

Social hierarchy shapes foraging decisions

Lisa Blum Moysé and Ahmed El Hady*

Centre for the Advanced Study of Collective Behaviour, University of Konstanz, Konstanz, Germany

(Dated: March 5, 2025)

Social foraging is a widespread form of animal foraging in which groups of individuals coordinate their decisions to exploit resources in the environment. Animals show a variety of social structures from egalitarian to hierarchical. In this study, we examine how different forms of social hierarchy shape foraging decisions. We developed a mechanistic analytically tractable model to study the underlying processes of social foraging, tying the microscopic individual to the macroscopic group levels. Based on a stochastic evidence accumulation framework, we developed a model of patch-leaving decisions in a large hierarchical group with leading and following individuals. Across a variety of information sharing mechanisms, we were able to analytically quantify emergent collective dynamics. We found that follower-leader dynamics through observations of leader movements or through counting the number of individuals in a patch confers, for most conditions, a benefit for the following individuals by increasing their accuracy in inferring patch richness. On the other hand, misinformation, through the communication of false beliefs about food rewards or patch quality, shows to be detrimental to following individuals, but paradoxically may lead to increased group cohesion. In an era where there is a huge amount of animal foraging data collected, our model provides a systematic way to conceptualize and understand those data by uncovering hidden mechanisms underlying social foraging decisions.

Introduction— Foraging is a fundamental decision-making behavior crucial for animal survival [1]. Often, animals forage socially and adopt a wide variety of social organizations, from the less to the more hierarchical. Some examples of animals adopting hierarchical structures are eusocial insects [2], birds [3], mice [4] or baboons [5]. While social hierarchy shapes a wide variety of behavioral facets [6], interactions with social foraging are not well understood in quantitative mechanistic terms. To investigate the underlying processes of social foraging, several theories and quantitative models have been developed, such as the ideal free distribution [7], the marginal value theorem for groups [8], agent-based models [9], reinforcement learning models [10][11], game-theoretic approaches [12][13], or Bayesian models [14]. However, these approaches are often coarse-grained, implying general interaction rules, or are too complex to be treated analytically. Another limitation is that uncertainty and noise in decision processes are often not taken into account, which can strongly shape the agents' patch departure statistics while foraging [15][16]. This situation calls for the development of an analytically tractable framework, including inherent stochasticity, to study the mechanisms underlying social foraging across a variety of social and environmental conditions. Models developed within this framework should also be able to be fitted to experimental data, to unravel mechanisms underlying the decision-making of animals while they forage socially. In this letter, we introduce a new framework to meet these expectations.

The new model developed in this paper is built on a widespread stochastic decision-making framework based

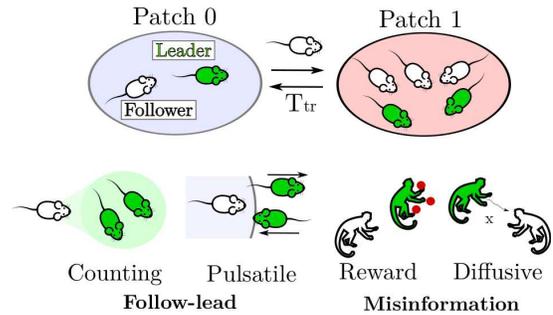


FIG. 1. (Top) Schema of a hierarchical foraging group in a two-patches environment, with a travel time T_{tr} . (Bottom) Followers collect social information from leaders in different ways, to follow leaders (counting, pulsatile) or be misinformed about patch qualities (reward, diffusive).

on the evidence accumulation process, drift-diffusion models, which can be applied to patch-leaving tasks [15]. So far, such mechanistic models have focused only on one [15] or two cooperative foragers [17]. We have previously extended this modeling approach to study coupled agents in an egalitarian group for a two-patches environment [18].

In this letter, we provide a quantitative analytically tractable model to unravel the mechanisms of how social hierarchy shapes social foraging decisions strategies. For simplicity, we focus on a hierarchical organization in a two-patches environment with leading and following agents. We analytically derive strategies under a variety of information sharing mechanisms, gathered in two categories, follow-lead and misinformation (see Fig. 1). The second category, also known as tactical deception, is more likely to occur mainly in primates [19]. Throughout this letter, numerical simulations and analytical analysis enable us to characterize how these different information

* Also at Department of Collective Behavior, Max Planck Institute of Animal Behavior, Konstanz, Germany

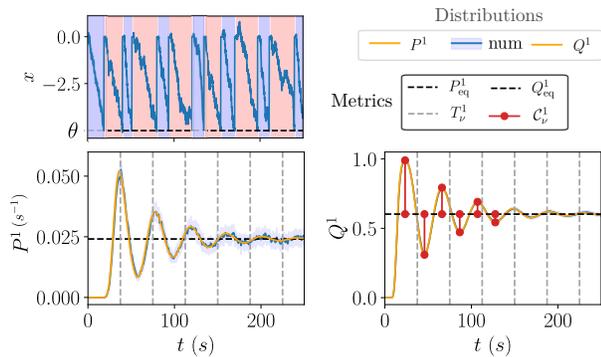


FIG. 2. Dynamics and quantification of the collective foraging task. (Top) Foraging agents accumulate evidence in a patch of food (blue: patch 0, red: patch 1), and leave it for the next patch when their decision variable x reaches the threshold θ . (Bottom) Distributions of leaving times $P^1(t)$ (Left) and fraction of agents $Q^1(t)$ (Right) in patch 1 show damped oscillations converging to an equilibrium state. Blue lines correspond to numerical simulations, and orange lines to theoretical predictions. The shaded area corresponds to standard deviation. The distributions can be quantified by their equilibrium values, patch residence time and measure of cohesion. Parameters are $\theta = -5$ and $\alpha = 1 \text{ s}^{-1}$, $B = 0.05 \text{ s}^{-1}$, $p^0 = 0.6$, $p^1 = 0.8$, $\tau_d = 0.1 \text{ s}$, $T_{\text{tr}} = 2 \text{ s}$, $N_L = 2500$, 20 simulations.

sharing mechanisms lead to various collective dynamics. We found that each information sharing mechanism shapes decision strategies in a specific way. For example, misinformation through false belief sharing decreases patch richness accuracy estimation, while increasing group cohesion. We think that this model can provide a benchmark for evolutionary comparison across species, opening up the ability for quantitative behavioral evolutionary studies.

Social foraging models— Agents move between two patches identified by a number k . $k = 0$ is the initial patch, and $k = 1$ is the highest quality patch. The decision to leave a patch for another one is taken through the underlying process of evidence accumulation. Groups are organized hierarchically, with a subgroup of leading (L) and following (F) agents. Leading individuals do not consider the information coming from other agents, they act like non-interacting agents. Following individuals have non-zero coupling parameters, and only follow leading individuals. In addition, they receive less food.

x_i , the decision variable of a foraging agent i in a patch k , evolves according to the following stochastic differential equation

$$dx_i(t) = (r_i(t) - \alpha)dt + \sum_{j \in L} c_{ij}(t - \tau_d)dt + \sqrt{2B}dW_i(t) \quad (1)$$

with the initial condition $x_i(0) = 0$ for all i s. The forager leaves a patch when x_i reaches the threshold θ ,

and x_i is reset to zero. The travel time between patches is T_{tr} . The drift-diffusion process is illustrated in Fig. 2. The decision variable evolves following an evidence accumulation process, where α represents the cost associated with foraging, and $r_i(t)$ the food rewards. Every time step Δ_r , an L agent has a probability p^k of receiving a food reward $r_i = 1 \text{ s}^{-1}$. Throughout this letter, we fix $\Delta_r = dt$. The following individuals have a reward probability βp^k , with $\beta \leq 1$. p^k is constant here, this is a non-depleting environment. $W_i(t)$ is the standard Wiener process, B represents the noise amplitude. For F agents, information sharing from a leading j to a following i individual in the same patch is expressed by the coupling term c_{ij} , with a delay τ_d . For L agents, this term is equal to zero. The different coupling types are illustrated in Fig. 1 and detailed below.

Follow-lead dynamics. Two different mechanisms may underlie the following dynamics. First, F agents could perceive the beliefs of L agents through pulses of information, corresponding to an agent's arrival (κ_a) or departure (κ_d) in a patch. This is the *pulsatile coupling*. If an L agent j leaves, $c_{ij}(t) = -\kappa_d(N_L dt)^{-1}\delta(x_j - \theta)$. If an L agent j leaves to join the current patch of a F agent i , $c_{ij}(t) = \kappa_a(N_L dt)^{-1}\delta(x_j - \theta)$.

Instead of being sensitive to the departure or arrival times of L individuals, F agents could also perceive how many L individuals of their group are in the same patch. This is the *counting coupling*. For clarity reasons, we write this term without the sum in eq. (1), $c_i(t) = \kappa_c(n_L^k(t)/N_L - \eta)$. If many L agents are in the same patch, $n_L^k(t) > \eta$, so $c_i > 0$, i.e. the F agent accumulates evidence to stay in the patch. Inversely for $n_L^k(t) < \eta$.

Misinformation L agents may communicate false information to mislead F individuals. L agents may continuously share their beliefs about patch quality, this is the *diffusive coupling*. $c_{ij}(t) = \kappa_{\text{diff}}^k N_L^{-1} x_{L,j}(t)$. Lying L agents may modify their communicated decision variable, to misinform F agents. If $\kappa_{\text{diff}}^1 \geq \kappa_{\text{diff}}^0$, this incites F agents to leave the best patch.

False information can also come from a wrong communicated number of catches that L agents in the same patch get. $c_{ij}(t) = \kappa_r N_L^{-1} r_j(t)$. κ_r corresponds to the *reward coupling* strength. The false communicated reward probability is ζ^k . If $\zeta^1 \leq \zeta^0$, this incites F agents to stay in the worst patch.

Pulsatile, diffusive and reward coupling results with normalization terms $n_L^k(t)$ instead of N_L are presented in the appendix.

Simulations details. The stochastic differential equation (1) of the decision variable is solved numerically with the Euler method. The parameters $\theta = -5$, $\alpha = 1 \text{ s}^{-1}$, $B = 0.05 \text{ s}^{-1}$, $p^0 = 0.6$, $p^1 = 0.9$, $\tau_d = 0.1 \text{ s}$, $T_{\text{tr}} = 0.5 \text{ s}$, $\beta = 0.3$ remain constant throughout the figures. The coupling parameters remain small enough

Follow-lead

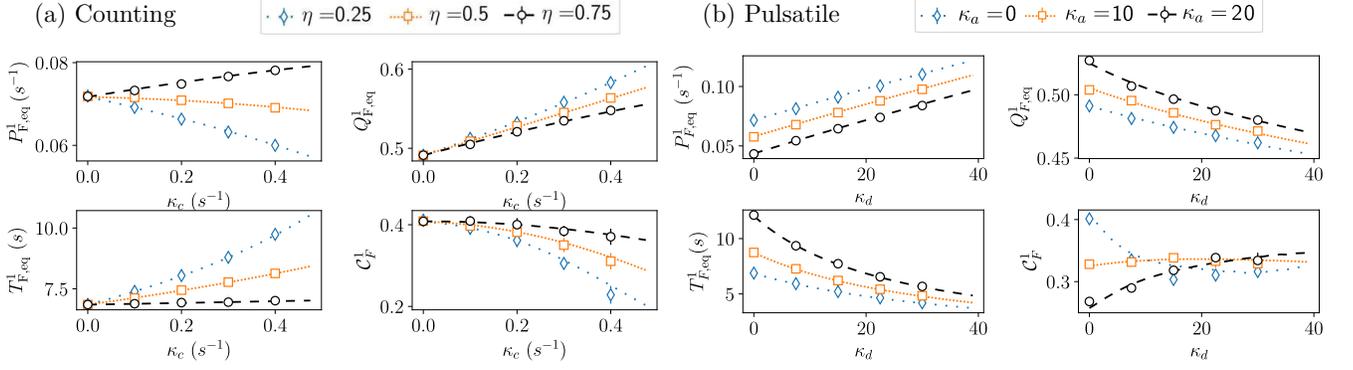


FIG. 3. Follow-lead dynamics of the best patch metrics: equilibrium leaving time distribution $P_{F,\text{eq}}^1$, equilibrium accuracy $Q_{F,\text{eq}}^1$, equilibrium patch residence time $T_{F,\text{eq}}^1$, and cohesion C_F^1 . (a) Counting coupling metrics as a function of the coupling strength parameter κ_c for different η values. (b) Pulsatile coupling metrics as a function of the departure strength parameter κ_d for different arrival strength parameter κ_a values. Lines correspond to theoretical predictions and dots to numerical simulations. Error bars correspond to the standard deviation.

compared to the asocial drift term. The duration of the simulation is equal to 250 s. Cohesion is quantified on the $\epsilon = 6$ first extrema. The L and F group sizes are $N_L = 200$ and $N_F = 200$, with $dt = 0.001$ s, 20 simulations, for all couplings except for pulsatile, with $N_L = 60000$ and $N_F = 2000$, 2 simulations. A larger L group size was needed for this condition, to get a correct match between theory and simulations. A larger time step $dt = 0.02$ s was implemented to keep tractable computation times. In fact, leaving probability densities $P_L^k(t)$ are included in the drift terms of pulsatile coupling, and they are more noisy than the fraction of agents $Q_L^k(t)$. Equilibrium metrics from numerical simulations were computed through the averaged last 100 s of the process.

The distributions and metrics used to quantify the process are first described, and then applied to the follow-lead and misinformation dynamics.

Distributions and metrics— The probability density of a group X ($X = \{L, F\}$) to leave a patch k at time t is given by

$$P_X^k(t) = \sum_{\nu=1}^{\infty} P_{X,\nu}^k(t - (2\nu - 2 + k)T_{\text{tr}}) \quad [18], \text{ with}$$

$$P_{X,\nu}^k(t) = \underbrace{(\Psi_{X,1}^k * \dots * \Psi_{X,1}^0 * \Psi_{X,1}^1 * \Psi_{X,1}^0)}_{2\nu-1+k}(t) \quad (2)$$

For $t \leq (2\nu - 2 + k)T_{\text{tr}}$, the convolution is equal to zero.

With $\Psi_{X,\nu}^k(t) = \frac{-\nu\theta}{\sqrt{4\pi Bt^3}} \exp\left(\frac{-(\nu\theta + \tilde{\alpha}_X^k t)^2}{4Bt}\right)$ [20] for $\nu \geq 1$, and $\Psi_{X,0}^k(t) = 1$ for all t . $\tilde{\alpha}_X^k$ is the effective drift term.

For L agents, the effective drift term is $\tilde{\alpha}_L^k = \alpha - p^k$. The mean rate is equal to p^k (with unit s^{-1}). For clarity reasons, we write directly p^k in effective drifts. Since it is constant over time, the L leaving probability density can be found after using the Laplace transform of $\Psi_{X,1}^k$ s to compute the convolution of functions:

$$P_L^k(t) = \sum_{\nu=1}^{\infty} (\Psi_{L,\nu}^0 * \Psi_{L,\nu-1+k}^1)(t - (2\nu - 2 + k)T_{\text{tr}}).$$

F agents dynamics are described below.

The probability to be in the patch at a time t is given by the time integral of the arrival minus the departure probability density in a patch k :

$$Q_X^k(t) = 1 - k + \int_0^t d\tau [P_X^k(\tau) - P_X^k(\tau - T_{\text{tr}})] \quad (3)$$

Throughout this letter, we refer to the fraction of agents in the best patch $Q_X^1(t)$ as accuracy.

Fig. 2 shows an example of distributions $P_L^k(t)$, $Q_L^k(t)$. After an oscillating period, they converge towards an equilibrium value. These equilibrium values can be calculated [18]: $T_{X,\text{eq}}^k = -\frac{\theta}{\tilde{\alpha}_X^k}$,

$$Q_{X,\text{eq}}^k = \left(-\frac{2\tilde{\alpha}_X^k T_{\text{tr}}}{\theta} + 1 + \frac{\tilde{\alpha}_X^k}{\tilde{\alpha}_X^k}\right)^{-1}$$

$$P_{X,\text{eq}}^k = \tilde{\alpha}_X^k \left(2\tilde{\alpha}_X^k T_{\text{tr}} - \theta \left(1 + \frac{\tilde{\alpha}_X^k}{\tilde{\alpha}_X^k}\right)\right)^{-1}$$

Cohesion is quantified as the deviation of extrema from equilibrium. For ϵ local maxima and minima $Q_{X,\nu}^k$, the cohesion metrics is defined as $C_X^k = \frac{1}{\epsilon} \sum_{\nu=1}^{\epsilon} Q_{X,\nu}^k - Q_{X,\text{eq}}^k$. The larger C_X^k is, the less damped the oscillations are, i.e. the more cohesive the group is.

Follow-Lead dynamics— To analytically predict the F distributions and metrics, it is possible to include social information in an effective drift term $\tilde{\alpha}_F^k$. If the L dynamics is slow enough compared to the F one and $N_L \gg 1$ (see supplementary Fig. S1(a) for theory-simulation fit as a function of N_L), the averaged term over the time spent in a patch k , $\langle \tilde{\alpha}_F^k \rangle(t, \tau)$, may be used as a quasi-continuous drift term in the distribution calculations. $\langle \tilde{\alpha}_F^k \rangle(t, \tau) = (t - \tau - T_{\text{tr}})^{-1} \int_{\tau+T_{\text{tr}}}^t \tilde{\alpha}_F^k(\tau') d\tau'$

Misinformation

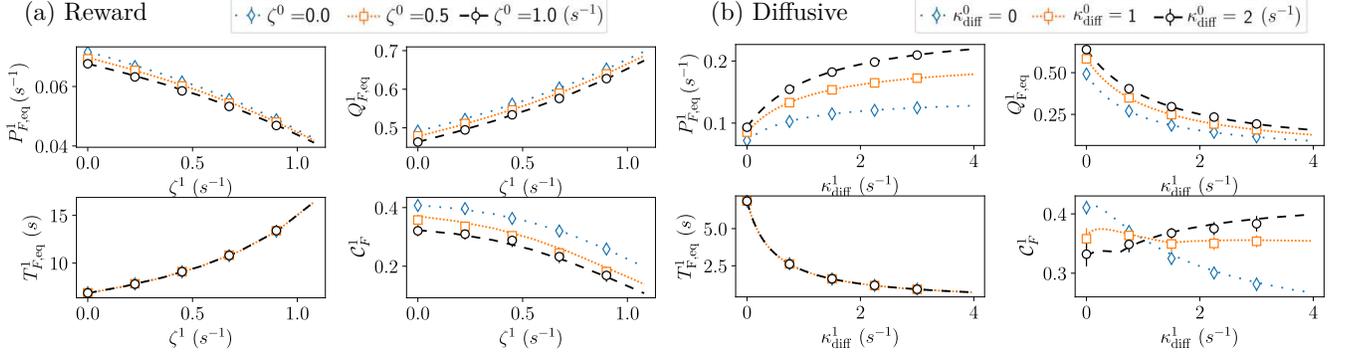


FIG. 4. Misinformation dynamics of the best patch metrics: equilibrium leaving time distribution $P_{F,\text{eq}}^1$, equilibrium accuracy $Q_{F,\text{eq}}^1$, equilibrium patch residence time $T_{F,\text{eq}}^1$, and cohesion C_F^1 . (a) Reward coupling metrics as a function of the communicated reward rates in patch 1 ζ^1 , for different communicated reward rates in patch 0 ζ^0 values. The reward strength parameter is $\kappa_r = 0.5$. (b) Diffusive coupling metrics as a function of the diffusive strength parameter in patch 1 κ_{diff}^1 , for different diffusive strength parameter in patch 0 κ_{diff}^0 values. Lines correspond to theoretical predictions and dots to numerical simulations. Error bars correspond to the standard deviation.

with the arrival (τ) and departure (t) times. At equilibrium, the coupling parameters are restricted by $\alpha_{F,\text{eq}}^k \geq 0 \text{ s}^{-1}$.

For counting coupling,

$$\tilde{\alpha}_F^k(t) = \alpha - \beta p^k - \kappa_c(Q_L^k(t) - \eta) \quad (4)$$

Fig. 3(a) shows that an increasing coupling parameter κ_c is associated with an increased equilibrium accuracy, patch residence time, and decreased cohesion. The effect is more important for small η values. The equilibrium leaving probability density increases with κ_c for large η values and decreases for small ones.

The effective drift term with pulsatile coupling depends on the departure and arrival probability densities.

$$\tilde{\alpha}_F^k(t) = \alpha - \beta p^k + \kappa_d P_L^k(t) - \kappa_a P_L^{k'}(t) \quad (5)$$

Fig. 3(b) shows that higher κ_d values are associated with a decrease in equilibrium accuracy, patch residence time, and an increased equilibrium leaving probability density. These effects are more important for small κ_a values. Please note that in the case of normalization by n_L^k instead of N_L , increased κ_d values are linked to an increased accuracy, see supplementary Fig. S1(b). Cohesion decreases for small κ_a values, and increases for large ones. All in all, departure and arrival coupling have opposite effects on accuracy, patch residence time and leaving probability, while their impact on cohesion is less trivial and depends on $\{\kappa_d, \kappa_a\}$ combinations. Cohesion decreases with increasing departure strengths κ_d for small κ_a values, and increases for larger κ_a values.

Lying: impact of misinformation— In this section, lying L agents display or communicate wrong information about the environment richness (reward coupling)

or their belief about a patch quality (diffusive coupling). To communicate wrong social information about patch quality, lying agents increase their communicated number of catches in the worst patch ($k = 0$) and decrease it in the best patch ($k = 1$). In this way, F agents are misinformed about the distribution of resources and perceive false reward rates ζ^k . The effective drift term is

$$\tilde{\alpha}_F^k(t) = \alpha - \beta p^k - \kappa_r \zeta^k Q_L^k(t) \quad (6)$$

Fig. 4(a) shows the impact of reward misinformation. A smaller ζ^1 value is associated with a decrease in equilibrium accuracy, patch residence time, and an increase in leaving density and cohesion. Larger ζ^0 values contribute to the decrease in accuracy, leaving density and cohesion, while having no impact on patch residence time.

To communicate wrong beliefs about patch quality, lying agents increase their decision variable in the worst patch ($k = 0$) and decrease it in the best patch ($k = 1$). In this way, F agents are misinformed about the leading agents' beliefs and perceive modified averaged decision variables $\kappa_{\text{diff}}^k \langle x_L^k \rangle(t)$. The effective drift term is

$$\tilde{\alpha}_F^k(t) = \alpha - \beta p^k - \kappa_{\text{diff}}^k \langle x_L^k \rangle(t) \quad (7)$$

See the appendix for $\langle x_L^k \rangle(t)$ calculation details.

At equilibrium, the averaged L decision variable can be estimated as $x_{L,\text{eq}}^k = Q_{L,\text{eq}}^k \left(\frac{\theta}{2} + \frac{B}{\alpha_L^k} \right)$. The first part corresponds to the mean and the second part represents the variance of the Ornstein-Uhlenbeck process [21].

Fig. 4(b) shows the impact of misinformation on patch quality belief. A larger κ_{diff}^1 is associated with a decrease in accuracy, patch residence time, and an increased leaving density. Smaller κ_{diff}^0 values contribute to

the decrease in accuracy, leaving density, while does not have an impact on patch residence time. The cohesion dynamics is more complex and depends on the $\{\kappa_{\text{diff}}^0, \kappa_{\text{diff}}^1\}$ combinations. Larger κ_{diff}^1 values decrease accuracy for small κ_{diff}^0 ones, and inversely for large κ_{diff}^0 values.

Discussion— This letter introduces a new framework to quantitatively understand collective patch foraging dynamics in a hierarchical group. Through analytical formal analysis and numerical simulations, we show that it is possible to characterize how different information sharing mechanisms lead to a variety of emergent collective dynamics. In particular, we found that being a following individual can be beneficial (i.e. increase accuracy) when agents count how many leaders are in their patch (counting coupling) or observe leader arrivals in their patch (arrival pulsatile coupling). Another result is that although misinformation from leading individuals is detrimental to the following group, other positive effects can emerge, such as greater cohesion. In fact, more cohesion can be interesting in the presence of threats, as it provides risk dilution [22][23]. Non-cooperative behaviors such as misinformation or sabotage (increased travel time, see [18]) could be especially advantageous for agents in case of competition for resources, e.g. depleting patches.

The diversity and non-trivial emergent processes found in our study highlight the mechanistic and flexible features of our model, which makes it able to understand the large diversity of animal behaviors and uncover their hidden mechanisms. In particular, social hierarchy changes the collective dynamics compared to an egalitarian group [18]. For example, we found a decreased accuracy with diffusive coupling in a hierarchical group, whereas no change was observed in an egalitarian organization. In addition, different normalization conditions, by the number of agents in a patch $n_L^k(t)$, instead of the group size N_L , may also change the emergent properties. i.e. whether agents keep in mind the size of the L group, without observing them, or adjust their behavior depending on the observable number of L agents. For example, accuracy decreases or increases with departure pulsatile coupling, depending on whether the reference is N_L or $n_L^k(t)$ (see supplementary Fig. S1).

Our study also suggests that although a widely used metric in foraging studies is patch residence time (a key variable in the Marginal Value Theorem [24]) accuracy and cohesion may be of critical importance to characterize group dynamics (as shown in some experimental studies [25][26]). As for pulsatile or diffusive coupling, the non-trivial variations of social cohesion strongly point out this metric to be a key descriptor to identify precisely underlying cognitive mechanisms of emergent group processes.

We may note that for some couplings, like pulsatile, large L groups were mandatory to have a theory-simulation match. The results may be different for smaller groups,

especially for cohesion which is more variable compared to equilibrium metrics (see supplementary Fig. S1(a)). Further directions include investigations of how different L dynamics (for example different cohesion features of the L group) may lead to different emergent processes in the F group. In addition, variations of existing couplings would be studied, such as repulsive counting and pulsatile coupling, as well as adding uncertainty, such as imprecise counting. Lastly, a gradual hierarchy with more than two groups would advance our understanding of the wide dynamic range of social organizations. Our aforementioned results and our mechanistic modeling framework (which has been applied in a variety of cases beforehand [15][17][18]) call for the reformulation of optimal foraging theories in a manner that accounts for the diversity of social dynamics. Our quantitative approach provides the foundations for a reformulated foraging theory.

Acknowledgments— We would like to thank Jacob Davidson and Zachary Kilpatrick for insightful discussions. All authors acknowledge support from the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy-EXC 2117-422037984.

Authors Lisa Blum Moyse and Ahmed El Hady have developed the study, derived the results and wrote the manuscript.

APPENDIX

Calculation of $\langle x_L \rangle$

The leading averaged decision variable $\langle x_L \rangle(t)$ is found through the average with the probability to get x at a time t , $U^k(x, t)$. So that $\langle x_L \rangle(t) = \int_{\theta}^0 x U^k(x, t) dx$.

The probability $U^k(x, t)$ is itself calculated through the convolution between the unit x probability for a single patch problem $u(x, t)$ [20](see eq. 8) and the arriving distribution from the patch k' to k , $P^{k'}(t)$. So that $U^k(x_L, t) = \int_0^t u^k(x_L, t - \tau) P^{k'}(\tau) d\tau$, with

$$u(x, t) = \frac{1}{\sqrt{4\pi Bt}} \left(\exp\left(-\frac{(x - \alpha t)^2}{4Bt}\right) - \exp\left(-\frac{\alpha\theta}{B} - \frac{(x + 2\theta - \alpha t)^2}{4Bt}\right) \right) \quad (8)$$

Results for normalization by n_L^k instead of N_L

In case of couplings normalized by the number of leading agents in a patch k , n_L^k , instead of the size of the leading group N_L , the effective drift terms are, for:

Reward coupling: $\tilde{\alpha}_F^k(t) = \alpha - p^k - \kappa_r \zeta^k$

Pulsatile coupling: $\tilde{\alpha}_F^k(t) = \alpha - p^k + \kappa_d \frac{P_L^k(t)}{Q_L^k(t)} - \kappa_a \frac{P_L^{k'}(t)}{Q_L^{k'}(t)}$

Diffusive coupling: $\tilde{\alpha}_F^k(t) = \alpha - p^k - \kappa_{\text{diff}}^k \frac{\langle x_L^k \rangle}{Q_L^k}(t)$

These effective drift terms are then used to compute the different distributions and metrics. See supplementary Fig. S1 for the dynamics of these couplings normalized by n_L^k .

-
- [1] D. W. Stephens, J. S. Brown, and R. C. Ydenberg, *Foraging: Behavior and Ecology* (University of Chicago Press).
- [2] C. Detrain and J.-L. Deneubourg, Collective Decision-Making and Foraging Patterns in Ants and Honeybees, in *Advances in Insect Physiology*, Vol. 35 (Elsevier) pp. 123–173.
- [3] M. C. Baker, C. S. Belcher, L. C. Deutsch, G. L. Sherman, and D. B. Thompson, Foraging success in junco flocks and the effects of social hierarchy, **29**, 137.
- [4] W. Lee, E. Yang, and J. P. Curley, Foraging dynamics are associated with social status and context in mouse social hierarchies, **6**, e5617.
- [5] A. Strandburg-Peshkin, D. R. Farine, I. D. Couzin, and M. C. Crofoot, Shared decision-making drives collective movement in wild baboons, **348**, 1358.
- [6] E. A. Tibbetts, J. Pardo-Sanchez, and C. Weise, The establishment and maintenance of dominance hierarchies, **377**, 20200450.
- [7] S. D. Fretwell and H. L. Lucas, On territorial behavior and other factors influencing habitat distribution in birds: I. Theoretical development, **19**, 16.
- [8] B. Livoreil and L.-A. Giraldeau, Patch departure decisions by spice finches foraging singly or in groups, **54**, 967.
- [9] E. Wajnberg, T. S. Hoffmeister, and P. Coquillard, Optimal within-patch movement strategies for optimising patch residence time: An agent-based modelling approach, **67**, 2053.
- [10] A. Falcón-Cortés, D. Boyer, and G. Ramos-Fernández, Collective learning from individual experiences and information transfer during group foraging, **16**, 20180803.
- [11] R. C. Löffler, E. Panizon, and C. Bechinger, Collective foraging of active particles trained by reinforcement learning, **13**, 17055.
- [12] L.-A. Giraldeau and T. Caraco, *Social Foraging Theory* (Princeton University Press).
- [13] R. Cressman, V. Křivan, J. S. Brown, and J. Garay, Game-Theoretic Methods for Functional Response and Optimal Foraging Behavior, **9**, e88773.
- [14] A. Perez-Escudero and G. De Polavieja, Collective Animal Behavior from Bayesian Estimation and Probability Matching 10.1038/npre.2011.5939.2.
- [15] J. D. Davidson and A. El Hady, Foraging as an evidence accumulation process, **15**, e1007060.
- [16] Z. P. Kilpatrick, J. D. Davidson, and A. El Hady, Uncertainty drives deviations in normative foraging decision strategies, **18**, 20210337.
- [17] S. Bidari, A. El Hady, J. D. Davidson, and Z. P. Kilpatrick, Stochastic dynamics of social patch foraging decisions, **4**, 033128.
- [18] L. Blum Moyse and A. El Hady, Social patch foraging theory in an egalitarian group.
- [19] R. W. Mitchell and N. S. Thompson, eds., *Deception, Perspectives on Human and Nonhuman Deceit*, SUNY Series in Animal Behavior (State University of New York Press).
- [20] D. R. Cox, *The Theory of Stochastic Processes* (Routledge).
- [21] C. W. Gardiner, *Handbook of Stochastic Methods for Physics, Chemistry, and the Natural Sciences*, 3rd ed., Springer Series in Synergetics (Springer-Verlag).
- [22] W. Siegfried and L. Underhill, Flocking as an anti-predator strategy in doves, **23**, 504.
- [23] G. Powell, Experimental analysis of the social value of flocking by starlings (*Sturnus vulgaris*) in relation to predation and foraging, **22**, 501.
- [24] E. L. Charnov, Optimal foraging, the marginal value theorem, **9**, 129.
- [25] N. Franks, T. Richardson, N. Stroeymeyt, R. Kirby, W. Amos, P. Hogan, J. Marshall, and T. Schlegel, Speed-cohesion trade-offs in collective decision making in ants and the concept of precision in animal behaviour, **85**, 1233.
- [26] N. Stroeymeyt, M. Giurfa, and N. R. Franks, Improving Decision Speed, Accuracy and Group Cohesion through Early Information Gathering in House-Hunting Ants, **5**, e13059.

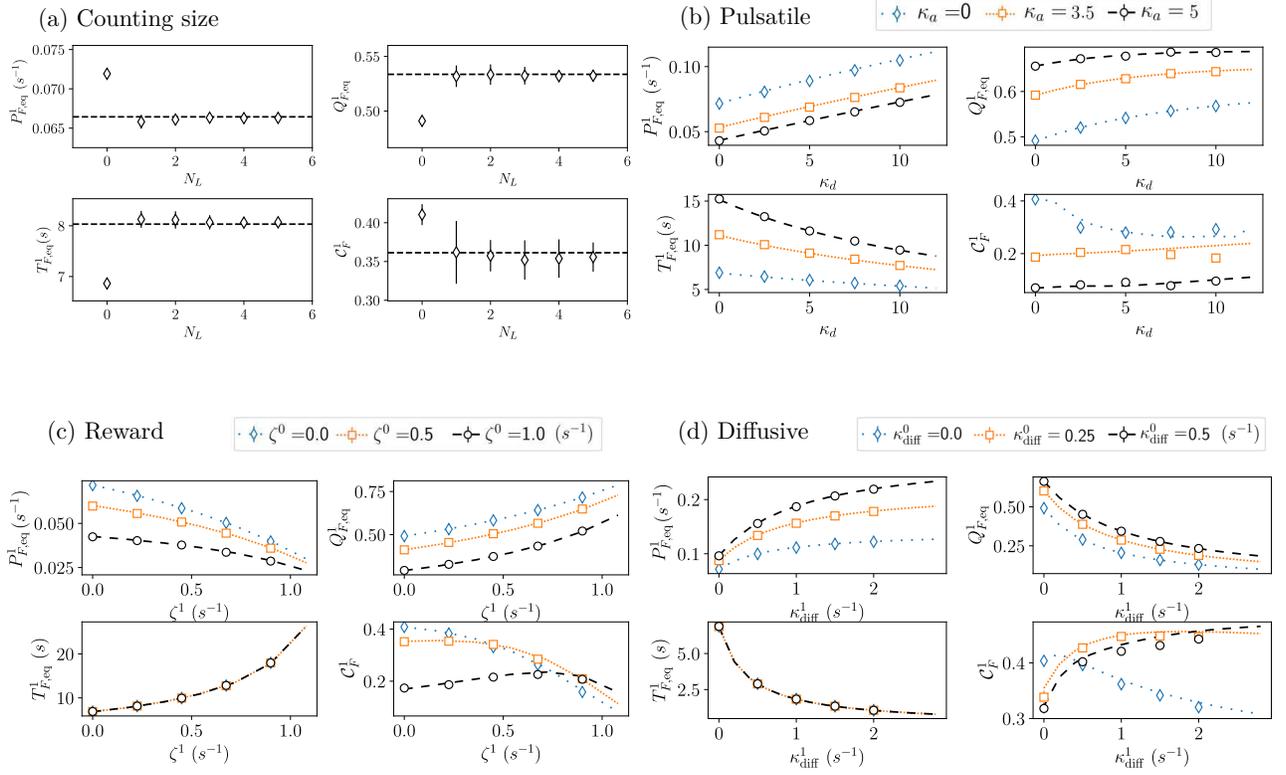


FIG. S1. Effect of L group size and normalization by n_L^k . Best patch metrics: equilibrium leaving time distribution $P_{F,\text{eq}}^1$, equilibrium accuracy $Q_{F,\text{eq}}^1$, equilibrium patch residence time $T_{F,\text{eq}}^1$, and cohesion C_F^1 . (a) Effect of L group size N_L for theory - numerical simulations match in counting coupling. Parameters are $\kappa_c = 0.2 \text{ s}^{-1}$, $\eta = 0.25$, $N_F = 200$, 20 simulations, $dt = 0.001 \text{ s}$. (b) Pulsatile coupling metrics as a function of the departure strength parameter κ_d for different arrival strength parameter κ_a values. Parameters are $N_L = 60000$, $N_F = 2000$, 2 simulations, $dt = 0.02 \text{ s}$. (c) Reward coupling metrics as a function of the communicated reward rates in patch 1 ζ^1 , for different communicated reward rates in patch 0 ζ^0 values. Parameters are $\kappa_r = 0.5$, $N_L = 2000$, $N_F = 2000$, 2 simulations, $dt = 0.02 \text{ s}$. (d) Diffusive coupling metrics as a function of the diffusive strength parameter in patch 1 κ_{diff}^1 , for different diffusive strength parameter in patch 0 κ_{diff}^0 values. Parameters are $N_L = 10000$, $N_F = 2000$, 2 simulations, $dt = 0.01 \text{ s}$. Lines correspond to theoretical predictions and dots to numerical simulations. Error bars correspond to the standard deviation.