bioRxiv preprint doi: https://doi.org/10.1101/170266; this version posted February 27, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity.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[1,2], Vittoria Colizza[3], Ephraim M. Hanks[4], Andreas P. Modlmeier[5], David P. Hughes[5], and Shweta Bansal[1]

[1]Department of Biology, Georgetown University, Washington DC, United States
[2]Roslin Institute, University of Edinburgh, Easter Bush, Midlothian, United Kingdom
[3]INSERM, Sorbonne Université, Institut Pierre Louis d'Épidémiologie et de Santé Publique (IPLESP UMRS 1136), F75012, Paris, France.
[4]Department of Statistics, Eberly College of Science, Penn State University, State College, United States
[5]Department of Entomology, College of Agricultural Sciences, Penn State University, State College, United States

Humans and other group-living animals tend to distribute their social effort disproportionately. Individuals predominantly interact with their closest companions while maintaining weaker social bonds with less familiar group members. By incorporating this behaviour into a mathematical model we find that a single parameter, which we refer to as social fluidity, controls the rate of social mixing within the group. We compare the social fluidity of 13 species by applying the model to empirical human and animal social interaction data. 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. For highly fluid social behaviour disease transmission is density-dependent. For species that form more stable social bonds, the model describes frequency-dependent transmission that is sensitive to changes in social fluidity.

SOCIAL behavior is fundamental to the survival of many species. It allows the formation of social groups providing fitness advantages from greater access to resources and better protection from predators [1]. Structure within these groups can be found in the way individuals communicate across space, cooperate in sexual or parental behavior, or clash in territorial or mating conflicts [2]. While animal societies are usually studied independently of each other, some questions about the nature of social living can only be answered by comparing behavior across a range of species [3, 4].

When social interaction requires shared physical space it can also be a conduit for the transmission of infectious disease [5]. For epidemic modellers it is vital to know what level of contact is necessary for host-to-host transmission as this determines how the density and structure of the population affect the rate at which the disease will spread [6, 7]. Typically, if the disease spreads through the environment then the transmission rate is assumed to scale proportionally to the local population density (density-dependence), whereas if transmission requires close proximity encounters that only occur between bonded individuals then we expect social connectivity to determine the outcome (frequency-dependence) [8].

In reality, however, animal-disease systems are not so easy to categorize [9]. For example, as social groups grow in size, new bonds must be created to maintain cohesiveness [10]. To manage their time and the increased cognitive effort required to maintain these bonds, individuals tend to interact mostly with their closest companions while weaker ties are maintained through infrequent contact [11–13]. This variability in the way social effort is distributed has been shown to affect contagion processes [14], and it leads us to the question motivating this study: can quantifying how group-living individuals choose to invest their social effort allow us to model the effects of population density on epidemic spread?

There is growing evidence for the disproportionate distribution of social effort in human communication [15–18]. Attempts to quantify this aspect of sociality in animal systems, however, are challenged by the fact that data on some individuals may be far richer than on others. These biases can be introduced in the data collection process, or result from behavioural differences across the sampled population [19]. Furthermore, while heterogeneous interaction frequencies and temporal dynamics such as circadian rhythms and bursty activity patterns have become common in social network models [20], little has been done to incorporate the way the individual chooses to distribute their social effort.

Here, we introduce a mathematical model founded on the concept of *social fluidity* which we define as variability in the amount of social effort the individual invests in each member of their social group. Using empirical data from previous studies, we estimate the social fluidity of 57 human and animal social systems. We use it in analytical and computational models of disease spread and show that the basic reproductive number defined on social fluidity is a better predictor of disease outcome compared to other social behavioral indicators. In addition, social fluidity emerges as a coherent mathematical framework providing the smooth connection between density-dependent and frequency-dependent disease systems, which have historically been studied in isolation.

Characterizing social behaviour

Our objective is to measure social behaviour in a range of human and animal populations. We start by introducing a model that captures a hidden element of social dynamics: how individual group members distribute their social effort. We mathematically describe the relationships between social variables that are routinely found in studies of

bioRxiv preprint doi: https://doi.org/10.1101/170266; this version posted February 27, 2020. 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.

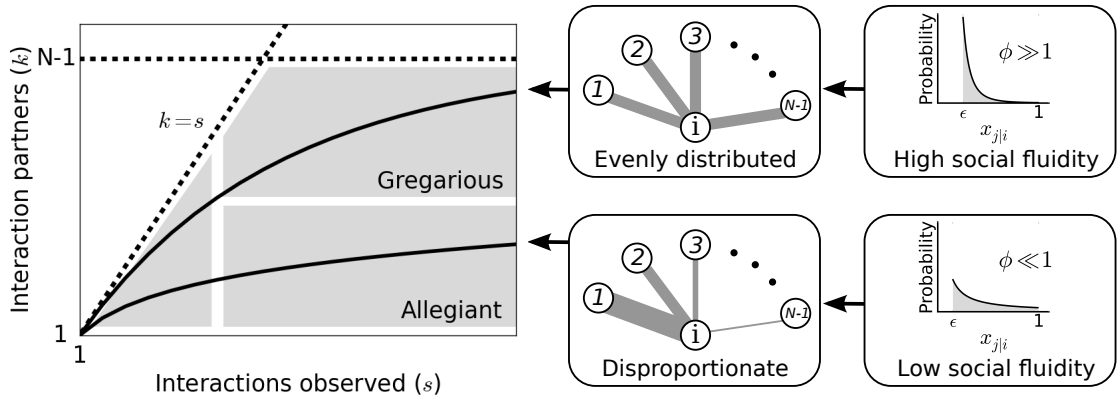

Figure 1: Left: Each individual can be represented as a single point on this plot. Dashed lines mark the boundary of the region where data points can feasibly be found. The mean degree is plotted for two values of ϕ representing two possible types of social behavior; as the number of observed interactions 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 group. Right: Probability distributions that correspond to the different levels of evenness in the contact propensities, both distributions are expressed by Eq.(2).

animal behavior, the number of social ties and the number of interactions observed, and apply the model to empirical data to reveal behavioural differences between several species.

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: each individual, i, is a *node*; an *edge* exists between two nodes i and j if at least one interaction was observed between them; the *edge weight*, $w_{i,j}$, denotes the number of times this interaction was observed. The total number of interactions of i is denoted *strength*, $s_i = \sum_j w_{i,j}$, and the number of nodes with whom i is observed interacting is its *degree*, k_i [21].

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 according to a probability distribution, $\rho(x)$.[1] Thus, if a node interacts s times, the marginal probability that an edge exists between that node and any other given node in the network 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 (Fig. 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 variability in the values of x, and ϵ simply truncates the distribution to avoid divergence. Combining (1) and (2) we find

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

where the notation ${}_2F_1$ refers to the Gauss hypergeometric function [22]. It follows from $\sum_j x_{j|i} = 1$ that

$$N = 1 + \frac{(1-\phi)(1-\epsilon^\phi)}{\phi\epsilon^\phi(1-\epsilon^{1-\phi})}, \tag{4}$$

which can be solved numerically to find ϵ for given values of N and ϕ. The expectation of the degree is $\kappa(s,\phi,N) = (N-1)\Psi(s,\phi,\epsilon)$.

Fig. 1 illustrates how the value of ϕ can produce different types of social behavior. As ϕ is the main determinant of social behaviour in our model, we use the term *social fluidity* to refer to this quantity. Low social fluidity ($\phi \ll 1$) produces 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 phenomenon could be similarly described as "social strategy" or "loyalty" [23, 24], here we use a different measure as it is consistent with previously studied social drivers of epidemic spread [25] establishing a direct connection with disease risk at the population scale.

Estimating social fluidity in empirical networks: To understand the results of the model in the context of real systems

[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 $P(X)$.

bioRxiv preprint doi: https://doi.org/10.1101/170266; this version posted February 27, 2020. 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.

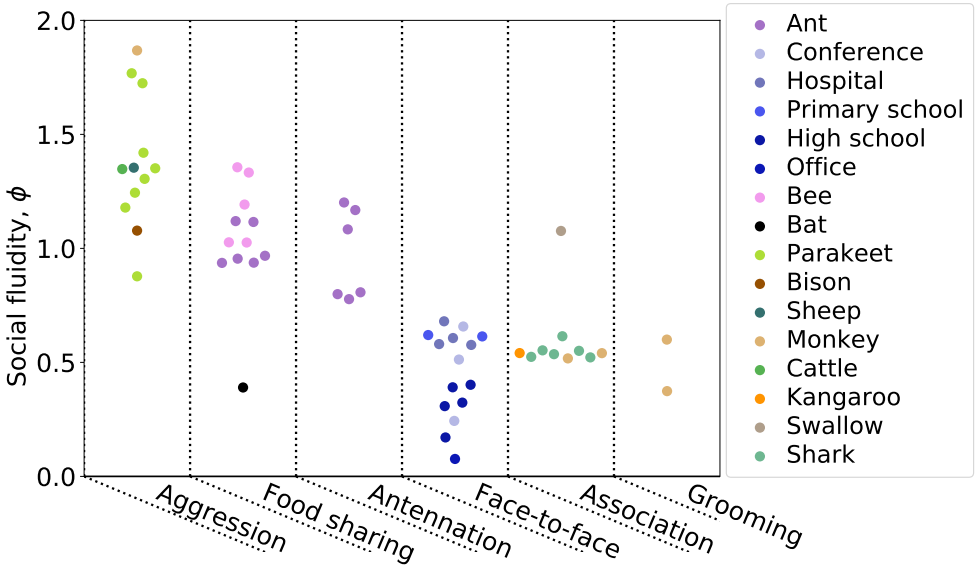

Figure 2: Each point represents a human or animal system for which social fluidity was estimated. Results are organized by interaction type: aggression includes fighting and displays of dominance, food sharing refers to mouth-to-mouth passing of food, antennation is when the antenna of one insect touches 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.

we estimate ϕ in 57 networks from 20 studies of human and animal social behavior (further details in the supplement) [26–46], focusing 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. We assume that the system is closed, and that the total network size (N) is equal to the number of individuals observed in at least one interaction. To estimate social fluidity we find the value of ϕ that minimizes $\sum_i [k_i - \kappa(s_i, \phi, N)]^2$ (the total squared squared error between the observed degrees and their expectation given by the model). Being estimated from the relationship between strength and degree, and not their absolute values, social fluidity is a good candidate for comparing social behavior across different systems as it is independent of the distributions of s_i or k_i, and of the timescale of interactions.

Fig. 2 shows the estimated values of ϕ for all networks in our study. We organize the measurements of social fluidity by interaction type. Aggressive interactions have the highest fluidity (which implies that most interactions are rarely repeated between the same individuals), while grooming and other forms of social bonding have the lowest (which implies frequent repeated interactions between the same individuals). Social fluidity also appears to be related to species: ant systems cluster around $\phi = 1$, monkeys around $\phi = 0.5$, humans take a range of values that depend on the social environment. Sociality type does not appear to affect ϕ; sheep, bison, and cattle have different social fluidity compared to kangaroos and bats, though they are all categorized as fission-fusion species [3].

There is no significant correlation between the mean number of interactions per individual ($\bar{s}$) and social fluidity (Pearson $r^2 = 0.02$, $p = 0.26$), which implies that sampling bias does not affect the estimation of social fluidity. Similarly, network size does not correlate with ϕ (Pearson $r^2 = 0.02$, $p = 0.33$). Larger values of ϕ correspond to higher mean degrees (Pearson $r^2 = 0.27$, $p < 0.001$) and lower variability in the distribution of edge weights (measured as the index of dispersion of $w_{i,j}$; Pearson $r^2 = 0.26$, $p < 0.001$). Weight variability and mean degree are uncorrelated in these data (Pearson $r^2 = 0.01$, $p = 0.59$) implying that ϕ combines these two entirely distinct features of social behavior.

Finally, the modularity of the network (computed by the Louvain method on the unweighted network [47]) is negatively correlated with ϕ ($r^2 = 0.57$, $p < 0.001$). This is expected as individuals tend to be loyal to those within the same module while maintaining weaker connections with the remaining network.

Characterizing disease spread with social fluidity

Our objective is to characterize how social behavior influences the susceptibility of the group to infectious disease in a range of human and animal social systems. We start by introducing a analytical transmission model that incorporates social fluidity. Using this model, we mathematically characterize the impact of social fluidity on density dependence, and apply the model to empirical networks to predict disease spread.

Disease transmission model: We consider the transmission of an infectious disease on the social behavior model introduced in the previous section. An infectious node i interacting with a susceptible node j will transmit the infection with probability β. The node will recover from infection with rate γ, assuming an exponential distribution of the length of the infectious period. The probability that the

bioRxiv preprint doi: https://doi.org/10.1101/170266; this version posted February 27, 2020. 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.

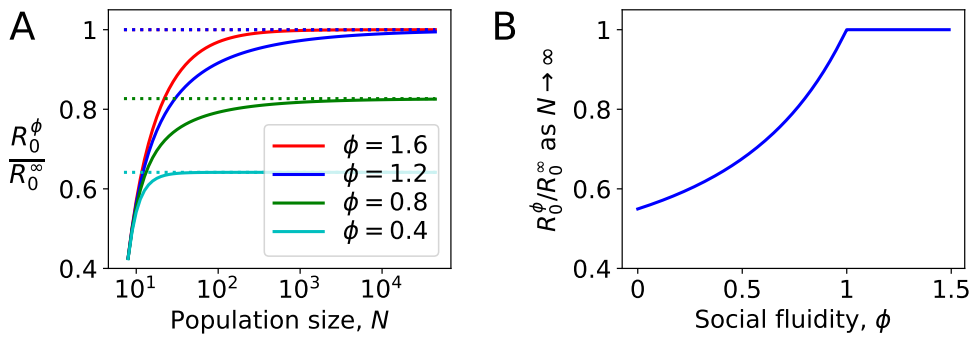

Figure 3: Density dependence in populations where every node has the same strength. A: For different values of social fluidity, ϕ, we show R_0^ϕ (from Eq.(6)) as a function of N (from Eq.(4)) through their parametric relation with ϵ. Dashed lines show the limit for large N. B: In large populations R_0^ϕ increases with ϕ up to $\phi = 1$. Beyond this value, infections occur as frequently as they would if every new interaction occurs between a pair of individuals who have not previously interacted with each other.

infection is transmitted from i to any given j is

$$T_{i\to j}(\beta, \gamma, s_i, \tau, x_{j|i}) = 1 - \exp(-s_i x_{j|i}\beta/\gamma\tau), \qquad (5)$$

assuming that the interactions s_i of i are distributed randomly across an observation period of duration τ.

By integrating Eq. (5) over all possible values $x_{j|i}$ and and infectious period durations and multiplying by the number of susceptible individuals $(N-1)$ we obtain the expected number of infections caused by individual i,

$$r(s_i) = \frac{1-\phi}{\phi(\epsilon^\phi - \epsilon)}\left[1 - \epsilon^\phi + \epsilon^\phi {}_2F_1(-\phi, 1, 1-\phi; -\beta s_i/\gamma\tau) - {}_2F_1(-\phi, 1, 1-\phi; -\epsilon\beta s_i/\gamma\tau)\right]. \qquad (6)$$

The basic reproductive number (usually denoted R_0) is defined as the mean number of secondary infections caused by a typical infectious individual in an otherwise susceptible population [48]. We will use the notation R_0^ϕ to signify the *social fluidity reproductive number*, that is the analogue of R_0 derived from our social behaviour model.

We assess the relation of the reproductive number with the population density by focusing on a special case where every node has the same strength, i.e $s_i = s$ for all i, so that $R_0^\phi = r(s)$. Furthermore, we choose $\beta = \gamma\tau R_0^\infty/s$ where R_0^∞ is R_0^ϕ as $\phi \to \infty$, i.e, a constant that represents what the basic reproductive number would be if every new interaction occurred between a pair of individuals who have not previously interacted with each other.

Fig. 3 shows the effect of social fluidity on the density dependence of the disease. At small population sizes, R_0^ϕ increases with N and converges as N goes to ∞ (Fig. 3**A**). The rate of this convergence increases with ϕ, and the limit it converges to is higher, meaning that ϕ determines the extent to which density affects the spread of disease. As $N \to \infty$, we find that $R_0^\phi \to R_0^\infty$ for $\phi > 1$. When $\phi < 1$, $R_0^\phi \to [(1-\phi)/\phi][{}_2F_1(-\phi, 1, 1-\phi; -R_0^\infty) - 1]$. At these values of ϕ the disease is constrained by individuals choosing to repeat interactions despite having the choice of infinitely many potential interaction partners (Fig 3**B**).

Estimating infection spread in empirical networks with heterogeneous connectivity: To apply this analogue of a reproductive number to an animal-disease system, we need to account for heterogeneous levels of social connectivity in the given population and thus the tendency for infected individuals to be those with a greater number of social partners [49]. For the basic reproductive number, this is often done using the mean *excess degree*, i.e. the degree of an individual selected with probability proportional to their degree [50]. Following a similar reasoning, we define R_0^{Est}, which incorporates the effect of social fluidity, as the expected number of infections $(r(s_i))$ caused by an individual that has been selected with probability proportional to their degree (k_i):

$$R_0^{\text{Est}}(\{s_i\}, \{k_i\}, \tau, \beta, \gamma) = \frac{\sum_i k_i r(s_i)}{\sum_i k_i}. \qquad (7)$$

Given the degree and strength of each individual in a network, the duration over which those interactions occurrred, and the transmission and recovery rates of the disease, we are able to estimate ϕ, compute Eq.(6) for each individual, and finally use Eq.(7) to derive a statistic that provides a measure of the risk of the host population to disease outbreak.

Numerical validation using empirical networks: We simulated the spread of disease through the interactions that occurred in the empirical data (materials and methods). We compute $R_0^{\text{Sim}}(g)$, defined as the ratio of the number of individuals infected at the $(g+1)$-th generation to the number infected at the g-th generation over 10^3 simulated outbreaks, for $g = 0, 1, 2$ ($g = 0$ refers to the initial seed of the outbreak).

Table 1 shows the Pearson correlation coefficient between $R_0^{\text{Sim}}(g)$ and its corresponding value R_0^{Est} obtained Eq.(7). For comparison, the correlation is shown for other indicators and network statistics. The results correspond to one set of simulation conditions, and are robust across a wide range of parameter combinations (see supplementary tables). Note that a different value of β was chosen for each network to control for the varying interaction rates between networks while keeping the upper bound (R_0^∞) constant (materials and methods). Thus, the mean strength does not have a significant effect on $R_0^{\text{Sim}}(g)$.

bioRxiv preprint doi: https://doi.org/10.1101/170266; this version posted February 27, 2020. 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.

Table 1: The Pearson correlation coefficient between quantities calculated on the network and the simulated disease outcomes (with $R_0^\infty = 3$). Results that are significant with $p < 0.01$ are labelled with *.

	Corr. with $R_0^{\mathrm{Sim}}(g=1)$)
R_0^{Est}	0.91*
Social fluidity	0.73*
Excess degree	0.64*
Mean degree	0.53*
Network size	0.47*
Mean strength	−0.07
Mean clustering	−0.15
Mean edge weight	−0.45*
Edge weight variability	−0.48*
Modularity	−0.60*

These correlations support a known result regarding repeat contacts in network models of disease spread: that indicators of disease risk that are derived solely from the degree distribution are unreliable and the role of edge weights should not be neglected [51,52]. After transmission has occurred from one individual to another, repeating the same interaction serves no advantage for disease (most directly-transmitted microparasites 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 basic reproductive numbers. Furthermore, variability in the distribution of weights concentrates a yet larger proportion of interactions onto a small number of edges, further increasing the number of repeat interactions and reducing the reproductive number.

Correlation between modularity and $R_0^{\mathrm{Sim}}(g)$ is partly due to the strong correlation between modular networks and those with high social fluidity. Consistent with other evidence [53], this suggests that transmission events occur mostly within the module of the seed node, with weaker social ties facilitating transmission to other modules. The effect of clustering (a measure of the number of connected triples in network [54]) correlates with smaller $R_0^{\mathrm{Sim}}(2)$, consistent with other theoretical work [51, 55].

Finally, we find the model estimate of the social fluidity reproductive number R_0^{Est} to be, on average, within 10% of the simulated value, $R_0^{\mathrm{Sim}}(g)$ at $g = 1$. At $g = 2$ the amount of error is larger (to up to 29% for some parameter choices). Prediction accuracy at this generation is negatively correlated with the mean clustering coefficient. This is not surprising as R_0^{Est} does not account for the accelerated depletion of susceptible neighbours that is known to occur in clustered networks [51,55]. No other properties of the network affect the accuracy of R_0^{Est} consistently across all parameter combinations (see supplementary tables).

Discussion

We proposed 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, contact heterogeneity and frequency, our methodology reduces 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 variability in the weight of those ties, though these quantities do not correlate with each other. Social fluidity is thus able to combine these two aspects seamlessly in one quantity.

By measuring social fluidity across a range of human and animal systems we are able to rank social behaviors. We identify aggressive interactions as the most socially fluid; this indicates a possible learning effect whereby each aggressive encounter is followed by a period during which individuals avoid further aggression with each other [56]. At the opposite end of the scale, we find interactions that strengthen bonds (and thus require repeated interactions) such as grooming in monkeys [57] and food-sharing in bats [33]. 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 consistently reinforce bonds with their kin [58].

Most studies that aim to describe and quantify social structure are met with a number of challenges, including ours. First, the degree of an individual, for example, is known to scale with the length of the observation period [59]. By focusing not on the absolute value of degree, but instead on how degree scales with the number of observations, our analysis controls for this bias. Second, observed interactions have been assumed to persist over time [60]. In our model, only the distribution of edge weights remains constant through time, an assumption consistent with growing evidence [24, 61]. Third, duration of contacts is known to be important for disease spread [52]. We did not include explicitly the duration of each contact in our model, since this information was only available in a fraction of the datasets [62]. There is therefore potential to improve the applicability of this model as more high resolution data becomes openly available.

Our estimate of reproductive number derived from social fluidity provides a better predictor for the epidemic risk of a host population, going beyond predictors based on density or degree only. To illustrate this point, the social network of individuals at a conference ($R_0^{\mathrm{Est}} = 1.60$; `conference_0`, supplementary document) is predicted to be at higher risk compared to the social network at a school ($R_0^{\mathrm{Est}} = 1.39$; `highschool_0`), despite having a smaller size and lower connectivity ($N = 93$ vs. $N = 312$, and $\bar{k} = 5.63$ vs. $\bar{k} = 6.78$, respectively). The discrepancy in the risk prediction comes from the lower frequency of repeated contacts between individuals in the conference, compared to the school. Interactions between infectious individuals and those they have previously infected are redundant in terms of transmission. This dynamic is nicely captured by the social fluidity, with $\phi = 0.66$ for the conference and $\phi = 0.40$ for the high school.

Unlike previous work that explores the disease consequences of population mixing [25, 63], our analysis allows us to investigate this relation across a range of social sys-

bioRxiv preprint doi: https://doi.org/10.1101/170266; this version posted February 27, 2020. 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.

tems. We see, for example, how the relationship between mixing and disease risk scales with population density. For social systems that have high values of social fluidity, R_0^ϕ is highly sensitive to changes in N, whereas this sensitivity is not present at low values of ϕ. This corroborates past work on the scaling of transmission being associated to heterogeneity in contact [64,65]. Going beyond previous work, our model captures in a coherent theoretical framework both density-dependence and frequency-dependence, and social fluidity is the measure to tune from one to the other in a continuous way. Since many empirical studies support a transmission function that is somewhere between these two modeling paradigms [7,66–68], the modeling approaches applied in this paper can be carried forward to inform transmission relationships in future disease studies.

Materials & Methods

A. Python libraries Mean clustering coefficients were computed using the *networkx* Python library. To evaluate the hypergeometric function in (3) we used the *hyp2f1* function from the *scipy.special* Python library. Numerical solutions to Eq.(4) using the *fsolve* function from the *scipy.optimize* Python library. All scripts, data, and documentation used in this study are available through `https://github.com/EwanColman/Social-Fluidity`.

B. Data handling Only freely available downloadable sources of data have been used for this study. Details of the experimentation and data collection can be found through their respective publications. Here we note some additional processes we have applied for our study.

Each human contact dataset lists the identities of the people in contact, as well as the 20-second interval of detection [26–29, 32]. To exclude contacts detected while participants momentarily walked past one another, only contacts detected in at least two consecutive intervals are considered interactions. Data were then separated into 24 hour subsets.

Bee trophallaxis provided experimental data for 5 unrelated colonies under continuous observation. We use the first hour of recorded data for each colony [46]. The ant trophallaxis study provided 6 networks: 3 unrelated colonies continuously observed under 2 different experimental conditions [30]. Ant antennation study provided 6 networks: 3 colonies, each observed in 2 sessions separated by a two week period. The bat study collected individual data at different times and under different experimental conditions [33]. For bats that were studied on more than one occasion we use only the first day they were observed.

Some data sets provided data for group membership collected through intermittent, rather than continuous, observation [34–38]. We construct networks from these data by recording an interaction when two individuals were seen to be in the same group during one round of observation. The shark data was divided into 6 datasets, each one constructed from 10 consecutive observations, and spread out through the full time period over which the data was collected.

For the grooming data [39, 40], if one animal was grooming another during one round of observations then this would be recorded as a directed interaction. Similarly for aggressive interactions [41–45, 56]. When an animal was determined to be the winner of a dominance encounter then this would be recorded as a directed interaction between the winner and the loser. We consider interaction in either direction to be a contact in the network.

We considered including two rodent datasets in which interaction is defined as being observed within the same territorial space [66,68]. We did not find this suitable for our analysis since the network we obtain, and the consequent results are sensitive to setting of arbitrary threshold values regarding what should, or should not, be considered sufficient contact for an interaction.

For data that did not contain the time of each interaction, contact time series were generated synthetically. For those datasets, the interactions between each pair were given synthetic timestamps in three different ways, Poisson: the time of each interaction is chosen uniformly at random from $\{0, 1, ..., 10^4\}$ seconds, Circadian: chosen uniformly at random from $\{0, 1, ..., 3333, 6666, ..., 10^4\}$, and Bursty: interaction times occur with power-law distributed inter-event times adjusted to give an expected total duration of 10^4 seconds.

C. Disease simulation Simulations of disease spread were executed using the contacts provided by the datasets. The the bat network was omitted from this part since these data were collected over a series of independent experiments carried out at different times and under different experimental treatments.

In one run of the simulation, one seed node is randomly chosen from the network and, at a randomly selected point in time during the duration of the data, transitions to the infectious state. The duration for which they remain infectious is a random variable drawn from an exponential distribution with mean $1/\gamma$. During this time any contact they have with other individuals who have not previously been infected will cause an infection with probability β.

The simulation runs until all individuals who were infected at the second generation of the disease, i.e. those infected by those infected by the seed, have recovered. The datasets are 'looped' to ensure that the timeframe of the data collection does not influence the outcome. In other words, immediately after the latest interaction, the interactions are repeated exactly as they were originally. This continues to happen until the termination criteria is met.

We set the parameters to normalise for the variation in contacts rates between networks. To achieve this we consider a hypothetical counterpart to each network in which the strength of every node is the same, but each interaction occurs between a pair of individuals who have not previously interacted. This is equivalent to $\phi \to \infty$. Under these conditions $x_{j|i} = 1/(N-1)$ for all pairs i, j. It follows that Eq. (5) becomes $T_{i\to j} \approx s_i\beta/\gamma\tau(N-1)$, then $r(s_i) \approx s_i\beta/\gamma\tau$, and, since $k_i = s_i$ for all nodes i, Eq. (7) gives

$$R_0^\infty = R_0^{\text{Est}}(\{s_i\}, \{s_i\}, \tau, \beta, \gamma) = \frac{\beta \sum_i s_i^2}{\gamma\tau \sum_i s_i} \tag{8}$$

The value of R_0^∞ can be chosen arbitrarily. Then, by setting $\gamma = 1/\tau$ and $\beta = R_0^\infty \sum_i s_i / \sum_i s_i^2$ we guarantee that Eq. (8) holds for every network. To test that our results hold over a range of disease scenarios we repeat our analysis with $R_0^\infty = 2$, 3, and 4.

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

bioRxiv preprint doi: https://doi.org/10.1101/170266; this version posted February 27, 2020. 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.

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, Janet Mann, and Shweta Bansal. Disease implications of animal social network structure: A synthesis across social systems. *Journal of Animal Ecology*, 87(3):546–558, 2018.

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

[5] 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.

[6] 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.

[7] Skylar R. Hopkins, Arietta E. Fleming-Davies, Lisa K. Belden, and Jeremy M. Wojdak. Systematic review of modelling assumptions and empirical evidence: Does parasite transmission increase nonlinearly with host density? *Methods in Ecology and Evolution*, n/a(n/a).

[8] 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.

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

[10] 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.

[11] Joan B Silk. The adaptive value of sociality in mammalian groups. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 362(1480):539–559, 2007.

[12] 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.

[13] Roslyn Dakin and T Brandt Ryder. Reciprocity and behavioral heterogeneity govern the stability of social networks. *Proceedings of the National Academy of Sciences*, 117(6):2993–2999, 2020.

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

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

[16] 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.

[17] 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.

[18] Ignacio Tamarit, José A. Cuesta, Robin I. M. Dunbar, and Angel Sánchez. Cognitive resource allocation determines the organization of personal networks. *Proceedings of the National Academy of Sciences*, 115(33):8316–8321, 2018.

[19] 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.

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

[21] A. Barrat, M. Barthlemy, R. Pastor-Satorras, and A. Vespignani. The architecture of complex weighted networks. *Proceedings of the National Academy of Sciences*, 101(11):3747–3752, March 2004.

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

[23] 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.

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

[25] 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.

[26] 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.

[27] 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.

[28] 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.

[29] 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.

[30] Andreas P Modlmeier, Ewan Colman, Ephraim M Hanks, Ryan Bringenberg, Shweta Bansal, and David P Hughes. Ant colonies maintain social homeostasis in the face of decreased density. *eLife*, 8:e38473, may 2019.

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

[32] MATHIEU GNOIS, CHRISTIAN L. VESTERGAARD, JULIE FOURNET, ANDR PANISSON, ISABELLE BONMARIN, and ALAIN BARRAT. Data on face-to-face contacts in an office building suggest a low-cost vaccination strategy based on community linkers. *Network Science*, 3:326–347, 9 2015.

[33] 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.

[34] 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.

[35] 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.

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

[37] 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.

bioRxiv preprint doi: https://doi.org/10.1101/170266; this version posted February 27, 2020. 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.

[38] 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.

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

[40] 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.

[41] 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.

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

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

[44] 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.

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

[46] Tim Gernat, Vikyath D. Rao, Martin Middendorf, Harry Dankowicz, Nigel Goldenfeld, and Gene E. Robinson. Automated monitoring of behavior reveals bursty interaction patterns and rapid spreading dynamics in honeybee social networks. *Proceedings of the National Academy of Sciences*, 115(7):1433–1438, 2018.

[47] Vincent D Blondel, Jean-Loup Guillaume, Renaud Lambiotte, and Etienne Lefebvre. Fast unfolding of communities in large networks. *Journal of statistical mechanics: theory and experiment*, 2008(10):P10008, 2008.

[48] Odo Diekmann, Johan Andre Peter Heesterbeek, and Johan AJ Metz. On the definition and the computation of the basic reproduction ratio r 0 in models for infectious diseases in heterogeneous populations. *Journal of mathematical biology*, 28(4):365–382, 1990.

[49] RM Anderson, GF Medley, RM May, and AM Johnson. A preliminary study of the transmission dynamics of the human immunodeficiency virus (hiv), the causative agent of aids. *Mathematical Medicine and Biology: a Journal of the IMA*, 3(4):229–263, 1986.

[50] Mark Newman. *Networks*. Oxford university press, 2018.

[51] Timo Smieszek, Lena Fiebig, and Roland W. Scholz. Models of epidemics: when contact repetition and clustering should be included. *Theoretical Biology and Medical Modelling*, 6(1):11, Jun 2009.

[52] Juliette Stehlé, Nicolas Voirin, Alain Barrat, Ciro Cattuto, Vittoria Colizza, Lorenzo Isella, Corinne Regis, Jean-François Pinton, Nagham Khanafer, Wouter Van den Broeck, and Philippe Vanhems. Simulation of an SEIR Infectious Disease Model on the Dynamic Contact Network of Conference Attendees. *BMC Medicine*, 9(87), jul 2011.

[53] 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.

[54] Duncan J Watts and Steven H Strogatz. Collective dynamics of small-world networks. *nature*, 393(6684):440, 1998.

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

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

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

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

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

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

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

[62] Alain Barrat, Ciro Cattuto, Alberto Eugenio Tozzi, Philippe Vanhems, and Nicolas Voirin. Measuring contact patterns with wearable sensors: methods, data characteristics and applications to data-driven simulations of infectious diseases. *Clinical Microbiology and Infection*, 20(1):10–16, 2014.

[63] 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.

[64] Michael Begon, Malcolm Bennett, Roger G Bowers, Nigel P French, SM Hazel, and Joseph Turner. A clarification of transmission terms in host-microparasite models: numbers, densities and areas. *Epidemiology & Infection*, 129(1):147–153, 2002.

[65] Matthew J Ferrari, Sarah E Perkins, Laura W Pomeroy, and Ottar N Bjørnstad. Pathogens, social networks, and the paradox of transmission scaling. *Interdisciplinary perspectives on infectious diseases*, 2011, 2011.

[66] 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.

[67] 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.

[68] 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.