SFI TRANSMISSION

COMPLEXITY SCIENCE FOR COVID-19

STRATEGIC INSIGHT: Group size matters when it comes to how many people should gather in one place. Let's use mathematical models to pin down consistent guidelines for complicated situations.

FROM: John Harte, Professor of Ecosystem Sciences, University of California, Berkeley; External Professor, Santa Fe Institute

DATE: 30 March 2020 **NO:** 002.1

Beginning in early March 2020, conflicting advice about COVID-19 emanated from local, state, and federal leaders, as well as from public health spokespeople. While there was unanimous agreement that some level of social distancing was critical to reducing the daily incidence of new COVID-19 cases, there was wide disagreement as to what group size should be allowed. Through the media, one learned that group size should be limited to 200, or maybe to 50, or maybe to 20, or maybe to 10. On the same day, you could learn that you could attend a large lecture but not a sports event or political rally, or you could go to a bar or restaurant but not a concert hall, or you could attend a small dinner party with friends but not a restaurant. Eventually, many regions of the US settled on home confinement, which implies group sizes of at most a handful.

So, what is the effect of group size on transmission rates of infectious disease? This question raises many secondary questions. How long does one stay within a group — perhaps two hours at a ball game, but all day in kids' classrooms — and how does that interact with group size? How thoroughly within a group does transmission occur? Surely somebody in the bleachers cannot directly infect someone in a box seat above home plate. And what about whether the group is indoors or outdoors; what about wind and humidity?

Clearly, it's complicated. But, to get at least a very simple insight, we can make some very simple assumptions and obtain a back-of-the-envelope result. Assume that a population (for example, the people in a big city) is segregated into N_0 groups each with n_0 people. The total population of the city is $P_0 = n_0 N_0$. Assume further that initially there are y_0 infected individuals, randomly distributed among the population. Moreover, if one of those individuals happens to be in a particular group, then everyone in that group becomes infected. Finally, assume there is no mixing among groups. Both of those last two assumptions are readily altered, but let's work out this simplest case first.

If the number of groups with infection is denoted N, it is easy to show using the rules of probability that

$$N = N_0[1 - (1 - \frac{n_0}{N_0})^{y_0}]$$

As a result of transmission of disease within groups, the number of infected people will become n_0N, and hence the fraction of the population that becomes infected, P/P_0, is

$$\frac{P}{P_0} = \frac{n_0N_0}{n_0N_0}\left[1 - \left(1 - \frac{n_0}{N_0}\right)^{y_0}\right] = \left[1 - \left(1 - \frac{n_0}{N_0}\right)^{y_0}\right]$$

How does this fraction depend on group size, n_0? If, plausibly, n_0 is much smaller than N_0, and y_0n_0/N_0 is also considerably less than one, which is also very plausible, then the equation above leads to

$$\frac{P}{P_0} \approx \frac{y_0n_0}{N_0}$$

But, because the fixed population size, P_0, equals n_0N_0, we can re-express this as

$$\text{fraction of population that is infected} \approx \frac{y_0n_0^2}{P_0}$$

In other words, a doubling of allowable group size results in a four-fold increase in infection. Group size matters a lot!

If the inequalities above do not hold, with for example y_0n_0/N_0 no longer considerably smaller than 1, then the dependence on group size weakens. At that point, most of the population is infected.

Suppose, instead of assuming that everybody in a group gets infected if one of the initially infected people is in the group, we assume that the number of infected in a group increases as the square root of the group size (the bleachers are far from home plate). Then it is easy to show that the infected fraction varies as $n_0 3/2$.

Suppose we allow inter-group mixing. Then, depending on the rate of mixing, the infection rate, the duration of infectiousness in a person, and other factors having to do the spatial pattern of mixing such as distance over which one mixes, the fraction infected can become much larger, but the dependence on group size probably does not ever become steeper than the quadratic dependence derived above.

Clearly, more sophisticated modeling is needed here. It is imperative, as we emerge from strict social distancing some months from now, that we don't go straight from group sizes of two or three or four to unlimited group gatherings, lest we trigger a resurgence in infection.

So what's the magic number? There isn't a single answer. However, because group size matters a lot, the precautionary principle urges us to err on the side of small group-size restrictions. If mathematics informs our decisions, then as we eventually ramp up our sociality and return to some approximation of normality, we can do so with more clarity than was available when we went into quarantine.