Behavioral Ecology

The official journal of the ISBE
International Society for Behavioral Ecology

Behavioral Ecology (2013), 00(00), 1–11. doi:10.1093/beheco/art085

Original Article

Social elites can emerge naturally when interaction in networks is restricted

Tamás Dávid-Barrett and R.I.M. Dunbar
Department of Experimental Psychology, University of Oxford, South Parks Road, Oxford OX1 3UD, UK

Received 26 January 2013; revised 5 August 2013; accepted 5 August 2013.

Animal (and human) societies characterized by dominance hierarchies invariably suffer from inequality. The rise of inequality has 3 main prerequisites: 1) a group in which inequality can emerge, 2) the existence of differences in payoff, and 3) a mechanism that initiates, accumulates, and propagates the differences. Hitherto, 2 kinds of models have been used to study the processes involved. In winner–loser models of inequality (typical in zoology), the 3 elements are independent. In division-of-labor models of inequality, the first 2 elements are linked, whereas the third is independent. In this article, we propose a new model, that of synchronized group action, in which all 3 elements are linked. Under these conditions, agent-based simulations of communal action in multilayered communities naturally give rise to endogenous status, emergent social stratification, and the rise of elite cliques. We show that our 3 emergent social phenomena (status, stratification, and elite formation) react to natural variations in merit (the capacity to influence others' decisions). We also show that the group-level efficiency and inequality consequences of these emergent phenomena define a space for social institutions that optimize efficiency gain in some fitness-related respect, while controlling the loss of efficiency and equality in other respects.

Key words: **behavioral synchrony, division of labor, elite formation, social hierarchy, social stratification.**

INTRODUCTION

At the core of every society lies a conflict between cooperation and competition among the members. As a result, all social animals face the problem of how to achieve coordinated group-level action (Conradt and Roper 2003; Rands et al. 2004; Bode et al. 2012). The efficiency of within-group coordination depends on communication ability, the costs associated with the network (Sueur et al. 2011), and the presence of leaders (Rands et al. 2003, 2008) or elites (Lusseau and Conradt 2009). In particular, it has been established that a hierarchical social network structure (Croft et al. 2008) can have a significant effect on the efficiency of collective action (Nagy et al. 2010; Bode et al. 2011b, 2012).

Hitherto, wisdom-of-the-crowd models have been the main focus of this literature. These were first formulated in the early 20th century (Galton 1907), and there has been a recent resurgence of interest in them in the animal behavior literature (Seeley 1995). One version of this is the many-wrongs principle, which shows that a set of poorly informed individuals can navigate more efficiently when travelling as a group rather than as independent individuals (Gould 2004; Conradt and Roper 2005). In effect, this is a wisdom-of-the-crowd model in the absence of any kind of leader (Codling et al. 2007). However, group-level coordination can also benefit group members even in cases where there is individual variation in some relevant trait. An example might be resource-holding potential (RHP), but we prefer the more general term *merit*, which we define as ability or knowledge that allows one individual to influence the decisions of another. In such cases, coordinated travel may emerge even when group members are ignorant (Degroot 1980; Sigg and Stolba 1981; Wright et al. 2003) or when group members benefit from passive collective action (e.g., predator deterrence in the absence of active defense, as appears to be the case in primates; Dunbar 2010). By the same token, competition leads to inequality naturally in every animal society: variation in strength (RHP), merit, social contacts, and luck can lead to differences in access to food and mates or exposure to predators and physical dangers.

The literature on the origin of social inequality can be divided into 2 strains: winner–loser and division-of-labor models. Winner–loser models (the typical approach in evolutionary biology) assume that a set of individuals form a group for the sake of some group-derived benefits (e.g., lowering predation risk) (summarized by Lindquist and Chase (2009)). This is essentially the simplest form of communal action: members of the community have nothing else to contribute than to be around. Thus, for the individual, the group's existence is a given, while there is internal competition for resources among the group members, manifested in a series of pairwise contests. These trigger a feedback loop: the winner gets better resources and thus will be more likely to be able to defeat others in the future. Winner–loser models tend to focus on the

Address correspondence to T. Dávid-Barrett. E-mail: tamas.david-barrett@psy.ox.ac.uk.

particularities of the resource feedback mechanism, as well as the geographic distribution of the positive and negative resources, showing that the consequent social structure is a hierarchical network. One of the advantages of winner–loser models is that social hierarchies emerge endogenously even if agents have no initial differences at all. Random outcomes of randomly assigned pairwise contests provide the winner with better access to resources, which in turn improves its chances in later contests. Thus, winner–loser models are models of accumulation, and the resulting social status hierarchy is independent of communal action. This is important for the purposes of our article, as our approach will see the hierarchy emerge as an integral part of the communal action. In contrast, division-of-labor models (Traniello and Rosengaus 1997; Theraulaz et al. 1998; Waibel et al. 2006) imply a more complex payoff function than winner–loser models. Here, the group is formed to actively do something *together*: a communal action with distinct and distinguishable parts. The members of the group specialize in these parts, with varying returns to different kinds of specializations. In some cases, this may be a consequence of trait differences of some kind (e.g., Degroot 1980).

The weakness of both these approaches is that they invariably assume that the population's social network is panmictic. In societies characterized by social panmixia, interactions are typically casual and of the moment (i.e., individuals interact with whoever they chance to come across while foraging). Societies of this kind can be found in many herd-forming species, such as grazing ungulates and fish schools, where individuals converge either for mutual protection or to access super-rich resource patches in highly patchy environments. However, in many other species, an individual's ability to interact with another member of its community may be restricted. In some populations, the spatial dispersion of the population makes it physically difficult to interact with anyone who is not an immediate geographical neighbor. Examples would include many territorial birds (Psorakis et al. 2012). In other species, including many group-living mammals, the constraint on interaction is social: individuals can interact only with a small number of socially adjacent members of the wider community. This occurs when dyadic relationships have longer term value, such as when animals form coalitions for mutual protection (e.g., cercopithecine primates; Harcourt and Waal 1992; Dunbar 2010), and when the strength (or functionality) of these relationships requires a significant investment of time (Dunbar 2012; Sutcliffe et al. 2012). There is empirical evidence to suggest that such a restriction on the number of social partners can lead to highly structured social groups in many mammals (examples include orcas, elephants, and many anthropoid primates, see Hill et al. (2008)) or elite societies (such as the high-status matrilines characteristic of some cercopithecine primates; Ehardt and Bernstein 1986; Samuels et al. 1987). Societies where subsets differ in power or access to resources are usually defined as "stratified" (Henrich and Boyd 2008), in which the relative ranking of one individual by all other individuals of the group converge (Henrich and Gil-White 2001). In this article, we differentiate between variation in social attention by others (i.e., status), positive assortment by status (i.e., social stratification), and the formation of an elite clique that has its social ties to the rest of the group substantially reduced.

Our model focuses on the interaction between coordinated group action and social hierarchy within a framework in which the number of social partners is restricted. In this, we build on both the agent-based (Hemelrijk 1999, 2000) and game theory (Hooper et al. 2010) literature of emergent dominance relationships, but use the behavioral synchrony approach to model restricted social networks (David-Barrett and Dunbar 2012, 2013). (There is recent empirical evidence supporting the use of behavioral synchrony models in these kinds of contexts; see Bahrami et al. (2010), Cwir et al. (2011), and Konvalinka et al. (2011).) We assume that the group of agents needs to achieve social coordination in order to be able to perform a collective action, and thus we focus on the synchronization of behavior within a community. The motivation behind using the behavioral synchrony framework is that it allows us to investigate the relationship between network structure and coordination efficiency (Bode et al. 2011a), where hierarchy is a particular form of social structure that has been shown to be essential for coordination efficiency (David-Barrett and Dunbar 2012, 2013). This theoretical finding is supported by the empirical evidence for a relationship between structure of the social network and the efficiency of information transfer (King and Cowlishaw 2009; Bode et al. 2012).

A MODEL WITH ENDOGENOUS STATUS

Our model is an agent-based synchrony model, in which the agents try to agree on a directional vector (Tsitsiklis et al. 1986; Couzin et al. 2005) via a series of dyadic (2-way) exchanges, during which the agents adjust their perception of the usefulness of the information received from others. A graphical summary of the basic behavioral synchrony model can be found in David-Barrett and Dunbar (2012). Although we regard this setup as an abstraction of all social decision problems, perhaps it is useful to think of it in terms of a concrete example in which a group of animals try to decide which direction they should all move toward. Convergence on a common cultural icon or behavioral trait (such as a vocal dialect) would provide other exemplars. In these examples, animals form dyadic synchronizations with each of their neighbors, such that the entire group will eventually synchronize once enough dyadic adjustment steps have been made. What we are interested in is whether this process results in the emergence of a socially structured group, and how that emergent structure affects the efficiency of the synchronization process when that has a (presumably fitness related) payoff.

We first spell out a simple version of the model in which the structure of the network is fixed. We use this to establish the basic properties of the model's behavior. In particular, in this first part of the article, we show how status can emerge autonomously.

The model

We assume that there is a set of n agents who are connected to each other, forming a network. The model makes 2 key assumptions about the way these networks are formed.

First, we assume that the number of connections, k, is the same for every agent. This assumption reflects the empirical observation that most social networks (whether human or nonhuman) are not fully connected (i.e., $k << n - 1$ for all agents), with only a limited variation in k among individuals (see Dunbar (2010) and Kudo and Dunbar (2001), as well as Supplementary Material). It also reflects the theoretical expectation and empirical observation that if maintaining relationships is costly, then agents are increasingly unlikely to be able to have links with every other member of the group as the group size increases (Dunbar et al. 2009; Lehmann and Dunbar 2009; David-Barrett and Dunbar 2012, 2013).

Second, we assume that the relationships among the agents are formed randomly: that is, the networks do not have any particular structure. (We relax this assumption in the second part of the article.) This assumption is important to allow us to show how status can emerge in a general network in which the structure is random and fixed. Later, we will allow agents to change their relationships, so as to examine the consequences of allowing network structure to change endogenously.

We are interested in the case where a group of agents (who are linked via these randomly generated *k*-degree networks) face a behavioral synchrony problem such as a coordinated social action that yields some resource payoff, which is shared by the entire group. We assume that to solve the synchrony problem, the agents go through a synchronization process, defined as follows.

Let us define a "synchronization run" as a series of T dyadic information exchanges, each of which takes place between 2 agents that have a network edge between them. Two agents take part in each of these information exchanges, both providing the other with their respective information, and then both updating their own information and using the new information they just have received from their meeting partner. As the agents update their own information, they necessarily associate a relative weight with the information they just have received. The agents use this weight in recalculating their own information value after receiving the information that their partner holds.

To put it formally, let us consider the full set of nonisomorphic *k*-regular connected graphs on n vertices. Let us draw, with uniform probability, an element of this set for each synchronization run. This way a group of n agents are linked to each other in a connected *k*-degree network. This group goes through S synchronization runs, each of which is made up of T information exchanges. The subject of the information exchange is an information variable, φ. At each point in time (where time is defined by the synchronization step number t, in the synchronization run number s), each agent's information variable takes a specific value, which in the case of agent i is $\varphi_{s,t,i}$. We assume that the synchronization takes place on a dial: that is the information variable takes a value between 0° and 360°: $\varphi_{s,t,i} \in [0°,360°]$ for all s, t, and i. (This assumption is in line with the directional choice problem of Couzin et al. (2005) and ensures that the postsynchronization value of the information variable cannot be predicted by the agents not holding the "true information." This is important for the optimization done by the agents: if the synchronization took place on a linear range, then the agents would be able to predict the outcome by simply observing 2 information points. This way, the problem would become essentially a noise-reducing wisdom-of-crowd problem rather than a synchronization problem, with dramatically different properties.)

We assume that the agents are not equal in terms of the quality of the information with which they start the synchronization (a property we define as their merit). (Note that information is as exemplar for some aspect of an individual's endogenous traits that provides leverage in the competition between individuals to persuade the rest of the community to adopt their view, i.e., to act as the focal point or model with whom everyone else coordinates (Boyd and Richerson 1985; Traulsen and Nowak 2006; West et al. 2011) or as a privileged source of information (Sigg and Stolba 1981; Dyer et al. 2009).) We select one node TI to be the recipient of the "true information," ϕ_{TI}, in the S synchronization runs. TI is selected uniformly from the set of integer values {1, …, n}, which we express as TI ~ {1, …, n}. The recipient of the true information, node TI, does not change her information variable during the synchronization process: $\varphi_{s,t,TI} = \varphi_{TI}$ for all t and s (recall that $s = 1, \ldots, S$ is the index of the synchronization run, and $t = 1, \ldots, T$ is the index of the meetings within a synchronization run). (The assumption of TI not updating her information during the synchronization process is intended to prevent computational time getting out of hand and does not have any consequences for the qualitative results.) In each synchronization run (defined thus as a series of T meetings), all other nodes receive an initial information value drawn randomly from the uniform distribution on the interval 0°–360°; $\phi_{s,0,i\neq TI} \sim U[0°,360°]$, and then alter these during a series of meetings with other nodes.

Formally (notice that for clarity, we temporarily drop the index s):

$$\mathrm{TI} \sim U\{1,2,\ldots,n\} \tag{1}$$

$$\phi_{\mathrm{TI}} \sim U(0°,360°) \tag{2}$$

$$\phi_{0,i} = \begin{cases} \phi_{\mathrm{TI}} & \text{if } i = \mathrm{TI} \\ \sim U(0°,360°) & \text{if } i \neq \mathrm{TI} \end{cases} \tag{3}$$

To define the information exchange, first we have to define the function that calculates the weighted average on a dial:

$$f(A,B,x) = \begin{cases} f_1 & 0° \leq f_1 \leq 360° \\ f_1 - 360° & \text{if } 360° < f_1 \\ f_1 + 360° & \text{if } f_1 < 0° \end{cases} \tag{4}$$

where

$$f_1(A,B,x) = \begin{cases} A - \dfrac{x}{1+x}(A-B) & \text{if } 0° \leq |A-B| \leq 180° \\ A - \dfrac{x}{1+x}(A-B-360°) & \text{if } 180° < A-B \\ A - \dfrac{x}{1+x}(A-B+360°) & \text{if } A-B < -180° \end{cases} \tag{5}$$

in which the parameter x serves as the weight, and the variables A and B are the agent's and the agent's partner's information, respectively. Notice that this function is a weighted average of A and B, with x being the weight (which would take the value of 1 for the case of the usual average). The only reason for introducing this complication is to allow weighted averages to be calculated in later versions. (Note that it is possible to express the same function in a trigonometric form; however, the resulting formula is less intuitive.)

Using this f function, we can now define an information exchange as a meeting in which 1) a pair of connected agents (a and b) are randomly selected providing they are connected to each other, 2) they exchange information, and 3) they update their respective information variables using the new information they just have received. Formally,

$$\{a,b\} \sim U\{1,2,\ldots,n\} \times \{1,2,\ldots,n\} \text{ s.t. } d(a,b) = 1 \tag{6}$$

$$\phi_{t+1,a} = f(\phi_{t,a}, \phi_{t,b}, \omega_{a,b}) \tag{7}$$

$$\phi_{t+1,b} = f(\phi_{t,b}, \phi_{t,a}, \omega_{b,a}) \tag{8}$$

where $\omega_{a,b}$ and $\omega_{b,a}$ are the weights they use when updating their respective information variables with the new information received from the other, and $d(i,j)$ is the network distance between nodes i and j. (i.e., if $d(i,j) = 0$, then $i = j$; if $d(i,j) = 1$, then they are connected; if $d(i,j) = 2$, then they are not connected, but there is a third agent that both are connected to; etc.)

Notice that if these information exchanges are repeated, the average distance from the true information converges to zero:

$$\lim_{t\to\infty} \frac{1}{n} \sum_{i=1}^{n} |\phi_{t,i} - \phi_{\mathrm{TI}}| = 0 \tag{9}$$

(It is possible to find rare special cases, for instance, the cycle graph, in which such convergence does not necessarily take place: it is initial value dependent. However, these are rare instances, and in practice, almost all randomly generated networks with $k > 2$ show convergence.)

Let us assume that the meetings are repeated until $t = T$, where $T = \tau n/2$ (i.e., all agents take part in τ information exchanges on average). At this end-of-synchronization point, the group can be characterized by the average distance of a group member's information value from the true information, a variable that can be interpreted as an efficiency measure of the group's synchronization. (Although the focus of this model is not the efficiency of the synchronization per se, we will return to this question in Efficiency and Inequality.) We have explored the relationship between n, k, and the average distance among the agents' information variables in our earlier study (David-Barrett and Dunbar 2012). In this article, we fix both n and k and then focus on the process whereby the agents choose the weights they associate with others.

To do this, we assume that the agents observe a series of synchronization runs and build up a memory, which then they use to find the weight that would minimize their distance from the true information (which they observe after each synchronization event). Formally, agent i builds up a memory

$$R_i = \left\{ \left\{ r_{s,t,i} \right\}_{t=1}^{T} \right\}_{s=1}^{S} \tag{10}$$

made of records

$$r_{s,t,i} = \begin{cases} \{a_{s,t}, b_{s,t}, \phi_{s,t,a}, \phi_{s,t,b}\} & \text{if } i = a_{s,t} \\ \{b_{s,t}, a_{s,t}, \phi_{s,t,b}, \phi_{s,t,a}\} & \text{if } i = b_{s,t} \\ \varnothing & \text{otherwise} \end{cases} \tag{11}$$

where $s = 1, \ldots, S$ is the index of the synchronization run, R_i is the memory of agent i, and the pair $\{a_{s,t}, b_{s,t}\}$ is the nodes chosen to participate in the information exchange, defined in Equation 1. Thus, $r_{s,t,i}$ is the record that agent i places in her memory at information exchange t in the synchronization run s.

The agents then use their memory to calculate the weight they would associate with the other agents. In effect, they attribute status to each agent as a function of some endogenous trait of the agent (that agent's charisma, performance on previous tasks, etc.). Formally,

$$\hat{\omega}_{i,j} = \arg\min_{\beta} \mathrm{E}\left[\left(\phi_{\mathrm{TI}} - \frac{A + \beta B}{1 + \beta} \right)^2 \middle| \{i, j, A, B\} \in R_i \right] \tag{12}$$

that is, agent i selects the part of her memory that recorded her information exchanges with agent j and chooses a weight that approximately minimizes the distance from the true information in this relevant part of her memory records. (In other words, to calculate the weight agent i associates with agent j, she first chooses the part of her memory that contains the records of her information exchanges with agent j and then chooses the weight coefficient that minimizes the least squares of the distance of the weighted average of the 2 information values from the true information.) The agents calculate this weight after enough full synchronization runs have taken place to build up a sufficient memory (we used $S = 200$ synchronization runs to build up the memory in our simulations). The agents use these updated weights in their next synchronization runs.

Simulation results

If we calculate the weights as set out above (using $\omega_{i,j} = 1$ for all i,j as initial weights), a particular pattern emerges: the weight agent i associates with the agent j is dependent on their network distances from the recipient of the true information, TI. In particular, we find that

$$\frac{\Delta\hat{\omega}_{j,i}}{\Delta d_{\mathrm{TI},i}} \begin{cases} < 0 & \text{if } d(j,\mathrm{TI}) - d(i,\mathrm{TI}) = -1 \\ = 0 & \text{if } d(j,\mathrm{TI}) - d(i,\mathrm{TI}) = 0 \\ > 0 & \text{if } d(j,\mathrm{TI}) - d(i,\mathrm{TI}) = 1 \end{cases} \tag{13}$$

The results are shown in Figure 1a (see Supplementary Material for details).

The intuitive view of these results is as follows. The best information at any point in time is obtained from the individual TI who had received the true information. As she passes on the information, the next best information will tend to be those nodes that are directly linked to TI. Then the next best information is held by the nodes that are not linked to TI but are linked to a node that is linked to TI, and so on. Hence, the quality of the information a node passes on during synchronization is determined by 2 factors: 1) the node's network distance from the recipient of the true information, node TI, and 2) the quality of information of the node's partners. Therefore, the weight node j assigns to node i is dependent on 1) i's network distance from node TI and 2) whether j is closer to TI than i.

Furthermore, notice that the relative valuations that node i receives from her partners correspond to a social status concept in the following way. Let us denote the list of ranked partners of node j as

$$e_j = \{ i \mid d(j,i) = 1 \} \tag{14}$$

such that

$$e_j = \{ i1, i2, i3, \ldots \} \mid \hat{\omega}_{j,i1} \hat{\omega}_{j,i2} \geq \ldots \hat{\omega}_{j,ik} \tag{15}$$

That is, the order in which the nodes are listed in e_j is based on the relative value of the weights that node j assigns to her partners. Let $\rho_{j,i}$ denote the position of node i in the partner list of node j, e_j. This preference ranking then allows the formal definition of the agent's status; node i's average position in her partners' ranking is

$$\bar{\rho}_i = \frac{1}{k} \sum_{j \in e_i} \rho_{j,i} \tag{16}$$

Notice that $\bar{\rho}_i$ corresponds to an inverse concept of social status: low $\bar{\rho}_i$ indicates that i is ranked high by its partners among their partners. Not surprisingly, the closer a node is to the receiver of the true information, node TI, the higher is its status (Figure 1b).

In sum, we have shown using this simple model that if the agents are optimizing the weights associated with their partners, then a communal action problem can lead to the emergence of social status, determined by the distribution of useful information and the structure of the network. In the next model, we examine the effect of allowing the agents to change their network connections based on the others' statuses.

A MODEL WITH STRATIFICATION

A particularity of the model we outlined above is the monopoly that one agent has on true information: one node is selected to receive

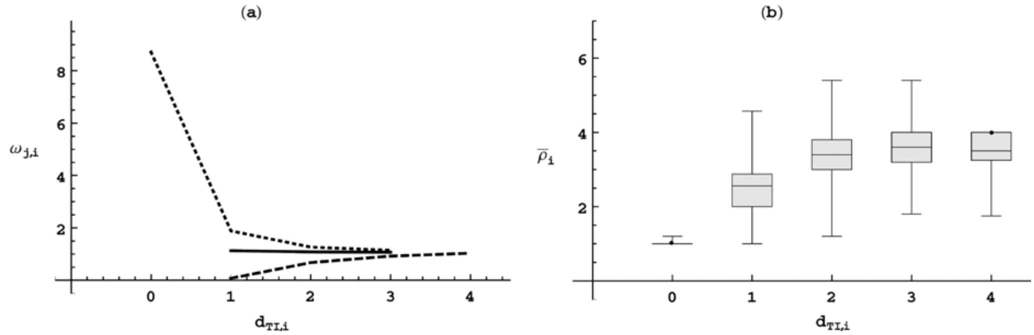

Figure 1
(a) Status is determined by the network distance to the true information. Continuous curve: the weight assigned to her by those equidistant to TI, dashed curve: by those partners of i that are closer to TI, and dotted curve: by those partners that are further away from TI. See Supplementary Table 1 for the statistics. (b) The social status of a node, $\bar{\rho}_i$, decreases with her network distance from TI. (Parameters used: $n = 20$, $k = 5$, $\tau = 40$, $S = 200$. The average of 300 repeats.)

the true information with 100% probability, whereas all the other nodes receive the true information with 0% probability. There is no real variation among the nodes in terms of their merit. As a consequence, in the first model, all weights associated with the agent TI were the highest weights. Hence, if the non-TI agents were to have a preference regarding whom they would like to acquire as a friend (i.e., a social contact with whom there is a pattern of repeated cooperative interactions), they would all put the agent TI as a top choice, while not being able to differentiate among the others. In this second model, we relax the assumption that only one agent can receive the true information, and thus allow for variation in merit. This generates the possibility of a new dynamics: an agent that prefers to receive the best quality information (i.e., one that is most likely to be close to the true information) would be able to rank all others, while they could in turn rank her. As the personal network size is fixed (which is a realistic assumption in all large animal and human social networks), the emergence of the ranking of others in terms of preference (befriending) can lead to a restructuring of the network edges. Only a limited number of nodes could befriend the high-merit nodes, which themselves will tend to favor other high-merit nodes. As a result, the low-merit nodes are left to form links only with other low-merit nodes. Hence, we show how the $k << n - 1$ assumption can lead to a socially stratified group, and the emergence of an elite.

The model

In this second model, we will leave most building blocks of the basic model unchanged but add the following new elements. First, at the outset of each synchronization run, there is a lottery to decide which agent becomes TI. Thus, it is this probability that corresponds to the concept of merit, where we assume that the probability of being chosen is uneven among the agents. So as to be able to track the differences in merit, we set the probability of a node being the recipient of the true information in any one synchronization run as a function of its index number in the following way:

$$\Pr(i = \mathrm{TI}) = \frac{i^{\sigma}}{\sum_{i=1}^{n} i^{\sigma}} \quad (17)$$

That is, the higher the index number of a node, the more likely it is to be the recipient of the true information. The parameter σ, thus, denotes the steepness of merit inequality: the higher the value of σ, the larger the difference between low- and high-merit agents. (For a further interpretation of (17), see Supplementary Material.) Notice that although the form of (17) limits the variation of the merit in the sense that merit differences are not random, the introduction of σ allows us to track the effect of any changes in the extent of merit inequality.

The second modification of the basic model concerns the memory component: unlike in (11), agents observe the other agents' information values when recording their memory. The new definition of a memory record is

$$r_{s,t,i} = \begin{cases} \{a_{s,t}, b_{s,t}, \phi_{s,t,a}, \phi_{s,t,b}\} & \text{if } i = a \\ \{b_{s,t}, a_{s,t}, \phi_{s,t,b}, \phi_{s,t,a}\} & \text{if } i = b \\ \{\{i, a_{s,t}, \phi_{s,t,a}, \phi_{s,t,b}\}, \{i, b_{s,t}, \phi_{s,t,a}, \phi_{s,t,b}\}\} & \text{if } i \neq b \end{cases} \quad (18)$$

That is, we assume that the memory collection (11) changes such that all information exchanges are observed by all other nodes, although the nodes other than the 2 nodes in the exchange will not alter their information values as a result of the exchange. This means that although all agents record the information exchange, it only matters for the 2 agents that are taking part in the particular meeting at this point in the synchronization. (This assumption is due to computational time constraints: to assess those others that the agent is not connected to at a given time, the agents would need to swap with every other agent and observe their merit so that they could build up a full-preference ranking over all the other agents. This takes so long that the computational time for these models would be unrealistic. Our robustness checks for subparts of the model suggest that there is no difference between the simulation outcomes of the 2 setups.)

The third element in the model is a partner-changing algorithm in line with the pairwise stability network concept of economics (Jackson and Watts 2001; Jackson 2008). Once the weights are estimated by the agents, they are able to rank their partners, as well as all those agents to whom they are not linked (it is the latter possibility that required the memory buildup to contain all other agents). Then an agent would be able to tell if she would prefer to swap one connection for another. If this desire is mutual, then we allow them to drop the undesired edge and acquire the new one. (Gould (2002) uses a stratification mechanics that is similar in principle.) Formally, let $m(i) = e_i(-1)$ denote the last element of the list of partners of agent i. As the e_i list is ranked according to the weight due to (8), $m(i)$ is the partner of i to whom i assigned the lowest weight. Then the following algorithm is applied:

(step 1) Select a random pair of unconnected individuals: $i, j \sim U(1,\ldots,n) | d(i,j) > 1$.

(step 2) If the weight assigned to i by j is larger than the weight j assigns to $m(j)$ (i.e., $\omega_{j,m(j)} < \omega_{j,i}$) and vice versa ($\omega_{i,m(i)} < \omega_{i,j}$), then they both drop their respective lowest ranked partners and add the current interactant to their list of partners: $e_i := j \cup e_i \setminus m(i)$ and $e_j := i \cup e_j \setminus m(j)$.

These 2 steps are repeated until there is no unconnected agent pair left that would prefer to be connected to each other. However, the above change in the network raises the possibility of the personal network of some agents diverging from the preset personal network size, k: if an agent is ranked last by all other agents, it is possible that she gets dropped by some contacts but does not get any new contacts to replace them. Hence, we need to add a third step, which corrects for this:

(step 3) Select a random individual i, then if $k_i < k-1$, add a random new edge to i: $e_i := e_i \cup j$ and $e_j := e_j \cup i$, where $j \sim U\left(j \middle| d(i,j) > 1\right)$. This in turn can make some agents have too many edges; hence, we also adopt the rule that if $k_i > k+1$, then drop the lowest ranked partner: $e_i := e_i \setminus m(i)$. That is, if there are too many friends, drop the lowest ranking one, and if there are too few friends, add a random one. (The correction algorithm hence ensures that the network stays near the condition $k_i = k \forall i$ during simulations.)

Notice that one of the consequences of the first model is that the quality of the information passed on from one agent to another is dependent on both the probability of the agent being TI (100% for one, and 0% for all others in the first model) and the network position of the 2 agents relative to the position of the TI. As a consequence, once the network's structure changes using the partner-changing algorithm, the optimal weights will also change, which in turn allows for further change in the weights, leading to further realignment of the network. Thus, the 3 new elements of the second model (varied merit driven by σ, the individuals' assessment broadened to every member of the group, and the partner-changing algorithm) allow us to run a set of simulations in which the synchronization runs, the weight optimization by the agents, and the network's restructuring go through several cycles:

(phase 1): generate a random network of n and k, and set all initial weights to 1;
(phase 2): run a sufficient number of synchronization runs for the agents to build up a big enough memory to estimate everyone else's true probability of being TI;
(phase 3): calculate the new weights;
(phase 4): run the partner-changing algorithm;
(phase 5): repeat phases 2–4 until the network structure does not change significantly between rounds.

Simulation results

If the merit inequality parameter, σ, is sufficiently large, 3 phenomena emerge. First, the weights assigned to the agents by the others reflect their merit (the probability of being the recipient of the true information). That is, if we let $\bar{\omega}(i) = \frac{1}{n-1}\sum_{j \neq i} \omega_{j,i}$ denote the average weight agent i receives at the end of the simulation cycles, then $\partial \bar{\omega}(i)/\partial i > 0$ for σ = 1 at $P < 0.001$. That is, due to (17), the agents' respective merits increase their index numbers, and the higher their index number, the higher the weights the other agents assign to them (Figure 2a).

Second, as the network structure changes, high-status agents tend to get connected with other high-status agents, and low-status agents are left with other low-status agents, and thus the network displays a stratified pattern. Formally, let $\varepsilon(i) = (1/k)\sum_{j \in e_i} j$ denote the average index number of agent i, then $\partial \varepsilon(i)/\partial i > 0$ for σ = 1 at $P < 0.001$. In other words, the network structure undergoes a stratification process, yielding an oval-shaped network (Figure 2b; see Figure 3 for an example of this process). Note that this stratification process is due to the fact that high-merit agents are more preferred by their partners than low-merit agents. The partner-changing algorithm that allows the alteration of the network ties works along preferences (as defined in Equations 14 and 15), resulting in a stratification along status lines.

Third, although the very highest index nodes would prefer even higher ones as their partners, there are no such nodes. Hence, they fill their partner slots with somewhat lower (but still high) ranking nodes. These high mid-ranking nodes will always prefer the top nodes as partners. This creates a pattern in which an elite set of high-index nodes delineates from the rest of the network (Figure 2b).

Importantly, the result presented above holds for all σ > 0. In fact, it is this merit inequality parameter, σ, introduced in (17), that is the key driver of both the stratification (Figure 4a,b), and the elite delineation (Figure 4c,d) processes. When there is no merit inequality (i.e., σ = 0), no weight difference emerges, and thus there is no social stratification either. As the merit inequality increases, the weight differences increase and so does the pressure toward social stratification. As the elite delineation process is a consequence of

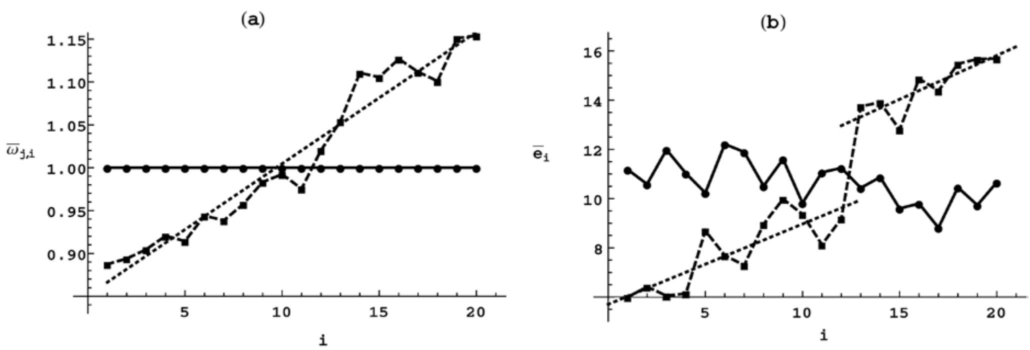

Figure 2
Stratification and elite delineation. Panel (a): the relationship between the index number and the average weight received by a node. Panel (b): the relationship between the index number and the average index number of the node's partners. In both graphs, the continuous curve is the average of initial values, the dashed curve is the average value after 10 rounds of iteration, and dotted line is the linear trend line. Notice that (b) displays a break in the average partner index at $i = 13$, a signal that an elite is beginning to be delineated. (Parameters used: $n = 20$, $k = 5$, σ = 1, τ = 40, $S = 200$. Average of 10 repeats.)

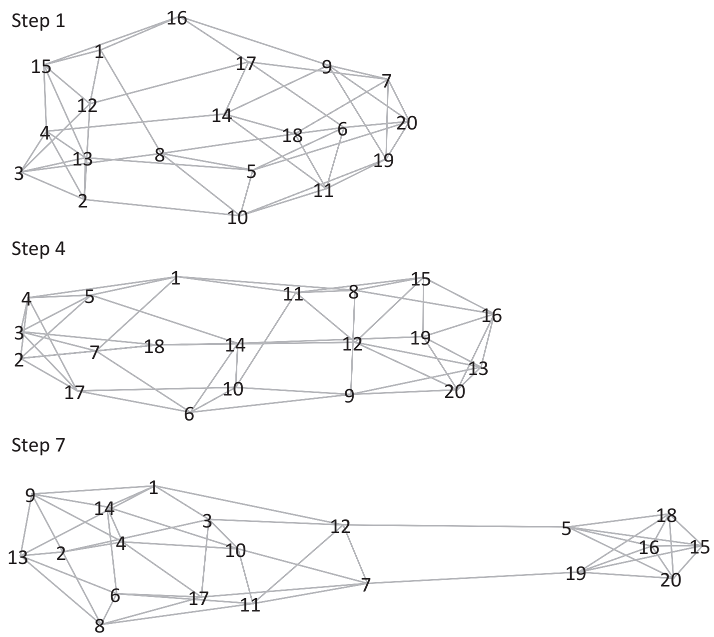

Figure 3
The structural changes of an example network during simulation. Each step shows the structure after the simulation round takes place, the name of each vertex is the index number of the corresponding agent. Notice that it takes several rounds for a mesh structure to be transformed into a stratified structure (i.e., high-index number agents link themselves to other high-index number agents, whereas low-index number agents get linked to other low-index number agents), which then leads to the delineation of an elite (i.e., the high-index number agents form a separated group).

social stratification, higher merit inequality also leads to stronger pressure toward the rise of an elite. (Notice that Figure 2 is a cross-section of Figure 4.)

A consequence of this is that increasing merit inequality leads to a change in the organization of the nodes. When σ is near 0, the calculated weights, $\hat{\omega}$, tend to be the same across the network. One could associate this situation with truly egalitarian societies, in which there is no status inequality in any form. However, increasing σ even marginally leads not only to stratification (which is bounded by the number of nodes) but also to an increase in the weight differences (which is essentially unbounded). As the weight differences increase, the network converges toward a hierarchical tree-like structure, in which unidirectional information flow is the most effective way of information diffusion.

EFFICIENCY AND INEQUALITY

In an earlier study (David-Barrett and Dunbar 2012), we have shown that status inequality may be an adaptive feature of groups that form social networks and synchronize their actions. However, it does not follow that social stratification and elite delineation are also adaptive. To examine this question, let us introduce 3 (artificially constructed) benchmark networks: a mesh network, a stratified network, and a stratified network with a delineated elite. We designed these networks to be representative of the 3 types of graph that emerged in the simulations of the previous section: the initial graph that tends to be a mesh network (i.e., a random *n*-node uniformly *k*-degree network), the stratified network (i.e., the average index numbers of an agent's partners are assortative with respect to the agent's index number, but without elite delineation), and an elite network (that is, a stratified network, with the agents with the highest index numbers delineated from the other agents in a distinct module). (See Supplementary Material for details of these graphs.) The purpose of these networks is to allow us to measure 2 components of social welfare: payoff for the group as whole and inequality of payoff among the members of the group.

To capture the group's payoff, we define a measure of synchronization efficiency, *h*, as the average distance of the agents' individual information values from the true information at the end of the synchronization run:

$$h(g,\omega,\sigma)=\sum_{i=1}^{n}\frac{\left|\phi_{T,i}-\phi_{TI}\right|}{n} \tag{19}$$

where *g* denotes the type of network (i.e., mesh, stratified, or elite), ω denotes the weight set, and σ is the merit inequality parameter as above. We use 2 sets of weights for each graph and σ value: a base case in which the weights are set uniformly to 1, and the estimated weight case in which we use the weights that are the outcome of the agents' optimization.

In linking the concept of payoff to synchronization efficiency, we use the same assumption as in the setup of the entire model: the group is performing some form of communal action that requires the behavioral synchrony of its members. The group's total payoff is dependent on the efficiency of the synchronization (defined in this case by how far the agents' end-of-synchronization information values are on average from the true information).

However, we also assume that although the agents' individual payoffs are dependent on the group's total payoff, they also depend on the distance of their individual end-of-synchronization information value from the true information. (This might, for instance, reflect a cost associated with distance from the true information, or the distribution of the group's payoff could be in proportion to the respective contribution to the group's synchrony.) To capture this concept, let *q* denote an economic inequality measure, defined as the average difference from the group average *h*:

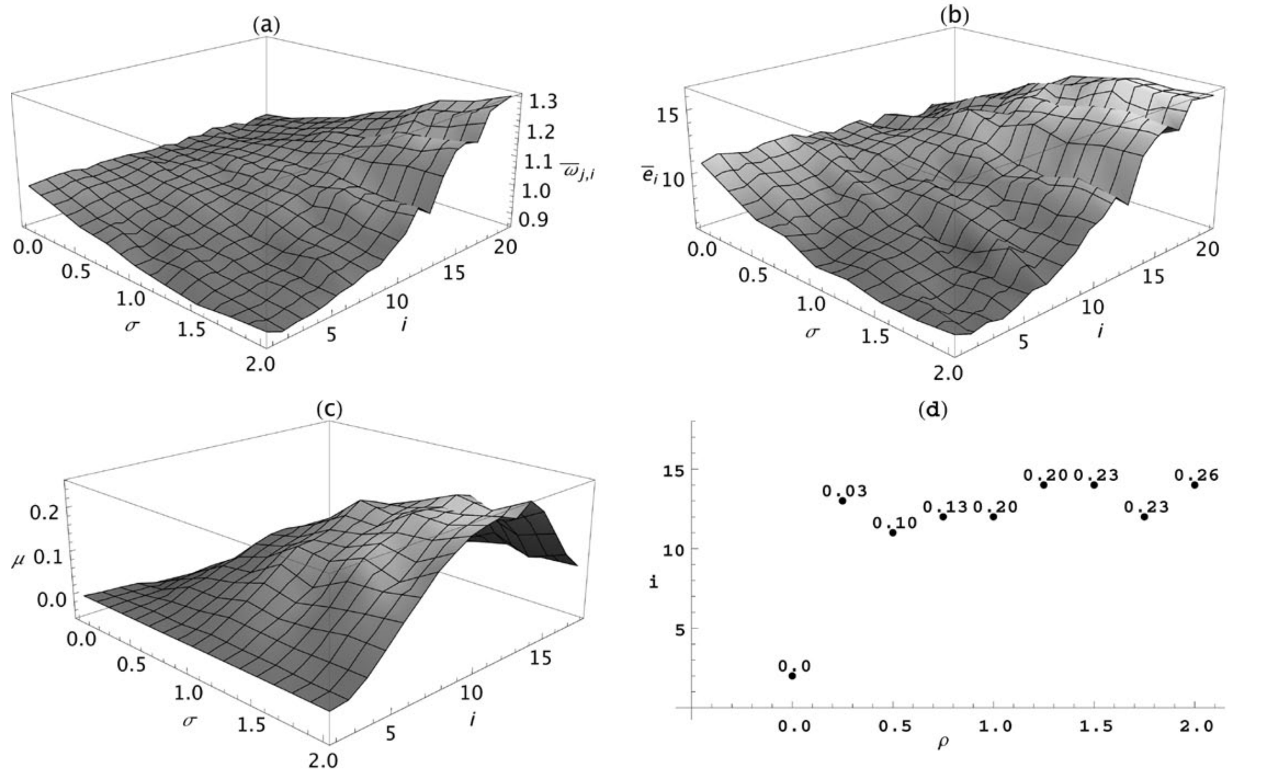

Figure 4
The effect of the merit inequality parameter on stratification and elite delineation: (a) the higher the merit inequality, the higher the social status steepness becomes; (b) the higher the merit inequality, the more stratified is the society; (c) the higher the merit inequality, the more pronounced is the elite's clique formation (μ denotes modularity); (d) *y* axis: *i* plots the location of the modularity peak, that is, the ridge of (c), and thus the numbers displayed are the level of modularity at each particular peak. (Parameters used: $n = 20$, $k = 5$, $\tau = 40$, $S = 200$. The average of 10 repeats.)

$$q(g,\omega,\sigma) = \sum_{i=1}^{n} \frac{\left| \left| \phi_{T,i} - \phi_{\mathrm{TI}} \right| - h(g,\omega,\sigma) \right|}{h(g,\omega,\sigma)} \Big/ n \qquad (20)$$

The simulation results are as follows (see also Supplementary Table 2 and Figure 4):

1. Allowing the nodes to optimize the weights they assign to their partners increases the group's synchronization efficiency and decreases inequality in all cases:

$$h(g,1,\sigma) > (g,\hat{\omega},\sigma) \qquad (21)$$

and

$$g(g,1,\sigma) > g(g,\hat{\omega},\sigma) \qquad (22)$$

for all $g \in \{\text{mesh}, \text{stratified}, \text{elite}\}$ and $\sigma \in \{0,1\}$, where $\hat{\omega}$ denotes the weights estimated by the agents.

2. Variation of merit does not necessarily affect synchronization efficiency, but always increases inequality:

$$g(g,\omega,0) < g(g,\omega,1) \qquad (23)$$

for all $g \in \{\text{mesh}, \text{stratified}, \text{elite}\}$ and $\omega \in \{1, \hat{\omega}\}$.

3. Stratification decreases efficiency and increases inequality:

$$h(\text{mesh},\omega,\sigma) < h(\text{stratified},\omega,\sigma) \qquad (24)$$

and

$$g(\text{mesh},\omega,\sigma) < g(\text{stratified},\omega,\sigma) \qquad (25)$$

for all $\omega \in \{1, \hat{\omega}\}$ and $\sigma \in \{0,1\}$. However, less so if weights are optimized:

$$\frac{h(\text{stratified},1,\sigma)}{h(\text{mesh},1,\sigma)} > \frac{h(\text{stratified},\hat{\omega},\sigma)}{h(\text{mesh},\hat{\omega},\sigma)} \qquad (26)$$

for all $\sigma \in \{0,1\}$.

4. Elite clique formation leads to a large loss of efficiency and equality, even if weights are adjusted. This effect is much larger than that arising from stratification:

$$\left(\frac{h(\text{elite},\hat{\omega},1)}{h(\text{stratified},\hat{\omega},1)} - 1 \right) \Big/ \left(\frac{h(\text{stratified},\hat{\omega},1)}{h(\text{mesh},\hat{\omega},1)} - 1 \right) = 6.92 \qquad (27)$$

and

$$\left(\frac{g(\text{elite},\hat{\omega},1)}{g(\text{stratified},\hat{\omega},1)} - 1 \right) \Big/ \left(\frac{g(\text{stratified},\hat{\omega},1)}{g(\text{mesh},\hat{\omega},1)} - 1 \right) = 8.71 \qquad (28)$$

That is, the loss of the synchronization efficiency and the increase in inequality is almost a magnitude worse in the elite case compared with the merely stratified case. (Perhaps, it is interesting to observe that most of the negative impact falls on the mid-ranked segment of the group. See Supplementary Figure 4.) Notice that the loss of efficiency of stratification is due to the lengthened information channels, whereas the additional loss of efficiency in the elite network is due to the information bottleneck emerging between the elite clique and the rest of the group.

These results suggest that we can separate social inequality into 3 markedly different types. Type I is *status inequality*, which emerges when there is variation in merit and inevitably increases the group's efficiency. Type II is *social stratification*, which requires status inequality but introduces an information bottleneck that leads to loss of efficiency and increased payoff inequality. Type III is *elite formation*, which requires stratification and results in a collapse in efficiency and is associated with extreme payoff inequality. Thus, the 3 different kinds of inequality, although closely linked, have opposite effects on societal efficiency and inequality. To our knowledge, this model is the first to differentiate between these 3 effects within 1 framework.

DISCUSSION

The rise of economic, resource, or fitness inequality has 3 main prerequisites in any theoretical framework: 1) a group in which inequality can emerge, 2) a basis for permanent differences in payoff, and 3) a mechanism that initiates, accumulates, and propagates these differences. In classic winner–loser models, the 3 elements are independent; in division-of-labor models, the first 2 elements are linked, whereas the third is independent. In this article, we propose a new model in which all 3 elements are linked.

We used a model of group synchronization with some general, nonspecific initial inequity in the usefulness of each individual toward the group's communal action. We show that, within an agent-based framework with restricted interaction, initial differences in merit can lead to emergent social status. We show that although these statuses originate from an individual-level optimization process, they increase the efficiency of the entire community's collective action. This is a confirmation of our earlier group-level results (David-Barrett and Dunbar 2012) and is in line with the empirical findings of Lu et al. (2009).

We show that the assignment of status increases the efficiency of the community and that if the network structure is fixed, then the accompanying inequality is relatively small. However, we also show how emergent status allows a new phenomenon: individual agents can now differentiate between their partners and other members of the group in terms of how useful they are. A recent set of experiments on vervet monkeys (Pansini 2011) provides empirical support for this. We show that if the personal network size is fixed (i.e., agents live in a 2-layered community), then the consequence of status-based friendship swapping is the restructuring of the community into a socially stratified group. During the process of stratification, social status differences increase further and so does payoff inequality. In contrast to the initial emergence of social status, social stratification worsens the group's synchronization efficiency. Stratification in turn opens up the possibility for the formation of an elite clique. In those animal societies where social differentiation occurs (e.g., dominance hierarchies in many animal species), social status is invariably power based rather than information based. And in those few cases where individuals may be accorded special status by virtue of the information they have (e.g., old males in hamadryas baboons; Kummer 1968; Sigg and Stolba 1981), the influence of their specialist knowledge is typically limited to very specific contexts. Elites form, however, if it is worthwhile for the highest ranking agents to choose each other in preference to any of the lower ranked agents. Something like this is found in the organization of both dominance hierarchies and kinship groupings in Old World monkeys, and it can lead to a form of segregation within social networks (Seyfarth 1977; Lehmann and Dunbar 2009). In many of these cases, segregation arises out of attempts to monopolize resources or protect personal interests (Dunbar 2010). However, when it happens in information contexts, it inevitably creates an information bottleneck, drastically reducing efficiency and increasing inequality. Because inequality has consequences for disadvantaged individuals' willingness to remain in a group (Dunbar et al. 2009), a natural optimization process will arise in respect of group size as a function of the advantages and disadvantages of the emergence of stratification.

Our model provides a mechanics for the process whereby status and stratification emerge from an initial difference in merit. Some elements of this process are similar to concepts in the winner–loser framework and the division-of-labor models: 1) we have a feedback loop related to the pairwise exchanges (the node that provides better information gets ranked higher and thus gets higher ranked friends and through this improved payoff), although, unlike in the winner–loser models, this feedback loop is the output of the model, rather than being an assumption; and 2) there is wealth accumulation in terms of network capital, an important feature of the division-of-labor models, although, crucially, in our model the network capital accumulation requires that status be assigned by others rather than being an assumption of the model. Winner–loser models with panmixia have been used to study the evolution of leader–follower relationships (e.g., van Vugt and Kurzban 2007; Hooper et al. 2010). Our model broadens these findings to nonpanmictic (i.e., network structured) populations and demonstrates that there are consequences beyond simple leadership in terms of the evolution of structured societies with formal elites.

As a consequence of having elements from both frameworks, our model offers insights into some unanswered questions that arise from the winner–loser and division-of-labor models. In particular, we provide a mechanics that allows more realistic complex social cognition to be incorporated into the winner–loser models. In our model, an agent perceives others as members of the group she cooperates with, as sources of information that improve her own standing in the community, and as competitors for the friendship of the highest status agents. Furthermore, our approach provides a mechanism for elite delineation, and a framework for class formation, perhaps the central question of the division-of-labor tradition in respect of humans. We show that the inequality stems from the recognition of varying merit in contributing to an increase in the group's collective action rather than decrements in the group's efficiency. At the same time, once merit is recognized in assigned status, social stratification takes place, which both worsens efficiency due to longer path lengths and increases inequality due to the merit associativity within the network. However, neither of these effects is large, unlike the impact of the delineation of an elite, which creates a bottleneck that dramatically worsens both efficiency and inequality. (Note that in some sense, elite formation is the extreme version of stratification. However, in a network framework, they are 2 distinct phenomena: the former increases the path lengths within the graph and hence increases the time to achieve behavioral synchrony, whereas the latter creates an outright bottleneck that amplifies the effect of longer path lengths.) Elite formation is not possible without the rise of a stratified society, which in turn is based on the statuses initially assigned on the basis of variation in merit. It does not escape our notice that the appearance of charismatic religious leaders in human societies provides a particularly germane example of this process.

It is important to emphasize that the results from our synchronization framework are entirely dependent on the presence of the multilayering of social relationships within groups (i.e., that the assumption of panmixia does not hold, leading to a natural layering of interaction patterns). In our models, we assume that 1) both group size and personal network size are fixed and that the latter is both 2) universal across the nodes and 3) substantially smaller than the former (see Supplementary Material). Assumptions 1) and 2) are not strictly necessary for the results of this study, and we deal with the emergence of multilayered groups elsewhere. (For a robustness check regarding the variation of the degree among the agents, see Supplementary Material.) However, assumption 3) that personal network size, k, is much smaller than the group size, n, is crucial. If the network is fully connected (i.e., panmictic), then initial statuses will emerge, but neither stratification nor elite formation takes place. In contrast, most models of both group coordination (Conradt and Roper 2003, 2007; Couzin et al. 2005) and cultural evolution (Boyd and Richerson 1985) assume panmixia. Because the validity of this assumption is crucial to the outcomes of the model, it is important to show that this assumption is realistic. In fact, this assumption is well supported in the literature: many species of mammals form multilayered societies (Hill et al. 2008), and, among both animals and humans, personal network size varies only to a small degree across individuals, species, and cultures (Kudo and Dunbar 2001; Zhou et al. 2005; Hamilton et al. 2007).

The consequence of groups being naturally multilayered in this way is the emergence of a social structure, reflecting the natural structuring of personal social networks (Granovetter 1973; Dunbar and Spoors 1995; Hill and Dunbar 2003) and the consequences this has for structuring natural communities (Zhou et al. 2005; Hamilton et al. 2007). It has been shown in evolutionary game theory in a network setting that this $k << n - 1$ (i.e., k is much smaller than n) result can be an optimal choice for the agents (Masuda 2007). There is ample evidence for the presence of the cost of relationship maintenance (Roberts and Dunbar 2011). Furthermore, the fact that we observe a variation of coordination strategies in itself can be seen as an evidence for the limited range of degrees of agents (King et al. 2011). For any species (human or otherwise), there is always a threshold of group size above which it would take too much time to fit in relationship maintenance with all others in the group. In this respect, our model has the neat property of providing a unified explanation for seemingly unrelated but well-studied phenomena in behavioral ecology and in economics.

SUPPLEMENTARY MATERIAL

Supplementary material can be found at http://www.beheco.oxfordjournals.org/

FUNDING

The authors are funded by a European Research Council Advanced grant to R.I.M.D.

The authors would like express their gratitude to A. Besudnov, K. Opie, A. Rotkirch, B. Szendroi, and S. West for their helpful comments.

Conflict of Interest

The authors do not have any conflict of interest.

Handling editor: Gil Rosenthal

REFERENCES

Bahrami B, Olsen K, Latham PE, Roepstorff A, Rees G, Frith CD. 2010. Optimally interacting minds. Science. 329:1081–1085.

Bode NWF, Franks DW, Wood AJ. 2012. Leading from the front? Social networks in navigating groups. Behav Ecol Sociobiol. 66:835–843.

Bode NWF, Wood AJ, Franks DW. 2011a. The impact of social networks on animal collective motion. Anim Behav. 82:29–38.

Bode NWF, Wood AJ, Franks DW. 2011b. Social networks and models for collective motion in animals. Behav Ecol Sociobiol. 65:117–130.

Boyd R, Richerson PJ. 1985. Culture and the evolutionary process. Chicago (IL): University of Chicago Press.

Codling EA, Pitchford JW, Simpson SD. 2007. Group navigation and the "many-wrongs principle" in models of animal movement. Ecology. 88:1864–1870.

Conradt L, Roper TJ. 2003. Group decision-making in animals. Nature. 421:155–158.

Conradt L, Roper TJ. 2005. Consensus decision making in animals. Trends Ecol Evol. 20:449–456.

Conradt L, Roper TJ. 2007. Democracy in animals: the evolution of shared group decisions. Proc R Soc B Biol Sci. 274:2317–2326.

Couzin ID, Krause J, Franks NR, Levin SA. 2005. Effective leadership and decision-making in animal groups on the move. Nature. 433:513–516.

Croft DP, James R, Krause J. 2008. Exploring animal social networks. Princeton (NJ): Princeton University Press.

Cwir D, Carr PB, Walton GM, Spencer SJ. 2011. Your heart makes my heart move: cues of social connectedness cause shared emotions and physiological states among strangers. J Exp Soc Psychol. 47:661–664.

David-Barrett T, Dunbar RIM. 2012. Cooperation, behavioural synchrony and status in social networks. J Theor Biol. doi: 10.1016/j.jtbi.2012.05.007.

David-Barrett T, Dunbar RIM. 2013. Processing power limits social group size: computational evidence for the cognitive costs of sociality. Proc R Soc B Biol Sci. 280:20131151. doi:http://doi:dx.doi.org/10.1098/rspb.2013.1151.

Degroot P. 1980. Information-transfer in a socially roosting weaver bird (Quelea-Quelea, Ploceinae)—an experimental-study. Anim Behav. 28:1249–1254.

Dunbar RIM. 2010. Brain and behaviour in primate evolution. In: Kappeler PH, Silk J, editors. Mind the gap: tracing the origins of human universals. Berlin (Germany): Springer. p. 315–330.

Dunbar RIM. 2012. Bridging the bonding gap: the transition from primates to humans. Philos Trans R Soc B. 367(1597):1837–1846. doi:http://dx.doi.org/10.1098/rstb.2011.0217.

Dunbar RIM, Korstjens AH, Lehmann J. 2009. Time as an ecological constraint. Biol Rev. 84:413–429.

Dunbar RIM, Spoors M. 1995. Social networks, support cliques, and kinship. Hum Nat. 6:273–290.

Dyer JRG, Johansson A, Helbing D, Couzin ID, Krause J. 2009. Leadership, consensus decision making and collective behaviour in humans. Philos Trans R Soc B. 364:781–789.

Ehardt CL, Bernstein IS. 1986. Matrilineal overthrows in rhesus-monkey groups. Int J Primatol. 7:157–181.

Galton F. 1907. Vox populi. Nature. 75:450–455.

Gould JL. 2004. Animal navigation. Curr Biol. 14:R221–R224.

Gould RV. 2002. The origins of status hierarchies: a formal theory and empirical test. Am J Sociol. 107:1143–1178.

Granovetter M. 1973. The strength of weak ties. Am J Sociol. 78:1360–1380.

Hamilton MJ, Milne BT, Walker RS, Burger O, Brown JH. 2007. The complex structure of hunter-gatherer social networks. Proc R Soc B Biol Sci. 274:2195–2202.

Harcourt AH, Waal FBMd. 1992. Coalitions and alliances in humans and other animals. Oxford: Oxford University Press.

Hemelrijk CK. 1999. An individual-orientated model of the emergence of despotic and egalitarian societies. Proc R Soc B Biol Sci. 266:361–369.

Hemelrijk CK. 2000. Towards the integration of social dominance and spatial structure. Anim Behav. 59:1035–1048.

Henrich J, Boyd R. 2008. Division of labor, economic specialization, and the evolution of social stratification. Curr Anthropol. 49:715–724.

Henrich J, Gil-White FJ. 2001. The evolution of prestige—freely conferred deference as a mechanism for enhancing the benefits of cultural transmission. Evol Hum Behav. 22:165–196.

Hill RA, Bentley RA, Dunbar RIM. 2008. Network scaling reveals consistent fractal pattern in hierarchical mammalian societies. Biol Lett. 4:748–751.

Hill RA, Dunbar RIM. 2003. Social network size in humans. Hum Nat. 14:53–72.

Hooper PL, Kaplan HS, Boone JL. 2010. A theory of leadership in human cooperative groups. J Theor Biol. 265:633–646.
Jackson MO. 2008. Social and economic networks. Princeton (NJ): Princeton University Press.
Jackson MO, Watts A. 2001. The existence of pairwise stable networks. Seoul J Econ. 14:299–321.
King AJ, Cowlishaw G. 2009. All together now: behavioural synchrony in baboons. Anim Behav. 78:1381–1387.
King AJ, Sueur C, Huchard E, Cowlishaw G. 2011. A rule-of-thumb based on social affiliation explains collective movements in desert baboons. Anim Behav. 82:1337–1345.
Konvalinka I, Xygalatas D, Bulbulia J, Schjodt U, Jegindo EM, Wallot S, Van Orden G, Roepstorff A. 2011. Synchronized arousal between performers and related spectators in a fire-walking ritual. Proc Natl Acad Sci USA. 108:8514–8519.
Kudo H, Dunbar RIM. 2001. Neocortex size and social network size in primates. Anim Behav. 62:711–722.
Kummer H. 1968. Social organization of hamadryas baboons. A field study. Basel (Switzerland): Karger.
Lehmann J, Dunbar RIM. 2009. Network cohesion, group size and neocortex size in female-bonded old world primates. Proc R Soc B Biol Sci. 276:4417–4422.
Lindquist WB, Chase ID. 2009. Data-based analysis of winner-loser models of hierarchy formation in animals. Bull Math Biol. 71:556–584.
Lu Y-E, Roberts S, Lió P, Dunbar RIM, Crowcroft J. 2009. Size matters: variation in personal network size, personality and effect on information transmission. Proceedings of IEEE International Conference on Social Computing, Vancouver, Canada.
Lusseau D, Conradt L. 2009. The emergence of unshared consensus decisions in bottlenose dolphins. Behav Ecol Sociobiol. 63:1067–1077.
Masuda N. 2007. Participation costs dismiss the advantage of heterogeneous networks in evolution of cooperation. Proc Biol Sci. 274:1815–1821. doi: 10.1098/rspb.2007.0294.
Nagy M, Akos Z, Biro D, Vicsek T. 2010. Hierarchical group dynamics in pigeon flocks. Nature. 464:890–894.
Pansini R. 2011. Induced cooperation to access a shareable reward increases the hierarchical segregation of wild vervet monkeys. PLoS One. 6: e21993. doi: 10.1371/journal.pone.0021993.
Psorakis I, Roberts SJ, Rezek I, Sheldon BC. 2012. Inferring social network structure in ecological systems from spatio-temporal data streams. J R Soc Interface. 9:3055–3066.
Rands SA, Cowlishaw G, Pettifor RA, Rowcliffe JM, Johnstone RA. 2003. Spontaneous emergence of leaders and followers in foraging pairs. Nature. 423:432–434.
Rands SA, Cowlishaw G, Pettifor RA, Rowcliffe JM, Johnstone RA. 2008. The emergence of leaders and followers in foraging pairs when the qualities of individuals differ. BMC Evol Biol. 8:51.
Rands SA, Pettifor RA, Rowcliffe JM, Cowlishaw G. 2004. State-dependent foraging rules for social animals in selfish herds. Proc R Soc Lond B Biol Sci. 271:2613–2620.
Roberts SGB, Dunbar RIM. 2011. The costs of family and friends: an 18-month longitudinal study of relationship maintenance and decay. Evol Hum Behav. 32:186–197.
Samuels A, Silk JB, Altmann J. 1987. Continuity and change in dominance relations among female baboons. Anim Behav. 35:785–793.
Seeley TD. 1995. The wisdom of the hive: the social physiology of honey bee colonies. Cambridge (MA): Harvard University Press.
Seyfarth RM. 1977. Model of social grooming among adult female monkeys. J Theor Biol. 65:671–698.
Sigg H, Stolba A. 1981. Home range and daily march in a hamadryas baboon troop. Folia Primatol (Basel). 36:40–75.
Sueur C, King AJ, Conradt L, Kerth G, Lusseau D, Mettke-Hofmann C, Schaffner CM, Williams L, Zinner D, Aureli F. 2011. Collective decision-making and fission-fusion dynamics: a conceptual framework. Oikos. 120:1608–1617.
Sutcliffe A, Dunbar R, Binder J, Arrow H. 2012. Relationships and the social brain: integrating psychological and evolutionary perspectives. Br J Psychol. 103:149–168.
Theraulaz G, Bonabeau E, Deneubourg JL. 1998. Response threshold reinforcement and division of labour in insect societies. Proc R Soc B Biol Sci. 265:327–332.
Traniello JFA, Rosengaus RB. 1997. Ecology, evolution and division of labour in social insects. Anim Behav. 53:209–213.
Traulsen A, Nowak MA. 2006. Evolution of cooperation by multilevel selection. Proc Natl Acad Sci USA. 103:10952–10955.
Tsitsiklis JN, Bertsekas DP, Athans M. 1986. Distributed asynchronous deterministic and stochastic gradient optimization algorithms. IEEE Trans Automat Contr. 31:803–812.
van Vugt M, Kurzban R. 2007. Cognitive and social adaptations for leadership and followership evolutionary game theory and group dynamics. Syd Sym Soc Psychol. 9:229–243.
Waibel M, Floreano D, Magnenat S, Keller L. 2006. Division of labour and colony efficiency in social insects: effects of interactions between genetic architecture, colony kin structure and rate of perturbations. Proc R Soc B Biol Sci. 273:1815–1823.
West SA, El Mouden C, Gardner A. 2011. Sixteen common misconceptions about the evolution of cooperation in humans. Evol Hum Behav. 32:231–262.
Wright J, Stone RE, Brown N. 2003. Communal roosts as structured information centres in the raven, *Corvus corax*. J Anim Ecol. 72:1003–1014.
Zhou WX, Sornette D, Hill RA, Dunbar RIM. 2005. Discrete hierarchical organization of social group sizes. Proc R Soc B Biol Sci. 272:439–444.