bioRxiv preprint doi: https://doi.org/10.1101/170266; this version posted November 1, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Social fluidity mobilizes contagion in human and animal populations

Ewan Colman*[1], Vittoria Colizza[2], Ephraim M. Hanks[3], Andreas P. Modlmeier[4], David P. Hughes[4], and Shweta Bansal[1]

[1]Department of Biology, Georgetown University, [2]INSERM, Paris, [3]Department of Statistics, Pennsylvania State University, [4]Department of Entomology, Pennsylvania State University

Humans and other group-living animals tend to distribute their social effort heterogeneously; individuals predominantly interact with their closest companions while maintaining weaker social bonds with less familiar group members. By incorporating this heterogeneity into a mathematical model we find that a single parameter, which we refer to as social fluidity, controls the level of social mixing in the population. Large values of social fluidity correspond to gregarious behavior whereas small values signify the existence of persistent bonds between individuals. To investigate how social behavior influences the likelihood of an epidemic outbreak we derive an analytical expression of the relationship between social fluidity and the basic reproductive number of an infectious disease. We compare social behavior across 12 species by applying the model to empirical human and animal social interaction data. For species that form strong social bonds, the model describes frequency-dependent transmission that is highly sensitive to changes in social fluidity. As social fluidity increases, animal-disease systems become increasingly density-dependent. Finally, based on a computational disease spread model on empirical social data, we find that social fluidity is a stronger predictor of disease outcomes than both group size and connectivity.

Social behavior is fundamental to the survival of many species such as ants, humans, and dolphins. It allows the formation of social groups providing fitness advantages from greater access to resources and better protection from predators [1]. Within these groups structure emerges from the interactions that occur when individuals communicate across space, cooperate in sexual or parental behavior, or clash in territorial or mating conflicts [2]. While many animal societies have been studied independently [3], some questions about the nature of social living can only be answered by comparing social behavior across a range of species [4]. This motivates the question: how can we compare the social behavior of one species to that of another?

Social animals typically have limited time and resources to invest in their relationships [5]. To compensate, the social effort of an individual tends to be distributed heterogeneously among the members of their group. Despite the growing evidence for this in human communication [6, 7, 8], attempts to quantify this aspect of sociality in animal systems are often challenged by unavoidable sampling biases [9]. Furthermore, while heterogeneous interaction frequencies and temporal dynamics such as circadian rhythms and bursty activity patterns have become common in social network models [10], realistic assumptions about how individuals distribute their social effort are rarely incorporated.

When social interaction requires shared physical space it can also be a conduit for the transmission of infectious disease [11]. It is generally thought that if transmission occurs through the environment then the risk of epidemic is driven by group size (density-dependence) [12], whereas if transmission requires close proximity encounters that only occur between bonded individuals then we expect social connectivity to determine the outcome (frequency-dependence) [13]. In reality, however, animal-disease systems are not so easy to categorize [14]. For example, social ties must be created to maintain cohesiveness in a growing social group [15], implying that transmission through direct contact may in fact depend on population density. On the other hand, as social effort is distributed heterogeneously, low levels of social mixing may constrain the spread of the disease [16].

Here, we introduce a mathematical model based on the concept of *social fluidity* which we define as heterogeneity in the distribution of social effort. Using empirical data from previous studies, we estimate the social fluidity of a number of human and animal social systems. Using analytical and computational models of disease spread we show that social fluidity predicts disease outcomes better than other social behavioral indicators and spans the distance between density-dependent and frequency-dependent disease systems.

Modeling social heterogeneity & disease transmission

Our objective is to measure social heterogeneity in a range of human and animal populations and provide an understanding of how social behavior influences the susceptibility of the group to infectious disease. We start by introducing a model that captures hidden elements of social dynamics (in particular, how individual group members distribute their social effort) and mathematically describe the relationships between social variables that are routinely found in studies of animal behavior (in particular, the number of social ties and the num-

Significance

The study of relationships within animal groups has provided many insights into the nature of social behavior. Despite this, finding a reliable measure of sociality that can be used to compare animal social systems has remained a challenge. We introduce *social fluidity*, a measure of heterogeneity in how individuals choose to distribute their social effort across the group. We demonstrate the use of this measure by applying it to 57 social networks from 12 animal species. Our results indicate that social fluidity is a better predictor of disease spread than measures typically used in social network analysis such as the mean number of social ties or the size of a social group.

*ec975@georgetown.edu

bioRxiv preprint doi: https://doi.org/10.1101/170266; this version posted November 1, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

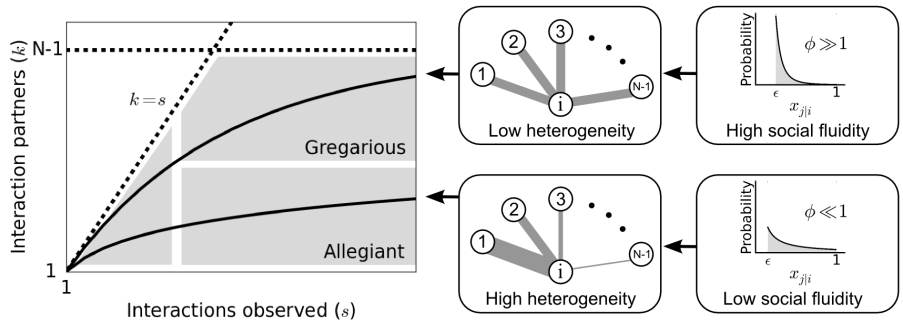

Figure 1: Left: Each individual can be represented as a single point on this plot. Dashed lines mark the boundary of the region where data points can feasibly be found. (4) is plotted for two values of ϕ representing two possible types of social behavior; as the number of observed interactions grows, the set of social contacts increases; the rate at which it increases influences how we categorize their social behavior. Middle: The weight of the edges between i and the other nodes represents the propensity of i to interact with each of the other individuals in the population. Right: Probability distributions that correspond to the different levels of heterogeneity in the contact propensities, both distributions are expressed by (2).

ber of interactions observed). We then couple this social behavior model with a general model of infectious disease spread to expound the link between social behavior and disease outcomes.

Social behavior model. Consider a closed system of N individuals and a set of interactions between pairs of individuals that were recorded during some observation period. These observations can be represented as a network: an individual, i, is a *node*; if at least one interaction between nodes i and j has been observed then we say that an *edge* exists between them; the *edge weight*, $w_{i,j}$, denotes the number of times this interaction was observed. The total number of times i was observed interacting is denoted by *strength*, $s_i = \sum_j w_{i,j}$, and the number of nodes with whom i is observed interacting is its *degree*, k_i.

We focus on one individual in the system, which we name the *focal* node, i, and consider the s_i interactions in which it participated. We define $x_{j|i}$ to be the probability that an interaction involving i will also involve node j. Therefore the probability that at least one of these interactions is with j is $1-(1-x_{j|i})^{s_i}$.

The main assumption of the model is that the values of $x_{j|i}$ over all i,j pairs are distributed heterogeneously according to a probability distribution, $\rho(x)$[1]. Thus, after s observed interactions of the focal node, the probability that an edge exists between the focal node and any other individual in the population is

$$\Psi(s) = 1 - \int \rho(x)(1-x)^s dx. \tag{1}$$

Our goal is to find a form of ρ that accurately reproduces network structure observed in real social systems. Motivated by our exploration of empirical interaction patterns from a variety of species (Figure S1), we propose that ρ has a power-law form:

$$\rho(x) = \frac{\phi\epsilon^\phi}{1-\epsilon^\phi} x^{-(1+\phi)} \text{ for } \epsilon < x < 1, \tag{2}$$

where ϕ (>0) controls the heterogeneity in the values of x, and ϵ simply truncates the distribution to avoid an asymptote at $x=0$. Combining (1) and (2) we find that

$$\Psi(s,\phi,\epsilon) = 1 - \frac{\phi\epsilon^\phi(1-\epsilon)^{s+1}}{(1-\epsilon^\phi)(s+1)} {}_2F_1(s+1, 1+\phi, s+2, 1-\epsilon) \tag{3}$$

[1] $x_{j|i}$ are subject to network interdependencies. Specifically, $AX = X^T A$ and $X\mathbf{1} = \mathbf{0}$, where X is a matrix whose i,j entry is -1 if $i=j$ and $x_{j|i}$ otherwise, A is any diagonal matrix with positive entries, and $\mathbf{0}$ and $\mathbf{1}$ are column vectors of length N containing only 0 and 1, respectively. Thus, $\rho(x)$ is the distribution of marginal $x_{j|i}$ values of the joint distribution $P(X)$.

where the notation ${}_2F_1$ refers to the Gauss hypergeometric function [17]. The value of ϵ is determined by $\sum_j x_{j|i} = 1$ and therefore depends on N and ϕ (*Materials and Methods* A.). It follows that the degree of the focal node is determined by its activity rate, the level of heterogeneity in partner choice, and the size of the population (*Materials and Methods* B.). More explicitly,

$$\langle k \rangle = (N-1)\Psi(s,\phi,N) \tag{4}$$

where $\langle k \rangle$ is the expected degree of the focal node.

The parameter ϕ controls the heterogeneity in interaction frequencies and is therefore the main determinant of social mixing in the model. We use the term *social fluidity* to refer to this quantity. Figure 1 illustrates how the value of ϕ can create different types of social behavior. Low social fluidity ($\phi \ll 1$) would produce what we might describe as "allegiant" behavior: interactions with the same partner are frequently repeated at the expense of interactions with unfamiliar individuals. As ϕ increases, the model produces more "gregarious" behavior: interactions are repeated less frequently and the number of partners is larger. While this phenomena could be described as "social strategy" or "loyalty" [18, 19], we feel that "social fluidity" is most appropriate for this work as it conveys an intuitive notion relevant to disease propagation [20].

Disease transmission model. Building on the social behavior model introduced in the previous section, we consider a model of disease transmission. We focus on the basic reproductive number R_0, defined as the mean number of secondary infections caused by a single infectious individual in an otherwise susceptible population. We first derive a formula for the individual reproductive number, $r(s_i)$, in terms of the relative rate of activity of the individual, and define R_0 to be the mean of $\{r(s_i)\}$ over the population. By calibrating the disease parameters in a way that controls for varying timescales of activity in different populations, we obtain a result that quantifies the relative significance of social fluidity across a range of social systems including those driven by group size (density-dependent transmission) and those driven by group connectivity (frequency-dependent transmission).

In our model, disease transmission requires an infected individual, i, who is social (i.e. has s_i interactions during an observation period of duration Δt) and may interact with any other individual, j (with probability $x_{j|i}$). When such an interaction occurs, i will infect j with probability β. The infectiousness of i continues for a period of duration τ, where τ is

bioRxiv preprint doi: https://doi.org/10.1101/170266; this version posted November 1, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

an exponentially distributed random variable with mean $1/\gamma$ (i.e. γ is the recovery rate).

Assuming the interactions of i are distributed randomly across the observation period, it follows that transmission events follow a Poisson process with rate $s_i x_{j|i}\beta/\Delta t$. Thus the probability that infection transmits from i to any given j is

$$T_{i\to j}(\tau, s_i, x_{j|i}) = 1 - \exp(s_i x_{j|i}\beta\tau/\Delta t). \tag{5}$$

The reproductive number for i, $r(s_i)$, is found by integrating (5) over all possible values of τ and $x_{j|i}$ then multiplying by the number of susceptible individuals, $N-1$ (SI Text).

Instead of choosing infection parameter values that pertain to a specific disease or social system, we select values for each system separately in a way that exposes the effects of population size and social fluidity. We achieve this by setting the mean infectious period in such a way that would cause R_0 to be equal to the constant R^* if the population was large and homogeneously mixed, i.e every interaction is with a new partner (*Materials and Methods* D.). Consequently, social systems with a high interaction frequency are coupled with diseases that have relatively short mean infectious periods.

After calibrating the parameters we can derive the following result for the reproductive number of an individual that was observed interacting s times,

$$\begin{aligned} r(s) = &\frac{1-\phi}{\phi(\epsilon^\phi - \epsilon)}\Big[1 - \epsilon^\phi + \epsilon^\phi {}_2F_1(-\phi, 1, 1-\phi; -R^* s/\langle s\rangle) \\ &-{}_2F_1(-\phi, 1, 1-\phi; -\epsilon R^* s/\langle s\rangle)\Big], \end{aligned} \tag{6}$$

where $\langle s\rangle$ is the mean of $\{s_i\}$ over the population. In all the analysis presented we arbitrarily choose $R^* = 2$.

To further expose the dependencies between population density, social fluidity, and disease transmission we evaluate N and R_0 over a range of ϵ and ϕ values (the formula for $N(\epsilon, \phi)$ is given in *Materials and Methods* A. while for R_0 we assume $s = \langle s\rangle$ in (6)). At small population sizes, R_0 increases with N and converges as N goes to ∞ (Figure 2**A**). The rate of this convergence increases with ϕ. When $\phi < 1$, the limit of R_0 is a function of ϕ (Figure 2**B**). At these values of ϕ, the individual will choose to repeat interactions despite having the choice of infinitely many potential interaction partners. When $\phi > 1$ and the population is large the probability of a repeated interaction falls to zero and we have that $R_0 = R^*$.

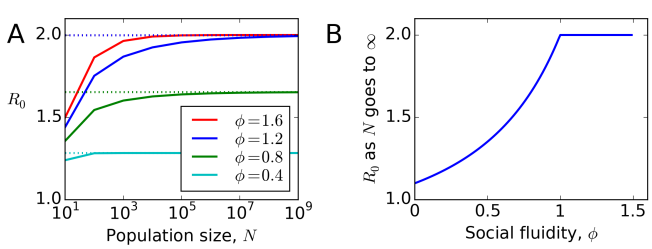

Figure 2: A: Theoretical results from (6); the relationship between R_0 and population size is shown to depend on the value of ϕ. Dashed lines show the limit for large N. B: In large populations R_0 increases with ϕ up to $\phi = 1$. Beyond this value, infections occur as frequently as they would in a homogeneously mixed population.

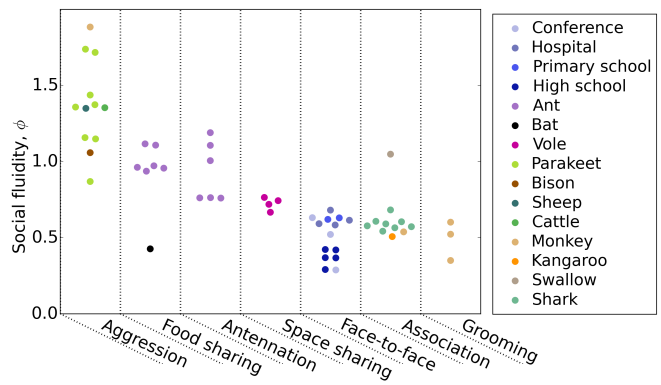

Figure 3: Each point represents a human or animal system for which social fluidity was estimated. Results are organized by interaction type: aggression includes fighting and displays of dominance, food sharing refers to mouth-to-mouth passing of food, antennation is when the antenna of one insect touches any part of another, space sharing interactions occur with spatial proximity during foraging, face-to-face refers to close proximity interactions that require individuals to be facing each other, association is defined as co-membership of the same social group.

Social behavior & disease spread in empirical systems

The previous section introduced a model of social behavior in which social fluidity, ϕ, quantifies heterogeneity in the way individuals distribute their social effort among other members of the population. To understand the results of the model in the context of real systems we estimate ϕ in 57 networks from 19 studies of human and animal social behavior (*Materials and Methods*) [21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39].

We focus our attention to those interactions which are capable of disease transmission (i.e. those that, at the least, require close spatial proximity). Each dataset provides the number of interactions that were observed between pairs of individuals (SI Text). We assume that the group size, N, is equal to the number of individuals observed in at least one interaction. Figure S1 shows the data and the distribution fitted using maximum likelihood estimation.

The model provides a good fit to every dataset; the worst fitting animal social network agrees with the model better than synthetic data generated 92% from model simulation and 8% from random noise (SI Text). For comparison, the model was also applied to 9 networks generated from social media data [40], 3 of which showed less agreement with the model than any of the non-electronic networks (SI Text). Since agreement between the model and the data is therefore not a certainty, it is remarkable that the social fluidity model is applicable to such a wide range of animal social networks.

Characterizing social fluidity. Figure 3 shows the estimated values of ϕ for all our study populations. We organize the measurements of social fluidity by interaction type. Aggressive interactions have the highest fluidity (which implies that most interactions are rarely repeated with the same individuals), while grooming interactions have the lowest (which implies frequent repeated interactions with the same individuals). Social fluidity also appears to be related to species: ant systems cluster around $\phi = 1$, voles around $\phi = 0.7$, humans around $\phi = 0.6$ (with the exception of lower values observed in the high school data and the last day of a conference). How-

bioRxiv preprint doi: https://doi.org/10.1101/170266; this version posted November 1, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

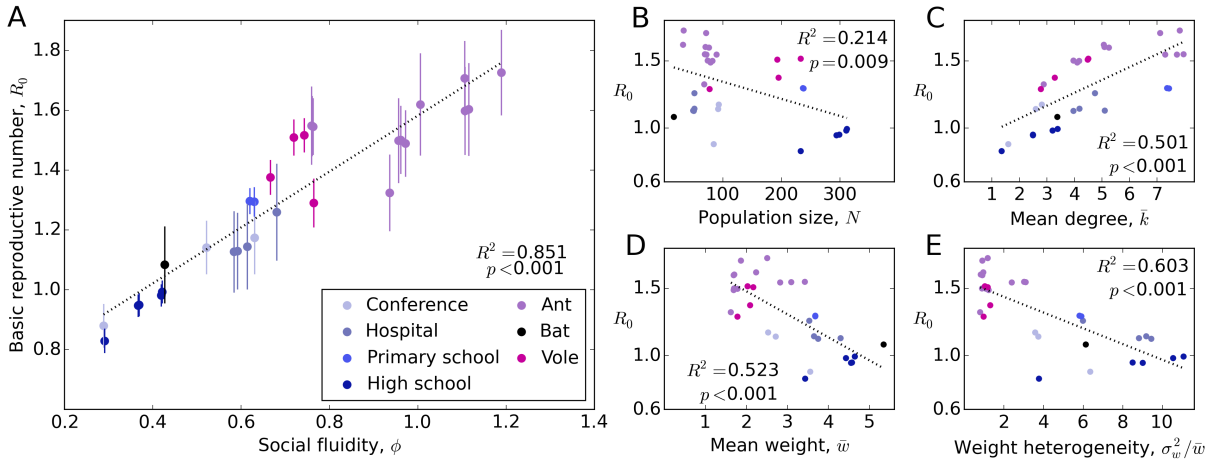

Figure 4: Each point represents a human or animal data-set for which the time of interactions appear in the data. For each individual, i, in a population, the mean number of secondary infections is obtained from simulation and the mean, R_0 ($\pm$ one standard error in part A), is plotted against metrics of social structure. Note that the parameters in the disease simulation have been calibrated with the rate of activity of each system to make these metrics comparable.

.

ever, we do not find that social fluidity is highly correlated with sociality type [41]. This is illustrated by the sizable difference in social fluidity between the sheep, bison, and cattle versus the kangaroos and bats all of which are categorized as fission-fusion species.

There is no significant correlation between the mean number of observations per individual, $\bar{s}$, and ϕ (Pearson $R^2 = 0.04$, $p = 0.14$), which implies that sample size does not affect the estimate (Figure S2**A**). Larger populations tend to have smaller social fluidity values (Pearson $R^2 = 0.205$, $p < 0.001$), however, this correlation is dependent on the presence of a few large populations in our data ($N > 200$) (Figure S2**B**). Larger values of ϕ correspond to higher mean degrees (Pearson $R^2 = 0.16$, $p = 0.002$) and less heterogeneity in the distribution of edge weights (measured as the variance divided by the mean $\sigma_w^2/\langle w \rangle$) (Pearson $R^2 = 0.339$, $p = 0.001$) (Figure S2**C** and **D**). Incidentally, weight heterogeneity and mean degree are uncorrelated in these data (Pearson $R^2 = 0.005$, $p = 0.568$) implying that ϕ combines two distinct features of social behavior.

Impact of social fluidity on disease transmission. To test the validity of the model and examine the utility of ϕ as an indicator of disease outcomes we simulated the spread of disease based on the interactions that occurred in the empirical data (*Materials and Methods* E.). We assume that all individuals are equally likely to introduce the infection to the group. In doing so we avoid biasing our choice of infection seed to those who are more socially active (our results are therefore not influenced by degree heterogeneity). Thus, the basic reproductive number, R_0, is defined simply as the mean of the individual reproductive numbers over the entire population.

For each individual, i, we compared the simulated number of secondary infections to the prediction for $r(s_i)$ given by (6). A small amount of error was observed (between 0.1 and 0.258 with one outlier at 0.344) and the prediction consistently overestimated R_0 in human systems (Table S3), possibly because of the bursty nature of human contact [42]. Despite this, our overall conclusions from the disease simulation are consistent with the predicted results.

Figure 4 shows that ϕ correlates with (simulated) R_0 better than other network metrics. Since R^* ($= 2$) represents the expectation of R_0 in a homogeneously mixed population of infinite size, the difference between R_0 and R^* can be attributed to the effects of social fluidity and population size. Because the association between ϕ and R_0 is strong (shown in Figure 4**A**), the relationship between each of the other metrics and R_0 is qualitatively the same as the their relationship with ϕ. Consequently, R_0 tends to be smaller in the largest populations (Figure 4**B**), suggesting that population size is not an important variable at such low values of ϕ. For $\phi > 1$, however, analytical results suggest that N plays a larger role in determining R_0 (Figure S3).

The relative weakness of the correlation between mean degree and R_0 (Figure 4**C**) suggests that the interaction dynamics taking place *on* each edge contribute significantly to the disease outcomes. For example, after transmission has occurred from one individual to another, repeats of the same interaction serve no advantage for the disease (most directly-transmitted infections are not dose-dependent). Since a large edge weight implies a high frequency of repeated interactions, networks with a higher mean weight tend to have lower values of R_0 (Figure 4**D**). Furthermore, heterogeneity in the distribution of weights concentrates a yet larger proportion of interactions onto a small number of edges, thus causing more repeat interactions and reducing R_0 (Figure 4**E**).

Discussion

We have defined a measure of fluidity in social behavior which quantifies how much mixing exists within the social relationships of a population. While social networks can be measured with a variety of metrics including size, connectivity, interaction heterogeneity and frequency, our methodology succeeds in reducing all such factors to a single quantity allowing comparisons across a range of human and animal social systems. Social fluidity correlates with both the density of social ties (mean degree) and the heterogeneity in the strength of those

bioRxiv preprint doi: https://doi.org/10.1101/170266; this version posted November 1, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

ties, yet they are not correlated with each other; this implies that social fluidity combines two distinct elements of social network analysis. We have also shown that this metric is a better predictor of R_0 than these other network metrics.

By measuring social fluidity across a range of human and animal systems we are able to rank social behaviors. In particular, we identify aggressive interactions as the most socially fluid; this indicates a possible learning effect whereby each aggressive encounter is followed by a period for which individuals avoid further aggression with each other [43]. At the opposite end of the scale, we find interactions that strengthen bonds (and thus require repeated interactions) such as grooming in monkeys [44] and food-sharing in bats [26]. The fact that food-sharing ants are far more fluid than bats, despite performing the same kind of interaction, reflects their eusocial nature and the absence of any need to maintain long-lasting bonds with their kin [45].

Most studies that aim to describe and quantify social structure are met with a number of challenges. The degree of an individual, for example, is known to scale with the length of the observation period [46]. By focusing not on the absolute value of degree, and instead focusing on how degree scales with the number of observations, our analysis controls for this bias. As with other network measures, however, social fluidity estimates lose reliability when only a sample of the group has been observed [47]. Another hurdle in previous work has been the assumption that observed interactions will persist into the future [48]. Our method assumes only that the distribution of relationships remains constant through time, an assumption that is consistent with growing evidence [19, 49].

In trying to elucidate the connection between social behaviors and the spread of disease the first question to ask is whether the type of interactive behavior that the species participates in enables transmission of the infection. If it does, then the next important variable is the frequency of these interactions. Beyond this, other proposed influences include the number of social ties [50], which we have shown to be less significant than social fluidity; modularity, which has previously been exposed as an unreliable predictor of disease outcomes [51]; and group size [52], which we have shown to only be relevant in highly fluid populations.

The relative significance of social fluidity becomes apparent when we compare the largest human face-to-face interaction system (`highschool_0` in Tables S2 and S3 $N = 312$, $R_0 = 0.99$, $\phi = 0.422$), to the smallest (`hospital_1`, $N = 49$, $R_0 = 1.129$, $\phi = 0.592$). Despite the former containing more than 6 times as many individuals as the latter, its lower social fluidity causes R_0 to be a smaller. Similarly, the mean degree of the hospital data ($\bar{k} = 5.1$) is more than double that of the first day of the conference (`conference_0`, $\bar{k} = 2.5$, $R_0 = 0.173$, $\phi = 0.631$), yet the relatively small decrease in ϕ is enough to reduce the value of R_0. Thus, the risks associated with group size and the density of social ties are easily offset by minor changes in the way individuals choose to distribute their social effort. If the survival of a group depends on its size or its connectivity then this social adaptation can be made with little impact on the disease risks.

Although social fluidity predicts the basic reproductive number R_0 better than alternative network-based metrics, this does not necessarily imply that it would be a good predictor of the final epidemic size in a population with higher-order social structure (e.g. clustering, degree heterogeneity) [53]. Even so, there are implications of this work that may improve the way epidemics are modeled. The growing evidence that human and animal interaction follows basic universal principals, such as constraints on the social capacity of an individual, ought to be encoded into infectious disease models that aim to achieve a coupling of social and disease dynamics [54, 55].

Unlike previous work that explores the disease consequences of population mixing [56, 20], our analysis allows us to compare its effect across a range of social systems. We see, for example, the relationship between mixing and how disease risks scale with population size. For social systems that have high values of social fluidity, R_0 is highly sensitive to changes in N, whereas this sensitivity is not present at low values of ϕ. Thus, both density dependent and frequency dependent disease dynamics are realized through the same model. Since many empirical studies support a transmission function that is somewhere between these two modeling paradigms [57, 58, 59], the modeling approaches applied in this study may be useful in informing transmission relationships in future disease studies.

Materials and Methods

Extended methods are provided in the SI Text

A. Computing the lower bound in (2). The value of ϵ is determined by the choice of ϕ and the value of N. Since interactions are pairwise, when i interacts exactly one other individual is involved. Thus, the sum of the $x_{j|i}$'s over all other members of the population, j, is equal to 1. We choose to express this as a constraint in expectation, with $(N-1)\langle x\rangle = 1$, where $\langle x\rangle$ denotes the mean of the distribution $\rho(x)$. This leads to

$$N = 1 + \frac{(1-\phi)(1-\epsilon^{\phi})}{\phi\epsilon^{\phi}(1-\epsilon^{1-\phi})}. \tag{7}$$

We can therefore find ϵ for any give N and ϕ by solving $(C+1)\epsilon^{\phi} - \epsilon - C = 0$, where $C = (1-\phi)/(N-1)\phi$, using the *fsolve* function from the *scipy.optimize* Python library (SI text).

B. Conditional degree distribution. The degree of i is determined by $N-1$ independent Bernoulli trials, each with success probability $\Psi(s_i)$. The probability that an individual will have degree k given that they have interacted s times, for any value of the global parameter ϕ is therefore given by

$$P(k|s,\phi) = \binom{N-1}{k}\Psi(s,\phi)^k\left(1-\Psi(s,\phi)\right)^{N-1-k}. \tag{8}$$

However, since this distribution gives non-zero probabilities for cases where $k > s$, which are invalid, we instead use

$$P(k|s,\phi) = \binom{s}{k}\left(\frac{(N-1)\Psi(s,\phi)}{s}\right)^k\left(1-\frac{(N-1)\Psi(s,\phi)}{s}\right)^{s-k} \tag{9}$$

when if $0 < s < N$. To evaluate the hypergeometric function in (3) we used the *hyp2f1* function from the *scipy.special* Python library.

C. Parameter estimation. For each individual in a study population we know the number of times they interacted, s_i, and the number of partners with whom they interacted, k_i, in vector notation $\mathbf{k} = \{k_1, k_2, ..., k_N\}$ and $\mathbf{s} = \{s_1, s_2, ..., s_N\}$. The marginal log-likelihood function is

$$\log\mathcal{L}(\phi|\mathbf{k},\mathbf{s}) = \sum_{i=1}^{N}\log[P(k_i|s_i,\phi)]. \tag{10}$$

We then compute the maximum likely estimate $\phi = \text{argmax}_{\phi}\log\mathcal{L}(\phi|\mathbf{k},\mathbf{s})$. Our estimation is based on the marginal distribution of the degree of each node and does not take into account the interdependencies of the network structure. As such, we do not report standard errors or confidence intervals for our point estimates. For discussion of the use of marginal distributions in likelihood estimation see [60].

The goodness-of-fit is calculated by comparing its likelihood to the likelihood of a null model in which the degree of each individual is a uniformly distributed random integer within the range of feasible values (SI Text).

bioRxiv preprint doi: https://doi.org/10.1101/170266; this version posted November 1, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

D. Disease model calibration. We choose γ for every system such that R_0 is equal to a fixed value, R^*, under the assumption of no effect from social fluidity and population size (i.e. in a large population with homogeneous mixing). Calibration is achieved when γ is chosen to be

$$\gamma = \frac{\beta\langle s\rangle}{\Delta_t R^*} \tag{11}$$

where Δ_t is the duration of the time-frame of the data and $\langle s\rangle$ is the mean of s_i over the whole population. The effect of this calibration is that the recovery rate, γ, is proportional to the mean activity rate.

E. Disease simulation. Disease simulations were only performed on data for which the time of every interaction is recorded. Since the transmission probability, β, and the rate of recovery, γ, are calibrated through (11), the choice of β does not affect R_0. We arbitrarily chose $\beta = 1/4$. We report mean absolute error $|e|$, which is the mean difference between the analytical and simulated values of $r(s_i)$, computed across 10^3 simulations for each node. An extensive description is provided in the supplement (SI Text).

Acknowledgments. This work was supported by NSF grant number 1414296. We are grateful for insightful feedback from Pratha Sah. We also thank all the researchers who have made their behavioral data openly accessible, making this study possible.

References

[1] Jens Krause and Graeme D Ruxton. *Living in groups*. Oxford University Press, 2002.

[2] R. A. Hinde. Interactions, relationships and social structure. *Man*, 11(1):1–17, 1976.

[3] Pratha Sah, José David Méndez, Janet Mann, and Shweta Bansal. Disease implications of animal social organization and network structure - a quantitative analysis. *bioRxiv*, 2017.

[4] Robin IM Dunbar and Susanne Shultz. Bondedness and sociality. *Behaviour*, 147(7):775–803, 2010.

[5] Cdric Sueur, Jean-Louis Deneubourg, Odile Petit, and Iain D. Couzin. Group size, grooming and fission in primates: A modeling approach based on group structure. *Journal of Theoretical Biology*, 273(1):156 – 166, 2011.

[6] P. Mac Carron, K. Kaski, and R. Dunbar. Calling dunbar's numbers. *Social Networks*, 47:151 – 155, 2016.

[7] Jari Saramäki, E Al Leicht, Eduardo López, Sam GB Roberts, Felix Reed-Tsochas, and Robin IM Dunbar. Persistence of social signatures in human communication. *Proceedings of the National Academy of Sciences*, 111(3):942–947, 2014.

[8] Bruno Gonalves, Nicola Perra, and Alessandro Vespignani. Modeling users' activity on twitter networks: Validation of dunbar's number. *PLOS ONE*, 6(8):1–5, 08 2011.

[9] Mario S. Di Bitetti. The distribution of grooming among female primates: Testing hypotheses with the shannon-wiener diversity index. *Behaviour*, 137(11):1517–1540, 2000.

[10] Petter Holme. Modern temporal network theory: a colloquium. *The European Physical Journal B*, 88(9):234, Sep 2015.

[11] Sonia Altizer, Charles L Nunn, Peter H Thrall, John L Gittleman, Janis Antonovics, Andrew A Cunningham, Andrew P Dobson, Vanessa Ezenwa, Kate E Jones, Amy B Pedersen, et al. Social organization and parasite risk in mammals: integrating theory and empirical studies. *Annual Review of Ecology, Evolution, and Systematics*, 34(1):517–547, 2003.

[12] Mart CM de Jong, O. Diekmann, and J.A.P. Heesterbeek. How does transmission of infection depend on population size? In *Epidemic models: their structure and relation to data*, volume 5, page 84. Cambridge University Press, 1995.

[13] Matthew J. Silk, Darren P. Croft, Richard J. Delahay, David J. Hodgson, Mike Boots, Nicola Weber, and Robbie A. McDonald. Using social network measures in wildlife disease ecology, epidemiology, and management. *BioScience*, 67(3):245–257, 2017.

[14] Jesse EH Patterson and Kathreen E Ruckstuhl. Parasite infection and host group size: a meta-analytical review. *Parasitology*, 140(7):803–813, 2013.

[15] J. Lehmann, A.H. Korstjens, and R.I.M. Dunbar. Group size, grooming and social cohesion in primates. *Animal Behaviour*, 74(6):1617 – 1629, 2007.

[16] Márton Karsai, Nicola Perra, and Alessandro Vespignani. Time varying networks and the weakness of strong ties. *Scientific Reports*, 4:4001, 2014.

[17] M. Abramowitz and I.A. Stegun. *Handbook of Mathematical Functions*. Dover, New York, 1975.

[18] Eugenio Valdano, Chiara Poletto, Armando Giovannini, Diana Palma, Lara Savini, and Vittoria Colizza. Predicting epidemic risk from past temporal contact data. *PLOS Computational Biology*, 11(3):1–19, 03 2015.

[19] Giovanna Miritello, Rubén Lara, Manuel Cebrian, and Esteban Moro. Limited communication capacity unveils strategies for human interaction. *Scientific reports*, 3, 2013.

[20] Timothy C. Reluga and Eunha Shim. Population viscosity suppresses disease emergence by preserving local herd immunity. *Proceedings of the Royal Society of London B: Biological Sciences*, 281(1796), 2014.

[21] Lorenzo Isella, Juliette Stehlé, Alain Barrat, Ciro Cattuto, Jean-François Pinton, and Wouter Van den Broeck. What's in a crowd? analysis of face-to-face behavioral networks. *Journal of theoretical biology*, 271(1):166–180, 2011.

[22] Juliette Stehl, Nicolas Voirin, Alain Barrat, Ciro Cattuto, Lorenzo Isella, Jean-Franois Pinton, Marco Quaggiotto, Wouter Van den Broeck, Corinne Rgis, Bruno Lina, and Philippe Vanhems. High-resolution measurements of face-to-face contact patterns in a primary school. *PLOS ONE*, 6(8):e23176, 08 2011.

[23] Rossana Mastrandrea, Julie Fournet, and Alain Barrat. Contact patterns in a high school: A comparison between data collected using wearable sensors, contact diaries and friendship surveys. *PLOS ONE*, 10(9):1–26, 09 2015.

[24] Philippe Vanhems, Alain Barrat, Ciro Cattuto, Jean-Fran?ois Pinton, Nagham Khanafer, Corinne R?gis, Byeul-a Kim, Brigitte Comte, and Nicolas Voirin. Estimating potential infection transmission routes in hospital wards using wearable proximity sensors. *PLoS ONE*, 8(9):e73970, 09 2013.

[25] Benjamin Blonder and Anna Dornhaus. Time-ordered networks reveal limitations to information flow in ant colonies. *PLOS ONE*, 6(5):1–8, 05 2011.

[26] Gerald G. Carter and Gerald S. Wilkinson. Food sharing in vampire bats: reciprocal help predicts donations more than relatedness or harassment. *Proceedings of the Royal Society of London B: Biological Sciences*, 280(1753), 2013.

[27] Stephen Davis, Babak Abbasi, Shrupa Shah, Sandra Telfer, and Mike Begon. Spatial analyses of wildlife contact networks. *Journal of The Royal Society Interface*, 12(102), 2014.

[28] T.R. Grant. Dominance and association among members of a captive and a free-ranging group of grey kangaroos (macropus giganteus). *Animal Behaviour*, 21(3):449 – 456, 1973.

[29] Iris I Levin, David M Zonana, Bailey K Fosdick, Se Jin Song, Rob Knight, and Rebecca J Safran. Stress response, gut microbial diversity and sexual signals correlate with social interactions. *Biology Letters*, 12(6):20160352, 2016.

[30] Lee Douglas Sailer and Steven JC Gaulin. Proximity, sociality, and observation: the definition of social groups. *American Anthropologist*, 86(1):91–98, 1984.

[31] Johann Mourier, Culum Brown, and Serge Planes. Learning and robustness to catch-and-release fishing in a shark social network. *Biology Letters*, 13(3):20160824, 2017.

[32] Jorg JM Massen and Elisabeth HM Sterck. Stability and durability of intra-and intersex social bonds of captive rhesus macaques (macaca mulatta). *International Journal of Primatology*, 34(4):770–791, 2013.

[33] DS Sade. Sociometrics of macaca mulatta i. linkages and cliques in grooming matrices. *Folia primatologica*, 18(3-4):196–223, 1972.

[34] ML Butovskaya, AG Kozintsev, and BA Kozintsev. The structure of affiliative relations in a primate community: allogrooming in stumptailed macaques (macaca arctoides). *Human evolution*, 9(1):11–23, 1994.

[35] Yukio Takahata. Diachronic changes in the dominance relations of adult female japanese monkeys of the arashiyama b group. *The monkeys of Arashiyama. State University of New York Press, Albany*, pages 123–139, 1991.

[36] Christine C Hass. Social status in female bighorn sheep (ovis canadensis): expression, development and reproductive correlates. *Journal of Zoology*, 225(3):509–523, 1991.

[37] Dale F Lott. Dominance relations and breeding rate in mature male american bison. *Ethology*, 49(4):418–432, 1979.

[38] Martin W. Schein and Milton H. Fohrman. Social dominance relationships in a herd of dairy cattle. *The British Journal of Animal Behaviour*, 3(2):45 – 55, 1955.

[39] Elizabeth A Hobson and Simon DeDeo. Social feedback and the emergence of rank in animal society. *PLoS Comput Biol*, 11(9):e1004411, 2015.

[40] Nathaniel Charlton, Colin Singleton, and Danica Vukadinović Greetham. In the mood: the dynamics of collective sentiments on twitter. *Royal Society Open Science*, 3(6), 2016.

[41] Pratha Sah, Janet Mann, and Shweta Bansal. Disease implications of animal social network structure: a synthesis across social systems. *Biorxiv*, page 10.1101/106633, 2017.

[42] M Karsai, M Kivelä, RK Pan, K Kaski, J Kertész, A-L Barabási, and J Saramäki. Small but slow world: How network topology and burstiness slow down spreading. *Physical Review E*, 83(2):025102, 2011.

[43] Geoffrey A Parker. Assessment strategy and the evolution of fighting behaviour. *Journal of theoretical Biology*, 47(1):223–243, 1974.

[44] Robert M Seyfarth and Dorothy L Cheney. Grooming, alliances and reciprocal altruism in vervet monkeys. *Nature*, 308(5959):541–543, 1984.

[45] Bert Hölldobler and Edward O Wilson. *The superorganism: the beauty, elegance, and strangeness of insect societies*. WW Norton & Company, 2009.

[46] Nicola Perra, Bruno Gonçalves, Romualdo Pastor-Satorras, and Alessandro Vespignani. Activity driven modeling of time varying networks. *Scientific reports*, 2, 2012.

[47] Matthew J. Silk, Andrew L. Jackson, Darren P. Croft, Kendrew Colhoun, and Stuart Bearhop. The consequences of unidentifiable individuals for the analysis of an animal social network. *Animal Behaviour*, 104(Supplement C):1 – 11, 2015.

[48] Charles Perreault. A note on reconstructing animal social networks from independent small-group observations. *Animal Behaviour*, 80(3):551–562, 2010.

[49] Simone Centellegher, Eduardo Lpez, Jari Saramki, and Bruno Lepri. Personality traits and ego-network dynamics. *PLOS ONE*, 12(3):1–17, 03 2017.

[50] Meggan E. Craft. Infectious disease transmission and contact networks in wildlife and livestock. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, 370(1669), 2015.

[51] Pratha Sah, Stephan T Leu, Paul C Cross, Peter J Hudson, and Shweta Bansal. Unraveling the disease consequences and mechanisms of modular structure in animal social networks. *Proceedings of the National Academy of Sciences*, page 201613616, 2017.

[52] Joanna L Rifkin, Charles L Nunn, and László Z Garamszegi. Do animals living in larger groups experience greater parasitism? a meta-analysis. *The American Naturalist*, 180(1):70–82, 2012.

[53] Joel C Miller. Spread of infectious disease through clustered populations. *Journal of The Royal Society Interface*, 2009.

[54] Livio Bioglio, Mathieu Génois, Christian L Vestergaard, Chiara Poletto, Alain Barrat, and Vittoria Colizza. Recalibrating disease parameters for increasing realism in modeling epidemics in closed settings. *BMC infectious diseases*, 16(1):676, 2016.

[55] Naoki Masuda and Petter Holme. Introduction to temporal network epidemiology. In *Temporal Network Epidemiology*, pages 1–16. Springer, 2017.

[56] Erik Volz and Lauren Ancel Meyers. Susceptible–infected–recovered epidemics in dynamic contact networks. *Proceedings of the Royal Society of London B: Biological Sciences*, 274(1628):2925–2934, 2007.

[57] Matthew J. Smith, Sandra Telfer, Eva R. Kallio, Sarah Burthe, Alex R. Cook, Xavier Lambin, and Michael Begon. Hostpathogen time series data in wildlife support a transmission function between density and frequency dependence. *Proceedings of the National Academy of Sciences*, 106(19):7905–7909, 2009.

[58] P. C. Cross, T. G. Creech, M. R. Ebinger, K. Manlove, K. Irvine, J. Henningsen, J. Rogerson, B. M. Scurlock, and S. Creel. Female elk contacts are neither frequency nor density dependent. *Ecology*, 94(9):2076–2086, 2013.

[59] Benny Borremans, Jonas Reijniers, Nelika K Hughes, Stephanie S Godfrey, Sophie Gryseels, Rhodes H Makundi, and Herwig Leirs. Nonlinear scaling of foraging contacts with rodent population density. *Oikos*, 2016.

[60] Cristiano Varin, Nancy Reid, and David Firth. An overview of composite likelihood methods. *Statistica Sinica*, pages 5–42, 2011.