH, Danchin E (2010) A taxonomy of biological information. *Oikos* 119(2):203–209.
11. Clark C, Mangel M (1986) The evolutionary advantages of group foraging. *Theor Popul Biol* 30(1):45–75.
12. Coolen I, van Bergen Y, Day RL, Laland KN (2003) Species difference in adaptive use of public information in sticklebacks. *Proc Biol Sci* 270(1531):2413–2419.
13. van Bergen Y, Coolen I, Laland KN (2004) Nine-spined sticklebacks exploit the most reliable source when public and private information conflict. *Proc Biol Sci* 271(1542):957–962.
14. Rieucou G, Giraldeau LA (2009) Persuasive companions can be wrong: The use of misleading social information in nutmeg mannikins. *Behav Ecol* 20(6):1217–1222.
15. Lima SL (1995) Collective detection of predatory attack by social foragers: Fraught with ambiguity? *Anim Behav* 50(4):1097–1108.
16. Nordell SE, Valone TJ (1998) Mate choice copying as public information. *Ecol Lett* 1(2):74–76.
17. Magurran AE, Higham A (2010) Information transfer across fish shoals under predation threat. *Ethology* 78(2):153–158.
18. Danchin E, Giraldeau LA, Valone TJ, Wagner RH (2004) Public information: From nosy neighbors to cultural evolution. *Science* 305(5683):487–491.
19. Dall SRX, Giraldeau LA, Olsson O, McNamara JM, Stephens DW (2005) Information and its use by animals in evolutionary ecology. *Trends Ecol Evol* 20(4):187–193.
20. Dugatkin LA (1992) Sexual selection and imitation: Females copy the mate choice of others. *Am Nat* 139(6):1384–1389.
21. Ward AJW, Herbert-Read JE, Sumpter DJT, Krause J (2011) Fast and accurate decisions through collective vigilance in fish shoals. *Proc Natl Acad Sci USA* 108(6):2312–2315.
22. Pérez-Escudero A, de Polavieja GG (2011) Collective animal behavior from Bayesian estimation and probability matching. *PLOS Comput Biol* 7(11):e1002282.
23. Gómez-Laplaza LM, Gerlai R (2011) Can angelfish (*Pterophyllum scalare*) count? Discrimination between different shoal sizes follows Weber's law. *Anim Cogn* 14(1):1–9.
24. Perna A, et al. (2012) Individual rules for trail pattern formation in Argentine ants (*Linepithema humile*). *PLOS Comput Biol* 8(7):e1002592.
25. Ward AJW, Sumpter DJT, Couzin ID, Hart PJB, Krause J (2008) Quorum decision-making facilitates information transfer in fish shoals. *Proc Natl Acad Sci USA* 105(19):6948–6953.
26. Sumpter DJT, Krause J, James R, Couzin ID, Ward AJW (2008) Consensus decision making by fish. *Curr Biol* 18(22):1773–1777.
27. Goss S, Aron S, Deneubourg JL, Pasteels JM (1989) Self-organized shortcuts in the Argentine ant. *Naturwissenschaften* 76(12):579–581.
28. Deneubourg J, Aron S, Goss S, Pasteels J (1990) The self-organizing exploratory pattern of the Argentine ant. *J Insect Behav* 3(2):159–168.
29. Meunier H, Leca JB, Deneubourg JL, Petit O (2006) Group movement decisions in capuchin monkeys: The utility of an experimental study and a mathematical model to explore the relationship between individual and collective behaviours. *Behaviour* 143(12):1511–1527.
30. Tero A, Kobayashi R, Nakagaki T (2007) A mathematical model for adaptive transport network in path finding by true slime mold. *J Theor Biol* 244(4):553–564.
31. Gallup AC, et al. (2012) Visual attention and the acquisition of information in human crowds. *Proc Natl Acad Sci USA* 109(19):7245–7250.
32. Sumpter DJT, Pratt SC (2009) Quorum responses and consensus decision making. *Philos Trans R Soc Lond B Biol Sci* 364(1518):743–753.
33. Jeanson R, et al. (2005) Self-organized aggregation in cockroaches. *Anim Behav* 69(1):169–180.
34. Bousquet CAH, Sumpter DJT, Manser MB (2011) Moving calls: A vocal mechanism underlying quorum decisions in cohesive groups. *Proc R Soc B Biol Sci* 278(1711):1482–1488.
35. Behrend ER, Bitterman ME (1961) Probability-matching in the fish. *Am J Psychol* 74(4):542–551.
36. Greggers U, Menzel R (1993) Memory dynamics and foraging strategies of honeybees. *Behav Ecol Sociobiol* 32(1):17–29.
37. Kirk KL, Bitterman ME (1965) Probability-learning by the turtle. *Science* 148(3676):1484–1485.
38. Wozny DR, Beierholm UR, Shams L (2010) Probability matching as a computational strategy used in perception. *PLOS Comput Biol* 6(8):7.
39. Fretwell S, Lucas H (1969) On territorial behavior and other factors influencing habitat distribution in birds. *Acta Biotheor* 19(1):16–36.
40. Houston A, McNamara J (1987) Switching between resources and the ideal free distribution. *Anim Behav* 35(1):301–302.
41. Vulkan N (2000) An economist's perspective on probability matching. *J Econ Surv* 14(1):101–118.
42. Staddon J (1983) *Adaptive Behavior and Learning* (Cambridge Univ Press, Cambridge, UK).
43. Gaissmaier W, Schooler LJ (2008) The smart potential behind probability matching. *Cognition* 109(3):416–422.
44. Green CS, Benson C, Kersten D, Schrater P (2010) Alterations in choice behavior by manipulations of world model. *Proc Natl Acad Sci USA* 107(37):16401–16406.
45. Engeszer RE, Ryan MJ, Parichy DM (2004) Learned social preference in zebrafish. *Curr Biol* 14(10):881–884.
46. Wright D, Krause J (2006) Repeated measures of shoaling tendency in zebrafish (*Danio rerio*) and other small teleost fishes. *Nat Protoc* 1(4):1828–1831.
47. Feigenson L, Dehaene S, Spelke E (2004) Core systems of number. *Trends Cogn Sci* 8(7):307–314.
48. Dehaene S, Izard V, Spelke E, Pica P (2008) Log or linear? Distinct intuitions of the number scale in Western and Amazonian indigene cultures. *Science* 320(5880):1217–1220.

Supporting Information

Arganda et al. 10.1073/pnas.1210664109

SI Text

Here we give the derivation of Eq. 1, the derivation of a more general equation for an asymmetric setup (Fig. 4A, *Middle* row), and the derivation of an equation for a symmetric setup but with different types of animals to follow (Fig. 4A, *Bottom* row). Additionally, we show that the model in ref. 1 is a particular case of Eq. 3. We also present the derivation of an expression for the point τ separating the low-number and high-number decision behaviors, and proof of the approximate ΔN rule for low N .

Derivation of Eq. 1. The following derivation follows similar steps to our derivation in ref. 1, except for the key difference that animals now estimate the probability that different options are good instead of the best. This simple difference makes the theory more general with previous results in ref. 1 only a particular case, as shown in a subsequent section.

Consider a focal individual making a decision among several options (x, y, z, \dots). To make this decision, it estimates the probability that each option is a good choice. “Good” may refer to the presence of food, shelter, absence of predators, or any other feature. To perform this estimation, the individual uses the information of the environment gathered directly by its sensors (nonsocial information, C), and the behaviors of the other individuals (social information, B). The probability that a given option (say, option x) is a good choice, given both nonsocial and social information is

$$P(X|C, B), \quad [S1]$$

where X stands for “ x is a good choice.” We can compute this probability using Bayes’ theorem,

$$P(X|C, B) = \frac{P(B|X, C)P(X|C)}{P(B|X, C)P(X|C) + P(B|\bar{X}, C)P(\bar{X}|C)}, \quad [S2]$$

where \bar{X} stands for “ x is not a good choice.” Dividing the numerator and denominator of Eq. S2 by the numerator, we get

$$P(X|C, B) = \frac{1}{1 + a_x S_x}, \quad [S3]$$

with

$$a_x = \frac{P(\bar{X}|C)}{P(X|C)} \quad [S4]$$

and

$$S_x = \frac{P(B|\bar{X}, C)}{P(B|X, C)}, \quad [S5]$$

where we use the subindex x to indicate that it refers to the estimation for option x . Each of the options has a set of equations like Eqs. S3–S5. Note that a_x only contains nonsocial information (C), so we call it a nonsocial term, whereas the social information (B) is contained in the social term, S_x . A practical version of Eq. S3 is obtained using the approximation that the focal individual does not take into account the correlations among the rest of individuals (however, see ref. 1 for a treatment of these correlations). This assumption implies

that the probability of a given set of behaviors is equal to the product of the probabilities of individual behaviors. We apply this to the probabilities needed to compute S_x in Eq. S5,

$$P(B|X, C) = Z \prod_{i=1}^N P(b_i|X, C), \quad [S6]$$

where B is the set of behaviors of the other N animals at the time the focal individual is choosing, $B = \{b_i\}_{i=1}^N$, and b_i denotes the behavior of individual i . Z is a combinatorial term counting the number of possible decision sequences leading to the set of behaviors B , that will cancel out below. Substituting Eq. S6, and an analogous expression for $P(B|\bar{X}, C)$, into Eq. S5, we get

$$S_x = \prod_{i=1}^N \frac{P(b_i|\bar{X}, C)}{P(b_i|X, C)}. \quad [S7]$$

A more useful expression is obtained if we consider, instead of the full individual behaviors (b_i) with all their details, a set of behavioral classes that group together the behaviors that contain similar information about the choice. For example, in a two-choice setup, useful behavioral classes might be “choosing x ” (denoted as β_x) and “choosing y ” (β_y). Consider in general L behavioral classes $\{\beta_j\}_{j=1}^L$. We do not here consider animals to have individual differences, so all have the same probabilities for each behavior; for example, the same $P(\beta_1|X, C)$ and $P(\beta_1|\bar{X}, C)$ for behavior β_1 , which means that if the first n_1 individuals are performing behavior β_1 , we have $\prod_{i=1}^{n_1} \frac{P(b_i|\bar{X}, C)}{P(b_i|X, C)} = \left(\frac{P(\beta_1|\bar{X}, C)}{P(\beta_1|X, C)} \right)^{n_1}$. We can then write Eq. S7 as

$$S_x = \prod_{j=1}^L s_{xj}^{-n_j}, \quad [S8]$$

where n_j is the number of individuals performing behavior β_j , and

$$s_{xj} = \frac{P(\beta_j|\bar{X}, C)}{P(\beta_j|X, C)}. \quad [S9]$$

To summarize, the probability that option x is a good choice is, using Eqs. S3 and S8,

$$P(X|C, B) = \left(1 + a_x \prod_{j=1}^L s_{xj}^{-n_j} \right)^{-1}, \quad [S10]$$

with a_x in Eq. S4 and s_{xj} in Eq. S9.

The zebrafish experiments in the main text were performed in a setup with two identical sites to choose from, except for the number of animals at each site, n_x and n_y . The focal animal can observe two types of behaviors: stay at x (β_x) and stay at y (β_y). Eq. S10 then reduces to

$$P(X|C, B) = \frac{1}{1 + a_x s_{xx}^{-n_x} s_{xy}^{-n_y}}. \quad [S11]$$

Similarly, for option y the estimation is

$$P(Y|C, B) = \frac{1}{1 + a_y s_{yy}^{-n_y} s_{yx}^{-n_x}}. \quad [\text{S12}]$$

The nonsocial information for the two sites x and y is identical by experimental design, so

$$\begin{aligned} P(X|C) &= P(Y|C) \\ P(\bar{X}|C) &= P(\bar{Y}|C). \end{aligned} \quad [\text{S13}]$$

The relations in Eq. S13, together with Eq. S4, mean that $a_x = a_y$. For notational simplicity, we then define

$$a \equiv a_x = a_y. \quad [\text{S14}]$$

The symmetry of the setup also implies the following relations

$$\begin{aligned} P(\beta_x|X, C) &= P(\beta_y|Y, C) \\ P(\beta_x|\bar{X}, C) &= P(\beta_y|\bar{Y}, C) \\ P(\beta_x|Y, C) &= P(\beta_y|X, C) \\ P(\beta_x|\bar{Y}, C) &= P(\beta_y|\bar{X}, C) \end{aligned} \quad [\text{S15}]$$

In an idealized situation in which the only possible behaviors were “stay at x ” and “stay at y ,” we would have that $P(\beta_x|X, C) = 1 - P(\beta_y|X, C)$. Because real behaviors are much more complex, and different behaviors can exist, these two probabilities will not sum 1 in general.

According to Eqs. S15 and S9, we have that $s_{xx} = s_{yy}$ and $s_{xy} = s_{yx}$. It is then useful to define

$$\begin{aligned} s &\equiv s_{xx} = s_{yy} \\ k &\equiv -\frac{\log(s_{xy})}{\log(s_{xx})} = -\frac{\log(s_{yx})}{\log(s_{yy})}. \end{aligned} \quad [\text{S16}]$$

Using Eqs. S14 and S16, we can write Eqs. S11 and S12 as

$$\begin{aligned} P(X|C, B) &= \frac{1}{1 + a s^{-(n_x - k n_y)}} \\ P(Y|C, B) &= \frac{1}{1 + a s^{-(n_y - k n_x)}} \end{aligned} \quad [\text{S17}]$$

obtaining Eq. 1. Note that $s = P(\beta_x|X, C)/P(\beta_x|\bar{X}, C) = P(\beta_y|Y, C)/P(\beta_y|\bar{Y}, C)$, that is, the probability of choosing one option when it is a good choice over the probability of choosing it when it is a bad choice. Therefore, parameter s measures how reliable are the choices of each of the other individuals.

The probability of choosing x or y is then obtained using probability matching, Eq. 2 to give Eq. 3,

$$P_x = \left(1 + \frac{1 + a s^{-(n_x - k n_y)}}{1 + a s^{-(n_y - k n_x)}} \right)^{-1}. \quad [\text{S18}]$$

Derivation of a More General Equation for an Asymmetric Setup. In the case of an asymmetric setup (as in Fig. 4A, *Middle* row), the nonsocial information for the two sites x and y is different, so

$$\begin{aligned} P(X|C) &\neq P(Y|C) \\ P(\bar{X}|C) &\neq P(\bar{Y}|C). \end{aligned} \quad [\text{S19}]$$

The relations in Eq. S19 mean that $a_x \neq a_y$, as is clear from its definition in Eq. S4.

In the symmetric case, we used the relations in Eq. S15,

$$\begin{aligned} P(\beta_x|X, C) &= P(\beta_y|Y, C) \\ P(\beta_x|\bar{X}, C) &= P(\beta_y|\bar{Y}, C) \\ P(\beta_x|Y, C) &= P(\beta_y|X, C) \\ P(\beta_x|\bar{Y}, C) &= P(\beta_y|\bar{X}, C) \end{aligned} \quad [\text{S20}]$$

Because the nonsocial asymmetry can modulate the probabilities for the behaviors, these relations need not be satisfied exactly. However, this effect is probably much weaker than the effect of the nonsocial asymmetry on the nonsocial term in Eq. S19. Therefore, for simplicity we use relations (Eq. S20) also for the asymmetric setup. The good fit with experimental data confirms that they are a good approximation.

According to Eqs. S20 and S9, we have that $s_{xx} = s_{yy}$ and $s_{xy} = s_{yx}$, and using the definitions in Eq. S16, we find that Eqs. S11 and S12 become

$$\begin{aligned} P(X|C, B) &= \frac{1}{1 + a_x s^{-(n_x - k n_y)}} \\ P(Y|C, B) &= \frac{1}{1 + a_y s^{-(n_y - k n_x)}} \end{aligned} \quad [\text{S21}]$$

The probability of choosing x or y is then obtained using probability matching (Eq. 2) to get

$$P_x = \left(1 + \frac{1 + a_x s^{-(n_x - k n_y)}}{1 + a_y s^{-(n_y - k n_x)}} \right)^{-1}, \quad [\text{S22}]$$

represented in Fig. S13.

Derivation of an Equation for a Symmetric Setup but with Different Types of Animals to Follow. When there are different types of animals to follow, as in Fig. 4A *Bottom* row, following the steps of *Derivation of Eq. 1*, we find that each type of animal has its own reliability s . For the particular case of the experiment of ref. 2, we have three different types of animals (real animals, the most attractive replica, and the less-attractive replica, with reliability parameters s , s_R , and s_r , respectively). When the most attractive replica goes to x and the less attractive one goes to y , Eq. S18 becomes

$$P_x = \left(1 + \frac{1 + a s^{-(n_x - k n_y)} s_R^{-1} s_r^k}{1 + a s^{-(n_y - k n_x)} s_R^k s_r^{-1}} \right)^{-1}. \quad [\text{S23}]$$

Demonstration That the Model in Ref. 1 Is a Particular Case of Eq. 3. The decision-making model we used in ref. 1 was developed for a case in which an animal has to choose using the probability that an option is the best one, whereas the model in this paper is for estimated good options. In ref. 1 we obtained that the probability of choosing x in a two-choice setup that can present an asymmetry as

$$P_x = \left(1 + a_{\text{old}} s^{-(n_x - n_y)} \right)^{-1}, \quad [\text{S24}]$$

with $a_{\text{old}} = 1$ for the symmetric case.

Multiplying and dividing inside the brackets of Eq. S24 by $\left(1 + \frac{1}{a_{\text{old}}} s^{-(n_y - n_x)} \right)$, we rewrite this expression as

$$P_x = \left(1 + \frac{1 + a_{\text{old}} s^{-(n_x - n_y)}}{1 + a_{\text{old}}^{-1} s^{-(n_y - n_x)}} \right)^{-1}, \quad [\text{S25}]$$

so that Eq. S22 reduces to Eq. S24 for

$$k = 1$$

$$a_x = a_y^{-1} = a_{\text{old}}, \quad [\text{S26}]$$

$$\tau_y = \frac{k \log(a_x) + \log(a_y)}{(1-k^2) \log(s)}.$$

as we wanted to demonstrate.

Derivation of an Expression for the Point τ Separating the Low-Number and High-Number Decision Behaviors, and Proof of the Approximate ΔN Rule for Low N . We now consider the general expression of the probability (Eq. S22)

$$P_x = \left(1 + \frac{1 + a_x s^{-(n_x - k n_y)}}{1 + a_y s^{-(n_y - k n_x)}} \right)^{-1}. \quad [\text{S27}]$$

For the reasons described below, the transition between the two regimes takes place when the following conditions are met

$$a_x s^{-(n_x - k n_y)} = 1 \text{ and } a_y s^{-(n_y - k n_x)} = 1. \quad [\text{S28}]$$

These conditions define a point (τ_x, τ_y) with (Fig. S13):

$$\tau_x = \frac{\log(a_x) + k \log(a_y)}{(1-k^2) \log(s)} \quad [\text{S29}]$$

This transition point is relevant because when the left-side terms of Eq. S28 are much lower than 1 they can be neglected, so P_x is always 0.5. Therefore, the region above the transition point (τ_x, τ_y) in which both left-side terms of Eq. S28 are lower than 1 (region 1 in Fig. S14) is the plateau of $P_x = 0.5$.

However, if the two left-side terms of Eq. S28 are much higher than 1, we can use the approximations $1 + a_x s^{-(n_x - k n_y)} \approx a_x s^{-(n_x - k n_y)}$ and $1 + a_y s^{-(n_y - k n_x)} \approx a_y s^{-(n_y - k n_x)}$ to write Eq. S27 as

$$P_x \approx \left(1 + a_x / a_y s^{-\Delta N(1+k)} \right)^{-1}, \quad [\text{S30}]$$

which only depends on ΔN . Therefore, the region below the transition point (τ_x, τ_y) in which both left-side terms of Eq. S28 are higher than 1 (region 2 in Fig. S14) corresponds to a ΔN rule for decision making.

For the case of symmetric nonsocial information, in which $a \equiv a_x = a_y$, Eq. S29 reduces to

$$\tau \equiv \tau_x = \tau_y = \frac{\log(a)}{(1-k) \log(s)}. \quad [\text{S31}]$$

1. Pérez-Escudero A, de Polavieja GG (2011) Collective animal behavior from Bayesian estimation and probability matching. *PLoS Comput Biol* 7(11):e1002282.

2. Sumpter DJT, Krause J, James R, Couzin ID, Ward AJW (2008) Consensus decision making by fish. *Curr Biol* 18(22):1773–1777.

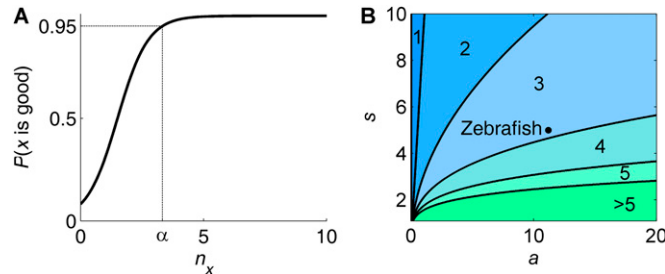


Fig. S1. Maximum number of individuals (α) that need to be counted according to the model for $k = 0$. (A) Probability that option x is good (Eq. 1) here plotted for parameters $a = 11.2$, $s = 5$, and $k = 0$. For $k = 0$, this probability only depends on the variable n_x , increasing as n_x increases to a value of 1. We compute α as the value of n_x for which the probability in Eq. 1 reaches 0.95, getting $\alpha = (\log(a) + \log(1/0.95 - 1))/\log(s)$. Because for $k = 0$ the probability to choose x (Eq. 3) only depends on n_x through $P(x \text{ is good})$, to make the decision the animals do not need to keep count of n_x beyond α . (B) The number up to which animals need to count, α , as a function of parameters a and s . For the parameters of the zebrafish dataset, animals only need to count up to 3.

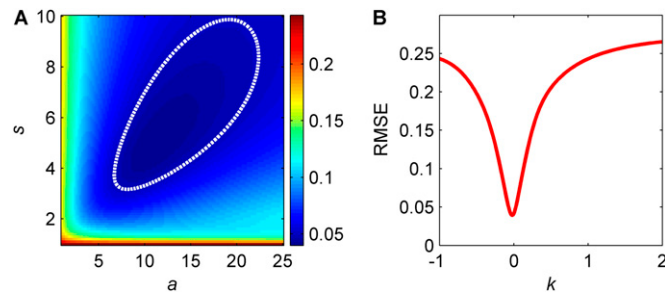


Fig. S2. Robustness of fit to zebrafish data. (A) Root mean-squared error between model predictions and data as a function of a and s ($k = 0$). The dotted line limits the region with error below 0.05. (B) Root mean-squared error between model and data as a function of k ($a = 11.2$, $s = 5$).

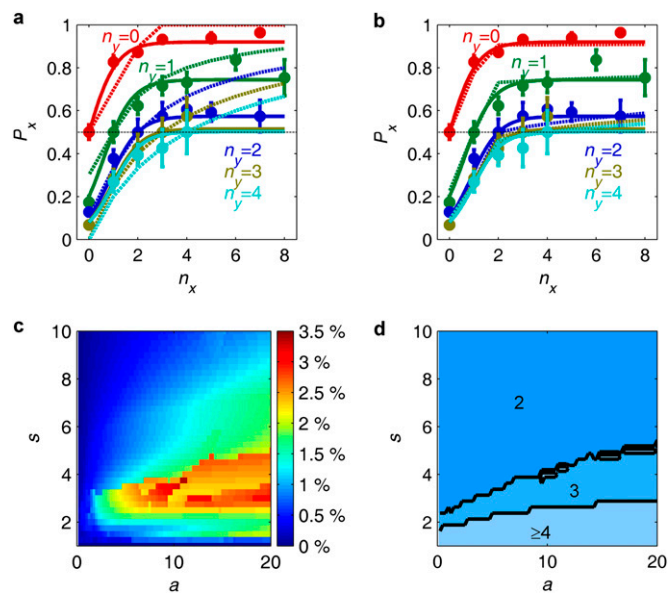


Fig. S3. A simple mechanistic model gives an approximation to Eq. 3. (A) Comparison between decisions using a simple stochastic model (dashed lines) and the model in Eq. 3 (solid lines). In the stochastic model, the focal fish either follows one of the other fish present in the setup (going to the zone where the followed fish is) or does not follow anyone (and therefore moves randomly). If there are N fish in the setup (apart from the focal one), the focal fish will follow any of them with equal probability P when $NP < 1$ and otherwise with probability $1/N$. The probability of not following another fish, and thus choosing at random, is then $\max\{(1 - NP, 0)\}$. We modeled the experiment as a series of repeated decisions following this rule, and calculated the time spent at each side in the limit of infinite decisions. Despite the simplicity of this simple stochastic model, it already shows some of the qualitative features of the data. (B) Same as A, but now the stochastic model considers that the focal fish has a different probability to follow close and far individuals. The implementation of the model was as follows. The probability of not following anyone is now $\max\{((1 - N_{\text{close}}P_{\text{close}} - N_{\text{far}}P_{\text{far}}), P_{\text{nf}})\}$, where N_{close} (N_{far}) is the number of fish in the same (opposite) zone as the focal fish, and P_{nf} is the minimum probability of not following anyone. When $N_{\text{close}}P_{\text{close}} + N_{\text{far}}P_{\text{far}} > 1 - P_{\text{nf}}$, P_{close} and P_{far} are renormalized so that $N_{\text{close}}P_{\text{close}} + N_{\text{far}}P_{\text{far}} = 1 - P_{\text{nf}}$, whereas $P_{\text{close}}/P_{\text{far}}$ remains constant. The model with $P_{\text{close}} = 0.71$, $P_{\text{far}} = 0.005$, $P_{\text{nf}} = 0.1$ (dashed lines) has a very good agreement both with the model in Eq. 3 (solid lines) and the experimental data (points). (C) Difference between the model in Eq. 3 and the mechanistic model in B as a function of a and s . For most values, there is a close agreement. (D) Maximum number of individuals that is necessary to count according to model in B when parameters are fitted to match the model in Eq. 3. For most parameter values, we can make $P_{\text{far}} = 0$ without a significant worsening of the fit. Then, the probability of not following any fish is $\max\{(1 - N_{\text{close}}P_{\text{close}}), P_{\text{nf}}\}$, which saturates when $N_{\text{close}} \geq (1 - P_{\text{nf}})/P_{\text{close}}$. Due to this saturation, the fish only needs to count up to $(1 - P_{\text{nf}})/P_{\text{close}}$. This model is consistent with the notion that for a very wide parameter range, animals only need to count up to a small number.

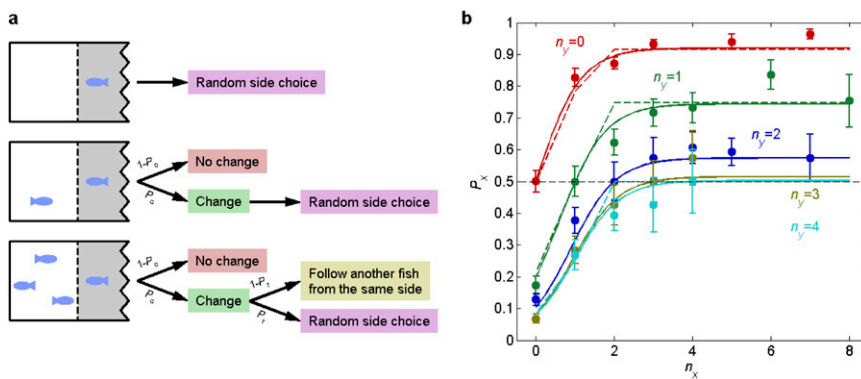


Fig. S4. A very simple mechanistic model gives an approximation to Eq. 3 for parameters corresponding to zebrafish data. (A) Schematic diagram of the model. The focal fish (the one on the shaded area) takes into account only the fish that are at the same side. If there are no other fish at the same side, the focal fish moves randomly, and therefore has probability one-half of choosing any side at the next time step (*Top*). If there are other fish at its side, the focal fish follows one of them; at the next decision, it chooses either to stay following the same fish (with probability $1 - P_c$) or to change (with probability P_c) and follow another fish, or not follow anyone. If there is only one fish at the same side, changing means necessarily not following anyone in the next time step, and therefore moving randomly (*Middle*). If there are more than one fish, then changing may lead to follow another fish and therefore remain at the same side, with probability $1 - P_r$, or not follow anyone, with probability P_r (*Bottom*). (B) Comparison among model in A (dashed lines), model in Eq. 3 (solid lines), and experimental data (points) for $P_c = 0.28$ and $P_r = 0.34$. The correspondence is good except for the $n_f = 2$ case (blue). The model corresponds to “counting up to 2”, whereas the data are best fitted with a “counting up to 3” model, as in the more complex model of Fig. S3B.



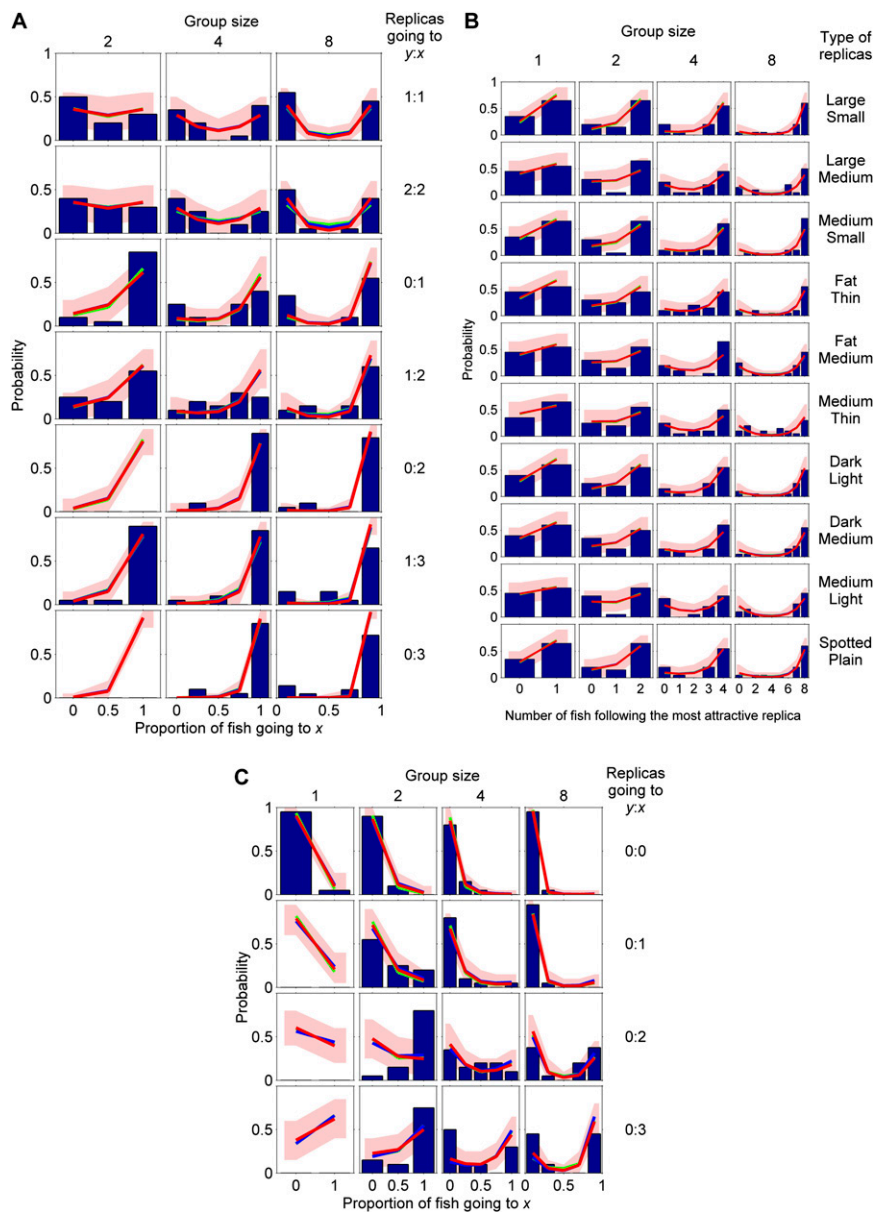


Fig. S7. Complete stickleback dataset (1, 2) and model fits. The three figures show experimental data as blue histograms and results for the $k = 1$ model (3) as red lines and green and blue lines for $k = 0.5$ and $k = 0$, respectively. In the three cases, $s = 2.5$ and a was refitted for each k . Pink regions limit the 95% confidence intervals for the $k = 1$ case. (A) Results for symmetric setup with different number of replica fish going to each side (e.g., 1:2 means one replica going to y and two replicas going to x). $a_x = a_y = 1$ for $k = 1$ (red line), $a_x = a_y = 5$ for $k = 0.5$ (green line), and $a_x = a_y = 224$ for $k = 0$ (blue line). (B) Results for symmetric setup and differently modified replica fish going to each side. We set the intermediate replica's reliability parameter equal to the one of the real fish ($s = 2.5$), and adjust the others to match the ratios found in ref. 3. We got $s_{\text{small}} = 1.25$, $s_{\text{medium}} = 2.5$, $s_{\text{large}} = 3.57$, $s_{\text{thin}} = 1.88$, $s_{\text{medium}} = 2.5$, $s_{\text{fat}} = 3.62$, $s_{\text{light}} = 1.95$, $s_{\text{medium}} = 2.5$, $s_{\text{dark}} = 4.55$, $s_{\text{plain}} = 2.5$, and $s_{\text{spotted}} = 5.81$. Parameter a as in A. (C) Results for setup with a replica predator at x . $a_x = 9.5$, $a_y = 1/9.5$ for $k = 1$ (red line), $a_x = 1.25$, $a_y = 31.5$ for $k = 0.5$ (green line), $a_x = 1,250$, and $a_y = 10,000$ (in this case, if we multiply these two parameters by any number greater than 0.1, the fit changes very little) for $k = 0$ (blue line).

1. Ward AJW, Sumpter DJT, Couzin ID, Hart PJB, Krause J (2008) Quorum decision-making facilitates information transfer in fish shoals. *Proc Natl Acad Sci USA* 105(19):6948–6953.
2. Sumpter DJT, Krause J, James R, Couzin ID, Ward AJW (2008) Consensus decision making by fish. *Curr Biol* 18(22):1773–1777.
3. Pérez-Escudero A, de Polavieja GG (2011) Collective animal behavior from Bayesian estimation and probability matching. *PLOS Comput Biol* 7(11):e1002282.

1. Ward AJW, Herbert-Read JE, Sumpter DJT, Krause J (2011) Fast and accurate decisions through collective vigilance in fish shoals. *Proc Natl Acad Sci USA* 108(6):2312–2315.
2. Pérez-Escudero A, de Polavieja GG (2011) Collective animal behavior from Bayesian estimation and probability matching. *PLOS Comput Biol* 7(11):e1002282.
3. Goss S, Aron S, Deneubourg JL, Pasteels JM (1989) Self-organized shortcuts in the Argentine ant. *Naturwissenschaften* 76:579–581.
4. Deneubourg J, Aron S, Goss S, Pasteels J (1990) The self-organizing exploratory pattern of the Argentine ant. *J Insect Behav* 3(2):159–168.
5. Meunier H, Leca JB, Deneubourg JL, Petit O (2006) Group movement decisions in capuchin monkeys: The utility of an experimental study and a mathematical model to explore the relationship between individual and collective behaviours. *Behaviour* 143(12):1511–1527.
6. Perna A, et al. (2012) Individual rules for trail pattern formation in Argentine ants (*Linepithema humile*). *PLOS Comput Biol* 8(7):e1002592.

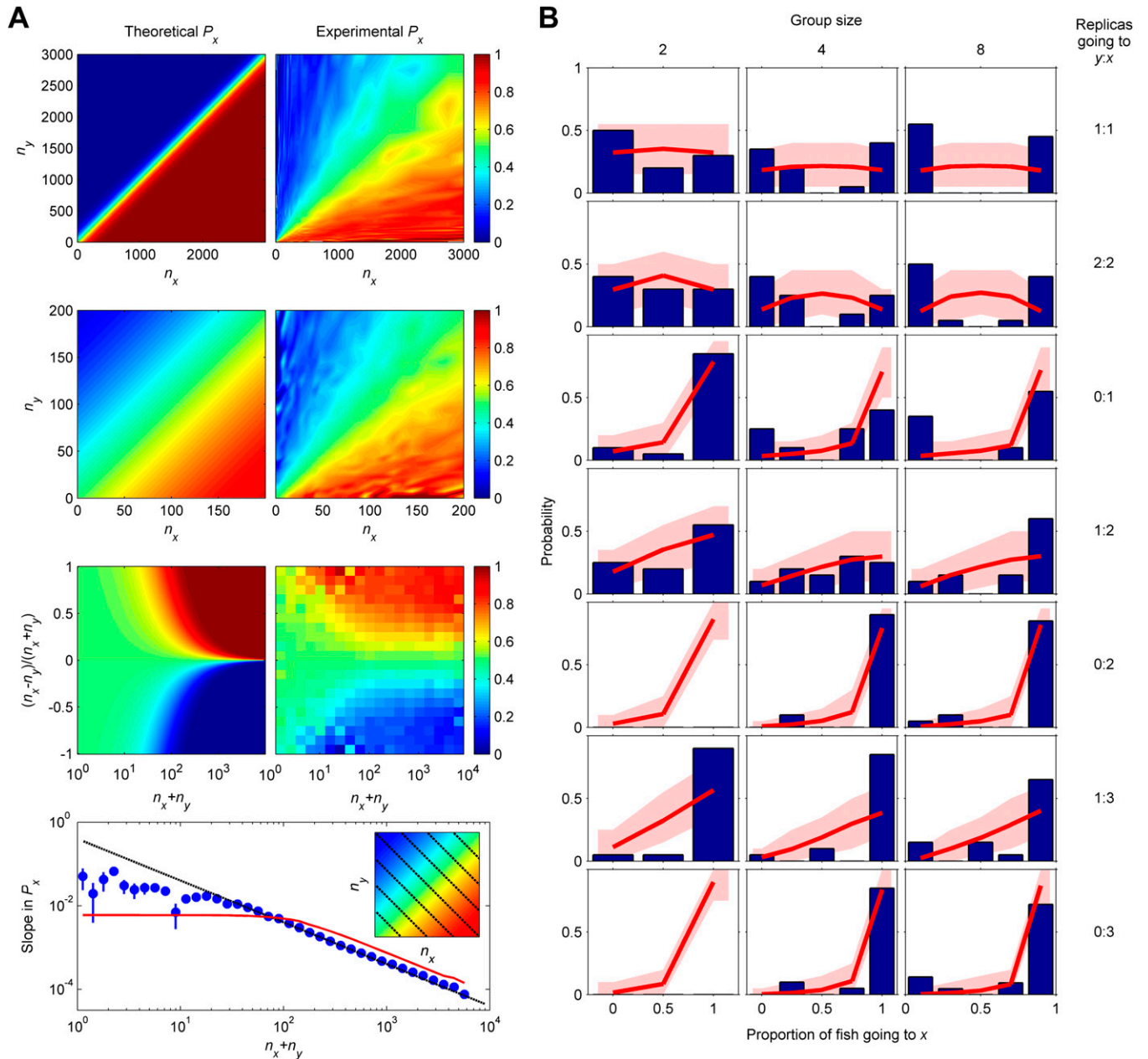


Fig. S10. The models in refs. 1 and 2 do not explain other datasets. (A) Same as Fig. 3, but using $P_x = (1 + s^{-(n_x - n_y)})^{-1}$, with $s = 1.012$. This model was used in ref. 1 to describe the stickleback dataset, and cannot describe the ant dataset. (B) Same as Fig. S7A, but with $P_x = 0.5 + A(n_x - n_y)/(n_x + n_y + T)$, with $A = 0.5$ and $T = 0.4$, which is the function used in ref. 2 to describe the ant dataset, with the 0.5 term added and with A restricted between 0 and 0.5, so that probabilities are between 0 and 1.

1. Pérez-Escudero A, de Polavieja GG (2011) Collective animal behavior from Bayesian estimation and probability matching. *PLOS Comput Biol* 7(11):e1002282.
2. Perna A, et al. (2012) Individual rules for trail pattern formation in Argentine ants (*Linepithema humile*). *PLOS Comput Biol* 8(7):e1002592.

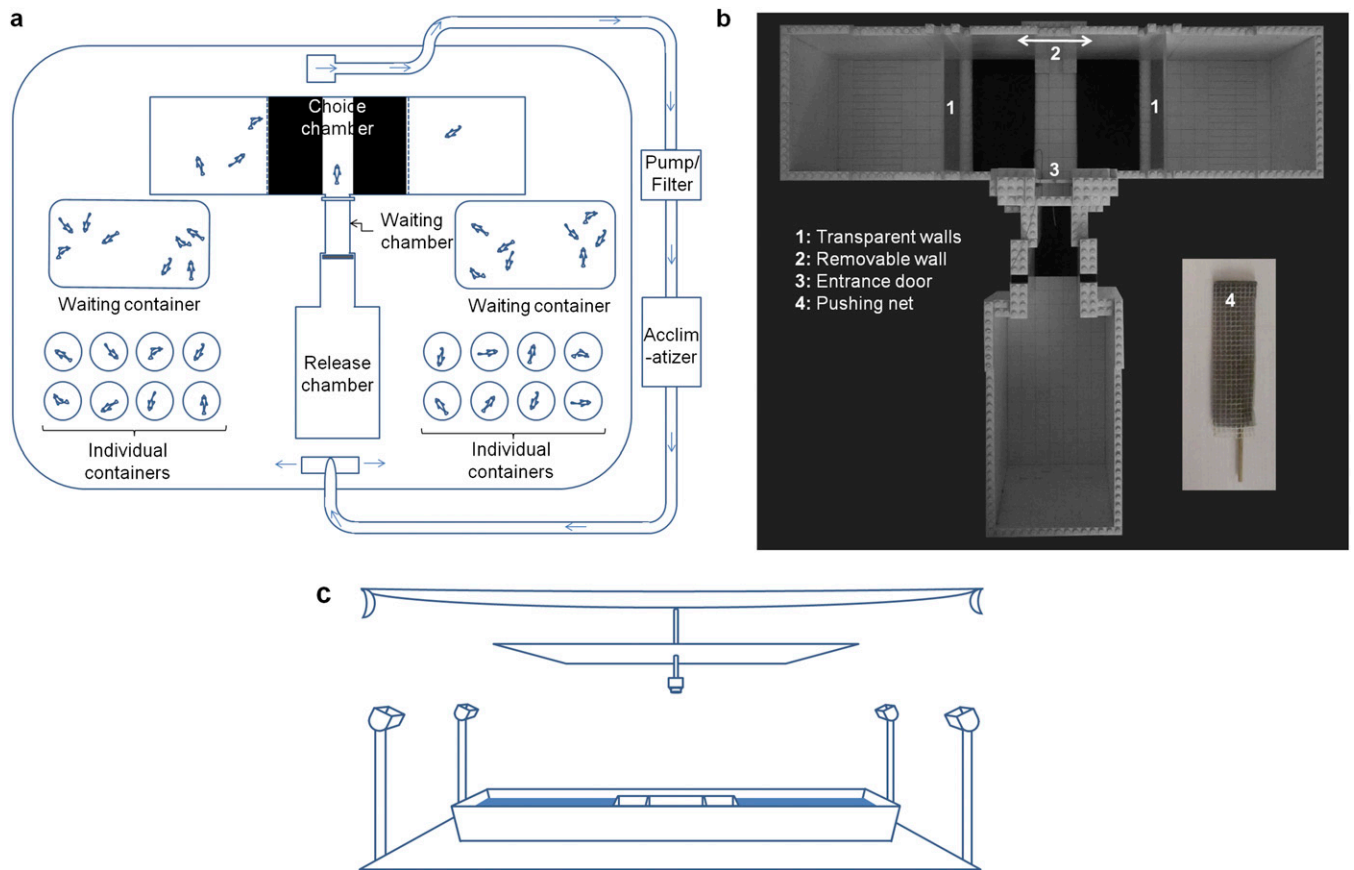


Fig. S11. Experimental setup for zebrafish. (A) The behavioral setup is inside a bigger tank so that fish are acclimated to the same water for 1 d before the experiment, housed in waiting containers in groups of 8–10 fish. At 1 h before the experiment, each fish is isolated and fed with frozen artemia in an individual container. The fish stays in the individual container until placed in the release chamber and gently pushed into the waiting chamber with a net that fits tightly between the walls to prevent the fish from going back to the release chamber. The door to the setup is then lifted and, once the fish enters the setup, it is closed. The camera records for 5 min from the opening of the door. After the experiment, the fish is pushed back to the release chamber, where it is caught. Then, a segment of wall opposite to the entrance door is removed, and water from outside is pumped into the central chamber so that odors are washed out. (B) The T-shaped setup is made of white LEGO bricks, with transparent walls separating the three chambers made of UV-transparent PLEXIGLAS (PLEXIGLAS GS 2458; Evonik Para-Chemie). The setup's central chamber (choice chamber) measures 20×13 cm. The floor of this central chamber has a central white zone 5 cm wide, and two black lateral zones 7.5 cm wide each. The two lateral chambers measure 14×13 cm each. Walls are 17-cm high, and water level is 6 cm. (C) Illumination is provided by four 500-W halogen lamps pointing to a white sheet on the ceiling. A Basler A622f camera records from above. An opaque roof just above the camera provides uniform shading on the setup.

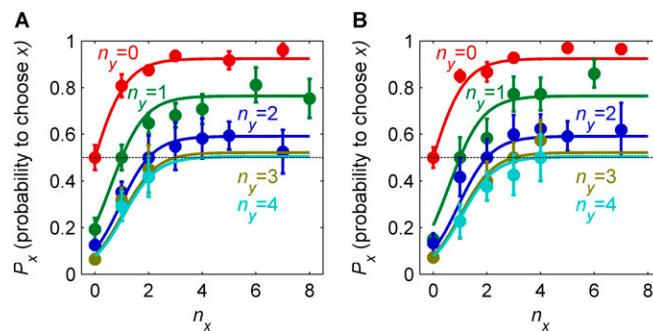


Fig. S12. Comparison of results using naive and nonnaive zebrafish. (A) Results for naive zebrafish, which have never seen the setup before the experiment. (B) Results for zebrafish that have been tested several times in the setup. Lines correspond to the theoretical model (Eq. 3) with same parameter values as for Fig. 2C.

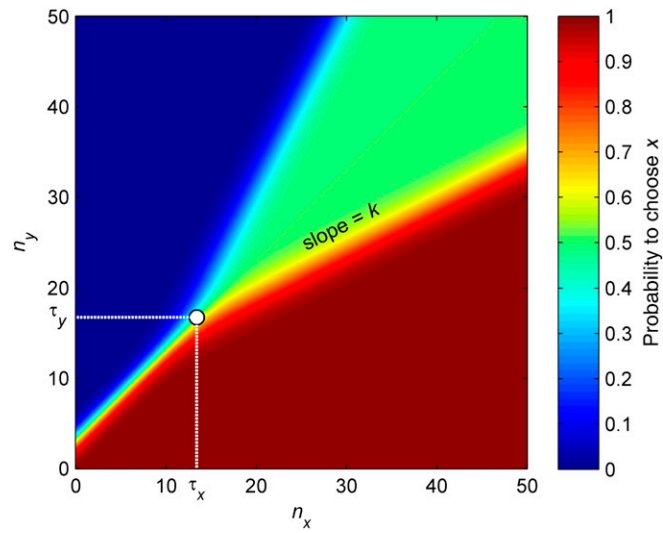


Fig. S13. Probability to choose x , P_x , for the general case of asymmetric nonsocial information (Eq. S22). Parameters: $s = 2.5$, $k = 0.5$, $a_x = 100$, $a_y = 10,000$. Compare this figure with the one corresponding to symmetric social information in Fig. 1B. See Eq. S29 for an analytical expression of (τ_x, τ_y) .

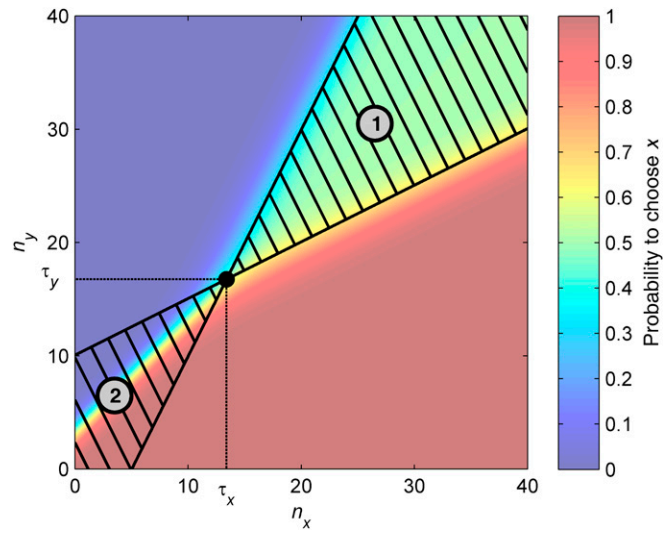


Fig. S14. Transition point τ between the low- and high-numbers regimes. Region 1 corresponds to the plateau with $P_x = 0.5$. The ΔN rule is approximately valid in region 2. Parameters are as in Fig. S13.