

## SOCIAL SCIENCES

# Conflicts of interest improve collective computation of adaptive social structures

Eleanor R. Brush,<sup>1,2,\*†</sup> David C. Krakauer,<sup>2</sup> Jessica C. Flack<sup>2</sup>

In many biological systems, the functional behavior of a group is collectively computed by the system's individual components. An example is the brain's ability to make decisions via the activity of billions of neurons. A long-standing puzzle is how the components' decisions combine to produce beneficial group-level outputs, despite conflicts of interest and imperfect information. We derive a theoretical model of collective computation from mechanistic first principles, using results from previous work on the computation of power structure in a primate model system. Collective computation has two phases: an information accumulation phase, in which (in this study) pairs of individuals gather information about their fighting abilities and make decisions about their dominance relationships, and an information aggregation phase, in which these decisions are combined to produce a collective computation. To model information accumulation, we extend a stochastic decision-making model—the leaky integrator model used to study neural decision-making—to a multiagent game-theoretic framework. We then test alternative algorithms for aggregating information—in this study, decisions about dominance resulting from the stochastic model—and measure the mutual information between the resultant power structure and the “true” fighting abilities. We find that conflicts of interest can improve accuracy to the benefit of all agents. We also find that the computation can be tuned to produce different power structures by changing the cost of waiting for a decision. The successful application of a similar stochastic decision-making model in neural and social contexts suggests general principles of collective computation across substrates and scales.

## INTRODUCTION

In biology, function emerges from the interactions of components making decisions with imperfect information. For example, behavior at the whole-organism level emerges from the firing activity of billions of neurons, each of which is responding to noisy input (1, 2). Quorum-sensing bacteria are able to detect the local density of conspecifics and change their behavior when density is sufficiently high, for example, producing a toxin or phosphorescing. This output depends on the individual cells' “decisions” to produce a signaling molecule when they detect other bacteria close by (3, 4). At a much larger spatial scale, in macaque and chimpanzee social groups, a distribution of power is the outcome of noisy decisions between pairs of individuals about which is subordinate (5–7). In several species of fish, the collective motion of a school emerges from the movements of individual fish, with movement decisions based on how fish perceive the environment and register the positions of neighbors (8–11). The ability of ant nests to keep out ants that do not belong can emerge collectively from decisions individual ants make when they can distinguish those they have met before from those they have not (12). These examples have inspired engineered networks: A group of robots can be designed to accomplish a task by interacting with each other and the environment according to specific rules (13).

In each of these systems, individuals gather information in noisy environments and change their behavior as they become better informed. Under some conditions, the joint behavior of individuals produces a stable aggregate-level pattern, which feeds back to affect the components' fitness. This two-part process constitutes a collective

computation (in table S1, we list the inputs and outputs of the individual-level and collective computations in each of these systems) (14–17). In many instances, these groups are able to produce collective computations that are beneficial for the individuals and the group, even though the individual group members are subject to noisy inputs from the environment, have conflicts of interest, and have finite time in which to make decisions. For example, fish schools are able to successfully navigate their environments (18–20), even though individual fish do not have perfect information, either about the environment or about the positions of the other members of their school, and different fish prefer different directions (19, 21).

## Principles of collective computation

We introduce a model of collective computation and ground our analyses in the collective computation of power structure in a primate society. To understand how groups collectively compute solutions, we partition collective computation into individual and collective phases (22). At the individual level, we ask how individuals make decisions based on accumulated information (22). At the collective level, we consider information aggregation, that is, how these decisions combine to produce a computation. Our goal is to understand how the quality of the computation varies as a function of both the strategies individuals use to make decisions and the mechanism required to aggregate the information encoded in individual decision-making.

At the individual level, three factors are important in decision-making: (i) making an accurate assessment of the environment, (ii) the time it takes to reach a decision, and (iii) one's self-interest, as opposed to the collective interest. In general, it is not possible to optimize all three of these factors. There is almost always a trade-off between speed and accuracy because taking more time to accumulate information results in a more accurate decision. This trade-off affects the strategies that animals use to reach decisions (23). There are also trade-offs between

Copyright © 2018  
The Authors, some  
rights reserved;  
exclusive licensee  
American Association  
for the Advancement  
of Science. No claim to  
original U.S. Government  
Works. Distributed  
under a Creative  
Commons Attribution  
NonCommercial  
License 4.0 (CC BY-NC).

Downloaded from <http://advances.sciencemag.org/> on March 2, 2021

<sup>1</sup>Program in Quantitative and Computational Biology, Princeton University, Princeton, NJ 08544, USA. <sup>2</sup>Santa Fe Institute, Santa Fe, NM 87501, USA.

\*Corresponding author. Email: [eleanor.brush@gmail.com](mailto:eleanor.brush@gmail.com)

†Present address: Department of Biology, University of Maryland, College Park, MD 20742, USA.

decision speed and realizing a preferred outcome. For example, in a fight, being stubborn and waiting for one's opponent to give up first makes winning more likely, but it also makes the fight last longer. In animal conflicts, human warfare, and economics, this trade-off has large effects on the strategies that individuals use to resolve differences (24–26). Finally, there is a trade-off between collective and immediate self-interest. For example, for groups to stay together, individuals must make compromises between the direction the group should be moving, to migrate successfully or to maximize the group's food intake, and individual movement preferences. The relative importance of group accuracy versus individual preferences affects both how individuals make decisions about how to move and the accuracy with which the whole group can navigate (18, 19, 21). Although there are many studies considering trade-offs between two of these three factors—accuracy, time, and individual preference—few consider the interaction between all three.

The question of how individual-level decisions are combined to compute an output is essentially a problem of consensus formation. There are two senses of consensus: The group can come to consensus with all individuals or components settling on a single decision, which can be called redundant consensus. In the brain, neurons can reach redundant consensus where they all “agree” about what is present in a visual stimulus, which can help the brain discriminate among alternatives (22). Fish schools navigate more quickly and accurately when there is a higher degree of redundant consensus about where to move (18–20). On the other hand, as each group member forms an opinion about, say, the value of another member of the group, the degree of agreement among group members about that value can be encoded collectively, with individuals retaining their own opinions, which can be called collective consensus. In the primate model system studied here, an individual's power depends on collectively encoded consensus among members of its group about its ability to use force successfully (5, 6, 27).

We have reviewed algorithms for computing collectively encoded consensus elsewhere (6, 27). Here, we focus on the functional utility of the collective computation, with emphasis on two properties in particular: the accuracy of the collective computation and the skewness of the consensus values that are the output of the computation. In the neural case, the accuracy of an individual's decision, given the environmental input, affects how successfully an individual will interact with its environment. If an individual is trying to decide among many alternatives, each alternative may be assigned a value that reflects the neurons' collective certainty about that alternative. It should be easier to discriminate among alternatives if the distribution of consensus values is right-skewed (28).

In the primate model system, “accurate” power scores—those that reflect the animals' fighting abilities—are useful because if a monkey can estimate another's power, then it can predict whether it will win in a fight against the other monkey and what the cost of interacting with it will be (6). A power structure that accurately reflects the monkeys' fighting abilities is also more stable because marked role reversals are less likely to occur (5). Right skewness is also valuable in this system. As in the neural example, right skewness in the power distribution can be interpreted to mean that there is high confidence within the group that a few individuals are disproportionately powerful. These individuals pay little cost during conflicts and can afford to engage in costly conflict management behavior that is beneficial to the group (see the detailed description of the model system below) (5).

### Models of individual and collective computation

A variety of models, including the leaky integrator model (2, 29–32) and the sequential probability ratio test (SPRT) (31, 33), have been devel-

oped to study how components (for example, individual animals or neurons) choose among alternatives in a noisy environment [reviewed in the study of Ratcliff *et al.* (34)]. For example, the leaky integrator model has been used to describe the firing of neurons during a motion coherence task in which a subject must decide whether it is seeing dots moving left or right (29–31, 35). Both the leaky integrator model and the SPRT keep track of the amount of accumulated evidence supporting alternative choices. The leaky integrator model has the advantage over the SPRT of allowing for memory loss, but application of the leaky integrator model to explain, for example, neural firing, has been largely phenomenological (more details on these decision-making models are provided in section S1).

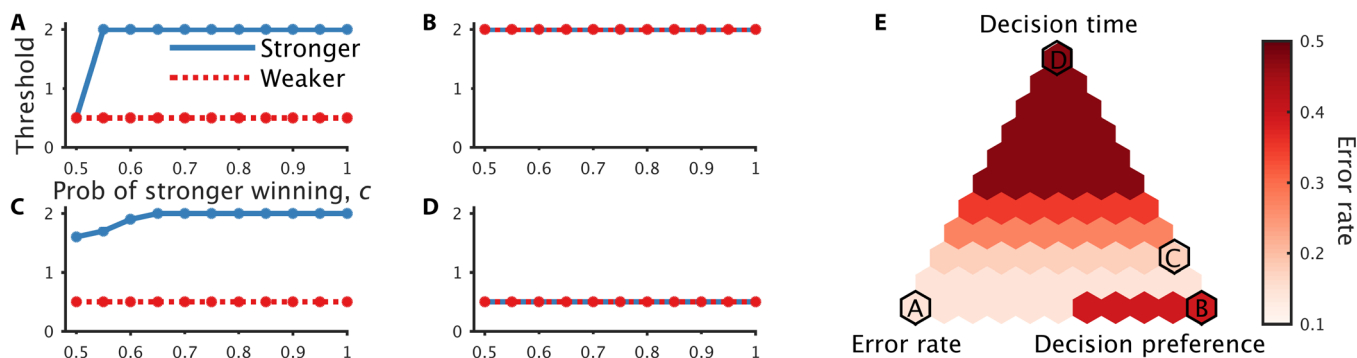
Here, we develop a leaky integrator model by deriving stochastic differential equations (SDEs) that mechanistically specify how information is accumulated by components and is used in decision-making. We use empirical and computational results from work on collective computation in a primate society model system to justify the form of our equations. Our model extends the standard leaky integrator model in two ways: In addition to considering the accuracy of the decision and the time it takes to make a decision, we introduce a game theoretic element describing an individual's preference to be dominant, and we use the stochastic model to generate a network of pairwise decisions between many individuals, rather than just two. These extensions allow us to study how the importance of the three decision properties—error rate, decision time, and individual preference about the decision—influence (i) the “correctness” of the collective computation of social structure, (ii) the accessibility of different social structures, and (iii) the best way to perform the collective computation.

## RESULTS

### Impact of conflict on signaling decision

The weight  $w_3$ , given to the probability of an individual's preferred outcome being reached, indicates the strength of the conflicts of interests between pairs of individuals. We start with a pair of individuals, that is, a group size of two, to build intuition for how the optimization weights affect the Nash equilibrium thresholds. If the error rate of the decision is important ( $w_1 = 1$ ), the Nash strategies are for the weaker individual to set its threshold as low as possible and the stronger individual to set its threshold as high as possible (Fig. 1A). This will, with high probability, lead to the weaker individual signaling, which is the correct outcome. When only decision preference matters and there is a strong conflict of interest between individuals ( $w_3 = 1$ ), the Nash strategies are for both to set their thresholds high because each prefers to wait for the other to signal and there is no incentive for the individuals to stop accumulating evidence (Fig. 1B). As the importance of decision time increases, the Nash thresholds of both individuals decrease, which enables them to reach the decision more quickly (Fig. 1, A to D). The error rate with which a pair using Nash thresholds can reach a decision is lowest when only error rate matters ( $w_1 = 1$ ) and increases as either decision time or decision preference becomes more important (Fig. 1E).

In a group with more than two individuals, the Nash thresholds respond to the optimization weights in a similar way (Fig. 2A). There is a significant change brought about by introducing additional individuals: As long as there are nonzero waiting costs, the average error rate of all decisions in a group using Nash thresholds decreases as error rate becomes less important and decision preference becomes more important (this can be seen by moving from left to right in Fig. 2B and directly in Fig. 2C). This can be explained as follows. Consider a case where only



**Fig. 1. The error rate of a pair using Nash equilibrium thresholds increases as the weight given to either decision time or decision preference increases. (A to D)** The lines show the Nash equilibrium thresholds for a pair of individuals as a function of the probability that the stronger animal wins,  $c$ , with blue indicating the stronger individual and red the weaker individual. When  $c = 0.5$ , the individuals are equally matched. In (C), when  $c = 0.5$ , the Nash thresholds are symmetric: Either individual could use either strategy. The optimization weights for each panel are indicated in the simplex with the corresponding letter. (E) The color in the simplex indicates the error rate of a decision made by a pair making a difficult decision ( $c = 0.55$ ) using Nash equilibrium thresholds, as a function of the optimization weights  $w_1, w_2$ , and  $w_3$ . In the lower left corner of the simplex, only error rate matters ( $w_1 = 1$ ). In the upper corner, only decision time matters ( $w_2 = 1$ ). In the lower right corner, only preference matters ( $w_3 = 1$ ). Parameters: in (A),  $w_1 = 1, w_2 = 0$ , and  $w_3 = 0$ ; in (B),  $w_1 = 0, w_2 = 0$ , and  $w_3 = 1$ ; in (C),  $w_1 = 0, w_2 = 0.2$ , and  $w_3 = 0.8$ ; and in (D),  $w_1 = 0, w_2 = 1$ , and  $w_3 = 0$ ; in all panels,  $b = 1, r = 1$ , and  $l = 0.1$ .

error rate and decision time matter (as in the blue curve in Fig. 2A, where  $w_1 = 0.9, w_2 = 0.1$ , and  $w_3 = 0$ ). When any individual raises its threshold, it becomes less likely to emit the subordination signal to all of its opponents. This is the correct behavior when it interacts with animals with lower ability, but an error when it interacts with animals with higher ability. Thus, when an animal increases its threshold, there are three effects: a lower error rate in its decisions with lower animals, a greater error rate in its decisions with higher animals, and an increase in its average decision time. Even if the animal's total error rate would be smaller with a higher threshold, the improvement will be small because of the increase in its error rate with higher animals, and it has to pay the cost of a greater average decision time. Being at Nash equilibrium means that it cannot improve its utility by increasing its threshold. However, if the animal starts to value receiving the subordination signal and decision preference starts to matter ( $w_3 > 0$ ), then the errors made by waiting for higher animals to signal are no longer perceived as costly and are instead perceived as beneficial. This provides an incentive for it to raise its threshold and pay the costs for waiting longer to make decisions. Therefore, increasing the weight given to decision preference encourages all individuals, except the very strongest and the very weakest, to raise their thresholds (Fig. 2A, compare iii and iv). (The strongest individual always uses the maximal threshold allowed. The weakest would have to wait so long to receive the signal that if there are any costs to waiting, it will not raise its threshold above the minimum allowed.) Because nearly all members of the group raise their thresholds, all decisions take longer, resulting in a decrease in the average error rate of all decisions. (We analyzed a group of  $n = 20$  individuals. In fig. S6, we show that our results are robust to increasing group size. In fig. S7, we show how the amount of time it takes for a pair to make a decision depends on the optimization weights and on the difference in their abilities.)

### Impacts of conflict and waiting costs on collective computation

For each consensus algorithm, the mutual information between the distribution of social power (DSP) and the underlying distribution of abilities increases as the pairwise error rate decreases (fig. S8). Hence, the information content of the consensus scores produced by every al-

gorithm is improved when decision preference is prioritized over error rate (Fig. 2D), as long as the group has more than two components and there are nonzero waiting costs.

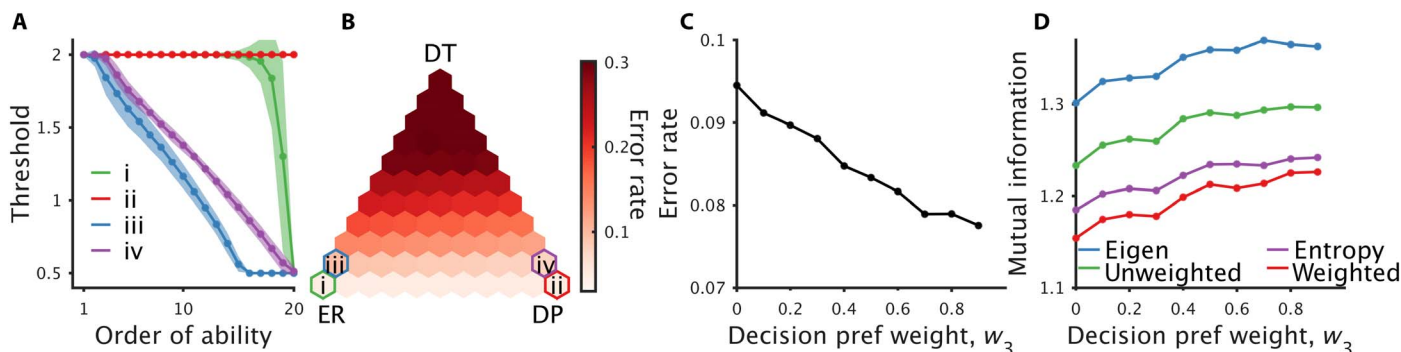
For each measure of consensus, skewness of the DSP is maximized at intermediate waiting costs and does not depend strongly on the trade-off between error rate and preference (Fig. 3) (see section S3 for more details). Example distributions are shown in fig. S9. Results for all algorithms are shown in figs. S10 and S11.

### Impacts of waiting costs on which consensus algorithm is most informative

For a fully developed network, when there are no waiting costs and decision preference matters ( $w_2 = 0$  and  $w_3 \geq 0.3$ ), weighted in-degree is the most informative algorithm (Fig. 4A). When there are no waiting costs and the error rate is very important ( $w_2 = 0$  and  $w_3 < 0.3$ ) or when there are small waiting costs ( $0 < w_2 \leq 0.3$ ), eigenvector centrality is the most informative algorithm (Fig. 4A). These more "global" consensus algorithms do well when waiting costs are low because, in these circumstances, the edges of the decision network tend to be accurate and these measures make use of more information in the network. When waiting costs are higher, unweighted in-degree and entropy are more informative than eigenvector centrality, but only by a very small margin (for further discussion, see section S10). When we consider the status-signaling network as it develops, we find that eigenvector centrality has the advantages of never losing information content and consistently performing well on networks that are not fully formed (Fig. 4, B to D).

### DISCUSSION

Conflicts of interest are a general feature of biological and social systems when resources are scarce or fates are not fully shared [for example, (36, 37)]. The dominant view in biology is that conflicts of interest are negative because conflict can lead to instability, gridlock, and increased mortality [for example, (25, 38–40)]. Yet, some data suggest that when conflicts of interest are expressed as controlled antagonisms (for example, fights) in which components can challenge one another at relatively low cost, this can foster invention and innovation (7, 41–43),



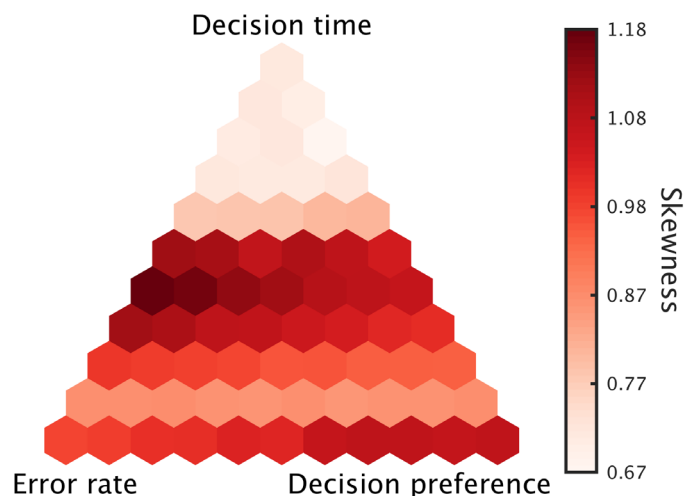
**Fig. 2. The mutual information of the power scores computed by a group using Nash thresholds increases as the weight given to decision preference increases, as long as there are nonzero waiting costs.** (A) The points show the average Nash thresholds as a function of position in the group for different optimization weights, where 1 is the strongest individual and 20 is the weakest individual. The shaded region around each curve shows the average plus or minus the SD across 1000 draws. The optimization weights for each line are indicated in the simplex with the corresponding letter. (B) The color in the simplex indicates the average error rate of all decisions made by members of a group using Nash thresholds, as a function of the optimization weights  $w_1$ ,  $w_2$ , and  $w_3$ . In the lower left corner of the simplex, only error rate matters ( $w_1 = 1$ ). In the upper corner, only decision time matters ( $w_2 = 1$ ). In the lower right corner, only preference matters ( $w_3 = 1$ ). (C) The points show the average error rate of decisions made by all individuals, as a function of the weight given to decision preference,  $w_3$ , when  $w_2 = 0.1$  [this is also shown by moving from left to right in the second row of hexagons from the bottom in (B)]. (D) The points show the mutual information between the power scores computed using each algorithm, as a function of the weight given to decision preference,  $w_3$ , when  $w_2 = 0.1$ . Each algorithm is shown in a different color: eigenvector centrality in blue, unweighted in-degree in green, weighted in-degree in red, and entropy in purple. Parameters: i indicates  $w_1 = 1$ ,  $w_2 = 0$ , and  $w_3 = 0$ ; ii indicates  $w_1 = 0$ ,  $w_2 = 0$ , and  $w_3 = 1$ ; iii indicates  $w_1 = 0.9$ ,  $w_2 = 0.1$ , and  $w_3 = 0$ ; and iv indicates  $w_1 = 0$ ,  $w_2 = 0.1$ , and  $w_3 = 0.9$ ; in all panels,  $n = 20$ ,  $b = 1$ ,  $r = 1$ , and  $l = 0.1$ .

facilitate information flow (44), allow components to test strategies (41), and even improve social cohesion (41). Hence, under some conditions, it appears that conflicts of interest can be beneficial.

We find support for this view. Specifically, we study a theoretical model of collective computation, based on data from a primate society model system. In the first stage of the model, pairs of components learn about each other through a stochastic process. Then, we quantify the collective computation of consensus by measuring the consensus encoded in the network of pairwise decisions about each individual's ability. The components in the model have a strategy that dictates how they make pairwise decisions. We study how the Nash equilibrium strategies depend on the importance of the accuracy of the pairwise decision, the amount of time it takes, and the desired outcome. We find that (i) conflicts of interest can improve the accuracy of the collective computation to the benefit of all individual competitors and (ii) the output of the collective computation, in particular the skewness of the distribution of consensus scores, can be tuned by manipulating properties of conflict dynamics. When there are conflicts of interest, each member of a pair desires a different outcome. Strengthening these conflicts essentially makes the components more stubborn, which leads to increased decision times, and, on average, improves the quality of information aggregation at the group level.

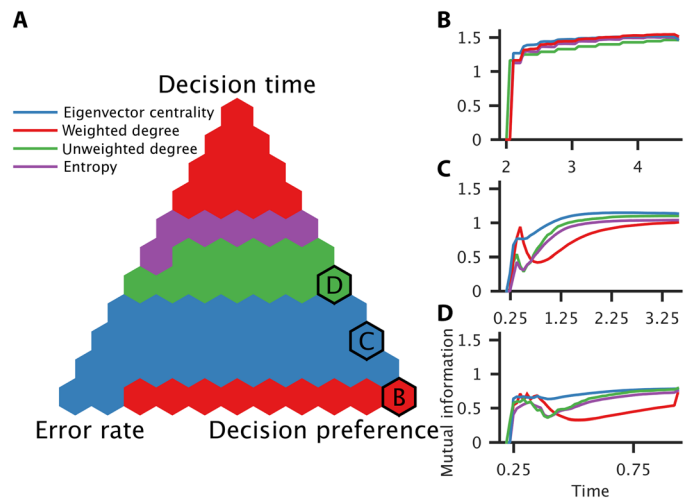
These findings improve our understanding of the emergence of power structure in primate social groups. In particular, they show how our primate model system may be able to construct a power structure that both accurately reflects the animals' fighting abilities and is right-skewed (5, 6, 27). The former can be achieved through the conflicts of interest inherent in the system. The latter can be achieved by changing the costs of waiting for a decision, for example, by fighting more aggressively and making it more likely for combatants to incur injuries (more details on how these costs can be tuned are provided in section S11).

More broadly, our findings provide a novel way of interpreting the widespread observation of competitive dynamics as a means of obtaining the most reliable information about components of a system, rather than solely as a mechanism by which components can gain access to



**Fig. 3. The average skewness of the distribution of eigenvector centrality is maximized at intermediate waiting costs.** The color indicates the average skewness of the distribution of power computed by a group using Nash thresholds, as a function of the optimization weights  $w_1$ ,  $w_2$ , and  $w_3$ . In the lower left corner of the simplex, only error rate matters ( $w_1 = 1$ ). In the upper corner, only decision time matters ( $w_2 = 1$ ). In the lower right corner, only preference matters ( $w_3 = 1$ ). Parameters:  $n = 20$ ,  $b = 1$ ,  $r = 1$ , and  $l = 0.1$ .

resources (for more details on how our work extends previous findings on animal conflict, see section S12). Together, previous results about the benefits of conflict in biological systems (7, 41–44), along with those presented here, have implications for understanding the evolution and social engineering of information aggregation mechanisms and collective computation. If this view of the utility of conflicts of interest is correct, then we predict, following the study of Stearns (42) and Krakauer and Mira (43), that when information processing is noisy and uncertainty is high, either strong regulatory mechanisms or conflict arenas will be a



**Fig. 4. The best measure of consensus in the decision network depends on the average error rate and the types of errors being made.** (A) The color indicates the most informative consensus measure to apply to a network constructed by a group using Nash thresholds, as a function of the optimization weights  $w_1$ ,  $w_2$ , and  $w_3$ . In the lower left corner of the simplex, only error rate matters ( $w_1 = 1$ ). In the upper corner, only decision time matters ( $w_2 = 1$ ). In the lower right corner, only preference matters ( $w_3 = 1$ ). (B to D) We show how the mutual information of each consensus algorithm changes over time as the network forms. The optimization weights for each panel are indicated in the simplex with the corresponding letter. Parameters: in (B),  $w_1 = 0$ ,  $w_2 = 0$ , and  $w_3 = 1$ ; in (C),  $w_1 = 0$ ,  $w_2 = 0.2$ , and  $w_3 = 0.8$ ; and in (D),  $w_1 = 0$ ,  $w_2 = 0.4$ , and  $w_3 = 0.6$ ; in all panels,  $n = 20$ ,  $b = 1$ ,  $r = 1$ , and  $l = 0.1$ .

general feature of collective computation in biological and social systems, from biofilms to financial markets.

Our finding also suggests that in the design of any collective computation, it could be advantageous to introduce conflicts of interest between agents. For example, imagine a group of robots tasked with evaluating the relative likelihood of various events, where each robot is responsible for gathering evidence that a given event will occur and pairs of robots compare how certain they are about their assigned events. Our results suggest that the network as a whole would come up with more accurate predictions if the robots were “rewarded” for being more confident in their assessments than if they were rewarded only on the basis of the accuracy of their individual assessments.

The stochastic learning model that we derived based on our primate model system is nearly equivalent to one that has been used to describe stochastic learning in neural populations [for example, (29–31, 35)]. Our model extends this leaky integrator model by incorporating game theoretic strategic decisions and a social context, thus addressing some challenges associated with collective computation of social structure. This raises the possibility of a reapplication of the extended framework to collective computation in populations of neurons, where conflicts are not assumed to be common (45, 46). Conflicts of interest, even among the cells within somatically clonal tissues, could make the brain a better decision-maker. This provides indirect support for neural Darwinism (47). When accuracy is not important for decision-making, competitive self-interest reduces to the “war of attrition” game in the game theory literature (for more details, see section S13) (24–26). The similarities between three frameworks—social decision-making, neural decision-making, and game theory—suggest that there are principles of

collective computation that could be applicable to a large class of decision problems in which information is distributed and noisy.

## MATERIALS AND METHODS

### Model system

Our model system was a well-studied captive group of pigtailed macaques (*Macaca nemestrina*;  $n = 48$ ). This system is characterized by social learning at the individual level, frequent non-kin interactions, multi-individual conflict interactions, and social structures that arise from nonlinear processes and feed back to influence individual behavior (see section S2 for more details) (5–7, 38, 41, 48, 49).

Each individual learns about the fighting ability of other members of the group through direct fighting and observation (for operational definitions, see section S2.1). If an individual loses many fights against an opponent, it will come to perceive a large asymmetry in fighting ability in its opponent’s favor and will perceive the cost of continued aggression with the opponent as greater than the cost of accepting the subordinate role (41). When this happens, the focal individual can decide to emit a silent-bared teeth display. The silent-bared teeth display communicates agreement to the subordinate role in a relationship when emitted during peaceful contexts (41). The signal is highly unidirectional, that is, if one individual emits a silent-bared teeth display to another in a peaceful context, then the second individual is highly unlikely to emit the signal to the first (6, 41). Fighting after signals are exchanged is reduced, continuing at a low level, so that the relationship can reverse if the fighting ability of the weaker individual improves. The decision to emit the signal constitutes an individual-level computation that involves integrating over a history of fight outcomes to estimate the magnitude of asymmetry between a pair of individuals. The output is a dominance relationship.

An individual’s power in the group depends on the degree to which it is collectively perceived as capable of using force successfully in fights (6). Individuals can estimate power by observing (some subset of) the network of subordination signals emitted between all pairs of individuals. The distribution of social power (DSP) in the group results from the collective computation by all individuals of their power scores (see also section S2) (6, 27). The functional significance of power is evident in how it changes social interactions. Individuals treat each other differently according to the power they perceive each other to have: Individuals solicit support in conflicts from powerful individuals more often, and powerful individuals use less aggression and receive less aggression when they do intervene in conflicts (6, 27). Consensus in the group about the capacity of an individual to successfully use force can be measured directly from the network of subordination signals. In previous work, we identified seven network metrics that assigned individuals scores that were significantly correlated with their power scores, as quantified by these social variables (6, 27).

Both the accuracy with which the DSP reflects “true” fighting abilities and the skewness of the DSP are functionally important. If the DSP has high mutual information with the underlying distribution of fighting abilities, it will be a reliable predictor of interaction cost (7). The fact that the distribution of power changes relatively slowly, marked changes in individuals’ power being relatively rare, and the fact that the individuals do seem to use information about the distribution of power to decide how to interact with each other both suggest that, in this system, the DSP is in fact highly informative about the individuals’ fighting abilities. The skewness of the DSP influences conflict management. Heavy-tailed distributions make otherwise costly conflict-management strategies,

such as policing, accessible to individuals who occupy the tail of the distribution of power (see section S2 for details on policing and section S3 for further explanation of the importance of skewness in primate social structure) (5). It therefore appears from the data that the study group was able to compute a DSP that was both accurate and structured in a beneficial way. Our goal was to understand the factors influencing the quality of the collective computation and how the group overcomes a noisy learning environment and the inherent conflicts of interest in this system.

**Stochastic approach**

Here, we developed a model describing the collective computation of social structure. First, we developed a stochastic model of individual decision-making—in this case, whether to signal subordination. The SDEs used to model noisy decision processes are typically presented without derivation. Here, we followed the mathematical derivation of SDEs in chemical systems, as given by Gillespie (50), to derive equations for how an individual learns about the fighting ability of each of its group mates. Each individual in the group accumulates evidence about its fighting ability relative to another individual by keeping track of the fights it has won and lost. For a given pair of individuals, A and B, A has a decision variable,  $X_1$ , indicating the evidence that it has accumulated about its ability relative to B, and similarly, B has a decision variable,  $X_2$ . In the absence of new information, the decision variables leak back toward 0 with rate  $l$  (Table 1 lists and defines all the variables used in the text). If there is no input, then over a period of length  $\tau$  each decision variable decreases as  $X_i(t + \tau) = (1 - l\tau)X_i(t)$ .

If they do fight and learn about each other, each individual incorporates this new evidence into its assessment. Specifically,  $X_1$  increases by an amount  $b$  when individual A wins a fight against individual B and decreases by  $b$  when it loses, and conversely for  $X_2$ . To calculate the variables at time  $t + \tau$ , we count how many times each type of input occurred in the time since  $t$  and add the changes resulting from these events to the background leaky estimate

$$X_i(t + \tau) = (1 - l\tau)X_i(t) + b \times \# \text{ times } i \text{ wins in } [t, t + \tau) - b \times \# \text{ times } i \text{ loses in } [t, t + \tau)$$

We ignored the possibility of individuals learning about each other by observing their fights with other individuals. This should mainly increase the rate at which they learn about each other and therefore should not greatly affect our results.

We assumed that individuals fight with each other at a constant rate. We also assumed that, even if one individual is stronger than another, there are random factors that affect which of the two will win a fight. Specifically, we described the number of each type of event—wins and losses—with a Poisson random variable,  $N_A$  and  $N_B$ , giving

$$X_1(t + \tau) = (1 - l\tau)X_1(t) + bN_A - bN_B$$

$$X_2(t + \tau) = (1 - l\tau)X_2(t) - bN_A + bN_B$$

If fights occur at a rate  $r$  and A wins with probability  $c$  and loses with probability  $1 - c$ , then the expectation of  $N_A$  and  $N_B$  in a period of length  $\tau$  are, respectively,  $\tau rc$  and  $\tau r(1 - c)$ . The parameter  $c$  ranges between 0 and 1 and is related to the strength of the asymmetry in the individuals' abilities: If A is stronger, then it is more likely to win and  $c > 0.5$ .

When enough events accumulate in an interval of time from  $t$  to  $t + \tau$ , we can approximate the Poisson random variables with normal random variables with mean and variance equal to the mean of the Poisson

random variables. Let  $Z_A$  and  $Z_B$  be independent standard normal random variables, that is, with a mean of 0 and an SD of 1, giving

$$X_1(t + \tau) = (1 - l\tau)X_1(t) + b(\tau rc + \sqrt{\tau rc}Z_A) - b(\tau r(1 - c) + \sqrt{\tau r(1 - c)}Z_B)$$

$$X_2(t + \tau) = (1 - l\tau)X_2(t) - b(\tau rc + \sqrt{\tau rc}Z_A) + b(\tau r(1 - c) + \sqrt{\tau r(1 - c)}Z_B)$$

Finally, as we make the period of time shorter, where  $\tau$  becomes infinitesimally small, these equations become SDEs

$$dX_1 = (-lX_1(t) + br(2c - 1))dt + b\sqrt{r}dW_t$$

$$dX_2 = (-lX_2(t) - br(2c - 1))dt - b\sqrt{r}dW_t$$

where  $dW_t$  is Brownian motion representing the wins and losses for individual A. We assumed that  $X_1(0) = X_2(0) = 0$ . The sensitivity of this model to initial conditions is discussed in section S4.

Nearly identical SDEs have been used to model decision-making in the brain (see section S5 for more details) (2, 28, 31). In that case,  $X_1$  denotes the firing activity of a neural population responding to one property in the environment, for example, left motion, and  $X_2$  denotes the firing activity of a neural population responding to its opposite, for example, right motion. In Table 1, we listed the inputs, outputs, and variables of the decision model and how they should be interpreted in both social and neural systems.

**Modeling pairwise decisions—individual level computation**

An individual decides to signal to another once it is sufficiently certain that it is the weaker of the two and that the costs of continued fighting are greater than the potential benefits of waiting for the other to signal. Here, B signals to A if  $X_2$  becomes very negative, and A signals if  $X_1$  becomes very negative. Specifically, there are two thresholds,  $T_1$  and  $T_2$ , such that if  $X_2 < -T_2$ , then B signals and the pair reaches the “decision” that A has higher ability and if  $X_1 < -T_1$ , then A signals and the decision is that B has higher ability. It can be shown that, regardless of how high  $T_1$  and  $T_2$  are, eventually one individual's decision variable will reach its threshold and it will signal.

In the empirical system, individuals do not emit subordination signals instantaneously (41), so in the model, their thresholds should be greater than zero. Conversely, in the empirical system, every individual except the strongest emitted a subordination signal to at least one other individual (41), so in the model, they should have finite thresholds. We restricted the thresholds that individuals in our model could use to be between 0.5 and 2. Our results depended on relative changes in threshold values, not the absolute values of thresholds, so our results should hold regardless of the actual range of thresholds allowed.

In the neural literature, it is usually assumed that the brain can evaluate the difference between two variables, which indicates the relative strength of evidence for each option. In most models of neural decision-making, it is assumed that, if  $Y = X_1 - X_2$ , then the brain decides on  $X_1$  when  $Y$  is large and positive and on  $X_2$  when  $Y$  is large and negative (2, 30, 31). Again, there are two thresholds,  $T_1$  and  $T_2$ , such that if  $Y > T_1$ , then the decision is that A has higher ability and if  $Y < -T_2$ , then the decision is that B has higher ability. This reduces the number of relevant variables from two to one.

Downloaded from <http://advances.sciencemag.org/> on March 2, 2021

**Table 1. Variables in the model and their interpretations in social and neural systems.**

Variable	Model definition	Social interpretation	Neural interpretation
$a_i$	Value of component $i$	Fighting ability of individual $i$	How much the part of the visual stimulus to which population $i$ responds is present in the input.
$b$	Change due to new evidence		
$c$	Strength of input	Degree of asymmetry between opponents	Coherence of moving dots
$l$	Leak rate		
$r$	Rate at which evidence appears	Rate at which fights occur	Rate at which dots appear
$T_1$ and $T_2$	Decision thresholds		
$w_1$	Error rate weight	Cost to both individuals when the stronger individual incorrectly signals, for example, because the relationship is less stable	Cost of making incorrect decision, imposed by experimenter
$w_2$	Decision time weight	Cost of prolonged fighting, for example, injury and opportunity cost of time	Penalty for taking a long time
$w_3$	Decision preference weight, captures strength of conflict	Cost of emitting a signal and agreeing to be subordinate, for example, because of limited access to resources	
$X_1$ and $X_2$	Decision variables	Evidence accumulated about relative dominance	Firing rates of neural populations

In social systems, the one-dimensional simplification implies a third party evaluating the difference in the evidence each individual has accumulated, which is not realistic. Hence, we used the two-dimensional system. In section S4 and in fig. S1, we show that, with our assumption that  $X_1(0) = X_2(0) = 0$ , the two-dimensional and one-dimensional processes are equivalent (we also show that, if  $X_1(0) - X_2(0)$  is not zero, then the two processes are nearly equivalent). Our results therefore also apply when the one-dimensional simplification is used.

### Assessing pairwise decisions

Once the decision thresholds  $T_1$  and  $T_2$  have been specified, the SDEs determine the probability that each individual's decision variable will reach its threshold before the other, that is, the probability that either of the two individuals would be the one to emit the subordination signal. They also determine the expected time it would take for one of the individuals to signal. We show in section S6 that each of these quantities—the probability of each individual signaling and the expected time until a signal is emitted—satisfies a partial differential equation that depends on the decision thresholds and the parameters of the model.

A “correct” decision is one that results in the weaker individual emitting the subordination signal. If individual  $i$  is the weaker individual and individual  $j$  is the stronger individual, then the probability that  $j$  incorrectly signals is the error rate (ER) of their decision. Each animal would prefer to receive the subordination signal. Thus, the error rate is equal to the probability of  $i$ 's decision preference being reached (DP<sub>*i*</sub>), whereas the probability of  $j$  receiving the signal and  $j$ 's decision preference being reached is 1 minus this quantity (DP<sub>*j*</sub> = 1 – DP<sub>*i*</sub>). The decision time (DT) for the pair is the expected time until either of the two individuals signals. These quantities determine each individual's utility from the decision-making process. Winning any particular fight can give an individual access to resources, which may improve its fitness. Once an individual emits the subordination signal, it agrees to be subordinate to its opponent and cedes access to resources when it comes into conflict with that in-

dividual in the future. We were interested in the formation of these stable relationships, so we focused on the benefits of receiving the subordination signal, rather than transient benefits from winning any particular fight.

In models of individual conflict and dominance relationships, it is often assumed that both individuals prefer to be dominant, regardless of whether they are in fact stronger [for example, (33, 51, 52)]. However, a correct subordination signal from a weaker to a stronger individual has been shown to lead to a more stable and affiliative relationship than would be the case had the signal been withheld (41, 53). Hence, error rate needs to be considered. Conversely, in models of neural decision-making, error rate, rather than decision preference, is assumed to be the currency driving decision-making strategies (30). Our model connects these two bodies of work by allowing for both error rate and decision preference to affect individual strategies.

To describe trade-offs between error rate, decision time, and preference, we quantified the utility of the decision process by introducing three weights,  $w_1$ ,  $w_2$ , and  $w_3$  such that  $w_1 + w_2 + w_3 = 1$ . These weights describe how the three quantities are prioritized. For individual  $i$  interacting with individual  $j$ , we defined  $i$ 's utility to be

$$U_{ij} = w_1(1 - \text{ER}) + w_2(1 - \text{DT}) + w_3\text{DP}_i$$

Each individual wants to maximize its utility. Note that it is impossible to minimize both ER and DT because waiting longer and accumulating more evidence will help the pair reach a decision more accurately. The weight  $w_2$  can be interpreted as the cost of fighting because, when  $w_2$  is higher, the time spent fighting until a decision is reached is more costly. The weight  $w_3$  captures the benefit from being the dominant individual in a pair and the extent to which each individual perceives agreeing to be subordinate to be more costly than continued fighting. The higher  $w_3$  is, the more stubborn the individuals will be about waiting for their desired outcome. It is impossible for both individuals to maximize DP<sub>*i*</sub>, so  $w_3$  indicates the strength of the conflict of interests between individuals.

We show outputs of the model that depend on the three weights,  $w_1$ ,  $w_2$ , and  $w_3$ , on a triangular simplex, as in Fig. 1E. In each such figure, points in the lower left corner indicate that  $w_1$  is high; points in the upper corner indicate that  $w_2$  is high; and points in the lower right corner indicate that  $w_3$  is high. Points on the edge of the triangle indicate that one of the three weights is equal to 0. For example, points on the bottom edge of the triangle indicate that  $w_2 = 0$ . Moving left means that  $w_1$  is increasing and  $w_3$  is decreasing; moving right means the opposite; and moving up means that  $w_2$  increases.

### Nash equilibrium thresholds

We assumed that if two individuals have equal fighting ability, then any fight between them is a toss-up, that is,  $c = 0.5$ , and that as the difference in their abilities increases, the probability of the stronger individual winning approaches 1. We assigned each individual in the model a fighting ability  $a_i$ . Then, we assumed  $c_{ij} = \frac{\exp(a_i - a_j)}{\exp(a_i - a_j) + 1}$ . We further assumed that each individual has the same decision threshold for all the decision processes with each of its peers. Given those thresholds and the importance of error rate, decision time, and decision preference ( $w_1$ ,  $w_2$ , and  $w_3$ , respectively), each individual  $i$  has a utility  $U_{ij}$  from its decision process with individual  $j$  and a total utility given by the average of these,  $\langle U_{ij} \rangle_j$ . For each set of abilities  $\{a_1, \dots, a_N\}$ , we found the Nash equilibrium thresholds  $\{T_1, \dots, T_N\}$  such that no individual has an incentive to choose another threshold to improve its total utility. Because the Nash thresholds depend on the abilities  $\{a_1, \dots, a_N\}$ , we drew a set of abilities from a uniform distribution 1000 times and found the Nash thresholds for each set.

To show how the weights encoding error rate, decision time, and decision preference ( $w_1$ ,  $w_2$ , and  $w_3$ , respectively) affect decision-making strategies, below we reported the average Nash threshold for an individual with the  $i$ th highest ability, for each rank  $i = 1, \dots, N$ , where 1 is the strongest and  $N$  is the weakest. However, we used the actual Nash thresholds, not their average, in our analyses. Specifically, to show how the weights affect how accurately individuals can make decisions, we considered 1000 groups of individuals using Nash thresholds and take the average error rate across all  $N \times (N - 1)/2$  pairs and all 1000 iterations. Similarly, we found the average decision time across all  $N \times (N - 1)/2$  pairs in all 1000 iterations of the model, where the individuals in each group used Nash thresholds.

We were interested in understanding the factors influencing whether the collective computation would produce a right-skewed distribution of power. To avoid building in a distribution with a long tail, we used a uniform distribution of fighting abilities because it has short tails. Our results did not change when we used a normal distribution of fighting abilities instead of a uniform distribution of fighting abilities (compare figs. S2 to S4 to Figs. 2 to 4). Our procedure for finding the Nash equilibrium thresholds is provided in section S7.

### Modeling collective computation at group level

We used the pairwise decision model to create a directed signaling network among all pairs of individuals. We drew the direction of the signal between each pair according to these probabilities. Once an individual decided to emit the subordination signal, it continued to do so regularly over a long period of time. Eventually, the relationship can reverse if an individual comes to perceive that it could win against an individual to which it has previously signaled (2). However, this would take longer than the period of time that we considered, and we do not allow for this possibility. To describe the accumulation of signals over

time, at each point in time  $t$ , we defined the weight of the edge from  $i$  to  $j$  to be

$$0 \text{ if } j \text{ signals to } i \text{ or } t < DT_{ij}$$

$$t \text{ if } i \text{ signals to } j \text{ and } t > DT_{ij}$$

where, as above,  $DT_{ij}$  is the time it takes the pair to reach a decision.

Individuals can estimate the degree of consensus within the group by observing (some subset of) the network of subordination signals emitted between all pairs of individuals. For these estimates to be good indicators of the individuals' true abilities, two things need to happen: First, the edges in the network need to be accurate indicators of the differences in abilities between pairs of individuals, and second, the algorithm that individuals use to quantify consensus needs to aggregate the relevant information appropriately. In previous work, we identified a set of algorithms that can be used to compute consensus about node state in any network (6, 27). In our system, consensus in the subordination signaling network reflected how much power a node was perceived to have (6, 27), as described above. For the sake of brevity, we considered four of the best-performing algorithms here.

Of these, the simplest algorithm was the unweighted in-degree of a node, that is, the number of individuals that have signaled to each individual. We also considered weighted in-degree, the sum of all signals an individual receives. The third algorithm that we considered was the entropy of the distribution of the numbers of signals that each individual receives, which gave a coarse measurement of the uniformity of all of the opinions in the group about a focal individual. The fourth was the eigenvector centrality of the network, which measured how central each node was in the global structure of the network (27, 54). We generated 1000 directed signaling networks as described above. To each of these networks, we applied the four network metrics, each of which generated a DSP.

### Assessing collective computation

We assessed the quality of the output of the collective computation—the DSP—by assessing the accuracy and skewness of the resulting distribution. We operationally defined an accurate distribution as one having high mutual information with the underlying distribution of fighting abilities. This allowed us to study how conflict influences DSP accuracy, as well as the relative ability of each algorithm to recover an accurate DSP [the question of how the system “knows” the desired output has been achieved, called the “halting problem” in computer science, is discussed by Flack (7, 17) and in section S8]. We measured the mutual information between the consensus scores produced by a given algorithm and the underlying distribution of fighting abilities, over all 1000 iterations of the model (see section S9 for details). We measured the skewness of the set of consensus scores given by each algorithm and took the average over the 1000 iterations of the model. A schematic of the whole model of collective computation, from pairwise decision-making to consensus computation, is provided in fig. S5.

### SUPPLEMENTARY MATERIALS

Supplementary material for this article is available at <http://advances.sciencemag.org/cgi/content/full/4/1/e1603311/DC1>  
 section S1. Background on models of decision-making  
 section S2. Study system  
 section S3. Skewness of DSP



section S4. Dimensionality and initial conditions  
 section S5. Analogous model of neural decision-making  
 section S6. Derivation of partial differential equations for decision time, error rate, and probability of reaching decision preference  
 section S7. Nash equilibria  
 section S8. A notion of correctness for biological computation  
 section S9. Calculation of mutual information  
 section S10. Most informative measures of consensus  
 section S11. Tuning waiting costs  
 section S12. Comparison of our model to previous studies of animal conflict  
 section S13. War of attrition  
 section S14. Supplementary table  
 section S15. Supplementary figures  
 table S1. Examples of collective computation.  
 fig. S1. Error rate decreases as decision time increases, as long as the initial conditions are not biased toward the correct decision.  
 fig. S2. The mutual information of the power scores computed by a group using Nash thresholds increases as the weight given to decision preference increases, as long as there are nonzero waiting costs.  
 fig. S3. The average skewness of the distribution of eigenvector centrality is maximized at intermediate waiting costs.  
 fig. S4. The best measure of consensus in the decision network depends on the average error rate and the types of errors being made.  
 fig. S5. Schematic of the model.  
 fig. S6. The error rate of a group using Nash thresholds decreases as the weight given to decision preference increases, regardless of the size of the group.  
 fig. S7. Pairs with similar and high abilities always take as long or longer to make a decision than any other pairs do.  
 fig. S8. The mutual information of each consensus algorithm is a decreasing function of the average pairwise error rate.  
 fig. S9. The average skewness of the distribution of unweighted in-degree is maximized at intermediate waiting costs.  
 fig. S10. The average skewness of the distribution of consensus scores from each measure is maximized at intermediate waiting costs.  
 fig. S11. The average skewness of the distribution of consensus scores from each measure is maximized at intermediate waiting costs.  
 fig. S12. When a pair of animals have equal fighting abilities,  $c = 0.5$ , there are asymmetric Nash equilibrium thresholds.  
 References (55–59)

## REFERENCES AND NOTES

- J. I. Gold, M. N. Shadlen, Neural computations that underlie decisions about sensory stimuli. *Trends Cogn. Sci.* **5**, 10–16 (2001).
- S. Feng, P. Holmes, A. Rorie, W. T. Newsome, Can monkeys choose optimally when faced with noisy stimuli and unequal rewards? *PLOS Comput. Biol.* **5**, e1000284 (2009).
- B. L. Bassler, How bacteria talk to each other: Regulation of gene expression by quorum sensing. *Curr. Opin. Microbiol.* **2**, 582–587 (1999).
- M. B. Miller, B. L. Bassler, Quorum sensing in bacteria. *Annu. Rev. Microbiol.* **55**, 165–199 (2001).
- J. C. Flack, F. B. M. de Waal, D. C. Krakauer, Social structure, robustness, and policing cost in a cognitively sophisticated species. *Am. Nat.* **165**, E126–E139 (2005).
- J. C. Flack, D. C. Krakauer, Encoding power in communication networks. *Am. Nat.* **168**, E87–E102 (2006).
- J. C. Flack, Multiple time-scales and the developmental dynamics of social systems. *Philos. Trans. R. Soc. B.* **367**, 1802–1810 (2012).
- I. D. Couzin, J. Krause, R. James, G. D. Ruxton, N. R. Franks, Collective memory and spatial sorting in animal groups. *J. Theor. Biol.* **218**, 1–11 (2002).
- J. E. Herbert-Read, A. Perna, R. P. Mann, T. M. Schaefer, D. J. T. Sumpter, A. J. W. Ward, Inferring the rules of interaction of shoaling fish. *Proc. Natl. Acad. Sci. U.S.A.* **108**, 18726–18731 (2011).
- Y. Katz, K. Tunström, C. C. Ioannou, C. Huepe, I. D. Couzin, Inferring the structure and dynamics of interactions in schooling fish. *Proc. Natl. Acad. Sci. U.S.A.* **108**, 18720–18725 (2011).
- S. B. Rosenthal, C. R. Twomey, A. T. Hartnett, H. S. Wu, I. D. Couzin, Revealing the hidden networks of interaction in mobile animal groups allows prediction of complex behavioral contagion. *Proc. Natl. Acad. Sci. U.S.A.* **112**, 4690–4695 (2015).
- F. Esponda, D. M. Gordon, Distributed nestmate recognition in ants. *Proc. R. Soc. B Biol. Sci.* **282**, 20142838 (2015).
- G. Punzo, G. F. Young, M. Macdonald, N. E. Leonard, Using network dynamical influence to drive consensus. *Sci. Rep.* **6**, 26318 (2016).
- J. C. Flack, D. C. Krakauer, Challenges for complexity measures: A perspective from social dynamics and collective social computation. *Chaos* **21**, 037108 (2011).
- M. Kearns, Experiments in social computation. *Commun. ACM* **55**, 56–67 (2012).
- A. M. Hein, S. B. Rosenthal, G. I. Hagstrom, A. Berdahl, C. J. Torney, I. D. Couzin, The evolution of distributed sensing and collective computation in animal populations. *Elife* **4**, e10955 (2015).
- J. C. Flack, Life's information hierarchy, in *From Matter to Life*, S. I. Walker, P. C. W. Davies, G. F. R. Ellis, Eds. (Cambridge Univ. Press, 2017).
- A. K. Shaw, I. D. Couzin, Migration or residency? The evolution of movement behavior and information usage in seasonal environments. *Am. Nat.* **181**, 114–124 (2013).
- I. D. Couzin, J. Krause, N. R. Franks, S. A. Levin, Effective leadership and decision-making in animal groups on the move. *Nature* **433**, 513–516 (2005).
- A. Berdahl, P. A. H. Westley, S. A. Levin, I. D. Couzin, T. P. Quinn, A collective navigation hypothesis for homeward migration in anadromous salmonids. *Fish. Fish.* **17**, 525–542 (2014).
- C. J. Torney, T. Lorenzi, I. D. Couzin, S. A. Levin, Social information use and the evolution of unresponsiveness in collective systems. *J. R. Soc. Interface* **12**, 20140893 (2015).
- L. Chittka, P. Skorupski, N. E. Raine, Speed-accuracy tradeoffs in animal decision-making. *Trends Ecol. Evol.* **24**, 400–407 (2009).
- B. C. Daniels, J. C. Flack, D. C. Krakauer, Dual coding theory explains biphasic collective computation in neural decision-making. *Front. Neurosci.* **11**, 313 (2017).
- J. M. Smith, G. R. Price, The logic of animal conflict. *Nature* **246**, 15–18 (1973).
- J. M. Smith, *Evolution and the Theory of Games* (Cambridge Univ. Press, 1982).
- B. Nalebuff, J. Riley, Asymmetric equilibria in the war of attrition. *J. Theor. Biol.* **113**, 517–527 (1985).
- E. R. Brush, D. C. Krakauer, J. C. Flack, A family of algorithms for computing consensus about node state from network data. *PLOS Comput. Biol.* **9**, e1003109 (2013).
- P. Cisek, Making decisions through a distributed consensus. *Curr. Opin. Neurobiol.* **22**, 927–936 (2012).
- R. Bogacz, E. Brown, J. Moehlis, P. Holmes, J. D. Cohen, "Optimizing reward rate in two alternative choice tasks: Mathematical formalism" (Technical Report 04-01, Center for the Study of Brain, Mind, and Behavior, 2004).
- E. Brown, J. Gao, P. Holmes, R. Bogacz, M. Gilzenrat, J. D. Cohen, Simple neural networks that optimize decisions. *Int. J. Bifurcat. Chaos* **15**, 803–826 (2005).
- R. Bogacz, E. Brown, J. Moehlis, P. Holmes, J. D. Cohen, The physics of optimal decision making: A formal analysis of models of performance in two-alternative forced-choice tasks. *Psychol. Rev.* **113**, 700–765 (2006).
- D. Pais, P. M. Hogan, T. Schlegel, N. R. Franks, N. E. Leonard, J. A. R. Marshall, A mechanism for value-sensitive decision-making. *PLOS ONE* **8**, e73216 (2013).
- A. J. De Froment, "Fighting for information: Decision-making, animal contests and the emergence of social hierarchy," thesis, Princeton University (2010).
- R. Ratcliff, P. L. Smith, S. D. Brown, G. McKoon, Diffusion decision model: Current issues and history. *Trends Cogn. Sci.* **20**, 260–281 (2016).
- M. N. Shadlen, W. T. Newsome, Neural basis of a perceptual decision in the parietal cortex (area LIP) of the Rhesus Monkey. *J. Neurophysiol.* **86**, 1916–1936 (2001).
- S. A. Frank, Perspective: Repression of competition and the evolution of cooperation. *Evolution* **57**, 693–705 (2003).
- A. Burt, R. Trivers, *Genes in Conflict: The Biology of Selfish Genetic Elements* (Harvard Univ. Press, 2008).
- J. C. Flack, D. C. Krakauer, F. B. M. de Waal, Robustness mechanisms in primate societies: A perturbation study. *Proc. R. Soc. B Biol. Sci.* **272**, 1091–1099 (2005).
- D. T. Bishop, C. Cannings, J. M. Smith, The war of attrition with random rewards. *J. Theor. Biol.* **74**, 377–388 (1978).
- D. C. Krakauer, K. Page, J. Flack, The immuno-dynamics of conflict intervention in social systems. *PLOS ONE* **6**, e22709 (2011).
- J. C. Flack, F. de Waal, Context modulates signal meaning in primate communication. *Proc. Natl. Acad. Sci. U.S.A.* **104**, 1581–1586 (2007).
- S. C. Stearns, The selection-arena hypothesis, in *The Evolution of Sex and its Consequences*, S. C. Stearns, Ed. (Birkhäuser, 1987), pp. 337–349.
- D. C. Krakauer, A. Mira, Mitochondria and germ-cell death. *Nature* **400**, 125–126 (1999).
- B. C. Daniels, D. C. Krakauer, J. C. Flack, Control of finite critical behaviour in a small-scale social system. *Nat. Commun.* **8**, 14301 (2017).
- X. J. Wang Probabilistic decision making by slow reverberation in cortical circuits. *Neuron* **36**, 955–968 (2002).
- K.-F. Wong, X.-J. Wang, A recurrent network mechanism of time integration in perceptual decisions. *J. Neurosci.* **26**, 1314–1328 (2006).
- G. M. Edelman, Neural Darwinism: Selection and reentrant signaling in higher brain function. *Neuron* **10**, 115–125 (1993).
- B. Thierry, M. Singh, W. Kaumanns, *Macaque Societies: A Model for the Study of Social Organization* (Cambridge Univ. Press, 2004).

49. J. C. Flack, M. Girvan, F. B. M. de Waal, D. C. Krakauer, Policing stabilizes construction of social niches in primates. *Nature* **439**, 426–429 (2006).
50. D. T. Gillespie, The chemical Langevin equation. *J. Chem. Phys.* **113**, 297–306 (2000).
51. C. K. Hemelrijk, An individual-orientated model of the emergence of despotic and egalitarian societies. *Proc. R. Soc. B* **266**, 361–369 (1999).
52. M. Enquist, O. Leimar, Evolution of fighting behaviour: Decision rules and assessment of relative strength. *J. Theor. Biol.* **102**, 387–410 (1983).
53. S. Preuschoft, Power and communication, in *Macaque Societies: A Model for the Study of Social Organization*, B. Thierry, M. Singh, W. Kaumanns, Eds. (Cambridge Univ. Press, 2004).
54. S. Allesina, M. Pascual, Googling food webs: Can an eigenvector measure species' importance for coextinctions? *PLOS Comput. Biol.* **5**, e1000494 (2009).
55. J. O. Caldecott, *An Ecological and Behavioural Study of the Pig-Tailed Macaque*. (S. Karger, 1986).
56. F. B. M. de Waal, L. M. Luttrell, The formal hierarchy of rhesus macaques: An investigation of the bared-teeth display. *Am. J. Primatol.* **9**, 73–85 (1985).
57. J. C. Flack, F. B. M. de Waal, Dominance style, social power, and conflict management: A conceptual framework, in *Macaque Societies: A Model for the Study of Social Organization*, B. Thierry, M. Singh, W. Kaumanns, Eds. (Cambridge Univ. Press, 2004), chap. 8, 157–181.
58. S. Preuschoft, Power and communication, in *Macaque Societies: A Model for the Study of Social Organization*, B. Thierry, M. Singh, W. Kaumanns, Eds. (Cambridge Univ. Press, 2004), pp. 56–60.
59. C. Gardiner, *Stochastic Methods: A Handbook for the Natural and Social Sciences* (Springer-Verlag, 2009).

**Acknowledgments:** We thank B. Daniels, C. Ellison, P. Poon, and E. Lee for helpful discussion. E.R.B. thanks S. Levin for helpful discussion and comments on the model during development. We also thank two anonymous reviewers for their comments, which have greatly improved the paper. **Funding:** This research was supported by two grants to the Santa Fe Institute from the John Templeton Foundation for the study of complexity, a grant to the Santa Fe Institute from the John Templeton Foundation to study the mind-brain problem, and a grant from the Templeton foundation to study niche construction (JTF number 60501, St. Andrews subaward number 13337), and by ARO (Army Research Office) contract W911NF-13-1-0340. J.C.F. acknowledges the Proteus Foundation. E.R.B. acknowledges support from NIH training grant 5T32HG003284. **Author contributions:** E.R.B. developed and analyzed the model and wrote the paper. D.C.K. helped develop the model and write the paper. J.C.F. designed the study and collected the data motivating the model, helped develop the model, and wrote the paper. **Competing interests:** The authors declare that they have no competing interests. **Data and materials availability:** All data about the model needed to evaluate the conclusions in the paper are present in the paper and/or the Supplementary Materials. Additional data related to this paper and the code for the model may be requested from the authors.

Submitted 29 December 2016

Accepted 14 December 2017

Published 17 January 2018

10.1126/sciadv.1603311

**Citation:** E. R. Brush, D. C. Krakauer, J. C. Flack, Conflicts of interest improve collective computation of adaptive social structures. *Sci. Adv.* **4**, e1603311 (2018).

## Conflicts of interest improve collective computation of adaptive social structures

Eleanor R. Brush, David C. Krakauer and Jessica C. Flack

*Sci Adv* 4 (1), e1603311.

DOI: 10.1126/sciadv.1603311

### ARTICLE TOOLS

<http://advances.sciencemag.org/content/4/1/e1603311>

### SUPPLEMENTARY MATERIALS

<http://advances.sciencemag.org/content/suppl/2018/01/12/4.1.e1603311.DC1>

### REFERENCES

This article cites 47 articles, 5 of which you can access for free  
<http://advances.sciencemag.org/content/4/1/e1603311#BIBL>

### PERMISSIONS

<http://www.sciencemag.org/help/reprints-and-permissions>

Use of this article is subject to the [Terms of Service](#)

---

*Science Advances* (ISSN 2375-2548) is published by the American Association for the Advancement of Science, 1200 New York Avenue NW, Washington, DC 20005. The title *Science Advances* is a registered trademark of AAAS.

Copyright © 2018 The Authors, some rights reserved; exclusive licensee American Association for the Advancement of Science. No claim to original U.S. Government Works. Distributed under a Creative Commons Attribution NonCommercial License 4.0 (CC BY-NC).