CORONAVIRUS

Losing the battle over best-science guidance early in a crisis: COVID-19 and beyond

Lucia Illari[1], Nicholas J. Restrepo[2], Neil F. Johnson[1]*

Ensuring widespread public exposure to best-science guidance is crucial in any crisis, e.g., coronavirus disease 2019 (COVID-19), monkeypox, abortion misinformation, climate change, and beyond. We show how this battle got lost on Facebook very early during the COVID-19 pandemic and why the mainstream majority, including many parenting communities, had already moved closer to more extreme communities by the time vaccines arrived. Hidden heterogeneities in terms of who was talking and listening to whom explain why Facebook's own promotion of best-science guidance also appears to have missed key audience segments. A simple mathematical model reproduces the exposure dynamics at the system level. Our findings could be used to tailor guidance at scale while accounting for individual diversity and to help predict tipping point behavior and system-level responses to interventions in future crises.

INTRODUCTION

Managing crises (*1*) such as the coronavirus disease 2019 (COVID-19) pandemic (*2*, *3*), climate change (*4*), and, now, monkeypox and abortion misinformation requires widespread public exposure to, and acceptance of, guidance based on the best available science (*5–12*) [where guidance is defined in the Oxford Dictionary (*12*) as "advice or information aimed at resolving a problem or difficulty"]. However, distrust of such best-science guidance has reached dangerous levels (*7–9*). The American Physical Society, like many professional entities, is calling scientific misinformation one of the most important problems of our time (*8*, *9*). During COVID-19's 2020 prevaccine period of maximal uncertainty and social distancing, many people went to their online communities for guidance about how to avoid catching it and proposed cures. Social media saw a huge jump (*13*) in users in 2020 (13.2%), taking the total to 4.20 billion (53.6% of the global population), with the top reason for going online given as seeking information (*13*).

Unfortunately, many people were exposed to guidance that was not best science (*3*) from their online communities of likely well-meaning but non-expert friends. Some even died as a result of drinking bleach or rejecting masks (*14*). This raises the following urgent questions that we address here: Who emitted guidance, and to whom? Who received guidance, and from whom? What went wrong and when? What does this tell us about how, where, and when to intervene in current and future crises beyond COVID-19?

This paper attempts to tackle these questions by mapping out empirically, and analyzing quantitatively, the network of emitted and received COVID-19 guidance among online communities. The period of study ranges from the outset in December 2019 until August 2020, which is several months before the December 2020 temporary authorizations for emergency use of any COVID-19 mRNA vaccine in the United States or the United Kingdom (*15*). We focus on the emitter-receiver dynamics, i.e., the extent to which sets of communities acted as dominant sources (i.e., emitters) of COVID-19 guidance and/or were exposed to such guidance (i.e., receivers) from each other. We supplement this with a mathematical model that reproduces the system-level dynamics. Hence, our study complements, but differs from, the many excellent existing studies, which

include discussions surrounding particular topics and sources or target groups (*16–56*).

Data collection and classification

We use data collected from Facebook because it is the dominant social media platform worldwide with 3.0 billion active users and is the top social network in 156 countries (*57*). Moreover, recent studies have confirmed that people (e.g., parents) tend to rely on Facebook's built-in community structure for sharing guidance (*58–60*). Hence, we choose our main unit of analysis to be built-in Facebook communities, specifically Facebook pages. We refer to each Facebook page simply as a community, but we stress that it is unrelated to any ad hoc community structure inferred from network algorithms. Each page aggregates people around some common interest; it is publicly visible, and its analysis does not require us to access personal information. Our starting point is the ecosystem of such communities that were interlinked on Facebook around the vaccine health debate just before COVID-19 (November 2019; see Materials and Methods and the Supplementary Materials). A link from community (page) *i* to community (page) *j* exists when *i* recommends *j* to all its members at the page level (*i* likes/fans *j*) as opposed to a page member simply mentioning another page: As a result, members of *i* can, at any time *t*, be automatically exposed to fresh content from *j*, i.e., *j* emits and *i* receives (see section S1). Although not all *i* members will necessarily pay attention to such content from *j*, recent work (*61*) has shown experimentally and theoretically that a committed minority of only 25% is enough to tip an online community to an alternate stance.

We collect these Facebook pages and the links between them using the same snowball-like methodology as our earlier work (*62*), and we then classify these pages in the same way as our earlier work (*62*). Because we review this methodology and classification scheme (*62*) in the Materials and Methods (and section S1), we will only summarize it here. We begin with a seed of manually identified pages that discuss vaccines/vaccination in some manner, and then, we obtain these pages' links to other pages using a combination of computer scripts and human cross-checking. This process is iterated several times to produce a network of Facebook pages (nodes) and links between them. Our trained researchers then classify each page as pro, anti, or neutral on the basis of its recent content. "Pro" characterizes a page whose content actively promoted best-science

[1]Physics Department, George Washington University, Washington, D.C. 20052, USA. [2]ClustrX LLC, Washington, D.C. 20007, USA.
*Corresponding author. Email: neiljohnson@gwu.edu

health guidance (pro-vaccination); "anti" is a page whose content actively opposed this guidance (anti-vaccination), and "neutral" is a page that had community-level links with pro/anti communities pre-COVID-19 but whose content focused on other topics such as parenting (e.g., child education), pets, and organic food (see Materials and Methods and section S1). Each researcher independently classified each page manually on the basis of its content and then they checked for consensus. When there was disagreement, they discussed, and agreement was reached in all cases. They also further categorized the neutral communities according to their declared topic of interest and found 12 categories, such as parenting, organic food lovers, and pet lovers (see section S2 for full discussion and examples). While further subdivision is possible, it would lead to categories with too few communities and blurred boundaries.

This data collection and classification methodology provided us with a list of 1356 interlinked communities (Facebook pages) comprising 86.7 million individuals from across countries and languages, with 211 pro communities (Fig. 1, blue nodes) comprising 13.0 million individuals, 501 anti communities (Fig. 1, red nodes) comprising 7.5 million individuals, and 644 neutral communities (Fig. 1A, gold nodes) comprising 66.2 million individuals. We can estimate a size for each community by its number of likes (fans) because a typical user only likes 1 Facebook page on average (13): This size typically ranges from a few hundred to a few million users, but we stress that our analysis and conclusions do not rely on us determining community

sizes. The public information that we gather about each page's managers (63) suggests that users come from a wide variety of countries (see section S2 for details). The most frequent manager locations are the United States, Australia, Canada, the United Kingdom, Italy, and France (section S2).

Accompanying mathematical model

To accompany our empirical analysis, we introduce a simple mathematical model that can mimic the collective dynamics of these online communities. We are not suggesting that it represents the (as-yet unknown) best possible mathematical model, but it does have some advantageous features. First, it has a minimal and transparent form, which is so simple that its output and predictions can be verified manually using standard calculus. Second, it can be derived systematically from first principles (see the Supplementary Materials for derivation) based on the empirical fact that individuals aggregate or "gel" into communities online. These communities then aggregate into communities-of-communities of a given type. Relatedly and third, the model can be applied at various levels of aggregation: For example, taking some measure of the activity for all the neutrals at time t [i.e., $G(t)$] together with the corresponding activity for all the pros [$B(t)$] and antis [$R(t)$] generates three coupled equations as in Eqs. 1 and 2 below, or for example, neutral subcategories could be included separately, which would add another equation for each [$G_1(t)$, $G_2(t)$, etc.]. Fourth, even at its crudest level of approximation
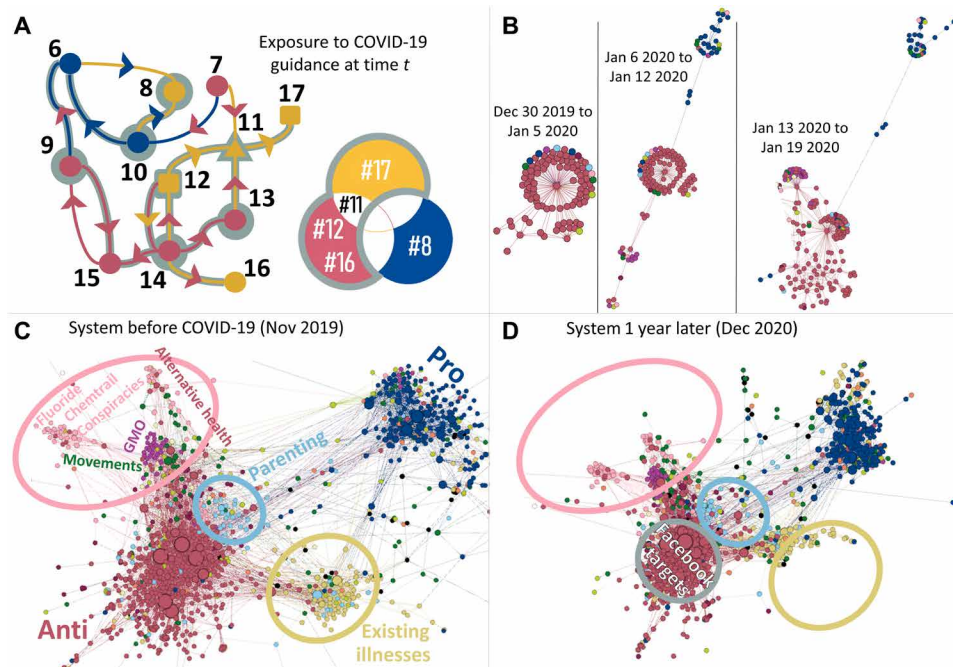
Fig. 1. Exposure dynamics. (**A**) Schematic illustrating emitter-receiver complexity. Each node is a community (Facebook page): Pro communities (blue) actively promote best-science guidance; antis (red) actively oppose it. Neutrals (gold) have a shape to denote their topic category (e.g., parenting). Link $i \rightarrow j$ means i "fans" j, which feeds content from page j to page i, exposing i's users to j's content. Link $i \rightarrow j$ color is that of node i; arrow color is node j, and arrow direction shows potential flow of COVID-19 guidance. Gray indicates appearance of COVID-19 guidance at time t. The Venn diagram shows the source of neutral communities' exposure to COVID-19 guidance at time t. (**B**) Early evolution of exposure to COVID-19 guidance. Non-red/non-blue nodes in (B to D) denote categories of neutral communities, e.g., parenting communities are turquoise (section S2 gives color scheme). Only links involving COVID-19 guidance during that time window are included [i.e., it is a filtered version of (A)]. (**C**) System pre-COVID-19, showing all potential links for exposure to COVID-19 guidance [unfiltered, as in (A)]. Layout is spontaneous (ForceAtlas2) with closer proximity indicating more mutual links. Node size indicates normalized betweenness centrality value of that node. (**D**) One year later, just before COVID-19 vaccine rollout. Nodes (pages) in gray ring were the main targets of Facebook's banners promoting best-science guidance (see section S3). Rings are in same position in (C) and (D) to show explicitly the increase in bonding. Both figures show only the largest component of the network. See the Supplementary Materials for the system 2 years later.

(Eq. 1), the model produces output curves that are similar to those observed empirically, and this agreement can be systematically improved using more sophisticated versions (e.g., Eq. 2). At the crudest level of approximation, our model's equations for the rate of change of $R(t)$, $B(t)$, and $G(t)$ (given by $\dot{R}$, $\dot{B}$, and $\dot{G}$, respectively) become

$$\dot{R} = r_R(R_0 - R) + r_B(B - R)$$

$$\dot{B} = b_B(B_0 - B)$$

$$\dot{G} = g_G(G_0 - G) + g_B(B - G) + g_R(R - G) \quad (1)$$

As is standard in ecosystem models, each of the three equations contains a self-interaction term [$g_G(G_0 - G)$, etc.] to account for each subpopulation's intrinsic growth or decay (G_0, etc., are constants). The coupling terms depend on the differences and, again, have a simple linear form. Positive (or negative) couplings imply positive (or negative) feedback, e.g., if $r_B > 0$ ($r_B < 0$), then having $B(t)$ exceed $R(t)$ will increase (decrease) the rate of change of $R(t)$ and hence increase (decrease) $R(t)$. Inspection of the online content supports the notion that (i) pro communities focus on emitting best-science guidance to the entire population, including the neutrals and antis, suggesting that the pros are not substantially influenced by the activity of the antis or the neutrals. Hence, the equation for $B(t)$ is not coupled to $R(t)$ or $G(t)$; (ii) the antis are influenced by the guidance emitted by the pros, in that they often turn it into their own versions (including misinformation) and then feed it to the neutrals to raise the neutrals' concern about best-science guidance. This suggests that the antis are not substantially influenced by the narratives of the neutrals. Hence, the equation for $R(t)$ is only coupled to $B(t)$; (iii) the neutrals are influenced by the guidance that they receive from the pros and the antis. Hence, the equation for $G(t)$ is coupled to both $B(t)$ and $R(t)$. These self-interaction and coupling terms are shown schematically in Fig. 2.

Equation 2 below provides a more sophisticated version of Eq. 1, in which we add back into Eq. 1 the feature from the full mathematical derivation whereby each gel (community or set of communities depending on the level of aggregation) has its own onset time t_c (see derivation in the Supplementary Materials); we also add decay terms to mimic loss of interest or moderator crackdown

$$\dot{R} = H(t - t_{c,R}) \left[r_R(R_0 - R) + r_B(B - R) - d_R R \right]$$

$$\dot{B} = H(t - t_{c,B}) \left[b_B(B_0 - B) - d_B B \right]$$

$$\dot{G} = H(t - t_{c,G}) \left[g_G(G_0 - G) + g_B(B - G) + g_R(R - G) - d_G G \right] \quad (2)$$

where each $H(\dots)$ is a Heaviside function that becomes nonzero at the respective onset time t_c. We use Eq. 2 to compare to the empirical data in Fig. 3 because it provides a better goodness of fit. However, as demonstrated in section S5, the inclusion of these onset and decay terms is not essential: Both Eqs. 1 and 2 produce similar shapes to the empirical curves because they both have the same core structure of coupling terms and self-interactions shown schematically in Fig. 2. Although the actual spread of information and rumors, like diseases, is stochastic, it is well known that such deterministic equations
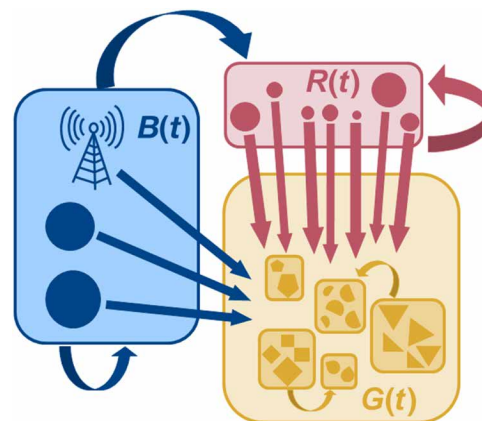
Fig. 2. Schematic of our model. This same schematic applies to both the crudest level approximation of our model (Eq. 1) and the more sophisticated version (Eq. 2). In Eqs. 1 and 2, pros (blue), antis (red), and neutrals (gold) have been aggregated over all communities (i.e., Facebook pages), as represented here by the shapes, with neutrals being further aggregated over all 12 categories as represented by the boxes.

can describe the behavior in time averaged over many such stochastic realizations. We have checked using stochastic simulations that our equations are similarly accurate. Moreover, parameter estimation and optimization can be difficult to perform with stochastic models (64, 65). We also investigated the effects of noise on our data by randomly deleting up to 15% of COVID-19–related links from the entire network to mimic links being missed or simply not existing (see the Supplementary Materials), and we found our main results and conclusions to be robust.

RESULTS
Complexity of the online exposure problem
Figure 1A shows that even with our highly simplified node and link classification scheme, the exposure dynamics at any single snapshot in time t can still be very complex, even for a small subset of nodes. The color of a link from node (i.e., Facebook page) i to node (Facebook page) j is that of node i, while the arrow's direction indicates potential flow of COVID-19 guidance, which is therefore j to i. The arrow's color is that of node j. If a node j posts COVID-19 guidance at time t, then we put a gray border around it and around any arrows emanating from it to indicate exposure of the linked nodes to j's COVID-19 guidance at time t. If there are no links going into node j, then it is not exposing any other node to its COVID-19 guidance. Each page (node) could link to various other pages, but irrelevant links get filtered out as explained in our earlier work (62) and section S1, yielding a network with a few links per node. The Venn diagram thus shows the COVID-19 guidance exposure of the neutral communities at time t (gold nodes, where each shape represents a separate topic). The gray border in the Venn diagram shows the neutral nodes at time t that are exposed to COVID-19 guidance that comes entirely from non-pro communities at time t, i.e., the COVID-19 guidance that they receive comes entirely from antis and/or other neutrals. For example, 12 is only exposed to COVID-19 guidance from anti node 14; hence, 12 is in the anti-only gray-border region. Twelve has a link into it from 11, so 11 is exposed to COVID-19 guidance from 12, but this does not affect 12 itself.
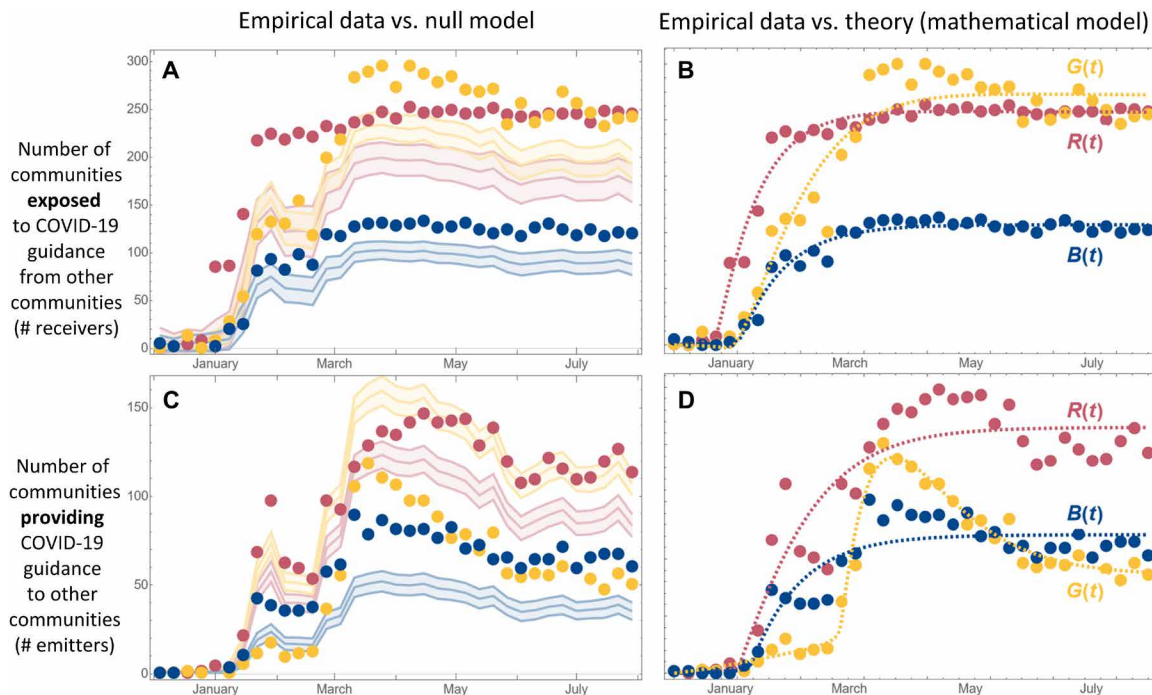
Fig. 3. Data versus null model and our model (Eq. 2). (**A**) Empirical data (circles) show number of pro (blue), anti (red), and neutral (gold) communities exposed to COVID-19 guidance (i.e., receivers). Lines show range of outputs from the null model, which provides a poor fit to the data. (**C**) Similar to (A) but circles show number of communities providing COVID-19 guidance (i.e., emitters). (**B**) and (**D**) compare the empirical data to our generative mathematical model Eq. 2 (dotted lines). The full dataset was used to estimate the model parameters. See section S6 and software files for full replication and comparison to parameter estimates using k-fold cross-validation.

Given the complexity already present in Fig. 1A, we do not further complicate this paper by attempting to assign a numerical value to the fraction of scientific truth in each piece of COVID-19 guidance. Such a number would, in any case, be unreliable because even the wildest anti content can contain truthful fragments. For example, the false story that semiconductor chips are being injected with COVID-19 vaccines is actually associated with a true piece of science: A 2019 peer-reviewed publication in a top scientific journal showed empirically that nanoscale semiconductor structures (quantum dots) can be used as injectable vaccine markers (66, 67). Hence, the falseness lies solely in the fact that they are not being used in this way, not that they scientifically cannot. Given this, we adopt a simpler approach: Our team's experience from analyzing all these communities' content on a daily basis shows that COVID-19 content emitted by the pros promotes best science, as expected, and that emitted by the antis opposes it. An initial post in a neutral community may sit between these two extremes but typically gets further downgraded by nonscientific comments and replies from its nonexpert page members and hence does not end up as definitive best-science guidance. This means that we can reasonably reserve the label "best-science guidance" for the guidance that comes from the pros. While this could be refined in the future, we note that even if a fraction of our classifications of communities and content is wrong, our main conclusions are unchanged because they only depend on relative numbers. We have checked the robustness of our results explicitly by simulating errors into our classifications. We randomly selected 1 to 15% of COVID-19 guidance links from the entire network to be deleted in 1% increments. On average, deleting 15% of the links only produced a 5% percent difference in the magnitude of the emitter-receiver curves, with a maximum 25% percent difference, and the general curve shapes were preserved. Furthermore, although some estimates of parameter values can vary substantially, the general curve shapes were preserved (see section S5 for details). Hence, any conclusions reliant on the model are robust to fluctuations because of noise and to large variations in parameter value estimates. We recognize that best-science guidance can change over time and may eventually be proven wrong, but that seems to be a rare occurrence.

Empirical features of the exposure dynamics

We now present our empirical findings for the observed exposure dynamics at various levels of aggregation. The data analyzed were collected from December 2019 to August 2020. Figure 1B shows how the initial conversations over COVID-19 guidance began primarily among the anti communities and well before the official announcement of the pandemic (11 March 2020). It is a filtered version of the construction in Fig. 1A: A link only appears in Fig. 1B when one of the nodes (communities) that it connects presents COVID-19 guidance in that time interval. It shows the largest connected component. Because we use the ForceAtlas2 layout algorithm, the observed segregation is self-organized, and proximity indicates stronger mutual links, i.e., the more links node i and its neighbors have with node j and its neighbors, the closer visually node i will be to node j (see section S4). It reveals how quickly anti communities (red nodes) influence the system, with neutrals (i.e., nodes that are not red nor darker blue, e.g., parenting communities are pale blue) also getting picked up or attaching themselves. Pro communities (darker blue) enter later on and form their own sphere. This pro-anti segregation suggests that the system strengthening observed when moving from Fig. 1C to Fig. 1D derives from this early 2020 bonding around COVID-19 guidance shown in Fig. 1B. In Fig. 1 (C and D), only the

largest component of the network is shown: It includes 91.96% of all nodes and 99.93% of all edges in the system in Fig. 2C and 87.24% of all nodes and 99.94% of all edges in the system in Fig. 4D.

Figure 1D provides a full system–level view just before COVID-19 vaccine rollout of all the links along which COVID-19 guidance can flow (panels C and D of Fig. 1 are equivalent to Fig. 1A, ignoring the yellow shading, and hence akin to a road network irrespective of the traffic, while Fig. 1B is the subset of roads carrying traffic). The observable changes from Fig. 1C to Fig. 1D indicate that not only did the subpopulation of anti communities tighten internally during the year-long period of maximal societal uncertainty before vaccines appeared, the neutrals were pulled and/or pulled themselves closer to the antis, and neutral categories such as the parenting communities (pale blue nodes) also tightened internally. The postvaccine version is visually similar to the network in Fig. 1D (see section S2).

This has the key consequence that by the time buy-in to the COVID-19 vaccines was becoming essential (i.e., December 2020), many parents who were responsible for health decisions about themselves, their young children, and also likely elderly relatives had become even closer in the network to the antis who held extreme views including distrust of vaccines and rejection of masks, as well as to other neutrals (see nodes within the pink ring) who were focused on nonvaccine and non–COVID-19 conspiracy content surrounding climate change, 5G, fluoride, chemtrails, GMO foods, and also alternative health communities that believe in natural cures for all illnesses (see section S2 for details). This increased closeness to more extreme communities was potentially very important for public health because proximity in the ForceAtlas2 network layout indicates stronger mutual links (see section S4). Hence, the closer those nodes appear spatially in the network, the more likely they are to share content and hence actually exert influence. In this case, this means a likely increased influence of extreme communities on these mainstream communities including parents.

Facebook conducted its own top-down promotion of best-science COVID-19 guidance by placing banners at the top of some pages (i.e., nodes) pointing to the U.S. Centers for Disease Control and Prevention for example (see section S3). However, our mapping in Fig. 1D shows explicitly that these banners appear primarily in anti communities (red nodes) and, moreover, that the antis that were targeted were primarily within the gray oval in Fig. 1D (section S3). Hence, many neutral communities were missed, yet this might have been avoidable using these maps.

We now take a closer look at what influence antis might have had on the neutral category parenting communities and compare that to the system-wide impact that antis might have had. The Venn diagram in Fig. 4A quantifies the extent to which non-pro communities acted as the dominant sources (i.e., emitters) of COVID-19 guidance to neutral communities during the period of maximal
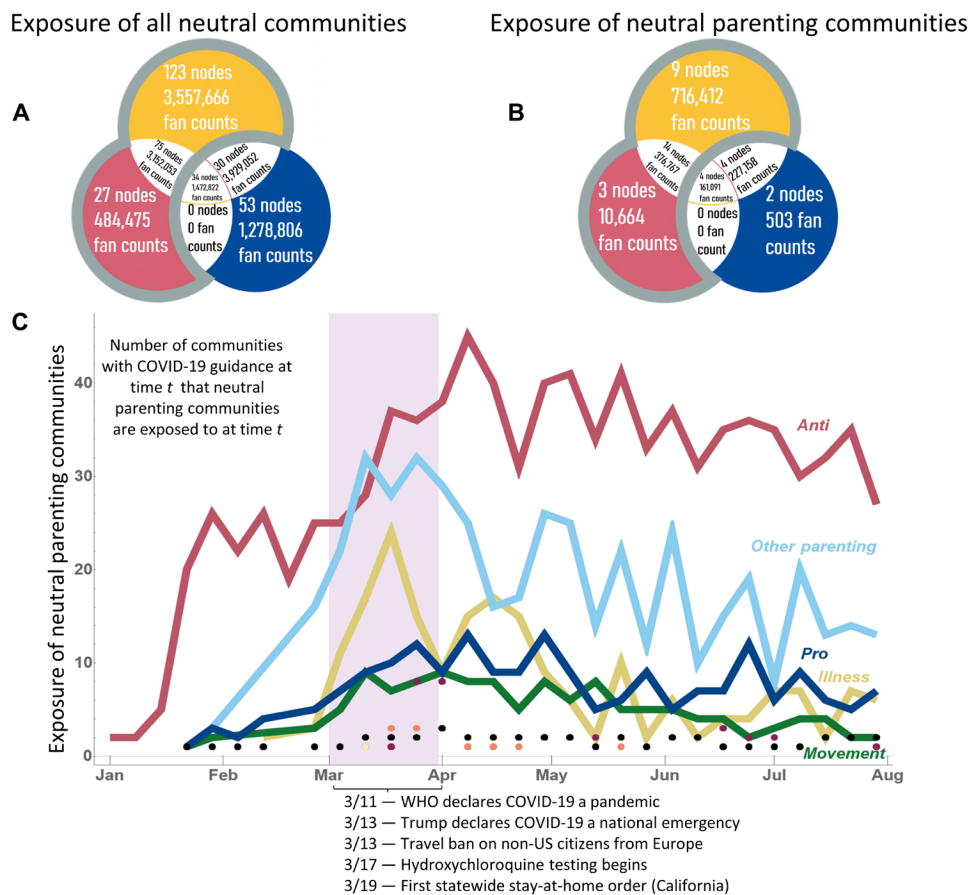
Fig. 4. Exposure of neutral communities to non-pro guidance. (**A**) The Venn diagram, as in Fig. 1A, shows sources of exposure to COVID-19 guidance for all neutral communities in the giant connected component of the system (Fig. 1, C and D) aggregated over January to August 2020. (**B**) Similar to (A) but just for the parenting community subset of all neutral communities. (**C**) Neutral parenting communities' exposure to COVID-19 guidance disaggregated over time and by source (i.e., emitter) type.

societal uncertainty before vaccine discovery. While 7.19 million individuals were exposed exclusively to COVID-19 guidance from non-pro communities, only 1.28 million were exposed exclusively to COVID-19 guidance from pro communities. The remaining 5.40 million were exposed to both, which may still have made them quite uncertain about what to think. Figure 4B reveals that this imbalance was even worse for individuals in parenting communities: A total of 1.10 million of these individuals were exposed exclusively to COVID-19 guidance from non-pro communities, but only 503 were exposed exclusively to COVID-19 guidance from pro communities.

Disaggregating this further, Fig. 4C shows the neutral parenting communities' exposure to COVID-19 guidance over time from the different types of communities. Starting in early January, the anti communities quickly generated COVID-19 guidance, which, when combined with the substantial number of links to them from parenting communities, generated the rapid rise in parenting communities' exposure from anti communities shown in Fig. 4C. This is followed by a rapid rise of exposure to guidance from other parenting communities, well before the official declaration of a pandemic, and a smaller rise in exposure from communities focused on preexisting, non-COVID-19 illnesses such as Asperger's syndrome and cancer. These high levels of exposure from anti and other parenting communities persisted for the entire period. In stark contrast, exposure from the pro communities never showed any strong response and remained low. Section S5 shows that these curves in Fig. 4C are statistically significant as compared to a null model in which a random network is chosen. This means that we can reject the hypothesis that the microstructure of the exposure network is not relevant. In short, the complexity of the exposure network (Fig. 1) is indeed key to understanding the exposure dynamics over time.

These findings paint the following picture of the prevaccine period: Individuals in the neutral parenting and other mainstream communities became aware of COVID-19 guidance from anti communities early in January 2020, which they then quietly deliberated over, perhaps interacting in private groups or apps such as WhatsApp or with others offline. By mid-February, they felt in a position to produce and share their own COVID-19 guidance with communities like theirs. Meanwhile, they only received minimal best-science guidance from the pro communities (dark blue curve is near zero). They did not have any strong tendency to create additional links to other pro communities, probably because they were already receiving guidance from other neutral communities that had similar interests (e.g., parenting) who they felt they could identify with and perhaps even trust more.

These findings also suggest a missed opportunity for intervention that arose very early in 2020. While the possibility of providing more direct messaging against antis during their January 2020 rise as guidance emitters (red curve, Fig. 4C) may not have been desirable given their active opposition and possible backlash, the observed February 2020 rise of other parenting communities as guidance emitters (pale blue curve) suggests that best-science COVID-19 guidance from the pros could have instead been tailored around popular topics within parenting communities at that time (which could have been read from their pages) and hence introduced at scale using the map in Fig. 1B.

Modeling the exposure dynamics using Eq. 2

What-if questions also arise about other types of potential interventions. For example, might a blanket intervention across all neutral categories have reduced the subsequent high peaks in exposure to COVID-19 guidance from non-pro communities and their subsequent persistence throughout 2020 (Fig. 4C)? The true impact of any intervention must ultimately be tested empirically. However, comparative discussions could benefit from an accompanying mathematical equation that reproduces the aggregate scale exposure dynamics and is also so transparent that it gives simple insight into such "what if" scenarios.

Equation 2 represents such an equation at the aggregate level of all neutrals, pros, and antis, as does its more approximate version given by Eq. 1. Figure 3 (C and D) confirms the suitability of Eq. 2 in terms of providing good agreement with the empirically observed exposure dynamics at this aggregate level. The more approximate version, Eq. 1, also yields acceptable agreement but, as expected, has lower goodness-of-fit statistics, so it is not shown. The full dataset was used to estimate the mathematical model parameters due to the number of points per curve; the use of k-fold cross-validation and holding out a validation set in estimating the model parameters is discussed in section S5. By contrast, the null model results in Fig. 3 (A and C) show what happens to Fig. 3 (B and D) if a randomly shuffled version of the empirical data is used. The observed poor agreement of this null model is notable not just because its predictions lie far from the empirical data but also because the null model's construction provides a rather demanding comparison: Instead of randomizing (shuffling) all the nodes across all types [which indeed produces curves that are very different from those in Fig. 3 (B and D)], we only randomize (shuffle) within types, i.e., we separately shuffle within the antis, within the pros, and within each of the 12 neutral subcategories. Hence, this null model contains exactly the same numbers of nodes within each subcategory as the empirical network: Therefore, the networks look visually the same for the null model as the real one because the colors of the categories are retained, but the node names are shuffled. Repeating this 1000 times yields the bands shown in Fig. 3 (A and C), which represent the mean and 1 SD. These bands are far from the empirical data (see the Supplementary Materials and mathematical and data replication files for full details). This shows the importance of the actual links and the full network in determining the online exposure dynamics, i.e., the observed dynamics are not a simple consequence of the number of nodes of each type. It also suggests that Eq. 2 is capturing real node-link characteristics of the empirical network, as opposed to simply reflecting relative subpopulation sizes and, hence, confirms the importance of understanding the real network when addressing questions about online exposure.

We can now use this mathematical equation to explore what-if interventions. Because we are only exploring qualitative outcomes, we adopt the cruder version of Eq. 2 (i.e., Eq. 1) because its behavior is exactly solvable and understandable without any need for a computer. Figure 5 (A to D) shows the predictions of future behaviors of Eq. 1 using a crude estimate of current $R(t)$, $B(t)$, and $G(t)$ values as the initial conditions and different coupling scenarios (see section S6 and figs. S13 to S15 for details). On the basis of the current persistence of hesitancy about vaccines and mask wearing, we assume that the pros have currently reached their maximum capability in terms of promoting best-science guidance; hence, $B(t)$ remains constant. Figure 5A shows what then happens when the future couplings between anti, pro, and neutral are all positive (i.e., positive feedback): $G(t)$ initially peaks before settling at a higher value. In Fig. 5B, neutral and pro have negative coupling (i.e., negative feedback): This causes $G(t)$ to escalate markedly. In Fig. 5C, all couplings are negative: $G(t)$ drops markedly. In Fig. 5D, neutrals and antis have the only positive
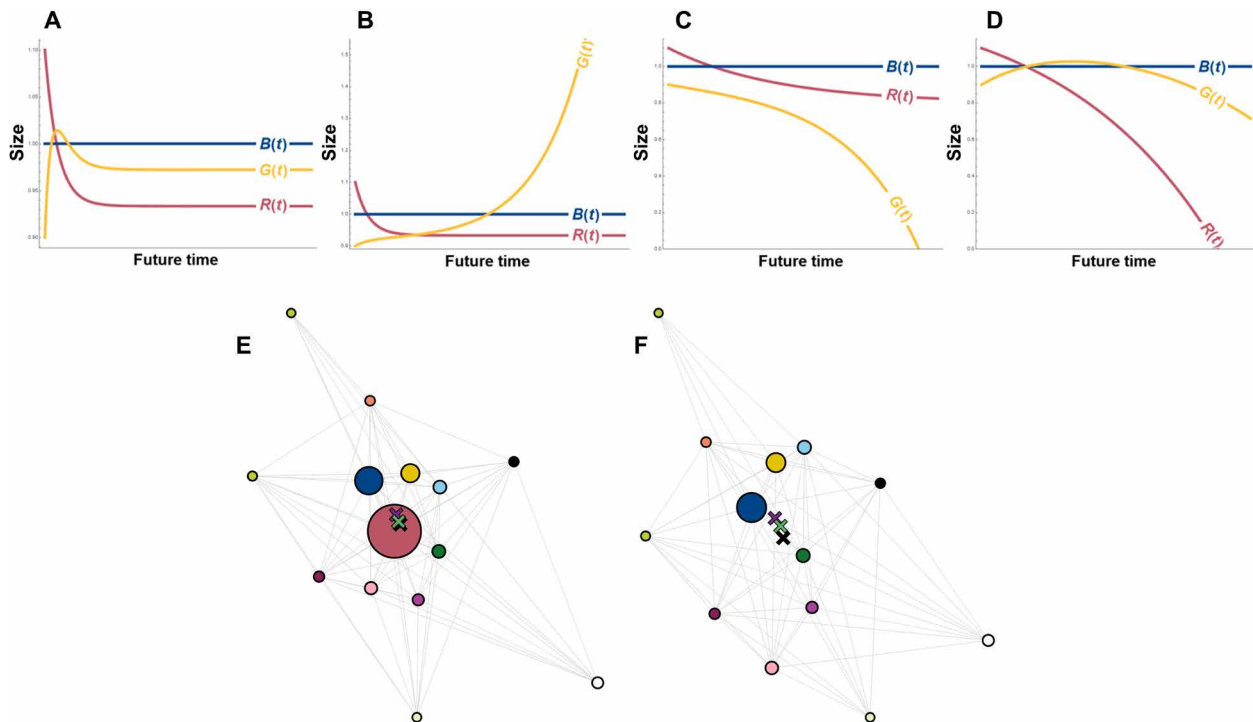
Fig. 5. Prediction of interventions on current system. (**A** to **D**) The four classes of future outcome predicted by the crudest form of our model (Eq. 1). Initial conditions crudely mimic the current situation (section S6 and figs. S13 to S15 show details and code). (A) All coupling terms are positive. (B) The coupling term between $G(t)$ and $B(t)$ is negative. (C) All coupling terms are negative. (D) The only positive coupling term is between $G(t)$ and $R(t)$. (**E**) A renormalized version of Fig. 1D in which nodes of a given type are aggregated into a single supernode with the corresponding weighted size/mass. The center of this online universe is shown using various definitions: spatial center (black "x" mark), center weighted by degree (purple x mark), and center weighted by number of clusters (green x mark). (**F**) The impact on (E) of removing the anti (red) supernode: The pro (blue) supernode still does not sit at the center of the new universe.

coupling: $R(t) \rightarrow 0$, but $G(t)$ remains high for an extended period of time. These different predicted futures, hence, provide a framework for comparing the advantages and disadvantages of different possible interventions.

To close our loop of analysis, these mathematical predictions can then be tied back to the network picture from Fig. 1 using the physics technique of renormalization in which the communities of the anti, pro, and the 12 neutral subcategories are each aggregated into their own community-of-communities "ball" (see Fig. 5E). Figure 5F then shows the impact of removing the antis from Fig. 5E, hence, mimicking Fig. 5D in which $R(t) \rightarrow 0$. The comparison is only valid at short times because we are not allowing the network as a whole to adapt or rewire after cutting all anti links. Figure 5F shows that the pro communities will still not sit at the center of this online universe because of the many-sided interactions with the neutral subcategory communities, particularly the movement communities (dark green ball). The fact that both neutrals and pros remain in play in Fig. 5F is broadly consistent with Fig. 5D in which $G(t)$ remains high and comparable to $B(t)$ for an extended period despite $R(t) \rightarrow 0$.

DISCUSSION
Our findings show that the anti communities jumped in to dominate the conversation well before the official announcement of the COVID-19 pandemic and that neutral communities (e.g., parenting) subsequently moved even closer to extreme communities and hence

became highly exposed to their content. Parenting communities first received COVID-19 guidance from anti communities as early as January 2020. This continued up to and beyond the official pandemic announcement, after which parenting communities felt confident enough to begin adding their own guidance to the conversation. Guidance from pro communities remained low throughout, which is consistent with parenting communities seeking other sources.

To complement our empirical analysis, we developed a simple, generative mathematical model that captures, at the system level, the interplay between the sets of communities that emit and/or receive guidance. It allows for easy exploration of what-if scenarios and, hence, for crude prediction of tipping point behavior responses to different intervention strategies. The combination of network mapping and model shows that there are more possible approaches to flipping the conversation than merely just removing all extreme elements from the system. The results in Fig. 5 (D and F) show that removing all extreme elements may not even be the most appropriate solution. Such removals can, in any case, be perceived as harsh; they run counter to the idea of open participation, and they can compromise the business model of maximizing user numbers. Figure 5 (D and F) shows the impact of removing antis on the pro and neutral communities. In Fig. 5F, we see that pros do not sit at the center of the system (i.e., the different measurements for system center are not contained near or within the pro supernode), and the system center is between pros and a neutral community (movement). In Fig. 5D, we see that while $R(t) \rightarrow 0$, $G(t) \rightarrow 0$ is likewise

occurring, which implies that the removal of anti-content may spur the self-removal of neutral communities, which suggests that there might be less opportunity for these self-removed neutral communities to be exposed to best-science COVID-19 guidance. Pro communities increasing their connections to other communities would have a positive impact on Fig. 5 (E and F) in pulling the other supernodes toward pros and shift the system center close to pro communities. Decreasing the impact of anti communities by removing their ability to outreach to other communities would also aid in moving the system center away from antis in Fig. 5E. Because our model can be interpreted at different scales including communities-of-communities, it can be applied across multiple platforms that have built-in community features and can be used to tackle the question of online misinformation more generally, beyond COVID-19 and vaccinations.

Of course, our study has limitations, which suggests opportunities for further studies. We were limited to pages in languages that our researchers could read; hence, we missed out on additional insights from languages such as Mandarin, Hindi, and Arabic. In addition, while Facebook may be the leading social network worldwide (57), its users may not be representative of a given country's population. It is an open question how well our results generalize to those with a low percentage use of the internet. For example, in Pakistan and Belize, Facebook is the leading social network (57); however, only 17 and 47%, respectively, of individuals use the internet (68). Furthermore, the exposure dynamics may be influenced by small groups of so-called chaos agents used by organizations or governments (61). However, we note that social media communities tend to self-police for troll-like behavior. A further limitation is that there are many other social media platforms on which such debates are held. Individuals may be taking the COVID-19 guidance that they see on Facebook and discussing it on any number of other social media websites. It is our belief, however, that similar behaviors will arise on any social media platform on which communities are able to develop, and Facebook is indeed the largest of these.

MATERIALS AND METHODS
Details of the data collection and categorization
The methodology here follows our earlier works (62). Facebook pages comprise the nodes in our data, and each link represents the occurrence of a page recommending another page to its members. This avoids needing to identify personal account information, which is forbidden by Facebook's public API terms of service. The process begins with a seed of manually identified pages that discuss vaccines/vaccination in some manner, and then, these pages' connections to other pages are indexed. These pages were identified by searching through Facebook pages in 2018 and 2019 using key words and phrases involving vaccines. The findings were vetted via human coding and computer-assisted filters, and then, at least two different researchers classified each node independently. When there was disagreement, they discussed, and agreement was reached in all cases. This process was repeated two more times to obtain the final list of candidate nodes and the links between them. To classify a page, the page's posts, the "About" section, and self-described category were reviewed. To be classified as either pro or anti, at least 2 of the most recent 25 posts had to deal with the vaccination debate or the page's title or About section self-identified the page as pro or anti vaccination. To be classified as neutral, 5 of the most recent 25

posts had to refer to the vaccination debate, but the page had not explicitly taken a pro or anti stance, or the About section explicitly declared the page neutral in the debate, or none of the 25 most recent pages dealt with vaccines but the page self-identified as an nongovernmental organization, a cause, a community, or a grassroots organization. Hence, our dataset only contains Facebook users. Our target population includes not only those solely dedicated to posting about vaccines but also nonprofit organizations, public figures, government organizations, medical companies, local businesses, etc. Of course, we could have defined nodes and links differently, and our dataset is ultimately an imperfect sample of some larger "correct" network. To help mitigate this, we repeated the process of manually identifying an initial seed of pages several times, with the goal of making that seed as diverse as possible by including pages posting in different languages, pages focused on different geographical locations, and pages with managers from a wide range of countries (section S2). Only those pages whose posts were in languages the researchers could read were included, e.g., English, French, Spanish, Italian, Dutch, and Russian. Determining whether a post is sarcastic or ironic, fake, or troll-like (61) is a very difficult task even for subject matter experts; it is also now possible for free, completely off-the-shelf machine learning language models to generate realistic vaccine misinformation (69). However, these social media communities tend to self-police for bot or troll-like behavior. The extremely difficult task of quantifying the realism and intent of these posts is left as the subject of another study.

For the purposes of determining who was emitting and who was receiving COVID-19 guidance, we first had to determine which posts were explicitly discussing it. For every post, the post message, description, image text, and link text were combined into one string in which cases were ignored. We then produced a list of strings to search through these posts, e.g., "corona virus," "covid," "19 ncov," and other terms. Because we wanted these terms to be as flexible as possible, we used regular expressions, so terms such as "corona virus" became "(c|k|[()]+orona(no|[[:punct:]]|\\s){,4}(virus\\>|vírus)," which catches common misspellings and punctuation marks, as well as non-English languages ("vírus" is used in Portuguese, and virus is used not only in English but also in Italian, Spanish, French, etc.). Thus, our approach captured deliberate misspellings, ignored punctuation, dealt with added spaces in between the words to avoid filters, and covered other languages in addition to English. We then selected out those posts that had been determined, via the use of filters, to be COVID-19-related, and we used that information to determine which pages were emitting COVID-19 guidance at some time t. With this information, we were able to produce the filtered construction of the network (e.g., Fig. 1A) where a link was only present at a given time if one of the nodes that it connects to was producing COVID-19 guidance at that time.

SUPPLEMENTARY MATERIALS

REFERENCES AND NOTES
1. R. Brown, Counteracting dangerous narratives in the time of covid-19, (Over Zero, 2022); www.projectoverzero.org/media-and-publications/counteracting-dangerous-narratives.
2. J. Zarocostas, How to fight an infodemic. *Lancet* **395**, 676 (2020).
3. H. J. Larson, Blocking information on COVID-19 can fuel the spread of misinformation. *Nature* **580**, 306–306 (2020).

4. S. van der Linden, A. Leiserowitz, S. Rosenthal, E. Maibach, Inoculating the public against misinformation about climate change. *Glob. Chall.* **1**, 1600008 (2017).

5. S. Gavrilets, Collective action and the collaborative brain. *J. R. Soc. Interface* **12**, 20141067 (2015).

6. M. J. Gelfand, J. R. Harrington, J. C. Jackson, The strength of social norms across human groups. *Perspect. Psychol. Sci.* **12**, 800–809 (2017).

7. B. Nogrady, 'I hope you die': How the covid pandemic unleashed attacks on scientists. *Nature* **598**, 250–253 (2021).

8. Trust science pledge calls for public to engage in scientific literacy, (News Direct, 2021); https://newsdirect.com/news/trust-science-pledge-calls-for-public-to-engage-in-scientific-literacy-737528151.

9. F. Slakey, American Physical Society (APS) message to members, (American Physical Society, 2021); https://aps.org/about/support/index.cfm.

10. P. Ball, Hot air by Peter Stott Review—The battle against climate change denial, (The Guardian, 2021); www.theguardian.com/books/2021/oct/09/hot-air-by-peter-stott-review-the-battle-against-climate-change-denial.

11. P. Stott, *Hot Air: The Inside story of the Battle Against Climate Change Denial* (Atlantic Books, 2021).

12. "guidance". Oxford Dictionaries. Oxford University Press. https://premium.oxforddictionaries.com/us/definition/american_english/guidance (accessed via Oxford Dictionaries Online on September 19, 2022).

13. S. Kemp, Hootsuite 2021, (Hootsuite, 2021); www.hootsuite.com/resources/digital-trends.

14. B. Hope, M. Ozdaglar, Parent assaults teacher over mask dispute at Amador County School, Superintendent says, (KCRA, 2021); www.kcra.com/article/sutter-creek-parent-assaults-teacher-over-mask-dispute-at-amador-county-school-superintendent/37295267?et_rid=934107578&s_campaign=fastforward%3Anewsletter.

15. Pfizer and BioNTech achieve first authorization in the world for a vaccine to combat COVID-19, (Pfizer, 2020); www.pfizer.com/news/press-release/press-release-detail/pfizer-and-biontech-achieve-first-authorization-world.

16. N. Calleja, A. AbdAllah, N. Abad, N. Ahmed, D. Albarracin, E. Altieri, J. N. Anoko, R. Arcos, A. A. Azlan, J. Bayer, A. Bechmann, S. Bezbaruah, S. C. Briand, I. Brooks, L. M. Bucci, S. Burzo, C. Czerniak, M. De Domenico, A. G. Dunn, U. K. H. Ecker, L. Espinosa, C. Francois, K. Gradon, A. Gruzd, B. S. Gülgün, R. Haydarov, C. Hurley, S. I. Astuti, A. Ishizumi, N. Johnson, D. J. Restrepo, M. Kajimoto, A. Koyuncu, S. Kulkarni, J. Lamichhane, R. Lewis, A. Mahajan, A. Mandil, E. McAweeney, M. Messer, W. Moy, P. N. Ngamala, T. Nguyen, M. Nunn, S. B. Omer, C. Pagliari, P. Patel, L. Phuong, D. Prybylski, A. Rashidian, E. Rempel, S. Rubinelli, P. Sacco, A. Schneider, K. Shu, M. Smith, H. Sufehmi, V. Tangcharoensathien, R. Terry, N. Thacker, T. Trewinnard, S. Turner, H. Tworek, S. Uakkas, E. Vraga, C. Wardle, H. Wasserman, E. Wilhelm, A. Würz, B. Yau, L. Zhou, T. D. Purnat, A public health research agenda for managing infodemics: Methods and results of the first who infodemiology conference. *JMIR Infodemiology* **1**, e30979 (2021).

17. P. S. Dodds, K. D. Harris, I. M. Kloumann, C. A. Bliss, C. M. Danforth, Temporal patterns of happiness and information in a global social network: Hedonometrics and Twitter. *PLOS ONE* **6**, e26752 (2011).

18. J. P. Onnela, J. Saramäki, J. Hyvönen, G. Szabó, D. Lazer, K. Kaski, J. Kertész, A. L. Barabási, Structure and tie strengths in mobile communication networks. *Proc. Natl. Acad. Sci. U.S.A.* **104**, 7332–7336 (2007).

19. A. Bessi, M. Coletto, G. A. Davidescu, A. Scala, G. Caldarelli, W. Quattrociocchi, Science vs conspiracy: Collective narratives in the age of misinformation. *PLOS ONE* **10**, e0118093 (2015).

20. D. Lazer, M. A. Baum, Y. Benkler, A. J. Berinsky, K. M. Greenhill, F. Menczer, M. J. Metzger, B. Nyhan, G. Pennycook, D. Rothschild, M. Schudson, S. A. Sloman, C. R. Sunstein, E. A. Thorson, D. J. Watts, J. L. Zittrain, The science of fake news. *Science* **359**, 1094–1096 (2018).

21. E. Chen, K. Lerman, E. Ferrara, Tracking social media discourse about the COVID-19 pandemic: Development of a public coronavirus Twitter data set. *JMIR Public Health Surveill.* **6**, e19273 (2020).

22. S. Havlin, D. Y. Kenett, A. Bashan, J. Gao, H. E. Stanley, Vulnerability of network of networks. *Eur. Phys. J. Spec. Top.* **223**, 2087–2106 (2014).

23. B. L. Hoffman, E. M. Felter, K. H. Chu, A. Shensa, C. Hermann, T. Wolynn, D. Williams, B. A. Primack, It's not all about autism: The emerging landscape of anti-vaccination sentiment on Facebook. *Vaccine* **37**, 2216–2223 (2019).

24. N. Smith, T. Graham, Mapping the anti-vaccination movement on Facebook. *Inf. Commun. Soc.* **22**, 1310–1327 (2019).

25. K. Starbird, E. S. Spiro, K. Koltai, Misinformation, crisis, and public health—Reviewing the literature. *Soc. Sci. Res. Council* 10.35650/md.2063.d.2020 (2020).

26. C. Miller-Idriss, *Hate in the Homeland: The New Global Far Right* (Princeton Univ. Press, 2020).

27. R. DiResta, Of virality and viruses: The ANTI-VACCINE movement and Social Media: Nautilus Institute for Security and Sustainability, *Nautilus Institute for Security and Sustainability | We hold that it is possible to build peace, create security, and restore sustainability for all people in our time* (2018); https://nautilus.org/napsnet/napsnet-special-reports/of-virality-and-viruses-the-anti-vaccine-movement-and-social-media/.

28. K. Starbird, Disinformation's spread: Bots, trolls and all of us. *Nature* **571**, 449–449 (2019).

29. A. Bechmann, Tackling disinformation and Infodemics demands media policy changes. *Digit. Journal.* **8**, 855–863 (2020).

30. Y. Theocharis, A. Cardenal, S. Jin, T. Aalberg, D. N. Hopmann, J. Strömbäck, L. Castro, F. Esser, P. Van Aelst, C. de Vreese, N. Corbu, K. Koc-Michalska, J. Matthes, C. Schemer, T. Sheafer, S. Splendore, J. Stanyer, A. Stępińska, V. Štětka, Does the platform matter? Social media and COVID-19 conspiracy theory beliefs in 17 countries. *N. Med. Soc.* 146144482110456 (2021).

31. A. M. Guess, B. Nyhan, J. Reifler, Exposure to untrustworthy websites in the 2016 US election. *Nat. Hum. Behav.* **4**, 472–480 (2020).

32. J. D. West, C. T. Bergstrom, Misinformation in and about science. *Proc. Natl. Acad. Sci. U.S.A.* **118**, e1912444117 (2021).

33. D. A. Salmon, M. Z. Dudley, J. M. Glanz, S. B. Omer, Vaccine hesitancy. *Am. J. Prev. Med.* **49**, S391–S398 (2015).

34. S. Vosoughi, D. Roy, S. Aral, The spread of true and false news online. *Science* **359**, 1146–1151 (2018).

35. G. Pennycook, D. G. Rand, Fighting misinformation on social media using crowdsourced judgments of news source quality. *Proc. Natl. Acad. Sci. U.S.A.* **116**, 2521–2526 (2019).

36. A. Gruzd, P. Mai, Inoculating against an Infodemic: A Canada-wide covid-19 news, social media, and misinformation survey. *SSRN Electron. J.* 10.2139/ssrn.3597462 (2020).

37. S. Lewandowsky, J. Cook, U. Ecker, D. Albarracin, M. Amazeen, P. Kendou, D. Lombardi, E. Newman, G. Pennycook, E. Porter, D. Rand, D. Rapp, J. Reifler, J. Roozenbeek, P. Schmid, C. Seifert, G. Sinatra, B. Swire-Thompson, S. van der Linden, E. Vraga, T. Wood, M. Zaragoza, *The Debunking Handbook 2020* (Boston Univ. Libraries OpenBU, 2020); https://open.bu.edu/handle/2144/43031.

38. J. Donovan, Social-media companies must flatten the curve of misinformation. *Nature* 10.1038/d41586-020-01107-z (2020).

39. T. Burki, Vaccine misinformation and social media. *Lancet Digital Health* **1**, e258–e259 (2019).

40. R. Smith, C. Wardle, S. Cubbon, Under the surface: COVID-19 vaccine narratives, misinformation and data deficits on social media, (First Draft, 2020); https://firstdraftnews.org/long-form-article/under-the-surface-covid-19-vaccine-narratives-misinformation-and-data-deficits-on-social-media/.

41. A. M. Guess, B. A. Lyons, Misinformation, disinformation, and online propaganda, in *Social Media and Democracy* (Cambridge Univ. Press, 2020), pp. 10–33.

42. H. Allcott, M. Gentzkow, C. Yu, Trends in the diffusion of misinformation on social media. *Res. Polit.* **6**, 205316801984855 (2019).

43. D. Allington, B. Duffy, S. Wessely, N. Dhavan, J. Rubin, Health-protective behaviour, social media usage and conspiracy belief during the COVID-19 public health emergency. *Psychol. Med.* **51**, 1763–1769 (2021).

44. E. Bakshy, S. Messing, L. A. Adamic, Exposure to ideologically diverse news and opinion on Facebook. *Science* **348**, 1130–1132 (2015).

45. A. J. Berinsky, Rumors and health care reform: Experiments in political misinformation. *Brit. J. Polit. Sci.* **47**, 241–262 (2017).

46. M. Cinelli, W. Quattrociocchi, A. Galeazzi, C. M. Valensise, E. Brugnoli, A. L. Schmidt, P. Zola, F. Zollo, A. Scala, The COVID-19 social media infodemic. *Sci. Rep.* **10**, 16598 (2020).

47. M. Del Vicario, A. Bessi, F. Zollo, F. Petroni, A. Scala, G. Caldarelli, H. E. Stanley, W. Quattrociocchi, The spreading of misinformation online. *Proc. Natl. Acad. Sci. U.S.A.* **113**, 554–559 (2016).

48. D. Freeman, F. Waite, L. Rosebrock, A. Petit, C. Causier, A. East, L. Jenner, A. L. Teale, L. Carr, S. Mulhall, E. Bold, S. Lambe, Coronavirus conspiracy beliefs, mistrust, and compliance with government guidelines in England. *Psychol. Med.* **52**, 251–263 (2020).

49. S. González-Bailón, N. Wang, Networked discontent: The anatomy of protest campaigns in social media. *Soc. Networks* **44**, 95–104 (2016).

50. H. O.-Y. Li, A. Bailey, D. Huynh, J. Chan, YouTube as a source of information on COVID-19: A pandemic of misinformation? *BMJ Glob. Health* **5**, e002604 (2020).

51. J. E. Oliver, T. J. Wood, Conspiracy theories and the paranoid style(s) of mass opinion. *Am. J. Polit. Sci.* **58**, 952–966 (2014).

52. O. Papakyriakopoulos, J. C. Medina Serrano, S. Hegelich, *The Spread of COVID-19 Conspiracy Theories on Social Media and the Effect of Content Moderation* (Harvard Kennedy School Misinformation Review, 2020).

53. D. Romer, K. H. Jamieson, Conspiracy theories as barriers to controlling the spread of covid-19 in the U.S. *Soc. Sci. Med.* **263**, 113356 (2020).

54. J. Strömbäck, Y. Tsfati, H. Boomgaarden, A. Damstra, E. Lindgren, R. Vliegenthart, T. Lindholm, News Media Trust and its impact on media use: Toward a framework for future research. *Ann. Int. Commun. Assoc.* **44**, 139–156 (2020).

55. J. R. Zaller, *The Nature and Origins of Mass Opinion* (Cambridge Univ. Press, 2011).

56. V. Cosenza, *World Map of Social Networks* (VincosBlog, 2022); https://vincos.it/world-map-of-social-networks/.

57. T. Ammari, S. Schoenebeck, Thanks for your interest in our Facebook group, but it's only for dads, in *Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work & Social Computing* (2016).

58. R. Y. Moon, A. Mathews, R. Oden, R. Carlin, Mothers' perceptions of the internet and social media as sources of parenting and Health Information: Qualitative study. *J. Med. Internet Res.* **21**, e14289 (2019).

59. R. Laws, A. D. Walsh, K. D. Hesketh, K. L. Downing, K. Kuswara, K. J. Campbell, Differences between mothers and fathers of young children in their use of the internet to support healthy family lifestyle behaviors: Cross-sectional study. *J. Med. Internet Res.* **21**, e11454 (2019).

60. D. Centola, J. Becker, D. Brackbill, A. Baronchelli, Experimental evidence for tipping points in social convention. *Science* **360**, 1116–1119 (2018).

61. N. F. Johnson, N. Velásquez, N. J. Restrepo, R. Leahy, N. Gabriel, S. El Oud, M. Zheng, P. Manrique, S. Wuchty, Y. Lupu, The online competition between pro- and anti-vaccination views. *Nature* **582**, 230–233 (2020).

62. What is the Page Transparency Section on Facebook Pages? (Facebook Help Center); www.facebook.com/help/323314944866264.

63. L. Gustafsson, M. Sternad, When can a deterministic model of a population system reveal what will happen on average? *Math. Biosci.* **243**, 28–45 (2013).

64. M. Chen, Y. Cao, L. T. Watson, Parameter estimation of stochastic models based on limited data. *ACM SIGBioinformatics Record* **7**, 1–3 (2018).

65. K. J. McHugh, L. Jing, S. Y. Severt, M. Cruz, M. Sarmadi, H. S. N. Jayawardena, C. F. Perkinson, F. Larusson, S. Rose, S. Tomasic, T. Graf, S. Y. Tzeng, J. L. Sugarman, D. Vlasic, M. Peters, N. Peterson, L. Wood, W. Tang, J. Yeom, J. Collins, P. A. Welkhoff, A. Karchin, M. Tse, M. Gao, M. G. Bawendi, R. Langer, A. Jaklenec, Biocompatible near-infrared quantum dots delivered to the skin by microneedle patches record vaccination. *Sci. Transl. Med.* **11**, eaay7162 (2019).

66. Weintraub, Invisible ink could reveal whether kids have been vaccinated, (Scientific American, 2019); www.scientificamerican.com/article/invisible-ink-could-reveal-whether-kids-have-been-vaccinated/.

67. M. Roser, H. Ritchie, E. Ortiz-Ospina, *Internet* (Our World in Data, 2015); https://ourworldindata.org/internet.

68. R. Sear, R. Leahy, N. Restrepo, Y. Lupu, N. F. Johnson, Machine learning language models: Achilles heel for social media platforms and a possible solution. *Adv. Artificial Intell. Mach. Learn.* **1**, 191–202 (2021).

69. N. F. Johnson, M. Zheng, Y. Vorobyeva, A. Gabriel, H. Qi, N. Valesquez, P. Manrique, D. Johnson, E. Restrepo, C. Song, S. Wuchty, New online ecology of adversarial aggregates: Isis and beyond. *Science* **352**, 1459–1463 (2016).

70. N. F. Johnson, R. Leahy, N. J. Restrepo, N. Valesquez, M. Zheng, P. Manrique, P. Devkota, S. Wuchty, Hidden resilience and adaptive dynamics of the Global Online Hate Ecology. *Nature* **573**, 261–265 (2019).