Power as the Integration of Information in a Signaling Network

Eleanor Brush

University of Chicago, Chicago, IL 60637 and

Santa Fe Institute, Santa Fe, New Mexico 87501

Jessica Flack, David Krakauer

Santa Fe Institute, Santa Fe, New Mexico 87501

Understanding the development of power structure is important in understanding the development of social structure. Power is predictive of social behaviors and affects the costs and benefits of living in a social group. We explore how individuals in a primate society use a network of subordination signals to determine the power structure of their group in hopes of understanding how this power structure arises. To do this, we develop two formalisms to measure the power distribution of the population. In addition, we are interested in how much and what kind of information is necessary for this process.

I. INTRODUCTION

First, we define what we mean by power. Power, simply put, is an individual's ability to obtain resources by coercing other individuals [1]. Three factors are important in our discussion of power. The first is the temporal stability of the power structure. Power should be predictive of other behaviors and should feed back to influence the behaviors that contribute to the power structure. However, if it changes as quickly as these underlying dynamics, it loses all predictive ability. Thus, power should change on a slower timescale than these other structures [2]. The second factor is that power is solely based on the perception that an individual can obtain the resources he needs. That is, it is more important for others to think an individual can fight well than that he actually be able to. As soon as an individual resorts to physical force to obtain the resources he wants, he loses legitimacy and thus power [1]. It follows that for an individual to be powerful requires consensus from the whole group since their (agreeing) perceptions about the individual are what create his power. Last, power structure is a group property. While it might be said that one individual in a pair is more powerful than another, we are more interested in an individual's power with respect to the entire population. Thus a powerful animal is one whom a large part of the population perceives as able to secure the resources he wants for a relatively long period of time.

In addition to these properties, power is meaningless if the individuals in a group are not aware of it. Therefore, every individual needs to be able to calculate his own and others' power in the group. Since any formalism we use to measure power should simulate calculations that individuals are actually performing, this constrains our formalism in two ways. First, our formalism should be calculatable using some heuristic. Second, it is unreasonable to expect that an individual will remember all of the interactions that have occurred in a group. Instead, individuals will preferentially pay attention to some of the information available. For example, he might remember the interactions between individuals with whom he spends the most time. Thus, our formalism should only require that individuals use local information. We are interested in how much information this power calculation requires.

We now turn our attention to power structure. The development of a power structure and the consequent reduction in social instability and uncertainty require the integration of information from lower level dynamics [2]. There are three dynamic networks that produce the power structure. The first is the interaction network. As individuals live together in a social group, they will participate in social activities like grooming, mating, sharing and others. Nodes in this network are individuals and edges between nodes indicate the frequencies of interactions between those individuals. Individuals interacting in a social group inevitably compete for resources such as food, mates, and partners in alliances. Competitions can escalate into fights. This integration from interactions into fights forms the aggression network. Again, nodes are individuals, but now edges represent the frequencies of fights between individuals. Over time as pairs of individuals fight, asymmetries develop. There are several factors that can determine the outcome of a particular fight. Some are temporally instable– health, weather, and the presence of coalition partners. Others are temporally invariant– body size, fighting ability, and coalition size. These temporally stable factors contribute to asymmetries, as it becomes evident that one individual in a pair is in general more likely to win.

Fighting is costly to both individuals: they both risk injury and expend energy during the fight. In a given pair, the weaker individual has a higher risk of injury or death. In order to avoid these costs, the weaker individual may try to prevent a developing conflict from escalating into an actual fight. One way to do this is for the weaker individual to signal his submission. Since fighting is costly to the stronger individual as well, he can accept this submission and consider the conflict won. The agonistic asymmetries may also lead to subordination signaling. The costs of fighting are not strictly physical. Conflict can disrupt the social system and perturb an individual's position within it. In order to maintain order and one's status, an individual may signal outside of a conflict setting to communicate that he is willing to be subordinate to stronger individuals in future interactions. This allows individuals to interact more peacefully since both individuals in the pair know that if they compete over a resource the signaler will cede it to his opponent. Subordination signaling is stronger than submission signaling in that it regards not just the present situation but future conflicts as well [2]. The integration of fighting asymmetries into a signaling network is the second level of integration. Here as above, nodes represent individuals, and now edges represent frequencies of signals sent between individuals. In distinction to the interaction and fight networks, the signaling network is directed in order to capture which individual is signaling and which is receiving.

Finally, as shown in Figure 1, the signaling network develops into a power network. The subordination signals are, in a sense, a kind of primitive social contract. They can be thought of as votes about a receiver's fighting ability. Since a subordination signal indicates a signaler believes the receiver to be able to win a fight, and power has to do with the perception of fighting ability, these subordination signals integrate to form the power network. As each level of structure develops, they in turn affect the lower level dynamics. As individuals are established as more or less powerful, their behavior and the behavior of the individuals around them will be affected. [2] It is the final level of integration that this work is focused upon.

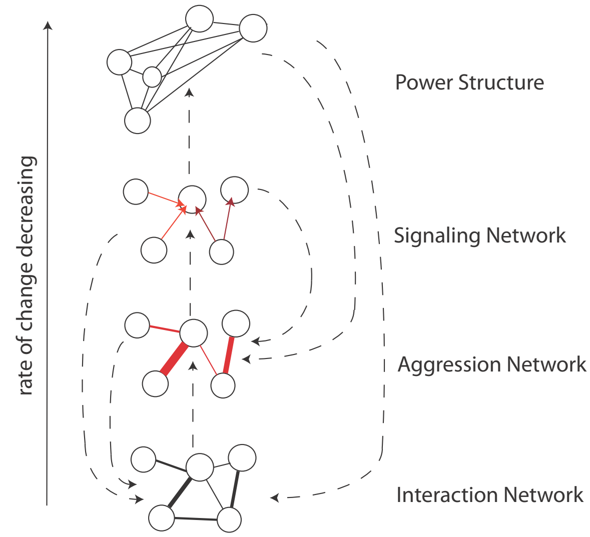

FIG. 1: Three social networks integrating to produce a power structure. As higher levels of structure develop, they in turn feed back to influence lower level dynamics.

II. METHODS AND DATA SET

Our model species is the pigtailed macaque (*Macaca nemestrina*). The subordination signal in this species is the silent bared teeth display (SBT). This signal is given in both conflict and peaceful situations, although it is thought to have evolved in the conflict setting [10]. In a conflict setting, the SBT indicates the signaler's submission in the particular conflict due to the signaler's wish to avoid a costly fight. However, the SBT indicates subordination when given in a peaceful context, i.e. when there is no danger of a fight breaking out. Our model system is a captive breeding group of pigtailed macaques at Yerkes National Primate Research Center in Lawrenceville, Georgia. There were 84 individuals in the group at the time of the study, including four adult males (6 years of age by study start), 25 adult females (4 years of age by study start), and 19 subadults (socially but not fully physically mature) (n=48). These 48 individuals were the only ones included in the study because younger animals have not established agonistic asymmetries. Data was collected between 1100 and 2000 hours from June until October 1998. Sampling of the population occurred every fifteen minutes. These sampling times gave rise to several types of data. The two that we will be primarily concerned with are signaling data and proximity data. The signaling data indicates the number of subordination signals sent during peaceful settings from each individual to each other individual over the course of the study. The proximity data indicates, for each pair, how many times the individuals were in arm's reach of each other. [8]

III. INTEGRATING SIGNALING NETWORK TO POWER STRUCTURE

We hope to develop formalisms to measure the degree to which a population agrees about an individual's fighting ability and how powerful they perceive him to be.

A. Original Power Measure

First, we describe a formalism that has already been developed [8]. This formalism makes several assumptions about the integration of signaling to a power structure. (1) Receiving more signals should increase an individual's power. (2) Transmitting more signals should decrease an individual's power. (3) An individual's power will be zero if there is no agreement or consensus in the population about his ability. (4) Maximal power requires there to be maximal agreement or consensus in the population about his ability. (5) Signals received from all individuals are weighted equally. (6) Receiving and signaling contribute independently to an individual's power. [8]

Based on these assumptions, we see that the number of signals an individual receives is related to the power he is perceived to have. Let M_{ij} be the number of signals sent from individual i to individual j. We define

$$R(j) = \sum_{i=1}^{n} M_{ij}$$

so that R is the sum of signals received. This certainly captures some of the information about power consensus in the group. However, a high R value might not indicate a high level of consensus. It could happen that an individual whose fighting ability in general is not very good has an anomalously successful fight history with one individual, leading the loser to signal many times to avoid more losses. It also might happen that an individual has a successful fight history with a particularly nervous individual, leading the nervous individual to signal more frequently than another might. In both cases, the winner will receive several signals from one individual, but not necessarily from any other individuals. This receiver will have a high R score, but there is no consensus in the population about his fighting ability.

This deficit in the R measure leads to another way of measuring power: the entropy of signals received. In information theory, the entropy of a probability distribution, $P = \{p(1), \ldots, p(n)\}$ is

$$H(P) = -\sum_{i=1}^{n} p(i) \log_2(p(i))$$

$H(P)$ is between 0 and $\log_2(n)$. Further, $H(P) = 0$ if and only p(i)=1 for some i and $p(j) = 0$ for all j not equal to i. Conversely, $H(P) = \log_2(n)$ if and only if $p(i) = \frac{1}{n}$ for all i. Entropy measures the uncertainty of a distribution: it is minimized when one event happens deterministically and is maximized when many events are equally probable. We now define

$$H_R(j) = -\sum_{i=1}^{n} \frac{M_{ij}}{R(j)} \log_2\left(\frac{M_{ij}}{R(j)}\right).$$

This is straightforward extension of the information theoretic definition of entropy. The signals, M_{ij}, have been normalized by $R(j)$ so that they sum to 1 and are thus a probability distribution as required in the definition of entropy. (If $M_{ij} = 0$ for some $1 \leq i \leq n$, we let $\frac{M_{ij}}{R(j)} \log_2\left(\frac{M_{ij}}{R(j)}\right)$ be 0. Similarly, if $R(j) = 0$, we define the entropy of signals received, $H_R(j)$, to be 0.) Now $H_R(j)$ is between 0 and $\log_2(k)$ where k is the number of individuals who signal to individual j. $H_R(j)$ is 0 if and only if individual j only receives signals from one individual and $H_R(j)$ is maximized when all signalers transmit the same number of times. The maximum possible value of $H_R(j)$, $\log_2(k)$, also increases with k. This measure increases as agreement in the population about an individual's fighting ability increases. Now suppose one individual receives one signal from all other individuals and another receives two signals from all other individuals. The entropy of the signals received will be $\log_2(n-1)$ for both. This shows correctly that the population agrees perfectly about the status of both individuals. However, the second individual should be regarded as more powerful since everyone sends him more signals. Therefore, entropy is not satisfactory either.

This formalizes our intuition about how the number of signals received can be translated into an idea of power. However, the signals transmitted by each individual should also affect power. If an individual transmits many signals, this indicates that he frequently agrees to subordination. If he transmits signals to many individuals, he agrees to be subordinate to a large part of the population. Both of these behaviors should subtract from his power in the group. We define $S(i)$ to be the total number of signals transmitted and $H_S(i)$ to be the entropy of signals transmitted. These should counteract $R(i)$ and $H_R(i)$.

Both the number of signals received (transmitted) and the entropy of signals received (transmitted) have advantages and disadvantages as measures of power. In order to capture the information that each measure captures, we combine the two. Let

$$P(i) = \alpha R(i) * H_R(i) - (1-\alpha) S(i) * H_S(i)$$

for some $0 \leq \alpha \leq 1$. The positive receiving termce is high if the number of signals received is high, the entropy of those signals is high, or both. Similarly for the negative signaling term. The constant α can be calibrated by comparing P to data that power should predict and seeing which value of α gives the highest correlation between this measure and these data. In previous work, it was shown that P has most predictive power when $\alpha = 1$. In other words, transmitting signals does not in fact subtract from an individual's power [8]. From now on, $H(i)$ will denote $H_R(i)$ and P will refer to $R(i) * H_R(i)$. Since P correlates well with data that power should predict, we will use P as a baseline measure of power later.

B. Random Walk and PageRank

We develop a different measure of power based on slightly different assumptions. In our previous formalism, signals from each individual were weighted equally. We change this to consider the fact that a signal from a powerful individual is more meaningful than one from a less powerful individual. Using this intuition, we recursively define power.

The principal eigenvector of a stochastic matrix S is informative about the structure of the network that S represents. We can represent a network G with n nodes by a stochastic matrix S where S_{ij} is the number of edges from vertex i to vertex j, normalized by the total number of edges from vertex i. There are two justifications for taking the principal left eigenvector of such a matrix to learn about G. The first is that the left eigenvector with eigenvalue 1 represents the stationary distribution of a random walk on G. That is, the i^{th} coordinate of such a vector is equal to the fraction of time spent on the i^{th} node during an infinite random walk on the graph. The second is called eigenvector centrality. Intuitively, a node in a graph should be considered more "central" or important if many other important nodes connect to it. Thus, the centrality of a node j should be the sum of the centralities of nodes i that connect to it weighted by the strength of the connection between each node i and node j. Letting c(j) denote the centrality of node j, we see that

$$c(j) = \sum_{i=1}^{n} c(i) S_{ij} \tag{1}$$

since S_{ij} is the strength of the connection from node i to node j. If $\vec{c} = (c(1), \dots, c(n))$ we find that

$$\vec{c} = \vec{c} \times S \tag{2}$$

So the centrality vector $\vec{c}$ is the left eigenvector of S with eigenvalue 1 (given that each $c(j) \geq 0$, $\sum_{j=1}^{n} c(j) = 1$), hence eigenvector centrality.

We can treat our signaling matrix as a weighted directed graph where an edge of weight k runs from vertex i to vertex j if the i^{th} individual signals the j^{th} individual k times. As a node is more important if other important nodes connect to it, an animal should be regarded as more powerful if many powerful individuals signal to it. We want to measure the power of each individual in the signaling network. Since eigenvector centrality measures the importance of each node in a graph, this is a relevant tool to use on our graph. However, before we look at eigenvector centrality we need to construct a stochastic matrix S. It does not suffice to normalize each row of the signaling matrix M by the sum of each row. Thinking about the random walk that such a matrix would represent, this would mean that from each node we walk away with probability 1 regardless of how many signals are sent. The probability of moving to a given node depends on the number of signals sent to that node, but the importance of a node should also depend on how many signals it transmits. Using this idea, we define

$$S_{ij} = \begin{cases} \frac{M_{ij}}{1218} & , \text{ if } i \neq j \\ 1 - \sum_{j=1}^{n} \frac{M_{ij}}{1218} & , \text{ if } i = j. \end{cases}$$

We normalize each entry in the signaling matrix by the total number of signals sent (1218). Further, we indicate that during a random walk on the graph, we stay at a given node with a probability that increases as the number of signals transmitted decreases. By construction, S is stochastic. The matrix S now defines a matrix and taking the principal left eigenvector as described above will capture the fraction of time spent at each node during a random walk and the eigenvector centrality of each node.

However, in this case, simply taking the principal left eigenvector does not suffice. In our system, there is one individual, with index 1, who does not transmit any signals so that during the random walk one stays at node 1 with probability 1. This makes node 1 an absorbing state in the random walk. Thus, the fraction of time spent at node 1 would be 1 and that spent on other nodes would be 0. This tells us correctly that individual 1 is the most powerful but is uninformative about the power of all other individuals.

There are two ways to fix the problem. One solution is to make the random walk finite. The length of the random walk thus becomes an important parameter and affects the distribution of power that we get from this metric. Given initial probability distribution $\overline{\Phi}$, the fraction of time spent at vertex i during a random walk of length m is the i^{th} coordinate of the n-dimensional vector

$$\left(\frac{1}{m+1}\right) * \left(\sum_{i=0}^{m} \overline{\Phi} \cdot S^{i}\right)$$

We let $\overline{\Phi} = \{\frac{1}{n}, \frac{1}{n}, \dots, \frac{1}{n}\}$, a uniform distribution over all vertices. Let F_m denote this vector representing the fraction of time spent at each node during a random walk of length m. Taking P as a baseline power metric, we can compare the distribution of power under P and that under F_m for various m. For a given power distirubtion, we can order individuals, giving them rank equal to the number of individuals with a higher power score plus one. For example, if a given individual has the highest power score, he gets rank $0 + 1 = 1$. This way of ranking allows for ties: if two individuals have the same power score, the same number of individuals will have higher scores than the two of them. To see how closely a given F_m matches P we can count the number of individuals whose rank is different under the two metrics. Similarly, we can calculate the Kendall rank correlation between the distribution of F_m and P. See Table I. Using this information, we can see for which m F_m best matches our baseline measure P. We do not pursue this here. Instead, we pursue another solution that sidesteps the absorbing state problem.

For another way to solve the absorbing state problem, we draw upon Google's PageRank algorithm. The World Wide Web can be seen as a network where websites are nodes and a directed edge runs from node i to node j if website i provides a link to website j. PageRank simulates a surfer on the web randomly clicking on links, or randomly walking on the WWW network, given that he has an equal probability of starting at every website. To find a site's PageRank, one calculates the fraction of time this random surfer will spend at a given website [14]. This is analogous to the random walk on the signaling network.

The network of websites has a problem similar to ours when attempting to calculate the fraction of time spent at

TABLE I: How the number of steps of the random walk affects (1) the number of animals whose rank is different under F_m and P and (2) the Kendall rank correlation between the P power vector and the random walk power vector. An animal's rank under a certain measure is the number of animals whose power score is strictly higher.

m	Changes in Rank	Kendall rank correlation .
1	36	0.818906
5	36	0.805967
10	38	0.804174
15	39	0.802381
20	39	0.802381
25	39	0.798795
30	40	0.797002
35	40	0.78983
40	40	0.788037
50	41	0.782658
150	44	0.757555
500	44	0.734246

a given site. Like the most powerful individual who signals no other individuals, there are websites that provide no links to any other site. Further, some pairs of websites only provide links to each other. If another website provides a link to one of these, the random walker will get stuck going back and forth between these two sites and will incorrectly increase the fraction of time spent on these websites. This "rank sink" makes the fraction of time spent at websites uninformative about their authority or power. PageRank compensates for this problem by providing a "rank source," a vector V (of length equal to the number of websites in the network) that adds some amount of rank to each website. In terms of the random walk, the rank source gives some probability that at any point in the random walk the walker will start over at another node. We can now modifiy (1):

$$c(j) = \sum_{j=1}^{n} c(i)S_{ij} + V(j) \quad (3)$$

Again, let $\vec{c} = (c(1), \ldots, c(n))$. Since $c(j)$ denotes a fraction of time, implying that $\sum_{j=1}^{n} c(j) = 1$, if $\vec{1}$ denotes a n-dimensional column vector consisting of 1s,

$$\vec{c} = \vec{c} \times (S + \vec{1} \times V) \quad (4)$$

With the addition of the rank source vector, the power vector is now an eigenvector of the matrix $S + \vec{1} \times V$, rather than just S as in (2). [14]

In the PageRank algorithm, the rank source was a uniform distribution over all websites. In our case, we would like a more principled distribution. In a sense, a random walk over the signaling network can be thought of as watching an individual, transferring attention to a receiver of this individual's signal, and continuing to transfer attention as the individuals signal other animals. During this process, the observer may lose track or stop paying close attention and start watching some other animal. This random transferral of attention is similar to the surfer on the web randomly starting over on a new website and can be simulated with the rank source distribution. So we must decide who an observer is most likely to switch his attention to.

One possibility is that he would preferentially pay attention to more social animals. Let $prox(i, j)$ denote the proximity score for the pair of individuals (i, j), i.e. the number of times the pair has been within arm's reach of each other. (Of course, $prox(i, j) = prox(j, i)$.) We now give each individual a sociability score. Let

$$soc(i) = \sqrt{\frac{\sum_{j=1}^{n} prox(i, j)}{151}}$$

We divide by 151 so that the maximum sociability score is 1. We can now let

$$V(i) = \frac{soc(i)}{\sum_{i=1}^{n} soc(i)}$$

Since we normalize by the sum of sociability scores, $\vec{1} \times V$ is stochastic. Thus $.9S + .1(\vec{1} \times V)$ is also a stochastic matrix and taking the principal left eigenvector of this matrix gives a power distribution. Henceforth $F(i)$ denotes the i^{th} coordinate of this eigenvector.

In Figure 2, we plot the five different measures of power consensus against a given rank ordering.

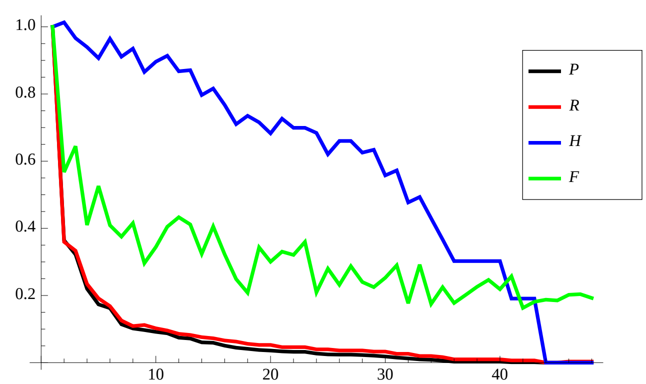

FIG. 2: Plot of four different measures of power consensus.

IV. HOW MUCH INFORMATION IS NECESSARY

We now modify the signaling matrix in different ways, keeping the convention that the ij entry of the matrix is the (modified) number of signals sent from individual i to individual j. All of the measures of power can be extrapolated without any difficulty to these modified matrices. For example, $R(i)$ applied to any modified matrix is still the sum of the entries of the i^{th} column. In applying $F(i)$ to other matrices, we construct the stochastic matrix S in the same way, dividing by 1218 because, while the total number of signals in a modified matrix may be different, the difference is slight and the important thing is to create a stochastic matrix. The rank source vector remains the same and we take the left eigenvector as above.

A. Coarse Graining the Signaling Matrix

All of the measures of power consensus we have considered so far use all of the information available in the signaling matrix. The goal is to produce a measure of power that mimics the power structure of the macaque society and perhaps duplicates calculations that the animals execute in order to determine the power of individuals around them. However, keeping track of the signals sent and received by all other individuals and yourself (in other words, all of the information in the signaling matrix) is computationally burdensome. It is unreasonable to expect that any individual is counting and remembering the number of signals sent and received by the entire population. Instead, individuals probably preferentially pay attention to some information in the network. We are interested in finding out how little information we can use in measuring power and still capture the true power structure. This may indicate how much information individuals must remember in order to determine the true power structure. We coarse-grain the signaling matrix repeatedly, each time systematically decreasing the information content of the matrix, and we see how this coarse-graining affects our power measures.

Here we explore different methods of systematically removing information from the matrix, in hopes of learning how much and what kind of information receivers pay attention to when calculation power. It is reasonable to expect that an individual might confuse two signalers who behave similarly towards him. Further, the more signals an individual receives and the more varied the number of signals different individuals transmit to him, the more easily the individual may confuse signalers. For example, if an individual receives one signal from two signalers, he is unlikely to confuse these individuals with each other (and if he were to, his calculations about his own power would not be affected). On the other hand, the most powerful individual in our model system received from two to twenty-eight signals and received signals from all other forty-seven individuals, so his confusing two signalers seems much more probable. Based on this intuition, we can construct another sequence of coarse-grained matrices, $D1, \ldots, D10$.

For each individual, we define an averaging threshold:

$$T(i) = \sqrt{\max\{\text{signals } i \text{ received}\} - \min\{\text{signals } i \text{ received}\}}$$

For example, for the most powerful individual,

$$T(1) = \sqrt{28-2} = \sqrt{26} = 5.09902$$

In a sense, this captures the spread or range of signals the individual receives. This can be scaled by a varying constant c, $0 < c < 1$, to change the severity of the threshold. M_{ji} as usual denotes the number of signals sent from individual j to individual i. For a given c and receiver i, each M_{ji} is averaged with all M_{ki} such that M_{ki} is within $c * T(i)$ of M_{ji}. Now for $1 \leq k \leq 10$, Dk denotes the original signaling matrix coarse-grained through this process using constant $c = k/10$.

Receivers might also pay less attention to individuals they recognize as less important. This would mean averaging the individuals who signal the most to a given receiver, since the tendency to signal could be a proxy for a lack of importance. For each receiver, we can average first the highest two signalers, then the highest three, etc., creating a sequence of coarse-grained matrices R1, . . . ,R(n-1).

Animals may have a tendency to ignore younger animals. First, younger animals have not had the time to interact much with the whole population yet. Animals may have a harder time distinguishing younger animals until they are more familiar with them. Second, younger animals are in the process of learning how to behave within adult society. They are thus much more likely to signal sporadically. A signal sent by a young animal who does not know any better is meaningless (in terms of conveying information about the fight history of the signaler and receiver and the power structure of the group). These signals should be ignored so that individuals do not miscalculate their own or others' power. This leads to another way of coarse-graining: ignoring signals sent by younger individuals.

Ignoring an individual of age i with probability $p_i \leq 1$ can be simulated by multiplying an individual's signaling vector by p_i. The oldest animals were 156 months old by the end of the study and the youngest were 32 months old. We assume that individuals pay full attention to the oldest animals, i.e. $p_{156} = 1$. With no coarse-graining, individuals pay full attention to the youngest animals, so that $p_{32} = 1$. In the extreme case, receivers completely ignore these animals, so that $p_{32} = 0$. Intermediate levels of ignoring can be achieved by having p_{32} run between 0 and 1.

What about p_i for $32 < i < 156$? Given that $p_{32} = p$ for some $0 \leq p \leq 1$, we can find the line f_p such that $f_p(32) = p$ and $f_p(156) = 1$. We can designate for $32 \leq i \leq 156$, $p_i = f_p(i)$. As p decreases from 1 to 0 the severity to which animals ignore younger individuals increases, and conversely as p increases. This defines a series of coarse-grained matrices, Ak, where Ak is the matrix created by using the constant $p = 1 - k/10$ for $1 \leq k \leq 10$.

Above we argued that an animal might not be able to distinguish between younger animals because they have not had a long interaction history. Similarly, an animal might have difficulty distinguishing between individuals with whom he has not spent a lot of time. We have proximity scores for each pair of individuals in the group that indicate how much time those two individuals spend together. If a given receiver spends no time with two signalers, he has no basis on which to distinguish them. Even if an individual has spent some time with a set of individuals it may be so little that he cannot distinguish between them.

We must clarify what “little time” means. This will vary from receiver to receiver. If an animal spends time with only two other animals, he is unlikely to confuse those two animals. On the other hand, if an animal spends a good amount of time with several animals, he is more likely to confuse those with whom he has only interacted a little. We can see that a receiver's sociability within the group will affect how he coarse-grains the signals transmitted to him. As above where we use sociability to modify the random walk measure of power, we let

$$soc(i) = \sqrt{\frac{\sum_{j=1}^{n} prox(i,j)}{151}}$$

For each individual and varying constants c where $0 \leq c \leq 1$ we define a threshold

$$T_c(i) = c * soc(i) * \max_j \{prox(i,j)\}$$

The sociability of an individual affects his threshold in two ways: with higher sociability in general $soc(i)$ increases and $\max_j\{prox(i,j)\}$ indicates the time an individual spends in proximity to his “best friend”. The justification for multiplying sociability by this maximum proximity score is that a receiver will compare signalers to the individual with whom he spends the most time as a baseline proximity score. For a given c, we create a coarse-grained matrix by averaging all signals sent to each receiver by individuals whose proximity scores with that receiver are less than $T_c(i)$. This gives a series of coarse-grained matrices, Pk, where Pk is the matrix created by using the constant $c = k/10$ for $1 \leq k \leq 10$.

B. Recovering Underlying Networks

We have pursued different ways by which to measure power consensus and different ways to mimic the cognitive deficits of the animals. Now we need to see the extent to which these different consensus measures can recover properties of the SBT and agonistic interaction networks, the extent to which the type of coarse-graining changes the ‘recovery capacity’, and the extent to which the different consensus measures, with and without coarse-graining, predict one another.

1. *Predicting Original PCI Entropy Measure*

A good method of measuring power consensus would be robust under coarse-graining: we would recover the same power distribution even if the animals can confuse signalers and forget information. There are two things to measure here– the rank of the animals and the relative differences in power between individuals– and there are several ways to measure how different the distributions become. For the following tests, I compare various measures of power (P,R,H,F) on coarse-grained data (through the various methods of coarse-graining) to P applied to the original signaling data.

1. For a given measure of power, we can give each animal a ranking equal to one more than the number of animals with higher power scores. In comparing one PCI on coarse-grained data to P on the original data, we can count the number of individuals whose rank is different under the two power distributions. We can also sum the absolute difference in rank for each individual. Kendall rank correlation is similarly informative about how the ranking changes.

2. To see generally how different a coarse-grained distribution is from the original, we can treat each distribution as a power vector, whose i^{th} coordinate is the power score for the i^{th} individual and find the distance between the two power vectors.

3. There are probabilistic tests we can perform between P on the original data and a coarse-grained power distribution to measure the distance between the two. Let X,Y be two distributions over n samples and let $\overline{X}, \overline{Y}$ represent their respective means,

   (a) t-test where

   $$t = (\overline{X} - \overline{Y}) * \sqrt{\frac{n*(n-1)}{\sum_{i=1}^{n}(X(i) - \overline{X} - Y(i) + \overline{Y})^2}}$$

   (b) χ^2 test where $\chi^2 = \sum_{i=1}^{n} \frac{(X(i)-Y(i))^2}{Y(i)}$

4. We can perform a linear regression of the coarse-grained power distribution against the original power distribution and use the R^2 value as a measure of how close the two distributions are.

We find the sum of the changes in rank between each measure of power on the last in each sequence of matrices (the most severe coarse-graining) and P on the original data. Further, we find the R^2 value for linear regression of each measure of power on the most severe in each sequence of matrices against P on the original signaling network. For example, in Table II we can look at how coarse graining affects the random walk measure of power. To take into account how different each measure of power on the original data is from P on the original data, I subtract the

TABLE II:

Coarse Graining	Sum of Changes in Rank	R^2 Value
None	249	0.824966
Similarity	247	0.825921
High Signalers	241	0.830316
Young	335	0.639718
Proximity	293	0.798779

original sum of rank changes (R^2 value) from each coarse-grained sum of rank changes (R^2 value). To get a mean score for how coarse-graining affects a measure of power, I average over the four methods of coarse-graining to get a score for each PCI. For example, for the random walk measure of power I average

$$247 - 249,\ 241 - 249,\ 335 - 249,\ 293 - 249$$

to get 30. See Table III.

TABLE III:

PCI	Mean difference in rank changes	Mean difference in R^2
P	28.75	0.00149618
R	16.25	0.00171482
H	141.26	0.0468465
F	30	0.0512821

A large mean difference in sum of rank changes (R^2) indicates that in general the measure of power is more affected by coarse-graining. We can therefore see that H is by far the most strongly affected when we look at the difference in the sum of rank changes. In considering changes in R^2 values, F is the most strongly affected, although it is not much more affected then H.

Entropy is maximal when all signals/ events occur with equal probability, so it stands to reason that as we average signals, the entropy of signals received should increase (perhaps quite a bit). Similarly, when we average, the sum of the averaged values should be close, if not equal, to the original signals, so that the total number of signals received by any individual will not change much. This explains why R is the most stable and why H is quite unstable.

V. FUTURE WORK

There are a number of questions that I would like to pursue in continuing this research. First, I would like to fine tune our methods of coarse-graining. We hope to capture the way that individuals actually preferentially pay attention to some parts of the signaling network. Before we can say that we are truly doing this, our methods need to be perfected. I would also like to pursue other methods of coarse-graining. It might be the case that receivers associate signalers with their matriline, or family unit, particularly when the signalers are quite young. Another way to remove information from the original signaling matrix might be to associate individuals from the same matriline. In addition, there are three basic methods of coarse-graining: hierarchically, by increasing group size, or by scaling. In coarse-graining hierarchically, receivers clump a couple of individuals and then clump these clumped individuals with others, and continue this way. When coarse-graining by increasing group size, a receiver may first average two individuals, then average three, etc., re-averaging each time. Coarse-graining by similarity, importance, and proximity are examples of this method. Finally, we can scale signals sent from certain individuals based on different criteria. Coarse-graining by age is an example of this method. These represent fundamentally different ways that individuals are preferentially paying attention to information and more thought should be given to which way is better.

P has shown to correlate well with independent data that power should predict. I compared our new formalism against P, but I plan to see how well all of the measures of power correlate with this independent data. This can give a better indication of how well each formalism captures power. I also plan to see how coarse-graining the matrices affects the predictive ability of different measures of power. A good measure of power would ideally predict independent data significantly well for all severities of coarse graining.

This research is part of a larger project. There are several larger questions that we hope to address in the future. These include how well power based on the signaling networks can capture the underlying signaling and fighting networks. The timescales over which the various networks- fighting, agonistic asymmetries, signaling, and power- change are also of interst.

Acknowledgments

E.B. received funding from the National Science Foundation's Research Experience for Undergraduates Program. Any errors in this work are due to E.B.

[1] Robert Bierstedt. An analysis of social power. *American Sociological Review*, 15(6):730–738, December 1950.

[2] Christopher Boehm and Jessica C. Flack. The emergence

of simple and complex power structures through social niche construction.

[3] Sergey Brin and Lawrence Page. The anatomy of a large-scale hypertextual web search engine. In *Computer Networks and ISDN Systems*, pages 107–117, 1998.

[4] Richard C. Connor, Rachel A. Smolker, and Andrew F. Richards. Dolphin alliances and coalitions. In Alexander H. Harcourt and Frans B. M. de Waal, editors, *Coalitions and alliances in humans and other animals*, chapter 15, pages 415–444. Oxford University Press, 1992.

[5] Larissa Conradt and Timothy J. Roper. Consensus decision making in animals. *TRENDS in Ecology and Evolution*, 20(8):449–456, August 2005.

[6] Carolyn L. Ehardt and Irwin S. Bernstein. Conflict intervention behaviour by adult male macaques: structural and functional aspects. In Alexander H. Harcourt and Frans B. M. de Waal, editors, *Coalitions and alliances in humans and other animals*, chapter 4, pages 83–112. Oxford University Press, 1992.

[7] Vincent S. E. Falger. Cooperation in conflict: alliances in international politics. In Alexander H. Harcourt and Frans B. M. de Waal, editors, *Coalitions and alliances in humans and other animals*, chapter 12, pages 323–350. Oxford University Press, 1992.

[8] J.C. Flack and D.C. Krakauer. Encoding power in communication networks. *The American Naturalist*, 168(3):87–102, September 2006.

[9] Jessica C. Flack and Frans B.M de Waal. 'any animal whatever': Darwinian building blocks of morality in monkeys and apes. *Journal of Consciousness Studies*, 7(1-2):1–29, 2000.

[10] Jessica C. Flack and Frans B.M de Waal. Context modulates signal meaning in primate communications. *Proceedings of the National Academy of Sciences of the United States of America*, 104(5):1581–1586, January 2007.

[11] Jessica C. Flack, David C. Krakauer, and Frans B.M de Waal. Robustness mechanisms in primate societies: a perturbation study. *Proceedings of the Royal Society B*, 2005.

[12] Karl Grammer. Intervention in conflicts among children: contexts and consequences. In Alexander H. Harcourt and Frans B. M. de Waal, editors, *Coalitions and alliances in humans and other animals*, chapter 10, pages 259–284. Oxford University Press, 1992.

[13] Terance D. Miethe. Types of consensus in public evaluations of crime: An illustration of strategies for measuring "consensus". *The Journal of Criminal Law and Criminology*, 75(2):459–473, 1984.

[14] Lawrence Page, Sergey Brin, Rajeev Motwani, and Terry Winograd. The pagerank citation ranking: Bringing order to the web. Technical Report 1999-66, Stanford InfoLab, November 1999.

[15] Joan B. Silk. Patterns of intervention in agonistic contests among male bonnet macaques. In Alexander H. Harcourt and Frans B. M. de Waal, editors, *Coalitions and alliances in humans and other animals*, chapter 8, pages 215–232. Oxford University Press, 1992.