

Estimation models describe well collective decisions among three options

Miller et al. (1) demonstrate, by confronting groups of fish with three options, that information can be effectively integrated, allowing consensus despite no individual being aware of the consensus option. The different ways in which the conflict can be resolved allow testing of collective decision-making theories.

The experimental results show more cohesion than predicted by a previous decision-making model based on Bayesian estimation (2) (Fig. 1A, thin solid lines). To account for this difference, Miller et al. add a multiplicative term producing extra aversion to be in small groups (Fig. 1A, thick solid lines).

Bayesian models (2, 3) were derived mathematically from the hypothesis that animals make estimations about the environment using both nonsocial and social information. This formalism includes, without explicit separation into different terms, many potential factors implying cohesion, for example foraging strategies and risk-aversion as estimation of presence of food and predators, respectively. However, there are other factors not included in the model, such as mating and competition, that make an individual's behavior relevant by itself, and not only for what it tells about the environment. We thus tested whether estimation models (2, 3) need to be supplemented with extra factors to explain the data in Miller et al. (1). We note that the version of the estimation model in ref. 1 is an approximation that only takes into account one degree of dependencies, neglecting the effect of past interactions among the animals that have already decided. We show here that a further version of the model that includes all dependencies (2), averaging their effect (so that it does not require any extra information or cognitive

abilities), corresponds well with the data (Fig. 1A, dashed lines).

We tested further whether more general Bayesian models, in which animals estimate whether options are good (and not the best) (3), can also account for the data. They perform well without dependencies (Fig. 1A, dotted lines), and even better when including them (Fig. 1A, dotted-dashed lines).

Thus, we find that, although the data reject significantly an approximated Bayesian model (2), they can neither reject the corresponding exact Bayesian model (2) nor the more general estimation models (3) (Fig. 1B). We therefore find that estimation alone, with no additional factors, can explain the data.

We also note that the function proposed by Miller et al. (1) obtains a slightly better fit to this dataset. This finding illustrates the complementary role of these two types of modeling strategies. Models derived from first-principles (2, 3) allow testing the relevance of these principles, have parameters with explicit biological and mathematical interpretation (such as reliability of behaviors and private information) that are experimentally measurable, and can be applied to different species. However, despite their success in this and other datasets (2, 3), these idealized models may in general not fit data perfectly, as they do not incorporate details of specific biological implementations. Heuristic models are obtained to fit particular datasets (1), having a better chance to capture the impact of specific implementations. Furthermore, although selection may integrate all selection pressures in the tuning of a single function, they can propose useful ways to explicitly separate different terms that with

further experiments may be associated to different biological factors.

ACKNOWLEDGMENTS. This work was supported in part by Spanish Plan Nacional Ministerio de Economía y Competitividad (MICINN) BFU2012-33448 (to A.P.-E.); National Science and Engineering Research Council of Canada postdoctoral Fellowship 388856 (to N.M.); National Science Foundation award PHY-0848755 (to I.D.C.); Office of Naval Research Grant N00014-09-1-1074 (to I.D.C.); Human Frontier Science Project Grant RGP0065/2012 (to I.D.C.); Army Research Office Grant W911NG-11-1-0385 (to I.D.C.); Spanish Plan Nacional MICINN BFU2012-33448 (to G.G.d.P.); and the ERASysBio+ initiative supported under the European Union European Research Area Networks Plus scheme in Framework Program 7 (to G.G.d.P.).

Alfonso Pérez-Escudero^{a,1}, Noam Miller^b, Andrew T. Hartnett^b, Simon Garnier^b, Iain D. Couzin^{b,1}, and Gonzalo G. de Polavieja^{a,1}

^aInstituto Cajal, Consejo Superior de Investigaciones Científicas, Madrid 28002, Spain; and ^bDepartment of Ecology and Evolutionary Biology, Princeton University, Princeton, NJ 08544

1 Miller N, Garnier S, Hartnett AT, Couzin ID (2013) Both information and social cohesion determine collective decisions in animal groups. *Proc Natl Acad Sci USA* 110(13):5263–5268.

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

3 Arganda S, Pérez-Escudero A, de Polavieja GG (2012) A common rule for decision making in animal collectives across species. *Proc Natl Acad Sci USA* 109(50):20508–20513.

Author contributions: A.P.-E. and G.G.d.P. designed research; A.P.-E. and G.G.d.P. performed research; N.M., A.T.H., S.G., and I.D.C. contributed new reagents/analytic tools; A.P.-E. and G.G.d.P. analyzed data; and A.P.-E., I.D.C., and G.G.d.P. wrote the paper.

The authors declare no conflict of interest.

¹To whom correspondence may be addressed. E-mail: alfonso.perez.escudero@cajal.csic.es, icouzin@princeton.edu, or gonzalo.polavieja@cajal.csic.es.

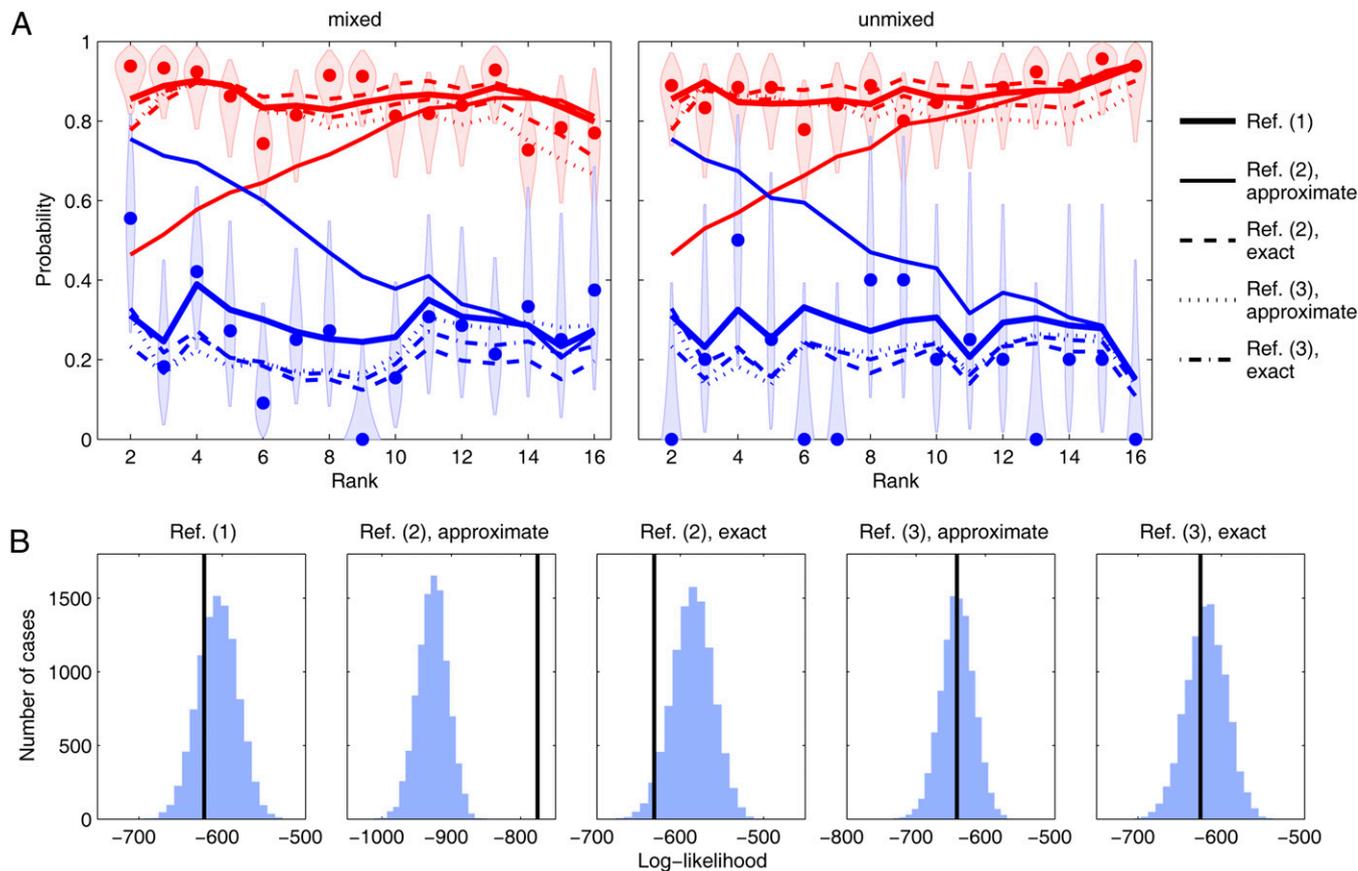


Fig. 1. Comparison of data from Miller et al. (1) and different models (from refs. 1–3). (A) Choice probability for each individual in the group, as a function of the order in which it made its decision (rank). Red: Probability of choosing the majority arm when it was one of the two preferred arms. Blue: Probability of choosing one of the preferred arms, when the majority arm was the nonpreferred arm (see ref. 1 for further details). Data are represented as dots. Pale colors show the uncertainty of the experimental data, with width proportional to the probability of the real value; they are truncated at 95% confidence intervals. All models correspond well to the data, except the one from ref. 2, assuming independent choices by previous animals. (B) Test of consistency of models in A with respect to the data. For each model, we built 10,000 datasets consisting of groups of fish with the same preferences and making the decisions in the same order as the experimental groups. Histograms show the log-likelihood of these synthetic datasets with respect to their corresponding model, indicating the range of values that are to be expected simply from sampling noise. Black lines correspond to the log-likelihood of the experimental data with respect to each model. Only the approximate model of ref. 2 is clearly rejected.