The interplay of social constraints and individual variation in risk tolerance in the emergence of superspreaders

Matthew J. Young, Matthew J. Silk, Alexander J. Pritchard, Nina H Fefferman

January 2023

Abstract

Individual host behaviors can drastically impact the spread of infection through a population. Differences in the value individuals place on both socializing with others and avoiding infection have been shown to yield emergent homophily in social networks and thereby shape epidemic outcomes. We build on this understanding to explore how individuals who do not conform to their social surroundings contribute to the propagation of infection during outbreaks. We show how non-conforming individuals, even if they do not directly expose a disproportionate number of other individuals themselves, can become functional superspreaders through an emergent social structure that positions them as the functional links by which infection jumps between otherwise separate communities. Our results can help estimate the potential success of real-world interventions that may be compromised by a small number of nonconformists if their impacted is not anticipated, and plan for how best to mitigate their effects on intervention success.

1 Introduction

Mathematical models of infectious disease dynamics are powerful tools for understanding and predicting the course of outbreaks [1]. They allow us to organize and quantify our conceptualization of the causes and influences of features of outbreak scenarios [1]. They also make it possible to interrogate hypothetical situations, considering particular interventions and evaluating their potential impacts [1, 2]. Among the features identified as potentially having a major impact on the course of outbreaks are superspreaders – individuals who are responsible for a disproportionate number of secondary infections should they become infected themselves [3, 4, 5]. The number and distribution of superspreaders have been shown to influence the accuracy of predicting overall outbreak dynamics (e.g., burstiness [6, 7, 8, 9] or dispersion [10, 11, 12, 13]).

One area that, until recently, has received less focus in epidemiological model formulation has been quantifying how factors that influence individual behavior affect emergent disease dynamics. The COVID-19 pandemic has provided many examples of how proposed interventions can be compromised by failing to incorporate individual-level effects (e.g., adherence to non-pharmaceutical interventions, or acceptance of novel vaccines; [14, 15]). While up until now, superspreading has been included in epidemiological models as an important feature in ongoing transmission dynamics [13, 16, 17, 18, 19], the growing realization that the motivations behind individual differences are (at least in part) due to behavioral heterogeneity [20, 21] suggests that the underlying motivations for behaviour that results in superspreading may itself provide additional points of intervention for altering the course of epidemics.

Another insight derived from epidemiological models has been the greater impact of some aspects of population-wide contact network structure than others on how fast infection spread (e.g., modularity; [22, 23]). From a networks perspective, individuals may function as a more indirect type of "superspreader" than strictly falls under the traditional definition. Individuals may be responsible for greater numbers of "descendant infections" (infections that occur due to ongoing transmission dynamics stemming from their own infection) if they critically connect otherwise isolated populations, even if they expose only relatively few others to infection themselves. We previously explored how individual discrete preferences regarding the value of social contact and infection risk tolerance might lead to the emergence of homophily in network community structure [24]. Here we extend this work to see if potential mismatches in the social networks available to individuals to satisfy their social and risk avoidance needs might lead to the endogenous emergence of individuals who act as these "transmission catalysts".

To address this question, we implement an explicit network structure to examine how a network with diverse individual priorities for disease avoidance and socialization might alter the transmission dynamics and epidemic severity of an infectious disease. We explore these dynamics with the acknowledgment that a majority of individuals ('conformists', hereafter) are likely to share similar priorities, e.g., centered around a population median. Some individuals, however, do not conform to these normalized bounds ('nonconformists', hereafter). Their behavior and social connectivity will be distinct, resulting in infrequent but disproportionately influential behaviors with regards to transmission (e.g., acting as superspreaders or transmission catalysts due to their unusual preferences).

Given known patterns of assortativity [24], individuals who depart from average behavioral responses when making decisions about social contacts under threat of socially-transmitted infections would be expected to impact emerging network connectivity patterns. We include multiple pathways by which individuals derive benefits from social contacts, ranging from purely psychological (i.e., satisfaction derived from in-person reinforcement of social bonds) to purely pragmatic (i.e., needing to maintain income from a job that requires in-person contact with clients and/or coworkers). This implies that societal function imposes a minimum bound for community-wide averages in socialization; individu-

als cannot fully disengage from social contact with their community, even if they are more risk averse or place far lesser value on social contact. We also include the potential for heterogeneity in infectious disease risk avoidance, where differences may be attributable to alternate perceptions of the severity of infection or in attitudes that determine the acceptability of undertaking those risks.

We formulate the question as one of individual fulfillment in a trade-off between social contact and risk tolerance, and then contextualize it within the available outlets for potential contact afforded by social organization. By doing so we show that emergent social patterns can lead non-conformist individuals to become transmission catalysts, even if they defy the traditional definition of a superspreader role. In particular, our model highlights that the impact of non-conformists in transmission dynamics peaks when constraints on socialization are intermediate.

2 Methods

Let P be a population of players with a bijection onto the vertices of a finite connected weighted graph G. We say two players are adjacent if their corresponding vertices on the graph are.

Game Formalism

Let $M > 0$ be a constant. At each time t, player x decides on a social action strategy $\vec{a}_x(t) = \{a_{x,y}(t) \in [0, M] : y \in N(x)\}$ corresponding to an amount of social contact they want to have with each player in N(x), the set of players adjacent to them. This social contact requires adjacency and mutual agreement, so the amount of socializing that actually occurs between any players x and y is

$$a'_{x,y}(t) := \begin{cases} min(a_{x,y}(t), a_{y,x}(t)) & \text{if } x \text{ and } y \text{ adjacent} \\ 0 & \text{if } x \text{ and } y \text{ not adjacent} \end{cases}$$

The utility of player x at time t is defined by

$$u_x(t) = \sigma_x f(a_x)\Delta t - \rho_x p(\vec{a'}_x, \Delta t) \tag{1}$$

where f is a nondecreasing concave function (such as $ln(x)$) which yields greater utility for higher inputs at diminishing rates, Δt is the amount of time that passes between each timestep, $a_x = \Sigma b_{x,y} a'_{x,y}$, and $p(\vec{a'}_x, \Delta t)$ is the expected probability of becoming infected (defined in the epidemiology portion of the model below). σ_x represents the value player x places on socializing overall: more social players will have higher values of σ_x and place more emphasis on socializing. ρ_x represents the value player x places on health: more health conscious (risk averse) players will have higher values of ρ_x, and place more emphasis on avoiding disease. $b_{x,y}$ is the weight of the connection between players x and y on the graph, representing the value they obtain from mutual socialization relative to the infection risk. Higher values allow players to obtain

more value with less risk. For simplicity we assume this is symmetric, with $b_{x,y} = b_{y,x}$.

Epidemiological Calculations within the Game Formalism

We model a discrete SIS disease on the graph structure with transmission varying based on the actions of the players. Since the force of infection can be expected to scale proportionally to the amount of social contact among the players [25], we model the probability of transmission from an infected player y to an uninfected player x in each timestep as $\Delta t \beta a'_{x,y}$, where β is the inherent transmissibility of the disease. Additionally, in order to avoid stochastic extinction of the disease both in the population as a whole and in socially isolated subgraphs, we add an additional small rate of infection g_a for each agent regardless of the social dynamics. (For figures presented in this work, g_a is set to 0.001.) It follows that if player x is uninfected, the probability of becoming infected in the next timestep is

$$p(\vec{a'}_x, \Delta t) = 1 - (1 - \Delta t g_a) \prod_{y \in N(x)} (1 - \Delta t \beta a'_{x,y} q_y) \tag{2}$$

where q_y is 1 if player y is infected, and 0 otherwise. For small transmission rates, this can be approximated as

$$p(\vec{a'}_x, \Delta t) = \Delta t (g_a + \beta \sum_{y \in G} a'_{x,y} q_y). \tag{3}$$

For simplicity we use Equation 3 for agents' utility functions.

To ensure that the transmission probability cannot exceed 1, we require the maximum pairwise socialization $M \leq 1/(\beta \Delta t)$, though in practice this limit is not reached in simulations with reasonably small β.

We assume that an infected player remains infected for a fixed amount of time, Γ, before returning to the susceptible state.

For Figures presented in this paper, we set $\beta = 0.125, M = 5, \Gamma = 10, \Delta t = 1$.

Social-risk Typology

We consider a population with distinct phenotypes of players. We define *low* and *high* values for the relevant parameters such that $\sigma_l < \sigma_h$, $\rho_l < \rho_h$, and assign players with $\sigma_x \in \{\sigma_l, \sigma_h\}$ and $\rho_x \in \{\rho_l, \rho_h\}$. This creates four types of players, which we denote with identifiers based on their combined parameters: AS (asocial, $\sigma_x = \sigma_l$), SO (social, $\sigma_x = \sigma_h$), RT (Risk-Taking, $\rho_x = \rho_l$), RA (Risk-Averse, $\rho_x = \rho_h$). Combining these identifiers leads to the four types: ASRA, ASRT, SORA, SORT.

For figures presented in this paper, we set $\sigma_l = 4, \sigma_h = 12, \rho_l = 30, \rho_h = 60$. Due to the scaling invariance of utility functions, the behavior of each player will be determined by their risk ratios σ_x/ρ_x, which will be $0.066, 0.133, 0.2, 0.4$ respectively.

Additionally, we assume that the type of each player is known to their neighbors. This is a realistic assumption because sociability and risk-aversion (in the context of health outcomes and protective actions taken) can be observable in many cases [26, 27, 28, 29]. We also assume players know the average frequency of infection within each type, but do not know the infection status of any particular individual. This is consistent with unreliable proximate cues of infection, especially for diseases with a pre-symptomatic phase; [30, 31, 32].

Since agents cannot observe the infection of status of their neighbors directly, they combine the known social-risk type of neighbors with known infection frequencies to estimate the expected utility of socializing with them. Players will estimate a local maximum for their utility with respect to each potential social partner based on that player's social-risk type. In particular:

$$\frac{\partial u}{\partial a'_{j,k}} > 0 \text{ iff } \Sigma a'_{x,y} < \frac{\sigma_x}{\rho_x \beta I_y}. \tag{4}$$

where I_y is the frequency of infections for agents of the same type as agent y. In essence, the nonlinear rate of returns on utility from socializing weighed against the linear decrease from infection risk will incentivize players to maintain medium values of total socialization: players with more or less total socialization than the RHS of equation 4 can increase their utility by decreasing or increasing their socialization respectively with that particular agent. The actual equilibrium point an agent reaches when all pairwise relationships are taken into account will depend on the (potentially complex) synergistic effects of each individual player's own utility function, the infection rate of each agent type, their neighbors' agent types, and the socialization desires of their neighbors over time. In general we expect each player to maintain intermediate levels of socialization rather than heading towards the boundaries of 0 or M; each player effectively having a soft quota for how much socialization they will maintain in the long term. We also expect players with higher risk ratio σ_x/ρ_x to maintain higher levels of socialization.

Community Structure

Unlike the well-mixed population we explore in [24], here we construct a community structure via constrained random graphs to extend the patterns observed in that work but also consider diversity of preference within the structural community, incorporating both individuals who are well-represented by a median valuation placed by their community on socialization and/or health, and individuals who are still members of a community of contacts, but whose preferences are not closely aligned with the median beliefs of the community as a whole.

1. We create a total of $N = N_c S_c$ agents subdivided into N_c communities of size $S_c \geq 4$, and choose a primary agent type from the four player types. Within each community, there are $S_c - 3$ agents of the primary type, which we refer to as "conformists" and 1 agent of each other type, which we refer to as "nonconformists". For all figures presented in this work, we set $N_c = 16, S_c = 10$.

2. We fix two probabilities $c_a \geq c_b$. For each pair of agents within the same community, they are connected with probability c_a. For each pair of agents in different communities, they are connected with probability c_b. This creates a form of stochastic block model in which any graph structure has nonzero probability, but as c_a increases and c_b decreases graphs with a clustered structure become increasingly common. For all figures presented in this work, we set $c_a = 2/3$, $c_b = 2/75$, which results in each agent having a mean of 6 connections in their community and 4 connections outside their community after this step.

3. For each nonconformist type, we partition all agents of that type into pairs and connect the agents in each pair if they were not already connected.

4. We require that every vertex have at least one neighbor, discarding any graph generated with isolated vertices - a rare occurrence with the implemented parameter set.

5. To incentivize interactions within a community we set $b_{x,y} = 1$ for connections within a community, and $b_{x,y} = 0.9$ for connections outside the community, so each agent will have a slight preference to socialize with their own community members.

We note that we also ran simulations of the model on random graph structures, which we briefly discuss in the Supplementary material.

Minimum Socialization

To model the tendency for individuals being required to interact with others in their community in a way that they cannot opt out of, we explore the impact of a minimum socialization value $\omega \geq 0$ and require that the socialization between every pair of mutually connected members within the same community always have socialization $a_{x,y}(t) \geq \omega$. We vary ω over simulations and explore its impact on the dynamics and outcomes. Low values of ω correspond to loosely bound, voluntary communities where it is easy to avoid undesired contact, while high values of ω correspond to close communities in which members interact frequently whether they prefer to or not. This affects not only disease transmission within each community, but also the interactions individuals will have outside their community, as individuals are more likely to seek socialization outside their community if they have less within it, and vice versa.

Simulations

Each agent x begins the simulation with $a_{x,y} = 1$ for each neighbor. At each time step we compute the actual socialization $a'_{x,y}$, then simulate the epidemiological dynamics that result for that time step. Each agent is then informed of the new infection frequency for each social-risk type, and updates their social preferences in the direction of best response. In particular, we fix an update speed $\Delta a = 0.01$, then for each neighbor y, agent x estimates u_x in the counterfactual situations where their socialization with agent y had been $a'_{x,y} + n\Delta a \Delta t$, where $n \in \{-1, 0, 1\}$, and all other socialization levels remained unchanged. The

player then increments $a_{x,y}$ by $n\Delta a\Delta t$ according to the n that yields the highest utility. This is computed simultaneously and independently for each y.

For each measurement process, we simulate the model for 2000 initial timesteps to get past the initial outbreak and reach a more representative long-term state. We then simulate the model for a measurement period of 10000 timesteps in which we track the number of new infections and who causes them. We choose values of ω ranging from 0 to 0.3 in intervals of 0.015 for Figures 4 and 5, and in intervals of 0.03 for all other figures. We run 10 realizations of each parameter set, yielding data from from 160 agents of each nonconformist type and 1120 agents of the conformist type, and average the values for each agent type.

For some measurements, we construct "parallel networks", in which we create one parent network according to the process defined earlier and create several copies of it. In particular, for each parent network we choose 5 communities in the network, and create 15 copy networks, one for each nonconformist in the chosen networks. In each copy, we change the corresponding nonconformist into a conformist, changing utility function and agent type that other players perceive them as, but we do not change their connections in the network structure. We then simulate all networks simultaneously with the same random seed, such that the same random event occurring in multiple networks on the same time step, with identical probabilities, will occur in all or none. This means any difference in outcomes reflects genuine differences in the probabilities caused by the network structure and agent behavior rather than random noise.

Measurements

To understand the social dynamics of the system, we consider how individuals shift in homophily along two axes of socialization: with their community and with their social-risk typology. We use a modified Krackhardt EI homophily index ([24], [33]). In particular, we partition the set of agents according to either their social-risk type or community membership. Then for each agent x, let the internal socialization I_x be the average socialization between x and each neighbor in their group, and let the external socialization E_x be the average socialization between x and each neighbor in another group.

Type-Homophily and **Community-Homophily** are defined to be

$$H_x = \frac{I_x - E_x}{I_x + E_x} \tag{5}$$

using the social-risk type partition and community partition respectively. In the case where individuals socialize exclusively within their own group, this yields $H_x = 1$, in the case where individuals socialize exclusively outside their own group this yields $H = -1$, and in the case where individuals express no preference between social partners this yields $H_x = 0$.

In the rare event that an agent has no neighbors in their own group or only has neighbors in their own group, then I_x or E_x respectively will be undefined, and we leave H_x undefined, omitting such agents from the computed averages

of that homophily score.

The effect an individual has on the spread of a disease is nontrivial to measure, given the complicated epidemiological dynamics. For instance, an individual might be considered to be responsible only for the infections that they directly transmit, or they might be considered partially responsible for infections caused by their epidemiological descendants. Thus, we use several different measures to compare the outcome and effects of individuals. We term these measures as: Direct Attribution, Epidemic Attribution, Global Attribution, and Community Attribution. For simplicity, we define each of these based on events occurring over a measurement period of $10,000$ timesteps. (Note: the values increase at an approximately linear rate, and so could be generalized to different lengths of times by normalizing based on the measurement period.)

Direct Attribution of agent x, D_x is calculated as the number of infections during the measurement period that were caused by a susceptible individual receiving the infection from agent x over the course of the measurement period. In the case where a susceptible agent is successfully infected by multiple infected individuals in the same timestep, the value is split evenly among them: the Direct Attribution of each infector increases by $1/n$ where n is the number of infectors. In this way, the total Direct Attribution of all agents is equal to the number of new infections during the measurement period.

While Direct Attribution is the most straightforward measure of the impact of an individual on the epidemic, it misses second-order effects caused by the epidemiological descendents of each agent. Observable differences in nonconformists could, however, be attributed to individual utility functions or topological network features. That is, the effect of an individual depends not only on their behavior, but also on their placement within the network structure. Individuals with more connections, or better placed in a position to bridge different communities or subgraphs in the network may have an outsized effect relative to their total socialization. Additionally, conformists and nonconformists differ not only their behavior, but in how other agents treat them, which will affect their role in the epidemic in a way not directly controlled by them.

Epidemic Attribution was introduced to assign credit to the originator of local outbreaks. The Epidemic Attribution E_x of agent x is defined as follows. If community c has no infected agents at timestep $t-1$, and has exactly one at timestep t, then we say an Epidemic Event $E_{x,c,t}$ has begun with the infected agent x as its originator. During the Epidemic Event, we track individuals who are infected directly by other agents within community c tracing back to agent x, while ignoring infections that come from outside the community during this time or that leave the community and come back. Effectively, we consider the graph of epidemiological ancestry and find the maximal subtree with agent x as the parent node at time t, with all members being part of community c. This means if an individual becomes infected from outside their community while

an Epidemic Event is ongoing, their infection, and any infections they cause, do not count as part of the Event. Only transmissions that trace back to the originator of the Event without leaving the community are counted. Additionally, if an individual becomes infected multiple times during this period they can be measured multiple times, so there is no upper bound on the number of infections an Event can record. The Epidemic Event ends once no more infections are present in the community, or when the measurement period ends. At this point, we attribute all infections caused during the Epidemic Event to the originator, agent x, and increase their Epidemic Attribution E_x by the total number of infections in the Event. Any epidemic still ongoing at the end of the measuring process is terminated and counted as if complete in order to avoid undercounting epidemics with extremely long durations.

Global Attribution was included to more directly quantify the actual impact of an individual's nonconformity. We construct parallel networks as described above, and define G_x of agent x to be the difference in the total number of infections between the parallel network with x recoded to be a conformist and the number of infections in original network.

Community Attribution C_x was defined analogously, but with infections restricted to the community of agent x, then multiplied by the number of communities, N_c, so that it's comparable in magnitude to Global Attribution.

For all of the above measures, we create an analogous measure for each type of agent by averaging the measure of multiple agents of that type over multiple networks.

We note that we also created a generalized version of Direct Attribution that accounts for epidemiological ancestry, which we call r-Attribution. However due to its similarity to Direct Attribution but increased complexity, it is presented in the Supplementary material.

3 Results

Homophily

Emerging homophily is dependent on the minimum socialization, ω, (see Figure 1 for results for a SORA-conformist network). For community-homophily, low values of minimum socialization result in nonconformists being defined almost exclusively by heterophily. This observation is expected given the mandated connections for conformists with an external community member of the same social-risk typology. Also as expected, increases in minimum socialization also increase homophily within communities.

With low values of minimum socialization, we observed similar patterns in assortativity for social-risk typology as previously published patterns in a model without defined structure (Young et al., 2022). As minimum socialization was

increased, however, all social-risk types converged towards heterophily with the exception of the conformists. This is an artifact of each community only having a single non-conformist of each type. When minimum socialization becomes high enough, each individual is socializing almost entirely due to the constraint and not their preferences, and thus socializes with all community members equally. For an agent with neighbors of every type in their community, this is measured as zero homophily. But nonconformists are the only agent of their type in their community, so socialize exclusively with other types, despite having neighbors of the same type outside their community. In larger communities with more nonconformists, all agents converge to zero type-homophily as minimum socialization increases.

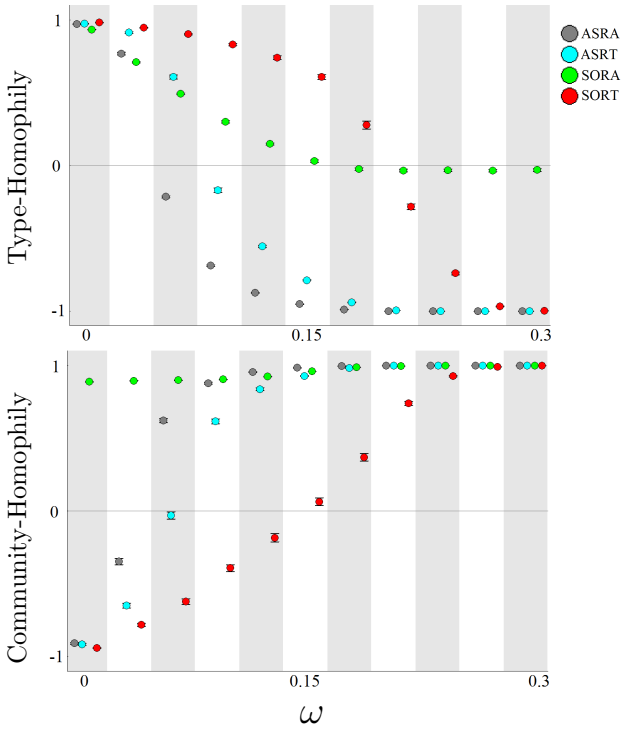

Figure 1: Homophily of each agent type (top) and community membership (bottom) as a function of minimum socialization (ω) for networks with SORA conformists. Points are colored by node type and error bars represent the standard deviation. We use a x axis jitter for visualisation purposes only, distinguishing between adjacent values of ω with gray/white shading.

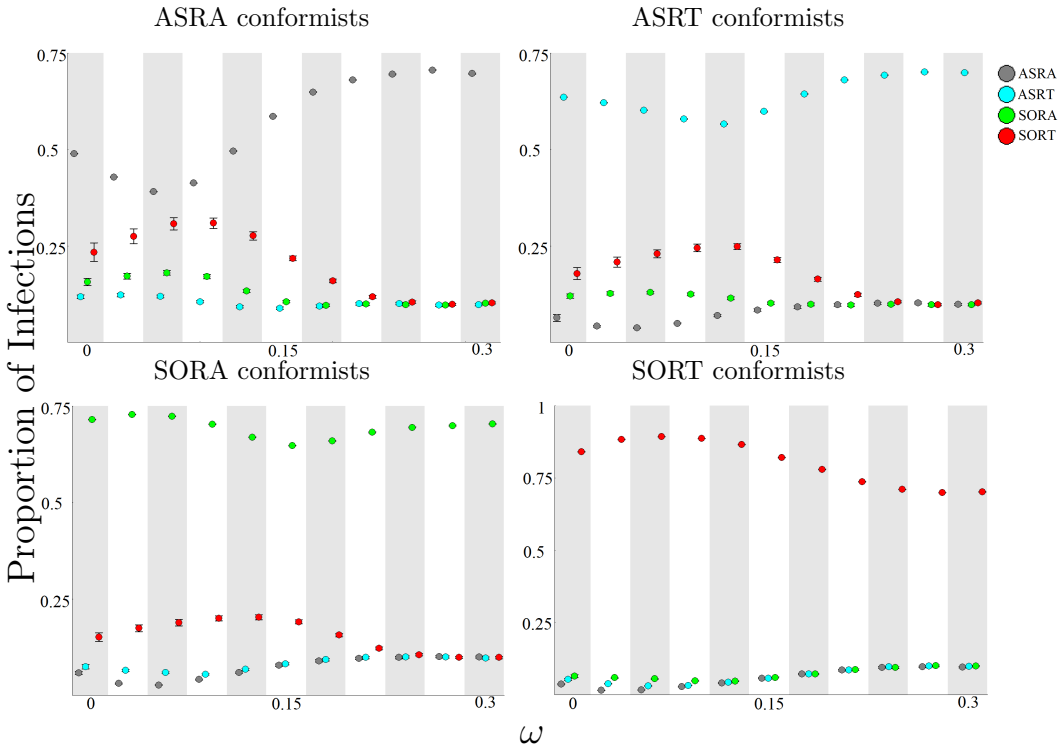

Figure 2: The proportion of Direct Attribution attributable to each agent type as a function of minimum socialization (ω) for the four networks with conformists of each type. Points are colored by node type and error bars represent the standard deviation. We use a x axis jitter for visualisation purposes only, distinguishing between adjacent values of ω with gray/white shading. Note that we do not correct here for the numbers of agent of each type; since the conformists in each network are seven times as frequent as each type of non-conformist, attribution of 0.7 for conformists and 0.1 for each non-conformist reflects representation rather than an emergent outcome. Each graph shows simulations for networks with a different conformist type.

Direct Attribution

At high values of minimum socialization, all networks produce proportional Direct Attribution that reflects the representation of each type within the network (i.e., 0.7 for conformists and 0.1 for each nonconformist type; see Figure 2). However, at lower values of minimal socialization, nonconformists with a higher risk-health ratio have a disproportionate impact relative to their frequency in the population in every case apart from with SORT conformists. The effect is greatest at intermediate values of ω. For SORA and ASRT conformists, only SORT non-conformists have a disproportionate impact, while for ASRA con-

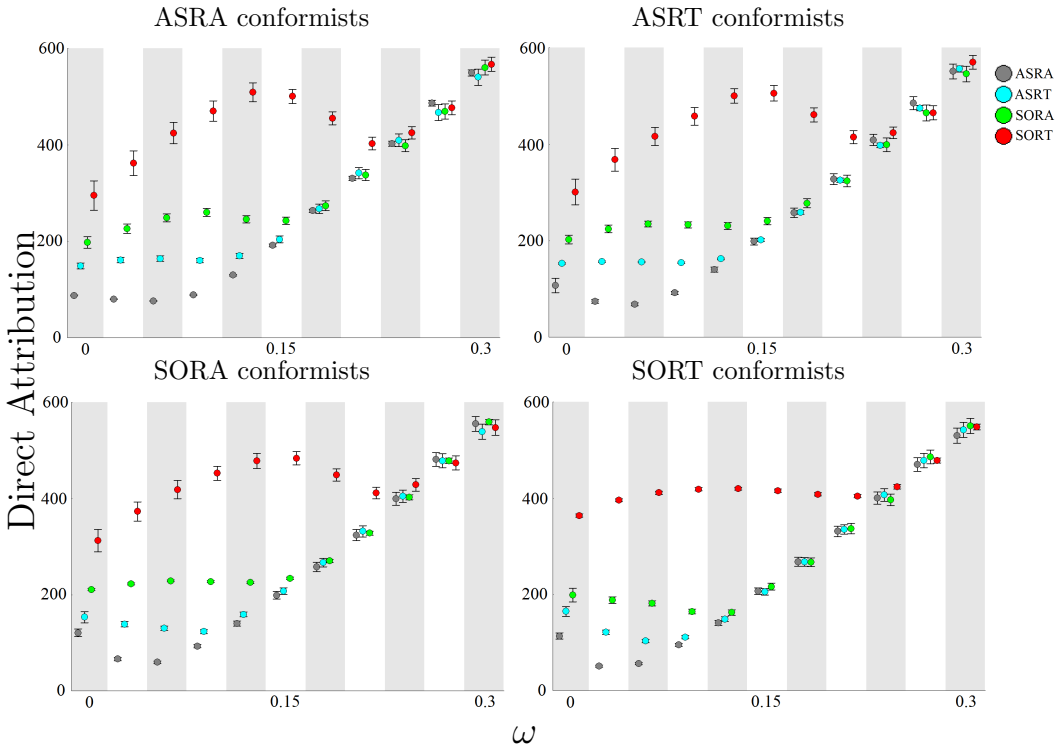

Figure 3: The mean Direct Attribution of agents of each type as a function of minimum socialization (ω) for the four networks with conformists of each type. Points are colored by node type and error bars represent the standard deviation. We use a x axis jitter for visualisation purposes only, distinguishing between adjacent values of ω with gray/white shading. Note that this normalization corrects for the numbers of agent of each type, such that the conformists and non-conformists are directly comparable.

formists this effect extends to all non-conformists, albeit to a lesser extent and at lower values of ω. For SORT conformists, the result is qualitatively different with conformists being more important than would be expected at low and intermediate values of ω, although the effect is small.

The peak in the importance of SORT non-conformists at intermediate minimum socialization values (ω) remains when we instead focus on the relative impact of the individual agent by normalizing with respect to the number of agents rather than the total amount of Direct Attribution (see Figure 3). In all cases, for low values of ω we see higher Direct Attribution for more risky agents and for high values of ω the Direct Attributions converge. In simulations with ASRA, ASRT and SORA conformists we observe a local maximum for SORT non-conformists for intermediate minimum socialization values around $\omega = 0.15$ (the peak is consistent across all three types). SORT Direct Attribution then

decreases again before converging to a concomitant increase with the other agent types at higher values of ω. For SORT conformists we see no similar peak at intermediate values of ω.

Note that the total Direct Attribution of all agents is equal to the total number of infections in the population. Thus, we see an increase in the total number of infections as the total socialization rate of agents increases, indicating that we have not yet hit epidemiological saturation (where everyone is constantly in contact with so many others that as soon as they return to being susceptible after infection, they are assured of being immediately reinfected) with even our largest considered value of ω.

Epidemic Attribution

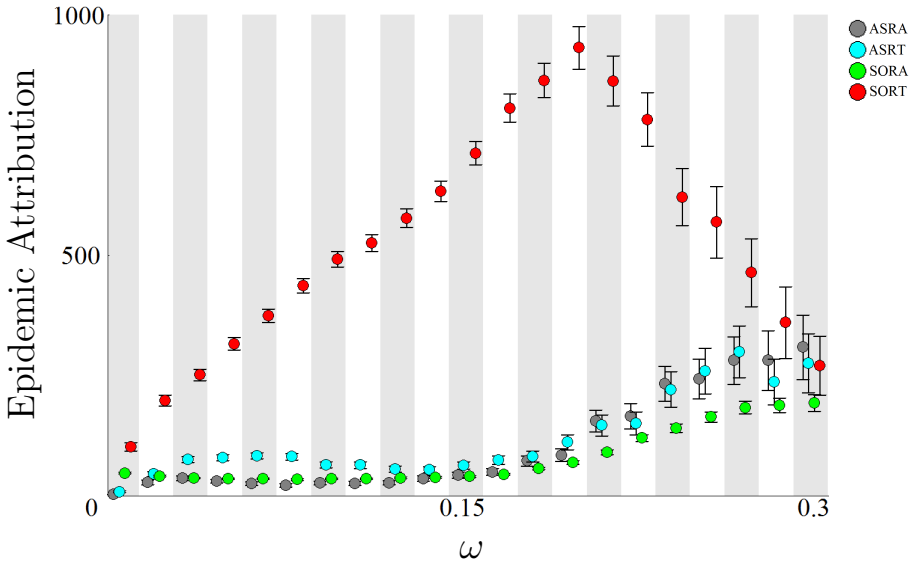

Figure 4: The mean Epidemic Attribution of each agent type as a function of minimum socialization (ω) for simulations with SORA conformists only. Points are colored by node type and error bars represent the standard deviation. We use a x axis jitter for visualisation purposes only, distinguishing between adjacent values of ω with gray/white shading. Note that this normalization corrects for the numbers of agent of each type, such that the conformists and non-conformists are directly comparable.

Figure 4 shows the Epidemic Attribution of each agent type in networks with SORA conformists as a function of ω. SORT agents again show the greatest impact, with substantially higher Epidemic Attribution for most values of ω. We also again observe a non-monotonic trend, where Epidemic Attribution for SORT agents is maximized at intermediate levels of ω rather than at the extremes.

To understand this, we separate Epidemic Attribution into two components: (1) the frequency of outbreaks introduced from outside the community via connection through that agent, and (2) the average size of an epidemic started by agents of each type, Figure 5.

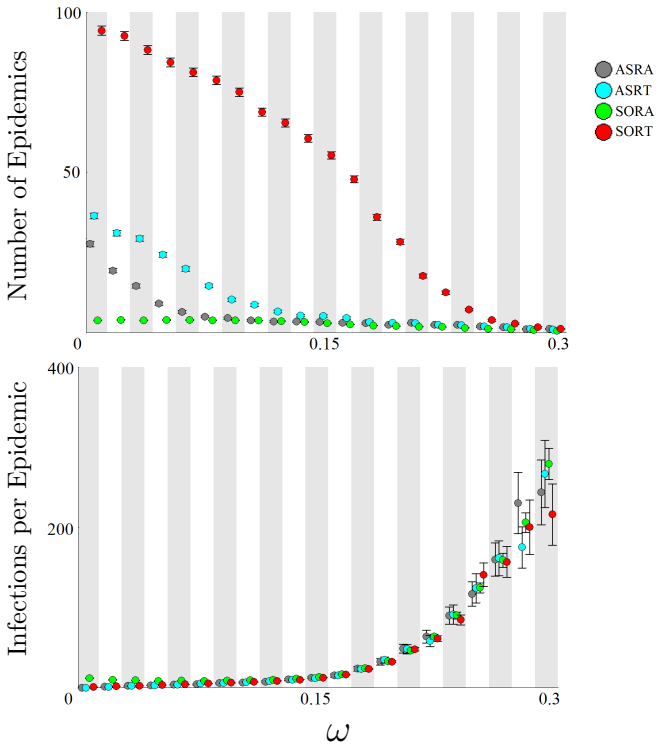

Figure 5: The mean number of epidemics (top) and mean epidemic size (bottom) for each agent type as a function of minimum socialization (ω) for networks with SORA conformists only. Points are colored by node type and error bars represent the standard deviation. We use a x axis jitter for visualisation purposes only, distinguishing between adjacent values of ω with gray/white shading. Note that this normalization corrects for the numbers of agent of each type, such that the conformists and non-conformists are directly comparable.

Figure 5 shows an example of these two components of Epidemic Attribution for networks with SORA conformists. We see that as ω increases, the number of epidemics decreases. To understand this, we consider each of two contributing factors. First, as agents are required to socialize more internally, they socialize less externally, thus being less likely to bring in a new infection when their community is uninfected. Second, our definition of Epidemic Events do not allow overlap, so the longer they tend to last the fewer opportunities there are for new ones to start. We also see that as ω increases, the total number of community members infected before the Epidemic Event ends also increases. Again, this

is expected, since increased socialization also increases the transmission rate of the infection (up until a saturation point), and an epidemic is consequently less likely to go locally extinct. The Epidemic Attribution for each agent type is the product of these two values, so we see it maximized for the SORT agents for intermediate values of ω where both of these factors are of moderate levels simultaneously rather than at the extremes. The patterns in Epidemic Attribution for simulations with ASRA and ASRT conformists are highly similar (with minor quantitative differences), while data for SORT conformists shows a strikingly different pattern, with significantly less Attribution for SORT agents, and slightly more for SORA at low values of ω. For concision, results for other conformist types are presented in the Supplementary Material.

Global and Community Attribution

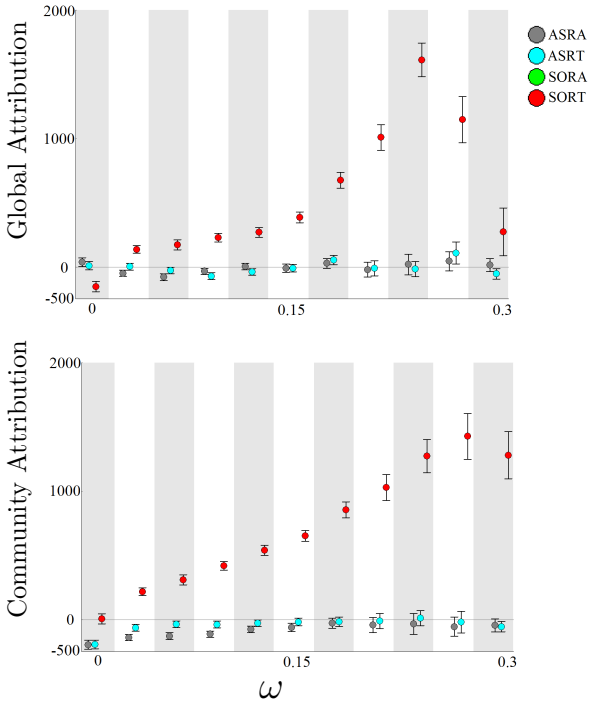

Figure 6: The mean Global (top) and Community (bottom) Attribution for each non-conformist type as a function of minimum socialization (ω) for networks with SORA conformists only. Points are colored by node type and error bars represent the standard deviation. We use a x axis jitter for visualisation purposes only, distinguishing between adjacent values of ω with gray/white shading., for networks with SORA conformists.

Figure 6 shows the Global and Community Attribution values for nonconformists as ω varies for networks with SORA conformists. Again, we observe

that SORT has much higher Attribution, and the value peaks for intermediate values of ω. This provides further evidence of the robustness of this peak in the epidemiological importance of nonconformists at intermediate values of ω, since the result is consistent across multiple different methods of measuring the role of agents in the epidemiological dynamics. However we note that the location of the peak is different from the other measures.

Interestingly, we see a small negative impact from ASRA and ASRT nonconformists at low to intermediate values of ω that suggests a protective community-level effect, though this seems less apparent in the Global population measure. Investigating over which region of the parameter space among transmissibility, local connectivity, and global connectivity this effect is present are questions for future consideration.

4 Discussion

Our results build on the growing awareness that individual preferences and behaviors can shape both the course of outbreaks and our ability to design effective interventions. For example, meaningful thresholds exist in population percentages for behaviors such as vaccine acceptance to prevent a widespread outbreak of infection [34, 35], yet there is still awareness of the importance of studying the impact of not only the distribution of behaviors, but their arrangement within the social fabric of a community [34, 36, 37, 38, 39]. Similarly, our work builds from an understanding that the distribution of behaviors can yield unexpected population-level outcomes [24], to the more nuanced understanding that rare individuals, who do not conform to the preferences and behaviors of their communities, can have a profound impact on epidemiological dynamics in these communities. Nonconformists who value socializing and are more risk tolerant than the rest of their community can emerge as non-traditional superspreaders (transmission catalysts), enabling broader transmission than would be possible in their absence. The similarities across multiple Attribution measurements helps demonstrate the robustness of the results.

The impact of more risk tolerant and socialization-valuing (SORT) nonconformists peaks at intermediate values of ω and then decreases before again increasing as the high level of minimum socialization overwhelms individual impacts. This change in the relative importance of the nonconformists as minimum socialization levels vary is qualitatively consistent with previous work showing that different patterns of connectivity in network-based infection dynamics yielded differential relative protection against ongoing transmission at different forces of infection [40]. While that work focused on the per contact transmission probability, our work suggests it is worthwhile to explore whether the emergent result is a composite effect of both connectivity and transmissibility.

From a global perspective, changes in the relative role of nonconformists can be explained by the assortment of agents into homophilous sets at extreme values of ω but not intermediate ones. For low values of ω agents mostly interact

with others of their same social-risk type, partitioning the graph into four sets with little cross-interaction. Outbreaks among one type of agent will mostly remain within that type, especially as other agents decrease their socialization with that type in response. For high values of ω agents mostly socialize within their own community, partitioning the graph into many smaller sets with higher minimum socialization causing local outbreaks to last much longer within each community, but less likely to spread to others. However for intermediate values of ω, neither partition is complete, so infections can spread across the entire network, and do so largely carried by the most social agent type.

From a local perspective, we can consider this same dynamic on the level of individual SORT agents and their capacity to act as transmission catalysts and/or superspreaders. Such agents are highly motivated to maintain substantially higher socialization values than others, leaving them more likely to be infected and making other agents avoid them when feasible. For low values of ω, other agents are able to avoid them, and are thus insulated from the impact of superspreader behavior. For high values of ω, all agents interact with all others within their community, meaning highly social agents are able to obtain sufficient socialization within their own community and do not seek it externally. Therefore, they are less likely to pick up new infections from outside and introduce them into their previously uninfected community. However, for intermediate values of ω, they engage in both internal and external socialization, causing them to acquire infections from other communities and spread them to their community.

By considering a simplified scenario in which one type of agent dominates as conformists and distributions of agent types are identical in different communities, we extract insight into emergent patterns that may be important in planning strategic outbreak mitigation. Minimum socialization is likely to be lowest when outbreaks involve deadly infections that make it acceptable to consider governmentally-imposed lockdowns [41, 42], or for populations of privileged people whose livelihoods and access to basic needs do not depend on public exposure (e.g., [43]). Minimum socialization is likely to be higher among poorer populations whose livelihoods are based on physical labor and other essential jobs that remain working in person even during lockdowns, as well as among groups that do not respect recommendations from public health services. A peak in superspreader impact at intermediate amounts of minimum socialization suggests the impact of a marginal change in socialization will vary depending on whether a community is already to the left or right of this peak. Policies intended to reduce infection risk may be counterproductive in communities with high levels of "minimum socialization" if they do not account for this.

We show that structural network properties combined with behavioural dynamics can put seemingly poorly connected individuals into critical transmission catalyst roles that connect otherwise disconnected communities. Consequently, policy-based and socially driven community norms designed to protect people may cause individuals whose beliefs are not well-represented to change their behaviour to compensate, linking between communities, acting as transmission

catalysts, and therefore putting others at risk. Without public health policies backed by strict enforcement on a population-wide scale (as have been imposed in some cases [44, 45]), policies will have to focus on individual decision making and the belief and value systems that motivate it. Community appeals and outreach efforts may have to be nuanced, or there may even be no practical meaningful way to prevent these sorts of connections from being made. At that point, it will likely come down to analyses of surge capacity and the timing of access to medical care as to whether implementing policies that delay, but not prevent, widespread community transmission will be beneficial to implement. These are not easy questions to answer - it may be that decreasing minimum socialization both effectively decreases the total number of infections while also creating new superspreaders who would otherwise have been unremarkable contributors. Shifting who is likely to be infected may also be a critical component of planning, especially if nonconformists in some communities are likely to be more medically vulnerable, as is sometimes the case (e.g., [46]). Additionally, understanding the source of superspreaders or transmission catalysts and which contexts are most likely to create them allows more targeted interventions targeted to mitigate their impact on an epidemic. Our results do not provide generalizable answers, but instead provide a generalizable framework by which to realize we should be asking these types of questions.

4.1 Limitations

There are some limitations in our study design that restrict its direct applicability to real world scenarios without additional refinement. By necessity we fix many model parameters rather than investigating its behavior over a broad parameter space. In particular, we consider a specific type of network with a fixed proportion of nonconformists with fixed social and epidemiological parameters. We selected this network structure because of the ubiquity of community structure in human societies and the potential importance of social "bridges" between communities in these types of network. While we expect our results to generalize outside these parameters, identifying the range of network structures over which this occurs requires further investigation.

Additionally, each of our measures of attribution have potential weaknesses. For example, Direct Attribution fails to capture the indirect effects of an individual as infections propagate further from them in the network and Epidemic Attribution assigns all credit to the originator of a somewhat arbitrarily defined "event" incorporating a definitional restriction that once an outbreak is ongoing, a new one cannot be initiated (a limitation if reinvigoration of outbreaks that would otherwise die out is important). Global and Community Attribution are more difficult to generalize to other features in alternative epidemiological models as they measure the effect of a change in parameters in a single individual. Further, they are impossible to measure in real world systems since they rely on counterfactual variations of the same network. However, by measuring the importance of individuals within the population using four different measurements that do not share the same issues, we mitigate these limitations of

each individual measure.

5 Conclusion

Our model highlights how mismatches between opportunities for social contact within a community can lead individuals who might otherwise not be described as superspreaders to act as critical connection pathways of transmission among communities. Researchers and policy makers should be aware of the potential for changes in socialization rates to incentivize some individuals to broaden their network of social partners and thus become superspreaders when previously they were not. Even this simplified model demonstrates how policies that may fail to include explicit consideration of the actions of such individuals (whether by preference or pragmatic requirement) might compromise the success of recommendations that consider only a "population average" description of population- and community-level structure. Our results show that incorporating underlying motivations that cause these individuals to depart from the expected "conformist" behaviours can drastically alter emergent network structure and consequently outbreak dynamics for an entire population. More effectively including these behavioral motivations will therefore be critical when considering or benchmarking policies that rely on individual behaviors to create (or adhere to) community standards of practice.

Acknowledgements

This work was supported by NSF DEB 2028710. MJS received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Sklodowska-Curie grant agreement No. 101023948

Code Availability

Project code is available on the code hosting platform GitHub at https://github.com/kazarraha/SocDistModel

References

[1] Eric T Lofgren, M Elizabeth Halloran, Caitlin M Rivers, John M Drake, Travis C Porco, Bryan Lewis, Wan Yang, Alessandro Vespignani, Jeffrey Shaman, Joseph NS Eisenberg, et al. Mathematical models: A key tool for outbreak response. *Proceedings of the National Academy of Sciences*, 111 (51):18095–18096, 2014. doi: https://doi.org/10.1073/pnas.1421551111.

[2] Oliver Morgan. How decision makers can use quantitative approaches to guide outbreak responses. *Philosophical Transactions of the Royal Society B*, 374(1776):20180365, 2019. doi: https://doi.org/10.1098/rstb.2018.0365.

[3] John T Kemper. On the identification of superspreaders for infectious disease. *Mathematical Biosciences*, 48(1-2):111–127, 1980. doi: https://doi.org/10.1016/0025-5564(80)90018-8.

[4] Sara H Paull, Sejin Song, Katherine M McClure, Loren C Sackett, A Marm Kilpatrick, and Pieter TJ Johnson. From superspreaders to disease hotspots: linking transmission across hosts and space. *Frontiers in Ecology and the Environment*, 10(2):75–82, 2012. doi: https://doi.org/10.1890/110111.

[5] Christina J Edholm, Blessing O Emerenini, Anarina L Murillo, Omar Saucedo, Nika Shakiba, Xueying Wang, Linda JS Allen, and Angela Peace. Searching for superspreaders: Identifying epidemic patterns associated with superspreading events in stochastic models. In *Understanding complex biological systems with Mathematics*, pages 1–29. Springer, 2018.

[6] Julius B Kirkegaard and Kim Sneppen. Superspreading quantified from bursty epidemic trajectories. *Scientific reports*, 11(1):1–7, 2021. doi: Superspreading quantified from bursty epidemic trajectories.

[7] Jessica Enright and Rowland Raymond Kao. Epidemics on dynamic networks. *Epidemics*, 24:88–97, 2018. doi: https://doi.org/10.1016/j.epidem.2018.04.003.

[8] Byungjoon Min, K-I Goh, and Alexei Vazquez. Spreading dynamics following bursty human activity patterns. *Physical Review E*, 83(3):036102, 2011. doi: https://doi.org/10.1103/PhysRevE.83.036102.

[9] Marco Mancastroppa, Alessandro Vezzani, Miguel A Muñoz, and Raffaella Burioni. Burstiness in activity-driven networks and the epidemic threshold. *Journal of Statistical Mechanics: Theory and Experiment*, 2019(5):053502, 2019. doi: https://doi.org/10.1088/1742-5468/ab16c4.

[10] Bjarke Frost Nielsen, Lone Simonsen, and Kim Sneppen. Covid-19 superspreading suggests mitigation by social network modulation. *Physical Review Letters*, 126(11):118301, 2021. doi: https://doi.org/10.1103/PhysRevLett.126.118301.

[11] Yan Gang, Zhou Tao, Wang Jie, Fu Zhong-Qian, and Wang Bing-Hong. Epidemic spread in weighted scale-free networks. *Chinese Physics Letters*, 22(2):510, 2005. doi: https://doi.org/10.1088/0256-307X/22/2/068.

[12] Camila Buono, Lucila G Alvarez-Zuzek, Pablo A Macri, and Lidia A Braunstein. Epidemics in partially overlapped multiplex networks. *PloS one*, 9 (3):e92200, 2014. doi: https://doi.org/10.1371/journal.pone.0092200.

[13] T Garske and CJ Rhodes. The effect of superspreading on epidemic outbreak size distributions. *Journal of Theoretical Biology*, 253(2):228–237, 2008. doi: https://doi.org/10.1016/j.jtbi.2008.02.038.

[14] A David Paltiel, Jason L Schwartz, Amy Zheng, and Rochelle P Walensky. Clinical outcomes of a covid-19 vaccine: Implementation over efficacy. *Health Affairs*, 40(1):42–52, 2021. doi: https://doi.org/10.1377/hlthaff.2020.02054.

[15] Yusheng Zhang, Liang Li, Yuewen Jiang, and Biqing Huang. Analysis of covid-19 prevention and control effects based on the seitrd dynamic model and wuhan epidemic statistics. *International Journal of Environmental Research and Public Health*, 17(24):9309, 2020. doi: https://doi.org/10.3390/ijerph17249309.

[16] Alison P Galvani and Robert M May. Dimensions of superspreading. *Nature*, 438(7066):293–295, 2005. doi: https://doi.org/10.1038/438293a.

[17] Ryo Fujie and Takashi Odagaki. Effects of superspreaders in spread of epidemic. *Physica A: Statistical Mechanics and its Applications*, 374(2): 843–852, 2007. doi: https://doi.org/10.1016/j.physa.2006.08.050.

[18] K Kawagoe, M Rychnovsky, S Chang, G Huber, LM Li, J Miller, R Pnini, B Veytsman, and D Yllanes. Epidemic dynamics in inhomogeneous populations and the role of superspreaders. *Physical Review Research*, 3(3): 033283, 2021. doi: https://doi.org/10.1103/PhysRevResearch.3.033283.

[19] Alex James, Jonathan W Pitchford, and Michael J Plank. An event-based model of superspreading in epidemics. *Proceedings of the Royal Society B: Biological Sciences*, 274(1610):741–747, 2007. doi: https://doi.org/10.1098/rspb.2006.0219.

[20] Istvan Szapudi. Heterogeneity in sir epidemics modeling: superspreaders and herd immunity. *Applied Network Science*, 5(1):1–12, 2020. doi: Heterogeneity in SIR epidemics modeling: superspreaders and herd immunity.

[21] Wei Duan, Xiaogang Qiu, Zhidong Cao, Xiaolong Zheng, and Kainan Cui. Heterogeneous and stochastic agent-based models for analyzing infectious diseases' super spreaders. *IEEE Intelligent Systems*, 28(4):18–25, 2013. doi: https://doi.org/10.1109/MIS.2013.29.

[22] Julian C Evans, David J Hodgson, Neeltje J Boogert, and Matthew J Silk. Group size and modularity interact to shape the spread of infection and information through animal societies. *Behavioral ecology and sociobiology*, 75(12):1–14, 2021. doi: https://doi.org/10.1007/s00265-021-03102-4.

[23] 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*, 114(16):4165–4170, 2017. doi: https://doi.org/10.1073/pnas.1613616114.

[24] Matthew J Young, Matthew J Silk, Alex J Pritchard, and Nina H Fefferman. Diversity in valuing social contact and risk tolerance leading to the emergence of homophily in populations facing infectious threats. *Physical Review E*, 105(4):044315, 2022. doi: https://doi.org/10.1103/PhysRevE.105.044315.

[25] Sergio Arregui, Alberto Aleta, Joaquín Sanz, and Yamir Moreno. Projecting social contact matrices to different demographic structures. *PLoS computational biology*, 14(12):e1006638, 2018. doi: https://doi.org/10.1371/journal.pcbi.1006638.

[26] Molly Leecaster, Damon JA Toth, Warren BP Pettey, Jeanette J Rainey, Hongjiang Gao, Amra Uzicanin, and Matthew Samore. Estimates of social contact in a middle school based on self-report and wireless sensor data. *PloS one*, 11(4):e0153690, 2016. doi: https://doi.org/10.1371/journal.pone.0153690.

[27] Thang Hoang, Pietro Coletti, Alessia Melegaro, Jacco Wallinga, Carlos G Grijalva, John W Edmunds, Philippe Beutels, and Niel Hens. A systematic review of social contact surveys to inform transmission models of close-contact infections. *Epidemiology (Cambridge, Mass.)*, 30(5):723, 2019. doi: https://doi.org/10.1097/EDE.0000000000001047.

[28] Rajiv N Rimal and J Douglas Storey. Construction of meaning during a pandemic: The forgotten role of social norms. *Health communication*, 35(14):1732–1734, 2020. doi: https://doi.org/10.1080/10410236.2020.1838091.

[29] Nehama Lewis. Priming effects of perceived norms on behavioral intention through observability. *Journal of Applied Social Psychology*, 43:E97–E108, 2013. doi: https://doi.org/10.1111/jasp.12038.

[30] Adam Catching, Sara Capponi, Ming Te Yeh, Simone Bianco, and Raul Andino. Examining the interplay between face mask usage, asymptomatic transmission, and social distancing on the spread of covid-19. *Scientific reports*, 11(1):1–11, 2021.

[31] Armanda Cetrulo, Dario Guarascio, and Maria Enrica Virgillito. The privilege of working from home at the time of social distancing. *Intereconomics*, 55(3):142–147, 2020. doi: https://doi.org/10.1007/s10272-020-0891-3.

[32] Andrea K Townsend, Dana M Hawley, Jessica F Stephenson, and Keelah EG Williams. Emerging infectious disease and the challenges of social distancing in human and non-human animals. *Proceedings of the Royal Society B*, 287(1932):20201039, 2020. doi: https://doi.org/10.1098/rspb.2020.1039.

[33] David Krackhardt and Robert N Stern. Informal networks and organizational crises: An experimental simulation. *Social psychology quarterly*, pages 123–140, 1988. doi: https://doi.org/10.2307/2786835.

[34] Matt Keeling, Mike Tildesley, Thomas House, and Leon Danon. The mathematics of vaccination. *Math. Today*, 49:40–43, 2013.

[35] Roy M Anderson and Robert M May. *Infectious diseases of humans: dynamics and control*. Oxford university press, 1992.

[36] Emma S McBryde. Network structure can play a role in vaccination thresholds and herd immunity: a simulation using a network mathematical model. *Clinical infectious diseases*, 48(5):685–686, 2009. doi: https://doi.org/10.1086/597012.

[37] Fengchen Liu, Wayne TA Enanoria, Jennifer Zipprich, Seth Blumberg, Kathleen Harriman, Sarah F Ackley, William D Wheaton, Justine L Allpress, and Travis C Porco. The role of vaccination coverage, individual behaviors, and the public health response in the control of measles epidemics: an agent-based simulation for california. *BMC public health*, 15 (1):1–16, 2015. doi: https://doi.org/10.1186/s12889-015-1766-6.

[38] Zhilan Feng, Andrew N Hill, Philip J Smith, and John W Glasser. An elaboration of theory about preventing outbreaks in homogeneous populations to include heterogeneity or preferential mixing. *Journal of theoretical biology*, 386:177–187, 2015. doi: https://doi.org/10.1016/j.jtbi.2015.09.006.

[39] Thomas House and Matt J Keeling. Epidemic prediction and control in clustered populations. *Journal of theoretical biology*, 272(1):1–7, 2011. doi: https://doi.org/10.1016/j.jtbi.2010.12.009.

[40] NH Fefferman and KL Ng. How disease models in static networks can fail to approximate disease in dynamic networks. *Physical Review E*, 76(3): 031919, 2007. doi: https://doi.org/10.1103/PhysRevE.76.031919.

[41] Jonathan P Caulkins, Dieter Grass, Gustav Feichtinger, Richard F Hartl, Peter M Kort, Alexia Prskawetz, Andrea Seidl, and Stefan Wrzaczek. The optimal lockdown intensity for covid-19. *Journal of Mathematical Economics*, 93:102489, 2021. doi: https://doi.org/10.1016/j.jmateco.2021.102489.

[42] Vadim A Karatayev, Madhur Anand, and Chris T Bauch. Local lockdowns outperform global lockdown on the far side of the covid-19 epidemic curve. *Proceedings of the National Academy of Sciences*, 117(39):24575–24580, 2020. doi: https://doi.org/10.1073/pnas.2014385117.

[43] Nina Fefferman, Chien-Fei Chen, Gregory Bonilla, Hannah Nelson, and Cheng-Pin Kuo. How limitations in energy access, poverty, and socioeconomic disparities compromise health interventions for outbreaks in urban settings. *Iscience*, 24(12):103389, 2021. doi: https://doi.org/10.1016/j.isci.2021.103389.

[44] Xuefei Ren. Pandemic and lockdown: a territorial approach to covid-19 in china, italy and the united states. *Eurasian Geography and Economics*, 61 (4-5):423–434, 2020. doi: https://doi.org/10.1080/15387216.2020.1762103.

[45] Tuo Ji, Hai-Lian Chen, Jing Xu, Ling-Ning Wu, Jie-Jia Li, Kai Chen, and Gang Qin. Lockdown contained the spread of 2019 novel coronavirus disease in huangshi city, china: Early epidemiological findings. *Clinical Infectious Diseases*, 71(6):1454–1460, 2020. doi: https://doi.org/10.1093/cid/ciaa390.

[46] Annette M La Greca, Mitchell J Prinstein, and Michael D Fetter. Adolescent peer crowd affiliation: Linkages with health-risk behaviors and close friendships. *Journal of pediatric psychology*, 26(3):131–143, 2001. doi: https://doi.org/10.1093/jpepsy/26.3.131.