Supplementary Material (SM)

This Supplementary Material (SM) provides comprehensive details to support the findings presented in the main manuscript and provide a broader context to thoroughly address common questions surrounding our social media analysis methodology and datasets. By extensively documenting our data collection processes, network construction, and prior related work upon which this study builds, we aim to preemptively dispel any potential doubts about the robustness of our approach. We have included full details in order to help explain this new application area to members of the complex systems research community who may not be familiar with all its details.

<u>CONTENTS:</u>

Section 1: Methodology: Collecting data, building network, use of term “glocal”, and analysis in the paper.

Section 2: Background: Global vaccine hesitancy crisis, the limits of content-focused interventions, and the network analysis approach

Section 3: Color scheme for neutral nodes in network plots. Breakdown of page admin country locations. Breakdown of local page country locations. The system before COVID-19, one year later, three years later, and four years later. Geolocalized network.

Section 4: Classification of neutral nodes.

Section 5: Example of Facebook banners promoting best-science COVID-19 guidance. Positions in network of the nodes that receive Facebook banners promoting best-science COVID-19 guidance.

Section 6: Topic filter. System in June 2025 without node labels filtered by topic.

Section 7: Chi-square test for topic-glocality. Top 20 nodes by betweenness centrality. Statistical results for fan count, topic count, and betweenness centrality scores.

Section 8: Relationship between the number of topics and glocality. Topic heatmaps.

Section 9: Details regarding the softening simulation in Figure 4.

Section 1: Methodology: Collecting data, building network, use of term "glocal", and analysis in the paper.

A list of communities (nodes) and links will be provided online with the Supplementary Material (SM) on publication.

Our **methodology for data collection and classification** follows the 2020 paper referenced in the main text [1]. We go further with the classification than in that paper, by sub-categorizing the neutral communities (the "greens", all non-red and non-blue nodes in Fig. 1B) according to their page's declared interest, e.g., parenting community. This is relatively easy since community pages are organized around a particular stated interest. As for the original classifications in [1], this entire process is carried out with the help of three subject matter experts working independently to classify the neutrals ('greens') by interest, and then cross-checking between them for any differences, as discussed below. There were no remaining contentious cases following this process. Here we refer to our research team, including the subject matter experts, simply as "we".

The process is as follows for **building the network of communities** (N.B. each community is a Facebook page). We start with a seed of manually identified Facebook pages discussing either vaccines, public policies about vaccination, or the pro-vs-anti vaccination debate. These were obtained by searching on Facebook's search engine in 2018 and 2019 for key words and phrases involving vaccines. Pages' selection was agnostic to the glocality of usernames, which is to say, there was no concerted effort to select pages with names that explicitly mention a geographical location. At that time, such pages were easy to find. Then we captured the list of outbound links from these pages (i.e., which pages did each of the seed pages recommend to its members at the page level, see below) using Facebook's GraphAPI endpoint for so-called 'fanned' pages. Some of these links went back to pages already in the list, while others did not. We repeated this process two times to obtain a final list of candidate communities (nodes). In the summer of 2021, Facebook removed this feature from its web interface, but as of fall 2021 the page-level links remained available in the background as an endpoint through Facebook's GraphAPI. Personal account information was avoided, as this is not allowed by Facebook's public API terms of service.

We then prune this link list manually since we are aiming at extracting meaningful links rather than the default of a nearly fully connected network with potential links on all possible topics. Specifically, we filtered out unrelated links by manually classifying the pages from the first step by their content and description. In the second step, we filtered candidate nodes using Facebook's built-in filter for page type, a variable descriptor selected by the page administrator, and then classified those filtered pages for content. A detailed description of the classification and filtering process is included below. In short, we only included pages that were talking about vaccines, or pages that weren't talking about vaccines but self-identified as a cause, community, or NGO and were connected within one step to pages active in the vaccine debate. A link from A to B means page A explicitly lists page B as one of the pages to which it links, not necessarily because it agrees with page B's content (or actually 'likes' it, using Facebook terminology) but because page B's content is of interest to page A. Such a link could have appeared because page A's users noticed it and then recommended it to page A's managers who then established the link. Such a link creates an information conduit feeding content from page B to page A, likely exposing A's users to B's content—and it serves as a flag to A's users for them to look at page B and perhaps get involved in B's ongoing debates. At each step, we vetted new findings through a combination of human coding and computer assisted filters.

We then take this list of nodes and classify them—based on their content—as being pro-vaccine (blue), neutral (green), or anti-vaccine (red). To do this, we reviewed the page's posts, about, and self-described category. Pro and anti-vaccine classifications required that either (a) at least 2 of the most recent 25 posts dealt with the pro or anti vaccine debate, or (b) the page's title or about section described it as a pro or anti vaccine page. The neutral (green) classification required that (c) either 1 of the most recent 25 posts referred to the vaccine debate, but that they did not explicitly took a side pro or against, or (d) the about section clearly classified the page as a neutral in the pro-vs-anti vaccine debate, or (e) none of the most recent 25 posts dealt with vaccines but the page self-classified as either an NGO, a cause, a community, or a grass roots organization. Most of the neutral nodes (green) appeared in the second iteration of fanned pages as many nodes in the first iteration were more explicitly focused on the vaccine debate. This makes sense since they are more removed from the debate, i.e., they do not actively engage in it as a focus but rather are pages that have become interconnected with the pros and the antis who do. The subject matter experts in our team classified each node independently, and when two disagreed on their suggested classification (which happened approximately 15% of the time, largely if the material was ambiguous) all three reviewers discussed these cases in more depth. Agreement was reached in every case. Our team collecting and classifying the communities (Facebook pages) consisted of analysts who have several years of experience in analyzing and classifying online community content on Facebook and other platforms, being trained through prior work on the content of their online communities in association with establishment health guidance as well as various types of extremism and hate [2,3]. We only included pages in languages understood by the researchers, such as English, French, Spanish, Italian, Dutch, and Russian. Identifying sarcastic or ironic posts, fake news, or troll behavior [4] is a challenging task, even for subject matter experts, and machine learning models can generate realistic vaccine misinformation [5]. However, social media communities are quite vigilant in self-policing against bot or troll-like behavior. Thus, the difficulty in measuring the realism and intent of these posts is best left as the topic of another study.

We now comment on Facebook's nomenclature of 'likes', follows and feed, since terms like 'like' can be confusing: 'like' does not literally guarantee liking in the sense of agreeing with, but rather having interest in. The information here is taken directly from online material, which is so common, with similar descriptions in many places, that we do not cite any particular source. The terms "you" and "your" in this paragraph can refer to the entire page. Facebook's website itself also has essentially the same following text, which we borrow from. A (Facebook) like is a person or page who has chosen to attach their name to your page as a 'fan' though this does not necessarily mean they agree with what is being said, rather it is that they have a keen interest in it and want to see content from it. The page will show up in the about section of your account under likes. A (Facebook) follow is a person or page who has chosen to see updates posted by a page on its news feed. If someone or another page follows a Facebook page, it means the content from the Facebook page will show up on their news feed. By default, when you like a page, you will automatically follow it, and this means the content from the page will show up on your newsfeed. The numbers of likes and follows are very close to each but not exactly the same because at some stage, some people who have liked your page may have manually unfollowed you. Your content won't show up on their timeline. We have never seen a significant difference between likes and followers, and the difference is typically a constant factor of about 5%. Thus, this technicality does not affect our conclusions in any way since we have tested adding up to 15% noise and our findings are robust. Finally, we stress again that the Facebook term 'like' should be taken to mean 'have interest in' rather than actually 'like' the content and hence agree with it -- for this reason, communities with very different opinions can end up 'liking' each other since they seek out each other's opposing content to disagree with it.

We now explain our usage of the term **"glocal"** in this paper. In terms of a page's geography, we use the term "global" to refer to a page that is not tied to a specific location or that specifically has a broad, worldwide focus, whereas "local" refers to a page that is location-centered or focused on a specific geographic area, such as a neighborhood, city, county, state, or country. For example, the page "Vaccine information for Los Angeles County parents" is considered local since an explicit location ("Los Angeles County") is mentioned, whereas "Vaccine information for parents" or "Global Trends" is considered global since the name implies a worldwide focus. In terms of a page's topic, we use the term "global" to refer to a page that features a broad-ranging discussion, whereas "local" refers to a page that features narrow and focused topic(s), such as pages that discuss specifically and only "COVID-19", "COVID-19 and mpox", etc. In the case of either geography or topic, we use **"glocal"** to refer to a situation where both global and local characteristics occur together.

In this paper, the study period for analyzing the topic entanglement was chosen to span from May 1, 2022, to October 5, 2023. The original time period of study was May 1, 2022, to October 17, 2023, which included significant events such as the first confirmed case of the 2022 mpox outbreak in the U.S.A. [6], the U.S. Supreme Court's reversal of Roe v. Wade [7], U.S. President Joe Biden signing into law the Inflation Reduction Act [8], and numerous primary and run-off elections in anticipation of the November midterm elections [9]. Facebook was chosen as it has 3.0 billion active users worldwide and is the top social network in 156 countries [10]. Recent studies have shown that Facebook's in-built community structure is a valuable tool for addressing information needs and supporting decision making [11–13]. Our main unit of analysis is Facebook pages, which aggregate people around a common interest and are publicly visible. While throughout the paper we refer to each Facebook page as a community, we stress that we do not use any ad-hoc community structure inferred from network algorithms. As explained above, our starting point is the ecosystem of interlinked communities around the vaccine health debate just prior to COVID-19 in November 2019. We recall that a link from community (page) i to community (page) j exists when i "likes" or "fans" j (which strictly speaking means that i has an interest in j's content, not that i actually agrees with j), hence indicating a recommendation to all i's members at the page level. A mere mention by a page member of another page is not enough to create a link. These communities can engage in cross-community chatter across different topics, analyze news articles, etc. As previously mentioned, these communities have been classified as pro, neutral, or anti-vaccination, and are part of a larger dataset of communities engaged in vaccination dialogues. Of the full dataset, 78.7% are still active, and 43% are producing topic-related posts. These still online communities are the ones that we see in Figs. 1 and 2. If we had included all communities that follow those nodes actively generating topic-related content but were not producing similar types of posts, the system would comprise 70.6% of all communities in the full dataset. Nonetheless, we chose to focus on the characteristics of actively involved communities.

In the rest of this section, we provide answers to potential questions and critiques that may be raised regarding our methodology and research in the paper. Though it repeats some of the material above, we feel this is necessary in the context of the potential question being addressed:

We could define nodes and links in another way, and we recognize that our dataset is ultimately an imperfect sample of some larger "correct" network. Any definition of nodes and links—including ours in the present work—can of course be criticized since we are reducing the many attributes of a real system down to the few necessary to build a network. We could instead for example analyze the content of the posts on the pages, identifying shares and URL links of posts from other pages to build a weighted, directed network that captures how users of one page are actually exposed to the content of another page and driven to it. However, this comes with its own downsides: the content may be in different languages and hence not so easy to identify as being shared, and the terminology can evolve quickly (e.g., the use of slang to avoid attracting the attention of Facebook moderators). This issue is worthy of a study to establish comparative advantages of the two approaches, but such a study is beyond the scope of the present paper. Ultimately, the "best" choice of links and nodes will depend on the questions being asked about the system, since the network is less useful if it has too few nodes and if it is either very sparse in terms of links or too dense with every node essentially connected to most others. The best choice will also depend crucially on the available data and the level of granularity at which this data is reliable. Even with this, the complexity of the data will require some form of simplification in order to make the analysis tractable and understandable. Fortunately, the simplification that we make of each node being a community (page) does bring some advantages: for example, it avoids the need for accessing individual-level information and makes the definition of each node unique since each page has its own unique identification number. Further, as the process we follow yields on the order 1000 nodes (communities), each containing on the order of 100,000 users, the network produced by our definition is visually manageable and yet interpretable at scale, as opposed to being overwhelmed with links or being too sparse. This means that the open-source software Gephi and its ForceAtlas2 algorithm that we use, which follows the principle of energy minimization, provides an uncluttered spatial representation. It further means the network is scalable to the population level, since 1000 nodes with 100,000 users each means we are potentially tapping into the behavior of 100 million users.

We also note that the set of nodes and links is obtained without regard for the specific classifications of the nodes (i.e., pro vs anti vs neutral parenting, etc.). Furthermore, they were obtained without regard for whether or not the node label (page username) explicitly mentioned a geographic location. Thus, the fact that the modular structures of such subpopulations emerge spontaneously in Fig. 1 lends support that the links we identify are meaningful. Taking a devil's advocate position, if the link methodology were so subjective that the links are not meaningful in a scientific sense, one would expect similar outcomes to a null model, which is not the case.

We have investigated that instead of using human experts, we train a supervised language model of the recent posts of each page, running it over newly discovered communities (pages) and including it in the network only if it can be considered relevant by the model. This gives a crudely similar list—however there are some glaring anomalies that are easily caught by a human expert, showing that while such automation improves volume and speed of analysis, it can also introduce glaring anomalies that a human would be very unlikely to let through. Also, it is hard to have a supervised language model treat material in Spanish, French, Russian, etc. on an equal footing to English. So, although such machine automation sounds desirable because it scales and promises to be less subjective perhaps, it is clearly a trade-off. Overall, the fact that we had already published a similar list of nodes and edges in [1] and the present list—obtained from scratch—was similar, combined with the fact that the number of nodes and links are not too large, and combined with the fact that the network gives results which are very different from a null network model, gives us confidence that our study, though imperfect, is capturing significant features of the actual online system.

Another approach would be for us to list a series of terms we consider relevant, and automatically include communities with sufficient posts using these terms. One might think that this could help ensure that the sampling is more systematic, transparent, and reproducible. However, it has a downside, in that we see the terms evolving in time—in part to avoid attracting the attention of Facebook and in part since the topics are evolving (e.g., the use of bleach then turned to other hot topics) and so such a list could also forever be playing catch-up. Further, there is the prior question of terms translating between different languages. Ultimately, whether this approach would perform well would need to be carefully studied and would represent a research project and paper in its own right.

We stress that we do not assume in any way that linked pages from a given page immediately influence the browsing and information sharing of the users subscribed to the linked pages, and hence that strong correlations in activity levels should accompany links between pages. The neutrals hardly ever discuss vaccines—but the fact that they have material from anti-vax pages appearing in their feed, will be noticed by many of them, and so when some later decision needs to be made such as taking their children to be vaccinated, they may think twice. Thus, it is not that the links carrying material automatically give rise to higher activity at that moment, but that they represent an influence. [4] showed experimentally and theoretically that an online community can suddenly tip to an alternate stance in a reproducible way if there is a committed minority of around 25%. No amount of prior analysis of levels of activity in the experiment of [4] would have predicted this. Having said this, one might wonder if there is some evidence of elevated activity levels. Though this is not important in our paper, we note that we do indeed see this, as shown by the following earlier work of Nicholas Gabriel in our group: consider two Pages i and j, among a collection of N Facebook Pages, and suppose that Page i likes Page j so in our analysis there is a link from Page i to Page j. We calculate a measure of correlation between Pages for unlinked and linked Pages, assigning 1 for a link and 0 for no link. The mean for the unlinked Pages is 0.115 and the mean of the linked Pages in 0.255. Since these samples are relatively large, and their means relatively far apart, this gives an extremely small p-value of $p = .0000037$ for the hypothesis that the means are the same. The reason the correlation is not much large is likely that (i) the links are not fully utilized by Facebook's algorithms to "forward" posts to the user feeds, e.g., all posts may not be shown from a linked Page if they are not interesting, or historically they haven't had success sharing along a particular link. In that way the links may be present, and hence the feed read and understood by users of the other community, but not necessarily responded to; (ii) people likely read and digest the content, but do not then go and post new activity in immediate response. In other words, the link has influence on the reader—but this influence likely sits passive until a later time. To extend this analysis, which again goes beyond the requirements and scope of the present paper, one could perform a Granger test of activity time series in pairs of linked and unlinked communities, including the link present as a variable in the times series models and assessing whether it can be considered significantly positive. A second approach would be to retrieve URL links and shares from the posts of all communities and run the test that the frequency of these links from one community to another is higher if they are linked as in the studied network.

One might query analysis and interpretation of the resulting visual shape of the network clusters with the ForceAtlas2 layout in Gephi. But the reason the layout can be interpreted in this way, is that the ForceAtlas2 layout is the result of a many-body physical calculation of energy minimization in which all nodes (regardless of their classification) repel each other with a force that decays with separation, while linked nodes have an additional attractive spring force. Hence sets of nodes that end up closer together do so not because of their hate classification, but because they share more links. It is certainly true that the final layout may be one of many many-body equilibrium states with similar overall energy and hence a local but not global minimum. However, as we show in Sec. 4, angles and lengths can still be used as a crude guide for changes in the network given that the system stays in this local minimum — and it is clear even visually from Supplementary Fig. S2 that there has been a strengthening of bonds (i.e., sets of links) however one chooses to measure this. In the future, again out of the scope of the present paper, we would operationalize this by developing an optimal network science measure of indirect bonding strength between communities, then performing a statistical analysis versus a null model that shows that there was indeed a reliable structural change between years.

Finally, we comment on the significant scientific drawback of using third-party alternatives to our approach of collecting data – focusing here specifically on a tool called CrowdTangle. The following text and discussion repeats our published critique of the use of tools like CrowdTangle for rigorous science research, that we published in "How Social Media Machinery Pulled Mainstream Parenting Communities Closer to Extremes and Their Misinformation During Covid-19", IEEE Access 10, 2330 - 2344 (2021). Despite this known criticism, there are still studies being published whose claims and findings rely entirely on the hope that CrowdTangle is accurate (see Broniatowski et al. Sci Rep 13, 15964 (2023); Sci. Adv. 9, eadh2132 (2023)). CrowdTangle is a commercial application tool owned by Facebook, but the crucial drawback is that researchers outside Facebook have little knowledge or quantitative explanation of how and why this tool returns the results that it does. In other words, it is effectively a black-box tool, which makes it unacceptable for academic science research in our opinion. Worse still, it is widely reported that the CrowdTangle tool is no longer being properly maintained and hence any accuracy that it may arguably have had has now likely degraded significantly. Although there was a verbal claim in the mainstream media that CrowdTangle is "the most effective transparency tool in the history of social media" (B. Smith, "A former Facebook executive pushes to open social media's 'black boxes'," The New York Times, 2 January 2022; www.nytimes.com/2022/01/02/business/media/ crowdtangle-facebook-brandon-silverman.html), that verbal claim is not backed up by rigorous scientific tests or analysis. So although it may be the best available tool for journalists, this gives no guarantees about it being good enough for rigorous science. Researchers outside Facebook cannot assess the accuracy of its results -- and they do not know how the results that it returns depend on Facebook's own algorithms, architecture and databases, which Facebook keeps secret. Nor can researchers outside Facebook control any of this: they simply input search prompts and the tool spits back data. Nor do they know about the completeness of the black-box search results returned. Nor do they know if there is any bias in the search process within the black-box tool, nor what that bias might be or how big it is. Our own investigations suggest that similar searches can produce quite different results. While a larger list of communities may indeed be obtained using such a black-box tool, that list may be significantly biased and hence less reliable than a smaller sample obtained using a non-black-box tool. The CrowdTangle search output is also not very precise. It can for example include results with different spellings that are unrelated in topic. Nor are the numbers that it returns checkable or proven to be accurate, which further calls into question the reliability of studies that use it for quantitative academic analysis. Nuancing the search terms can produce very different results, adding to concerns about how complete and robust the output is for academic research. Also, doing searches about the past cannot easily reveal communities that have removed themselves or were removed by Facebook, or have changed their name. It therefore remains unproven that such black-box tools are suitable for rigorous, reproducible scientific research as opposed to simply being used as a search tool for businesses and for qualitative exploration of a particular story. Without systematic, quantitative studies against ground truth lists, one cannot assume that findings obtained using black-box tools are reliable. Nor is the number of candidate communities that emerges an indicator of a larger sample and hence a broader or more reliable study, since it is always possible to capture many more communities by using a coarser net to capture many less relevant and potentially biased examples.

We also note that after classification of the communities, whether through CrowdTangle or otherwise, different studies with slightly different classification schemes may end up with very different numbers of communities in a given category, e.g. more anti communities. Again, this does not mean that a study with a larger list is better or has more reliable results, since the classification criteria are not identical for the categories, i.e. starting from the same bag of candidate communities, the researcher-chosen criteria for the 'anti' label in any given study could simply allow for more objects from that bag to be assigned the anti label.

Other studies such as those cited above, may also use different ways of defining links between nodes (communities), e.g. URLs listed in the content. But there is no guarantee that these URLs represent any meaningful connection between the majority of users in one community and another, nor that it influences their subsequent behavior in any significant way. This needs to be proven before any subsequent network analysis can be regarded as meaningful. By contrast, in our study the links between nodes (communities) are better defined, i.e. if page A 'likes' page B then this community A links to community B which creates an information conduit from B into A and hence exposes A's users to B's content (e.g. new posts from Facebook page B can appear on Facebook page A). So the difference between our network versus those of other studies, lies in the fact that our study is a follower/like/social network based on the formal page follow/like connections on Facebook itself whereas those other studies use a mention/hyperlink network created by pages linking to one another through URLs in their posts. So our network is a true follower network whereas their network is a mention network. Technically, we also note that our kind of network cannot easily be made using third-party tools such as CrowdTangle.

Obtaining the network in the way that we do, with this choice of links and nodes, is what enabled us to study in a meaningful, rigorous way how the network has changed over time. This led us to analyze and report for the first time the remarkable robustness of the vaccination-debate network despite Facebook mitigation measures and policies (e.g. implementation around November 2020) and hence provide the first report of the inefficacy of Facebook's vaccine misinformation policies and architecture during the COVID-19 pandemic. Specifically, our finding of the network's robustness despite Facebook's policies and ramped-up mitigations was presented in early 2022 in "Softening online extremes organically and at scale" https://doi.org/10.48550/arXiv.2207.10063; Elvira Maria Restrepo, Martin Moreno, Lucia Illari, Neil F. Johnson; also see "How Social Media Machinery Pulled Mainstream Parenting Communities Closer to Extremes and Their Misinformation During Covid-19", IEEE Access 10, 2330 - 2344 (2021); and "Losing the battle over best-science guidance early in a crisis: Covid-19 and beyond". Science (Advances) 28 Sep Vol 8, Issue 39 (2022). DOI: 10.1126/sciadv.abo8017; L. Illari, N. J. Restrepo, N.F. Johnson; as well as public presentations in 2021 and 2022 (e.g. Countering Misinformation session at the Annual Leadership Meeting of the American Physical Society, 2022; also see lab website https://donlab.columbian.gwu.edu).

<u>Section 2:</u> Background: Global vaccine hesitancy crisis, the limits of content-focused interventions, and the network analysis approach

Global confidence in vaccines has reached a critical inflection point. A comprehensive study spanning 55 countries revealed that perceptions of vaccines' importance for children declined in 52 nations during the COVID-19 pandemic, with some countries experiencing drops up to 44 percentage points[14–16]. In the United States, approximately 20% of parents now express vaccine hesitancy[14,17]. This sustained hesitancy has contributed to the resurgence of vaccine-preventable diseases, most notably measles. The 2025 outbreak has already claimed three unvaccinated lives and reached nearly 1,300 confirmed cases, marking the largest resurgence since measles was declared eliminated in the US in 2000[14,18,19]. Meanwhile, Canada experienced an even more dramatic surge with over 3,800 confirmed cases[20]—nearly three times the US total despite having a little more than one-tenth the population. This outbreak, concentrated among unvaccinated individuals in communities with historically lower vaccination rates, made Canada the only Western nation ranking in the global top 10 for measles outbreaks[20]. This crisis signals a broader erosion of the public health infrastructure that has saved an estimated 154 million lives over the past half-century[21].

This erosion of trust coincides with—and is amplified by—the appointment of prominent vaccine skeptics to the highest levels of health governance. Robert F. Kennedy Jr., who has repeatedly promoted the debunked connection between vaccines and autism, now leads the Department of Health and Human Services[22,23], the very agency responsible for vaccine approval and recommendations. His first acts included announcing plans to study the long-discredited vaccine-autism link and moving to restrict access to COVID-19 vaccines[14,22–26]. Yet the focus on prominent figures may obscure deeper structural forces: Canada's measles outbreak occurred despite having no equivalent high-profile vaccine skeptic in government. As public health expert Maxwell Smith notes, Canada lacks "a prominent RFK Jr-like figure"[20], yet still experienced nearly three times the infection rate. This suggests that vaccine hesitancy operates through mechanisms that transcend visible leadership.

Meanwhile, the scientific infrastructure needed to understand these underlying dynamics is being systematically dismantled. The National Institutes of Health has cancelled or scaled back over 40 research grants, related to vaccine hesitancy and promoting vaccine acceptance, including critical studies tracking how misinformation spreads through online communities[27–29]. At the precise moment when understanding vaccine hesitancy becomes most crucial, the tools to study it are disappearing.

The consequences extend far beyond personal health decisions. The same online machinery that fans vaccine falsehoods now cross-pollinates narratives about election integrity, climate policy, and social movements[30–34]. During 2024's unprecedented election cycle—with over 60 countries holding votes[35–37]—vaccine misinformation became weaponized as a tool for broader political manipulation. In multiple contests, disinformation campaigns braided anti-vaccine myths with voter suppression hoaxes and "government control" conspiracy theories[38–40]. The result is a crisis of institutional trust: only about one in five Americans trust the federal government to do what is right most of the time[41], while four in ten people worldwide admit they have already regretted a health decision made on the basis of misinformation[42]. This erosion of trust creates a vicious cycle: as institutions lose credibility, people become more vulnerable to the very disinformation that undermined faith in expertise, often finding refuge in the same online communities that cast vaccines, elections, and public health measures as tools of state overreach and hidden danger.

The challenge facing public health authorities is daunting. As researchers studying vaccine confidence note, scientists remain "light years behind" anti-vaccine activists who operate as "highly organized, orchestrated, highly funded people trying to undermine vaccination as a programme"[14]. But where does this battle for public trust actually unfold? While broadcast and official channels remain influential, the day-to-day skirmishes over vaccine safety, treatment fads, and public-health guidance now play out mainly inside social-media feeds and peer groups, making those platforms the "primary battlefield" for contemporary health debates[43–47]. Facebook, with its 3 billion monthly active users[48], hosts a large concentration of these communities, from official CDC pages, conspiracy theory groups, concerned parent forums, and organized anti-vaccine networks[1]. Within these digital spaces, competing narratives don't simply coexist; they form complex webs of influence that determine which messages reach mainstream audiences and shape family health decisions[49–52].

Faced with this mounting crisis, the response from platforms, governments, and public health organizations has been unprecedented in both scale and coordination. November 2020 marked a watershed moment when Facebook fundamentally shifted its approach from informational interventions to active content removal. Prior to this point, the platform had relied primarily on what could be characterized as soft interventions: between March and October 2020, Facebook displayed warnings on approximately 167 million pieces of COVID-19 content based on fact-checker debunking articles, while removing 12 million pieces deemed to pose "imminent physical harm"[53,54]. These soft interventions operated on a seemingly reasonable premise: provide users with accurate information alongside misinformation, and rational choice would prevail.

But as vaccine misinformation continued to flourish despite these warning labels, Facebook escalated to systematic community removal. By August 2021, the platform had "removed over 3,000 accounts, Pages and groups for repeatedly violating our rules against spreading COVID-19 and vaccine misinformation" and eliminated "more than 20 million pieces of content"[55]. The company explicitly targeted not just individual posts but entire networks, promising to remove "any Pages, Groups, Events or Instagram Accounts" that violated COVID-19 policies and to pursue those "that instruct or encourage users to employ code words when discussing vaccines or COVID-19 to evade our detection"[56].

This platform-level crackdown was just one prong of a much broader assault on vaccine misinformation, where a myriad mitigation strategies have been introduced and implemented across platforms[57–76]. The Social Science Research Council's Mercury Project, enabled by the Rockefeller Foundation, Robert Wood Johnson Foundation, Craig Newmark Philanthropies, committed $10 million to behavioral science campaigns designed to "build vaccine demand and counter misinformation"[65,66,77]. These weren't simple public service announcements but coordinated interventions run by teams of psychologists, data scientists, and communication experts working across five continents. In parallel, the National Institutes of Health launched ambitious research programs with dozens of grants aimed at understanding vaccine hesitancy, mapping misinformation networks, and developing evidence-based counter-messaging strategies[28,29,78,79]. At the implementation level, public health agencies adopted the CDC's "Vaccinate with Confidence" framework[80] while high-visibility campaigns like the Ad Council's "It's Up to You" flooded traditional and social media with pro-vaccine messages[81]. Meanwhile, Facebook invested $100 million in supporting journalism and fact-checking infrastructure during the pandemic[54]. Even frontline clinicians received training in motivational interviewing techniques to address patient concerns in one-on-one conversations[82,83]. Depending on funding sources and organizational mandates, these mitigation efforts typically focused on specific topics (COVID-19, elections, climate change) and operated at scales ranging from local community interventions to national campaigns and global initiatives.

The logic underlying these diverse efforts appeared unassailable: (1) algorithmically amplify authoritative health sources like the CDC and WHO; (2) provide the public with timely, accurate, evidence-based information; (3) remove accounts and communities repeatedly spreading false information; and (4) fund research to understand the psychology of hesitancy and develop targeted interventions. This approach treated misinformation as a content problem requiring content solutions: eliminate problematic sources, elevate credible voices, and ensure accurate information reaches audiences. The strategy drew on established principles of public health communication and assumed that better information would naturally outcompete false claims in the marketplace of ideas.

Yet here we are in 2025, and by most meaningful metrics, the situation has worsened. Despite platform removals of thousands of problematic accounts and millions of pieces of content, vaccine hesitancy has increased. Despite behavioral science campaigns reaching tens of millions of users, trust in medical expertise continues to erode. Despite coordinated fact-checking efforts, false narratives often spread faster than corrections. Most tellingly, the research infrastructure built to understand this problem is being systematically dismantled just when it's needed most.

This presents a profound puzzle that challenges fundamental assumptions about information ecosystems and human behavior. Why does misinformation resilience persist despite massive resource expenditure across platform interventions, research programs, and public messaging campaigns? If the problem were simply the presence of false information requiring correction with accurate information, these comprehensive efforts should have produced measurable improvement. Instead, we observe a system that appears to have strengthened rather than weakened.

Understanding this paradox requires reconceptualizing the problem at its most fundamental level. The persistent failure of content-focused interventions hints at a critical misunderstanding: online misinformation is not primarily a content problem but a network engineering problem. The issue isn't the presence of individual bad actors spreading false claims, but rather the structural properties of the network that enable certain types of information to flow efficiently to mainstream audiences while others remain confined to the periphery. When Facebook removes an anti-vaccination page, the assumption is that its influence disappears with it. But networks are not simple addition problems where removing nodes proportionally reduces influence. They are complex systems with emergent properties, redundant pathways, and adaptive capacities that transcend individual components. The question becomes not "how many bad pages can we remove?" but "what network structures enable misinformation to persist despite node removal?"

To reveal these hidden dynamics, we conducted a longitudinal network analysis of Facebook's vaccine ecosystem from 2019 through 2025, tracking how 1,356 interconnected communities encompassing 70–100 million users evolved in response to platform interventions. Using the same rigorous classification methodology established in Ref.[1], we mapped pro-vaccination, anti-vaccination, and neutral communities through multiple temporal snapshots, capturing both the November 2020 intervention watershed and its aftermath.

Our analysis reveals a striking paradox: despite Facebook removing one-third of anti-vaccine pages and over half their connections, the network's functional architecture remains visually and structurally intact (**Error! Reference source not found.**). The spatial arrangement of communities—anti-vaccine hubs, pro-vaccine peripheries, and neutral bridges—persists unchanged even as individual nodes disappear. Moreover, we document how the network evolved defensive adaptations that explain this resilience. Pages increasingly blend vaccine discussions with climate change, elections, abortion rights, and other topics, creating topic-agnostic information highways that span multiple issue domains (**Error! Reference source not found.**). Simultaneously, local neighborhood pages form connections with global conspiracy networks, establishing cross-scale pathways that enable misinformation to flow from hyperlocal communities to worldwide audiences (**Error! Reference source not found.**). This "glocal evolution"—the simultaneous dissolution of topic boundaries and geographic scales—creates redundant pathways that render targeted interventions ineffective.

The main contributions of this paper are thus:

- We demonstrate the structural persistence of misinformation networks despite massive platform interventions. Our longitudinal analysis reveals how the Facebook vaccine ecosystem maintains its core architecture even after losing substantial nodes and connections, challenging assumptions about content moderation effectiveness.
- We identify the specific mechanisms enabling network resilience: topic and geographic "glocality." Pages routinely discuss multiple ostensibly unrelated subjects while connecting local communities to global movements, creating robust information pathways that transcend traditional content categories and geographic boundaries.
- We show how the same network properties that confer resilience can be leveraged for positive outcomes. Through agent-based simulations, we demonstrate that small mixed-opinion deliberation groups can achieve large-scale opinion moderation within weeks without removing content or suppressing voices (**Error! Reference source not found.**).
- We thus propose network engineering approaches that work with rather than against existing topology. Rather than fighting network structure through removals, our findings suggest interventions that harness structural properties for constructive dialogue and opinion softening.

By revealing both the stability of structural properties and the adaptability of network routing, our analysis explains why current content-focused approaches fail and points toward more effective strategies for managing health misinformation at scale.

<u>Section 3:</u> Color scheme for neutral nodes in network plots. Breakdown of page admin country locations. Breakdown of local page country locations. The system before COVID-19, one year later, three years later, and four years later. Geolocalized network. Preventing node overlap in Fig. 1 networks.

Color scheme for neutral nodes in network plots:

Supplementary Table 1. Color scheme for neutral node categories in Fig. 2A:

Neutral category	Color
AltHealth	■
Conspiracy	■
GMO	■
Health	■
Illness	■
Movement	■
Organic	■
Organization	■
Other	■
Parent	■
Pet	■
X	■

Breakdown of page admin country locations:

Additionally for interest, it is possible to extract a list of admin top countries (i.e., the country in which the largest number of page admins reside) for all the pages collected. This information was retrieved for 834 of the pages in the dataset, or 62.17% of all pages, and the top 10 countries are as follows:

Supplementary Table 2. Top 10 counties in which the largest number of a Facebook page's admins reside

Country	Number of pages
United States	352
Australia	33
Canada	27
United Kingdom	20
Italy	9
France	8
New Zealand	6
Sweden	4
Germany	4
Belgium	4

Other countries such as: India, Ireland, Norway, Switzerland, Mexico, Uruguay, South Africa, the Netherlands, Slovenia, Thailand, Malaysia, Brazil, the United Arab Emirates, Croatia, Israel, Austria, Slovakia, Bulgaria, the Philippines, Singapore, Romania, Belize, Denmark, Serbia, Czechia, Portugal, Pakistan, Poland, and Costa Rica also appear in these results, though with diminishing quantities.

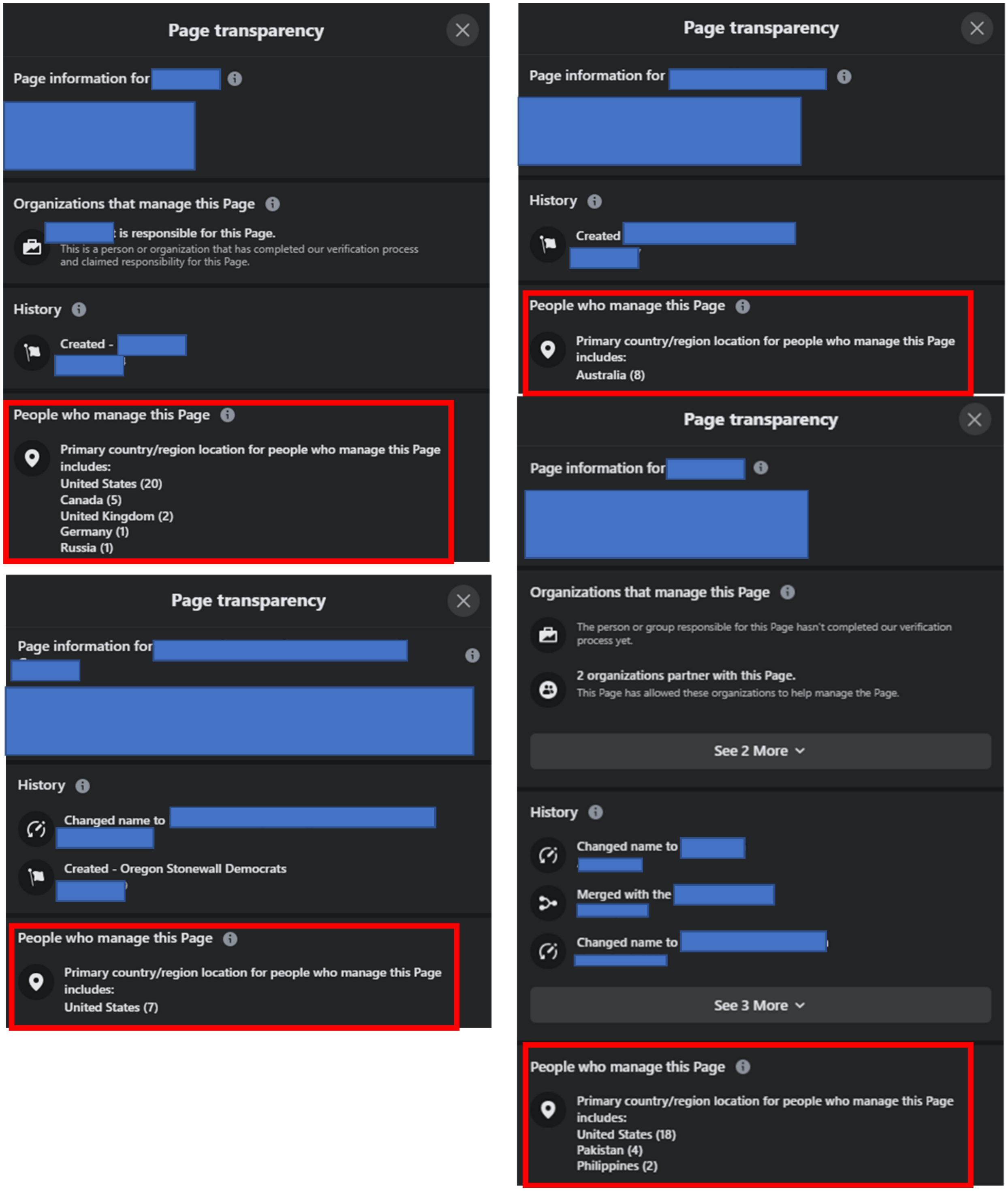
Supplementary Table 3. Top 10 counties mentioned in Facebook Page usernames

Country	Percentage of local pages
United States	66.33%
Australia	6.08%
Canada	4.81%
France	4.05%
Italy	3.04%
United Kingdom	1.77%
New Zealand	1.77%
Mexico	1.01%
Spain	0.76%
South Africa	0.76%

Other countries such as: Malaysia, Ireland, Croatia, Thailand, Poland, Norway, Japan, Germany, Egypt, Denmark, Uruguay, United Arab Emirates, Slovakia, Serbia, Peru, the Netherlands, Nepal, North Macedonia, India, Iceland, Costa Rica, Bulgaria, Brazil, Bosnia and Herzegovina, and Belgium also appear in these results, with diminishing quantities.

These results, however, can only give a partial or incomplete understanding of the true breakdown of admin locations, which a page's Transparency section can shed further light on. The Transparency section is part of Facebook's effort to provide more information regarding the page and the people who manage it as an effort to increase the accountability and transparency of pages and can be read about in further detail on the Facebook website. Importantly, this information includes the primary country locations where the page is managed and provides a more complete breakdown of where page admins are located.

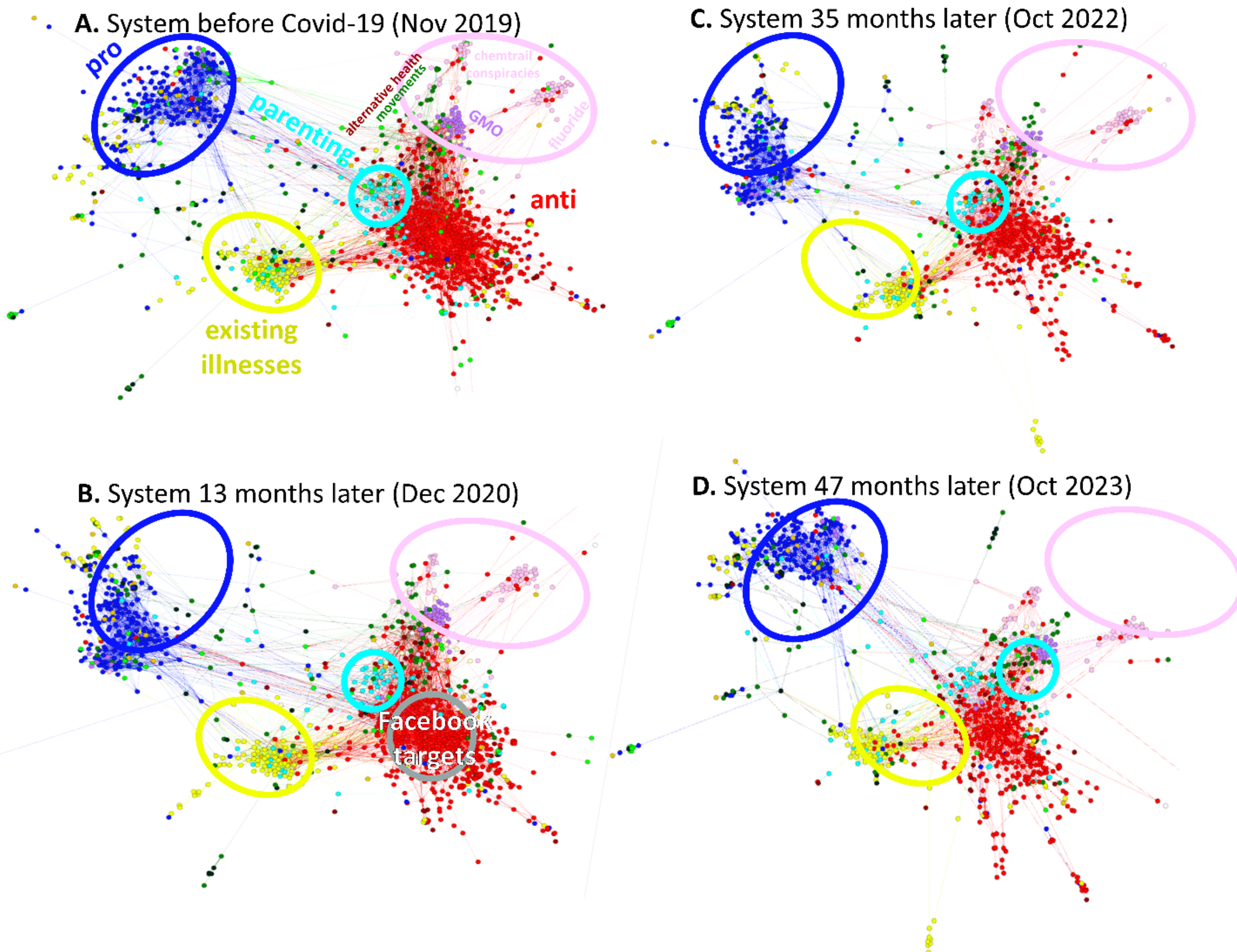
Some examples of what the Transparency section contains are provided in what follows, with identification information blocked out, and admin information boxed in red. The following page in the upper left-hand corner, for example, would be considered exactly the same as the two, smaller pages on the bottom left and right, which only have admins from one country. Thus, the dataset contains pages where the admins might all come from the same country (e.g., the pages on the right-side of the image and the lower left-hand), but it also contains pages where there is a mix of countries, and that includes countries where the dominant or official language is not English, even if the page's contents are primarily or only in English.



Supplementary Figure 1. Examples of the Transparency section for several Pages in the dataset. The Pages on the left and bottom right would all be classified as having top number of moderators in the U.S., obscuring in data extraction the true moderation team composition.

The system before COVID-19, one year later, three years later, and four years later.

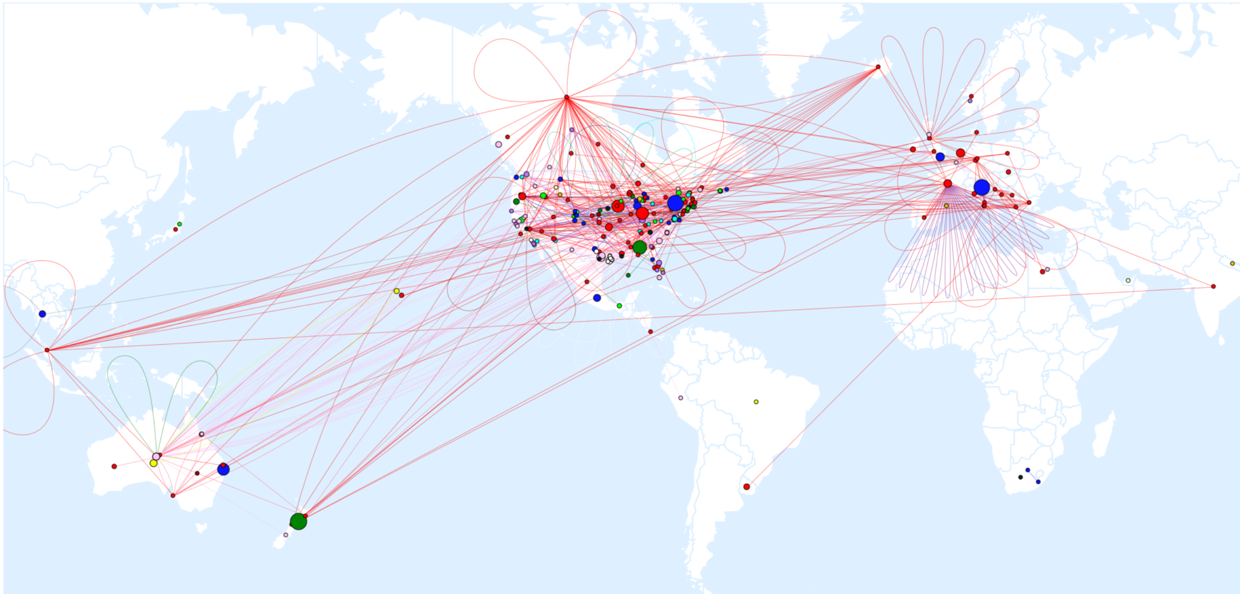
Similar to Fig. 1 in the paper, this figure below shows what the system looked like at 4 different times: pre-vaccine in Nov. 2019, post-vaccine in Dec. 2020, then Oct. 2022, then Oct. 2023. 257 nodes had been removed from this system (deleted) and 32 pages had gone private. The rings provide a visual aid in seeing the increase in node bonding due to the changing node-link structure and show how similar the system looks post-vaccine as it did pre-vaccine.



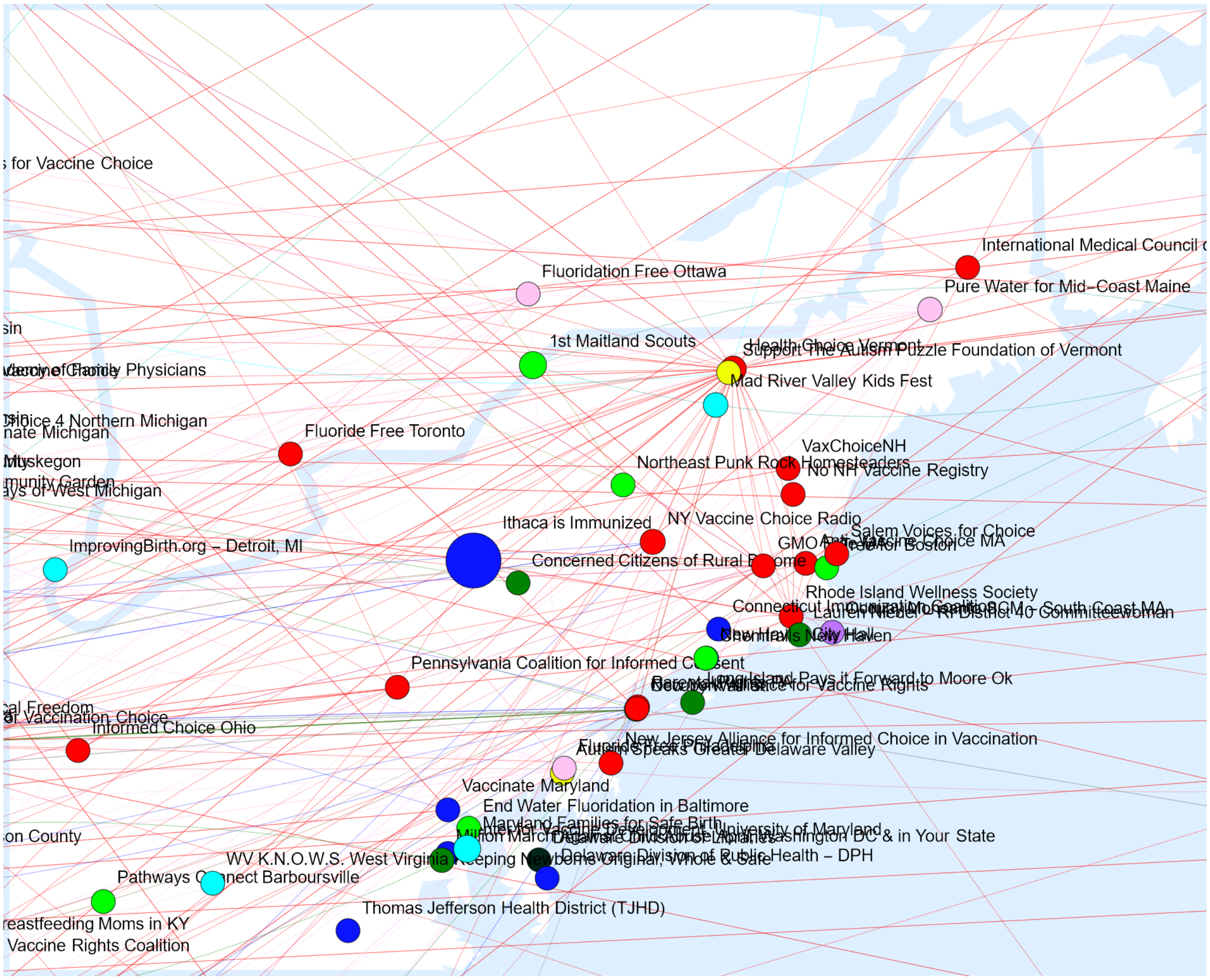
Supplementary Figure 2. The system before COVID-19, one year later, three years later, and four years later

Geolocalized network:

Network of local Facebook communities localized to their real geographic location. Sometimes owners of Facebook pages provide a real address on their page, though in other instances the location has been determined from the Facebook page's username, bio, and other information.



Supplementary Figure 3. The entire world map with geolocalized Facebook communities; node size is proportional to community betweenness centrality values in full network. Only edges between local Facebook communities are shown.



Supplementary Figure 4. Zoom in on the Northeast and Mid-Atlantic regions of the United States with geolocalized Facebook communities; this is the labelled version of the cut-out in Supplementary Fig. 20. Node size is proportional to community betweenness centrality values in full network, and only edges between local Facebook communities are shown.

<u>Section 4:</u> Classification of neutral nodes.

The following pages provide an explanation of the 12 neutral categories of neutral node used in the main paper.

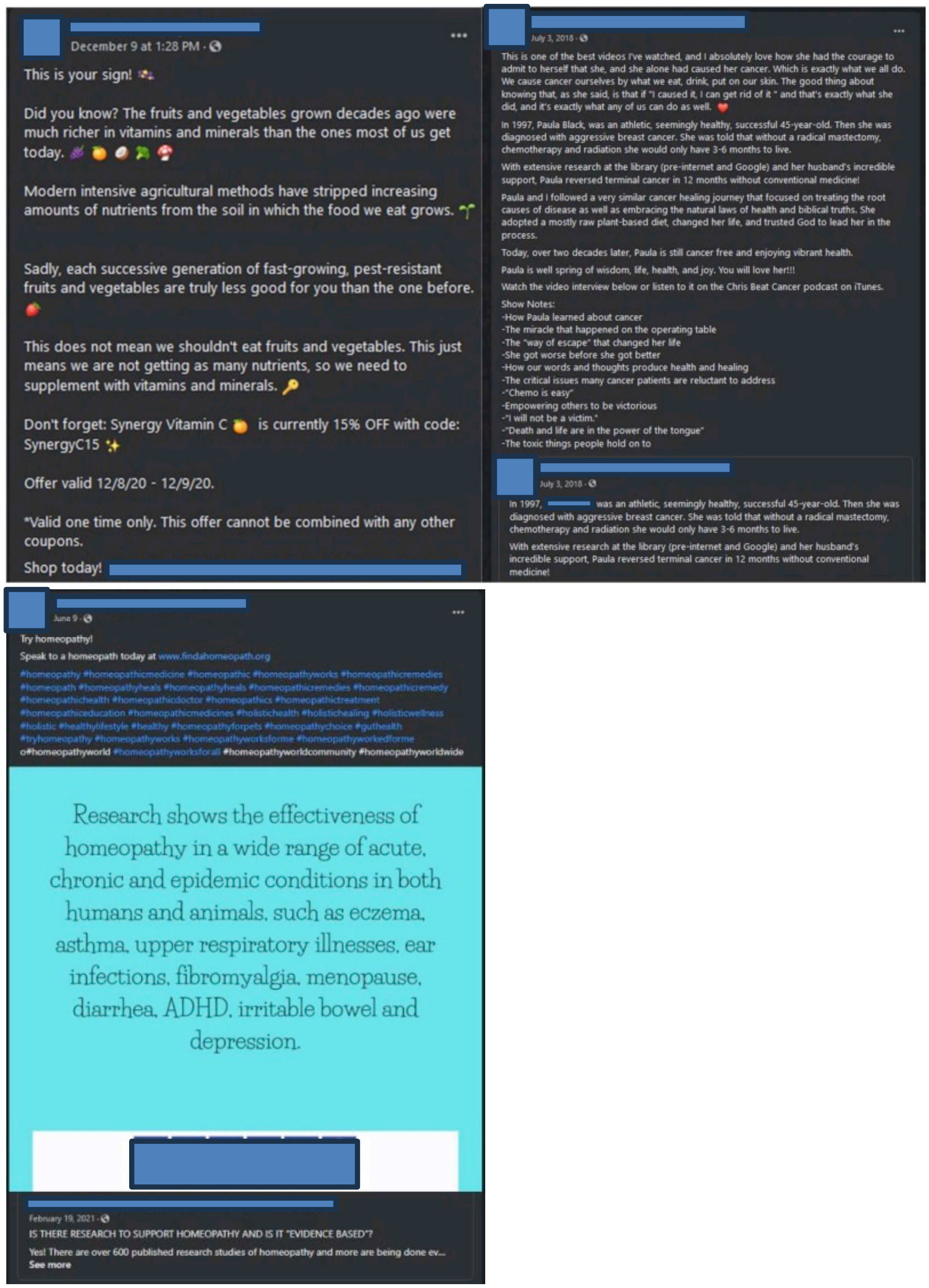
<u>AltHealth</u>

An Alternative Health (AltHealth) community (i.e. node) is a Facebook Page that promotes, discusses, or features content centered around alternative cures and practices, as opposed to traditional medical practical. This includes homeopathy, naturopathy, and spiritual healing. These communities focus on anything from more common conditions such as headaches, indigestion, and general wellness, up to serious illnesses/conditions such as cancer and genetic disorders.

Some of these communities promote and market "remedies" such as essential oils, herbal supplements, or unconventional medicines. They sometimes do this by addressing these products in their posts/pictures and/or sharing links to websites where they can be purchased. Other communities might share anecdotal remedies: this includes recommending alternative diets and practices such as eating "raw".

While not the largest category in terms of number of communities, this is by far the largest in terms of total user size. This can be attributed to the large size of the top few communities in this category.

For example, the largest green community in the network, "Sun Gazing", is an AltHealth community.



Supplemental Figure 5. Examples of AltHealth pages

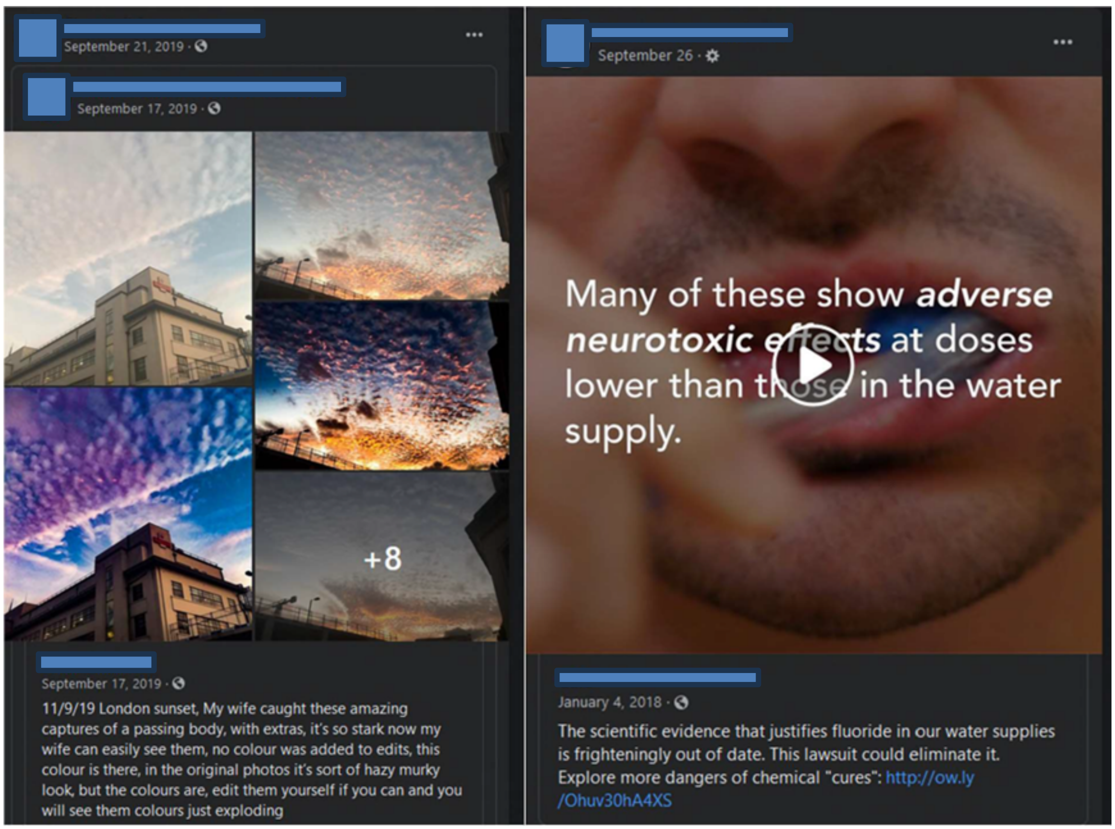
Conspiracy

A Conspiracy community is a Facebook page that promotes or discusses fringe or extreme theories based on unfounded claims that covert actors are responsible for events or circumstances.

This category is dominated by two main conspiracies: fluoride in water and chemtrails. The fluoride conspiracy theory claims governments use it to control the population, lower individuals' intelligence levels, affect fertility levels, and cause health problems. The nomenclature tends to include terms such as "fluoride free" or "clean water". Chemtrail conspiracy communities share the idea that planes/aircraft under the direction of the government and shadowy organizations are spraying chemicals in the sky to affect the health and mental capacity of the population below. Although these are two different conspiracies, they share a similar theme of malign actors controlling the population through chemical poisoning.

While there are many Conspiracy communities, they tend to be small in terms of users. The names of these communities tend to include geographical references, such as "Fluoride Free Kansas" or "Chemtrails Global Skywatch of Oklahoma", which perhaps limit the potential user base.

Nonetheless, these users tend to hold extreme views, which makes them susceptible to other conspiracies. These pages might also serve as a starting point for radicalizing users toward more extreme views.



Supplemental Figure 6. Examples of Conspiracy pages

<u>GMO</u>

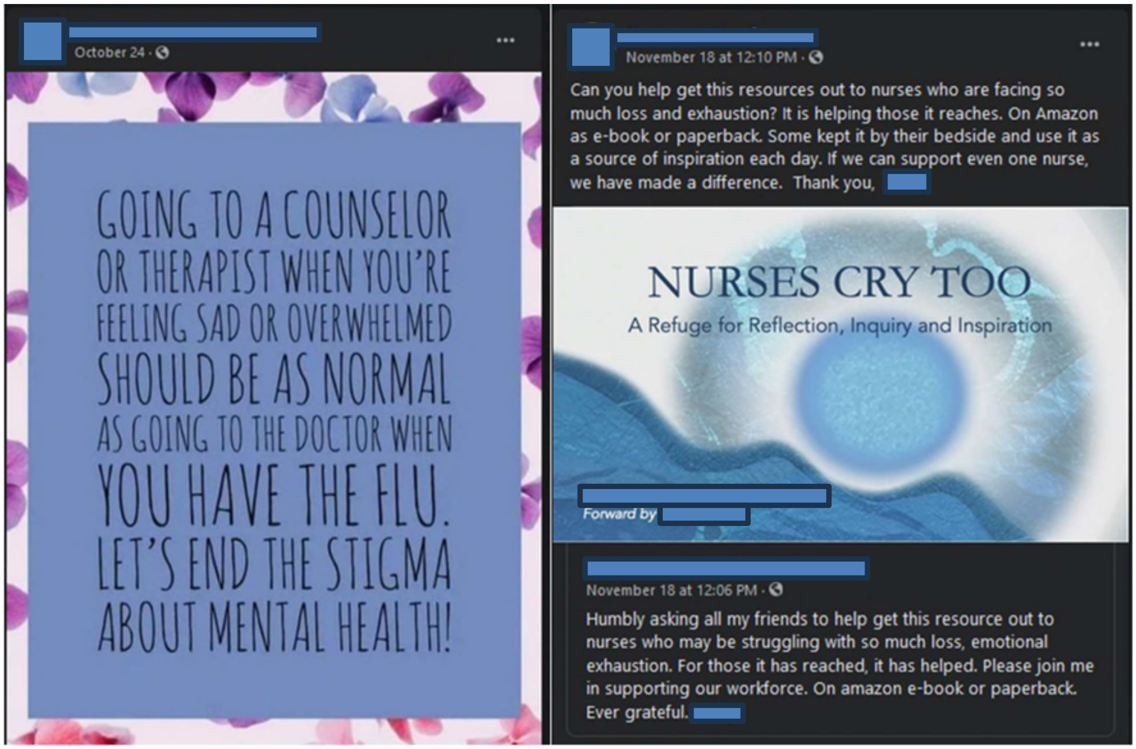
A GMO community is a Facebook page that debates or is against the use of genetically modified organisms in food and medicines. Posts generally attempt to raise awareness of products that contain GMOs and argue that these cause harmful effects. These communities often call for a boycott of certain brands or products and/or a requirement to label GMOs. Users often focus on Monsanto as the main antagonist; the company is often named in posts and the names of communities.

There are relatively few GMO communities, and they tend to be small in terms of users. The names of these communities tend to include geographical references, such as “March Against Monsanto Fort Meyers” or “GMO Free Canada”, which perhaps limit the potential user base.

Supplemental Figure 7. Examples of GMO pages

<u>Health</u>

A Health community is a Facebook page that discusses general health matters and medical institutions. In contrast to the AltHealth communities, the Health communities focus on traditional medicine, practitioners, and institutions. These include clinics, pharmacies, mental health services, medical staff, and health initiatives. The average Health community has relatively few users.

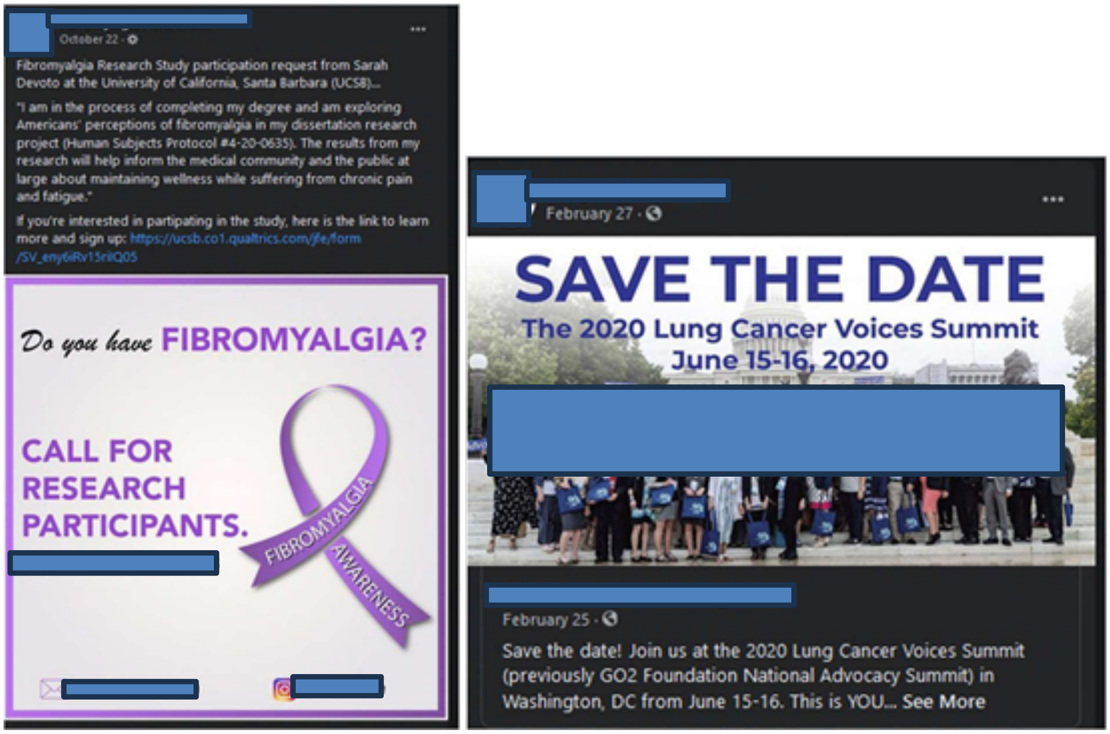


Supplemental Figure 8. Examples of Health pages

<u>Illness</u>

An Illness community is a Facebook page that aims to raise awareness of, discuss, or serve as a support group for certain medical and mental illnesses/conditions. Fibromyalgia, cancer, and HIV/AIDS play a significant role in this category, but by far the most discussed condition is autism.

Due to the serious nature of these medical conditions, the role of these pages as support groups leads to an interesting dynamic. This is the most common category of green cluster in the network, perhaps because they tend to narrowly focus on a specific illness. Nonetheless, the total number of users in this category is toward the middle of the pack, which means that the average number of users per Illness community is relatively small. These are cohesive communities, featuring discussions that are highly salient to a small number of users. They have relatively active discussions and are made up of users who have the condition or who have family and friends who are affected.

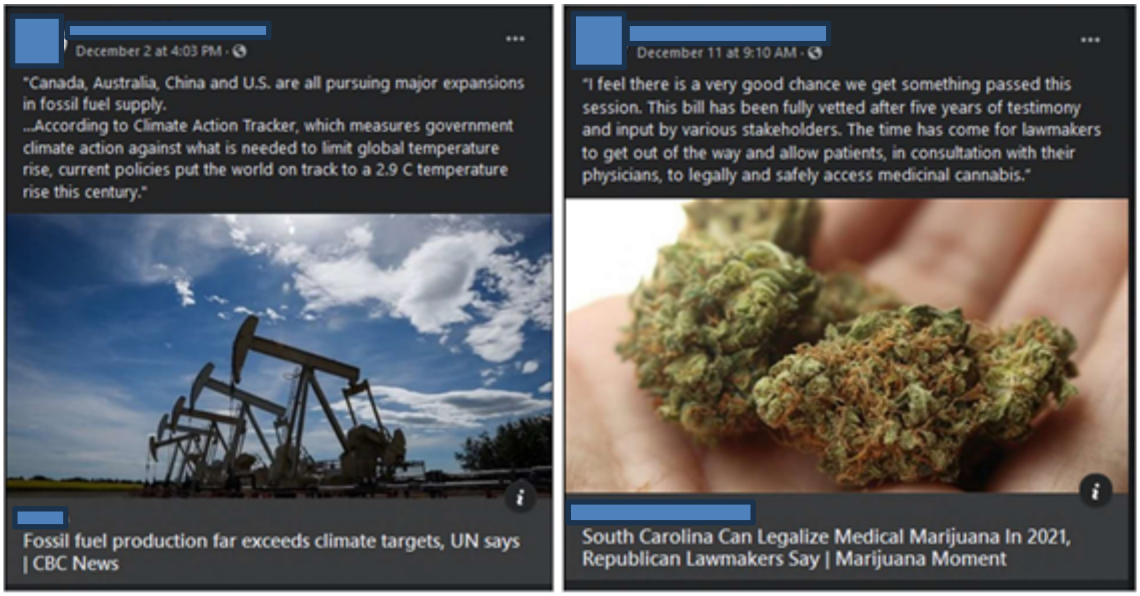


Supplemental Figure 9. Examples of Illness pages

<u>Movement</u>

A Movement community is a Facebook page that advocates for a specific cause or political objective. The main topics include cannabis legalization, universal equality, domestic violence/human trafficking victim protection, the environment, and mental health awareness.

In terms of both total number of nodes and user size, Movement is the second largest. The wide variety of causes tends to increase the total user size, but not all communities agree with each other. Some even hold opposing views on specific issues. Two of the three largest communities advocate for the legalization of cannabis. Victim protection and awareness is also a common theme among many in this category, including victims of bullying, domestic violence, human trafficking, child abuse, and mental health issues.

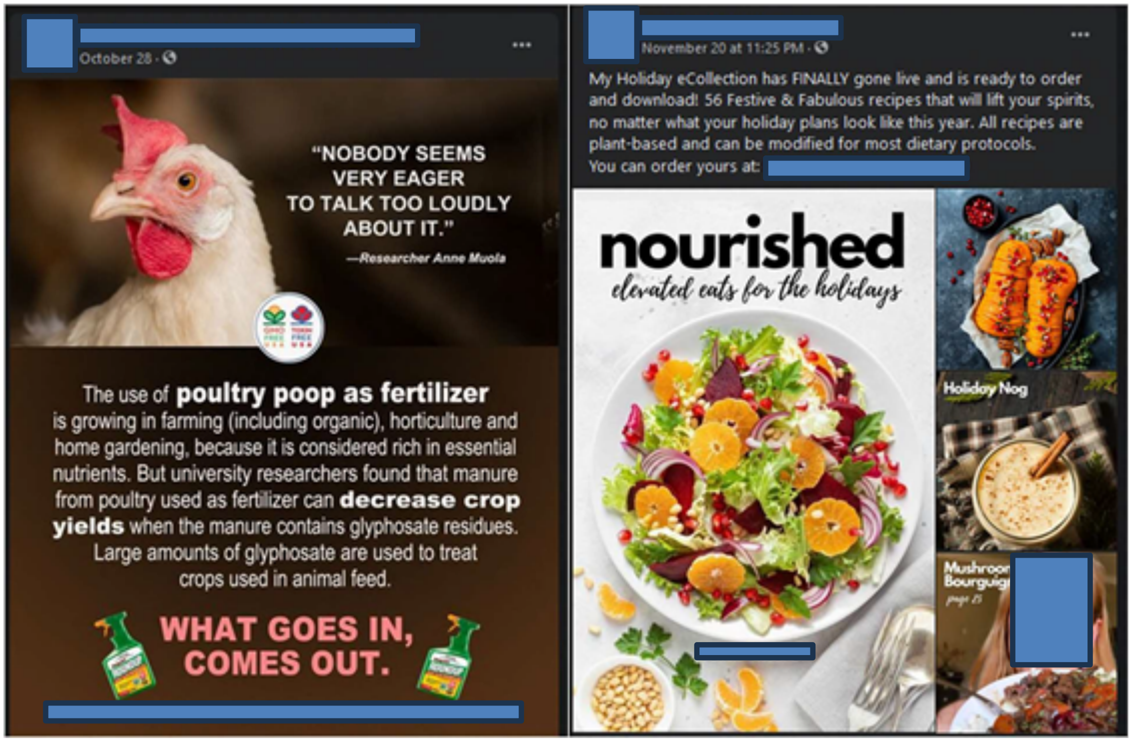


Supplemental Figure 10. Examples of Movement pages

Organic

An Organic community is a Facebook page that promotes an organic diet and lifestyle. Users tend to share recipes, diet plans, and information about particular organic ingredients. Growing your own food is an important activity that most of these pages advocate.

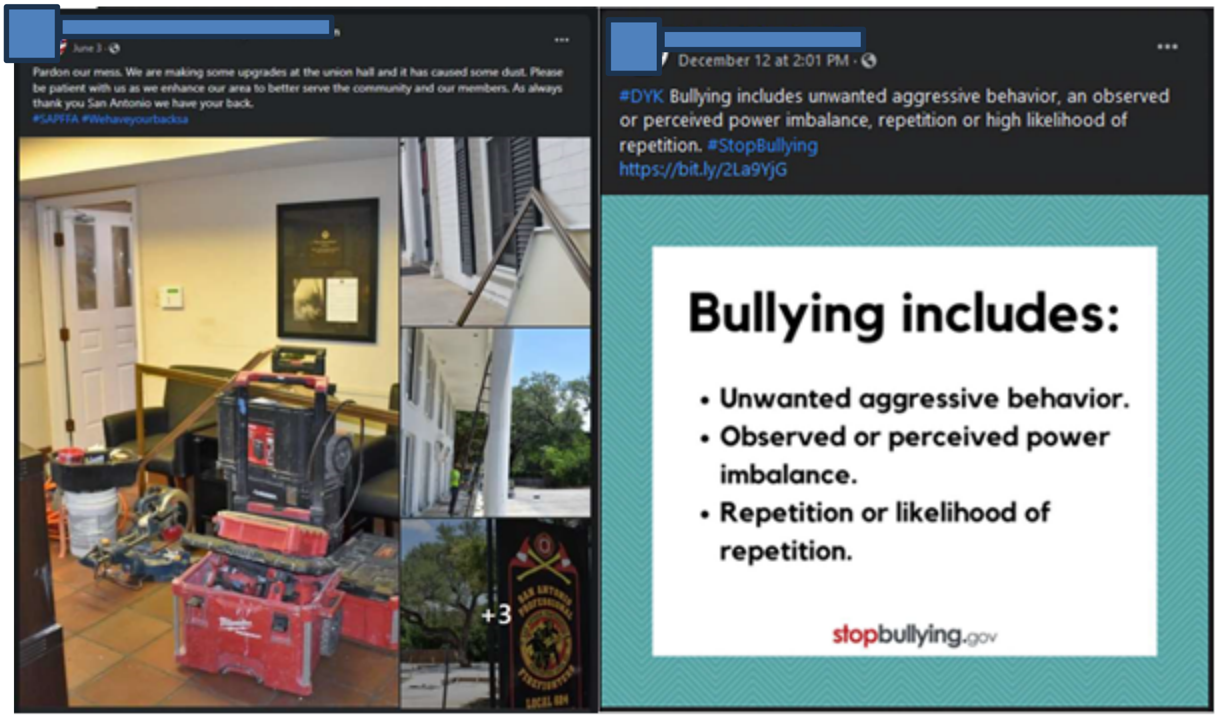
While this category is the smallest in terms of the number of communities, in terms of total users it is twice as large as the GMO and Health categories, indicating that the Organic communities are fairly popular.



Supplemental Figure 11. Examples of Organic pages

<u>Organization</u>

An Organization community is a Facebook page focused on a formal institution, including both government and non-governmental institutions. Communities include the U.S. government's Stop Bullying program, foreign government institutions, assistance projects, TEDx Change, and county initiatives such as "Get Healthy Knox County" and "Pathways Connect Central". The largest Organization communities are news organizations, such as the American Independent and LGBT News.



Supplemental Figure 12. Examples of Organization pages

<u>Parent = "Parenting"</u>

A Parent community is a Facebook page that discusses or offers advice and support for parenthood. The most commonly discussed issues include parental rights, breastfeeding, homeschooling and birthing, and raising children with special needs. Although there are communities focused on issues particular to the role of fatherhood, most of these communities seem to be focused on motherhood and frequented mostly by women.

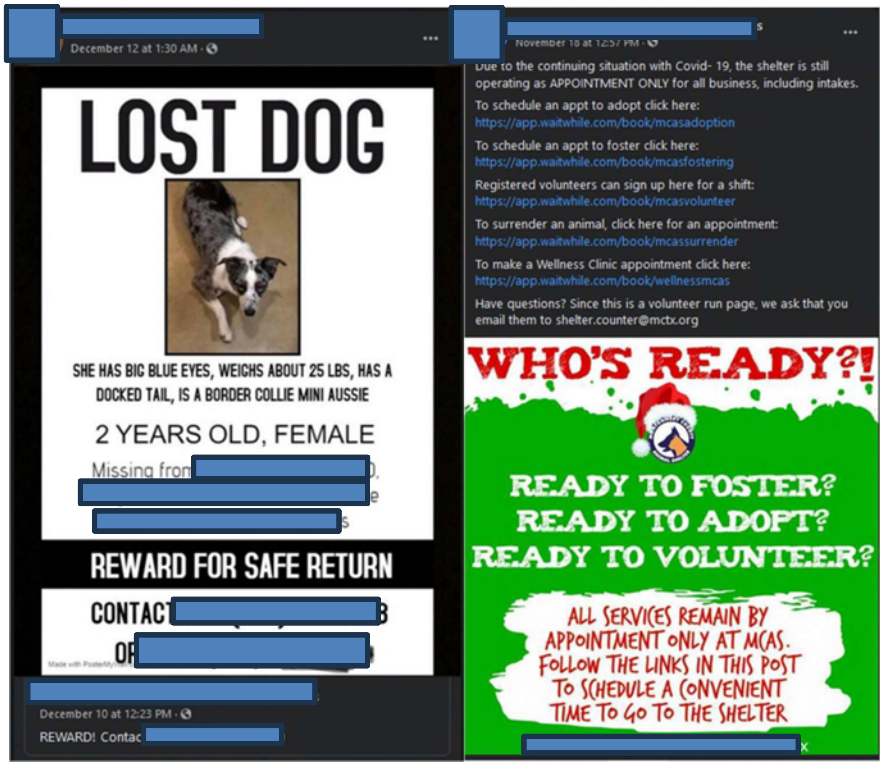
There are relatively many Parent communities, and they tend to be relatively large, making this one of the most important substantive categories in the network.

Supplemental Figure 13. Examples of Parenting pages

<u>Pet</u>

A Pet community is a Facebook page that is centered around pets (typically dogs and/or cats). The category is almost exclusively made up by shelters, pet rescue/adoption, and lost-and-found organizations. A substantial share of these clusters are based in Texas.

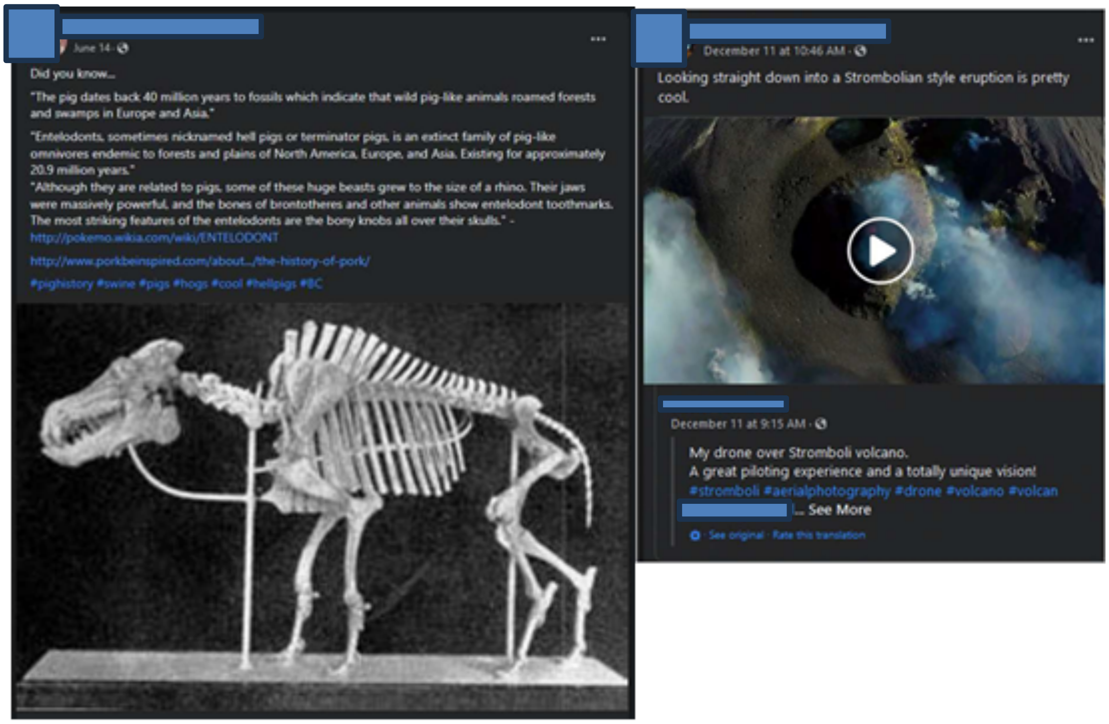
There are relatively many Pet communities, and they tend to be relatively large, making this one of the most important substantive categories in the network.



Supplemental Figure 14. Examples of Pet pages

<u>Other</u>

Some communities did not fit in the categories we created, and thus we classify them as Other. These include communities focused on specific churches, spiritual issues, farming, meme sharing, and community-building. The largest of these is an “earth lover” page, which shares photos of beautiful landscapes and natural features. There are also two large spiritualism communities that tend to share inspirational sayings and general spiritual content.



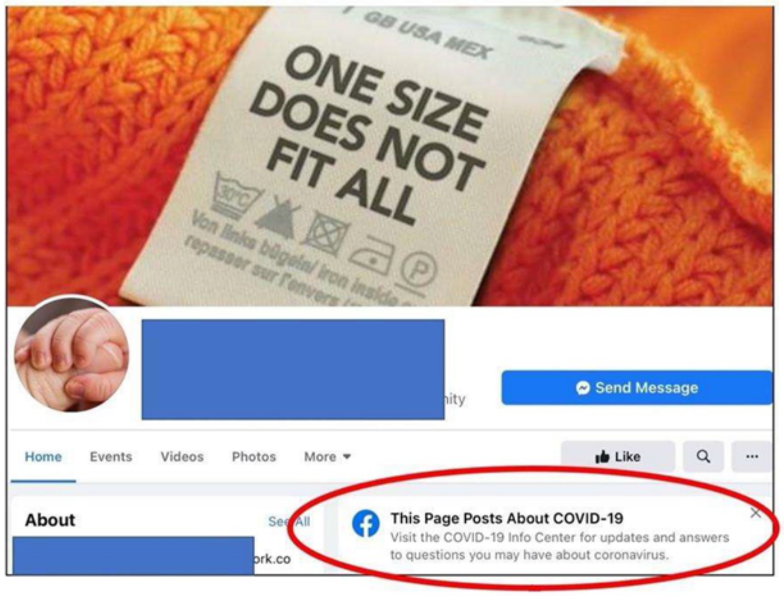
Supplemental Figure 15. Examples of Other pages

X

X are communities that did not have an interest focus, or it was too ambiguous. They are a minor, unimportant category in our study.

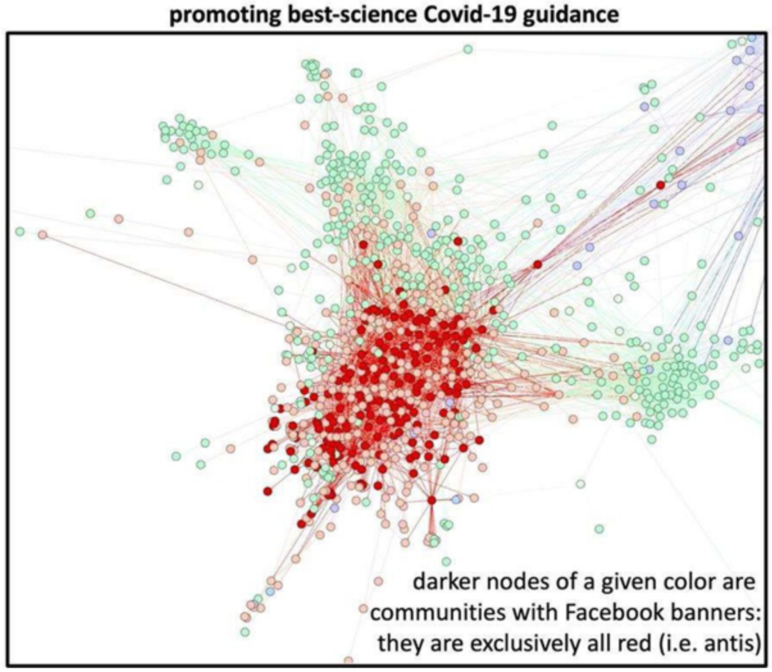
<u>Section 5:</u> Example of Facebook banners promoting best-science Covid-19 guidance. Positions in network of the nodes that receive Facebook banners promoting best-science Covid-19 guidance.

Example of Facebook banners promoting best-science Covid-19 guidance



Supplementary Figure 16. Example of Facebook banner promoting best-science COVID-19 guidance

Positions in network of the nodes that receive Facebook banners promoting best-science Covid-19 guidance



Supplementary Figure 17. Category and location of nodes (i.e., Facebook pages) that received a banner promoting best-science COVID-19 guidance.

<u>Section 6:</u> Topic filter. System in June 2025 without node labels filtered by topic.

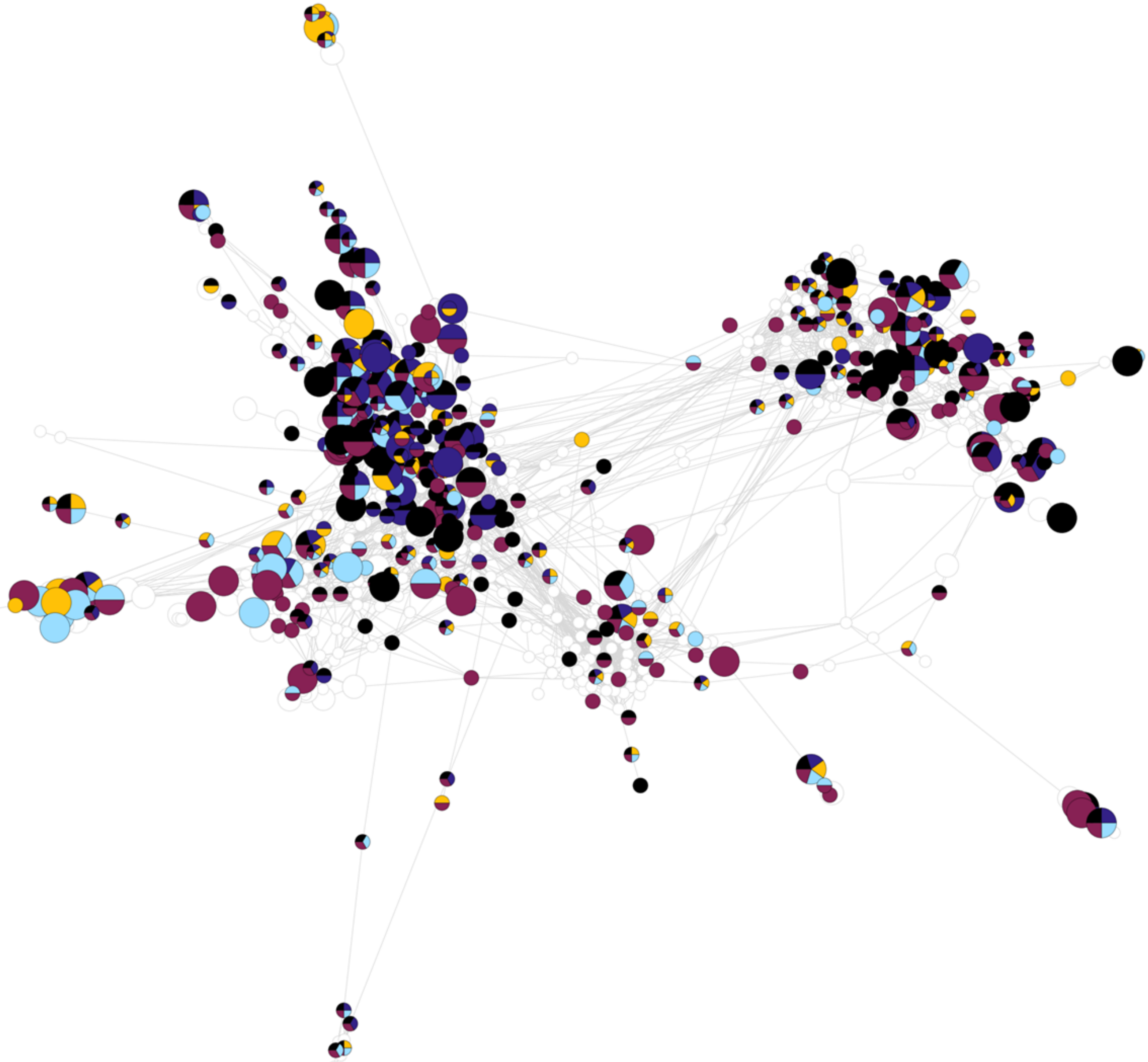
To identify COVID-19, mpox, abortion, elections, and climate change dialogues, we developed word filters that combined post content, post descriptions, image descriptions, and link text, while ignoring cases. For each topic, we created a list of search terms, including equivalents in other languages, such as “aquecimento global”, “globale erwärmung”, “calentamiento global”, “calentamiento mundial”, “réchauffement climatique”, “riscaldamento globale”, “riscaldamento climatico”, and “surriscaldamento climatico” for "climate change". We used regular expressions to catch misspellings, punctuation, and non-English languages. For example, “corona virus” became “(c|k|[(])+orona(no|[[:punct:]]|\\s){,4}(virus\\>|vírus)”. Additionally, we included emojis, specifically the monkey emoji, as often users would attempt to self-censor or evade detection by employing the usage of such terms as “🐒pox”. Word filters were augmented via searches on various social media websites, and identifying how users may try to self-censor their posts. Our flexible approach thus captured deliberate misspellings and word replacements, ignored punctuation, dealt with added spaces in-between the words to avoid filters, and other languages.

Our word filters have limitations. Although we can include various languages, we are restricted to those our researchers can read, potentially resulting in missed insights from other languages like Bengali, Hindi, Indonesian, Vietnamese, and Arabic. Furthermore, machine word filters have a "Scunthorpe problem" as they struggle to understand words in context, which makes it challenging to create a filter that can determine the intent of the text. For example, early in the coronavirus pandemic, “China flu” was coded language to refer to COVID-19, however a filter for “China flu” will also determine posts discussing “China flu season” or “Chinese flu vaccine” as containing COVID-19-related content. Thus, developing such a filter that can determine the intent of the human behind a social media post is a complex task that requires further research, and best left as the subject of another paper.

Supplementary Fig. 1A shows a portion of the network, with examples of local and global community names enclosed in boxes. Global or worldwide usernames are boxed in purple, while countries are boxed in green, states in blue, cities in red, and regions like counties or neighborhoods in orange. This represents the large online universe that consists of over 4 billion Facebook users, from which we sampled the communities that actively post about one or more of the topics being studied. Supplementary Fig. 1B is a magnification of the giant connected component of this system, consisting of 100 million users and only those communities that are actively posting. The communities are shown as nodes colored according to their vaccination debate status, with vaccine categories labeled next to where that category is primarily clustered. The size of the nodes indicates the geographic global/local aspect of the community's username, with larger circles representing local communities and smaller circles representing global communities. We can see that the anti (red) and pro (blue) communities are primarily clustered on different poles in this network, with many neutral communities (non-red and non-blue) clustering near the antis. By examining Supplemental Fig. 1B, we can quickly see that there are only a couple of local Illness nodes, the Anti core is predominantly comprised of local nodes, local Conspiracy nodes cluster together, and there are fewer local Parenting nodes when compared to global Parenting nodes. A detailed breakdown of the system at three different timepoints can be found in SM Section 2.

Pie charts can be used to illustrate the topics of engagement for each node in the network represented in Supplementary Fig. 1B, and the complete network is displayed in SM Section 6. Figs. 1C–G provide close-ups of various segments. Supplementary Fig. 1C zooms in on an outcropping of Anti nodes, where "abortion" is a minor topic compared to other topics. In Supplementary Fig. 1D, which focuses on the Illness communities, only 3 out of 36 communities in the neutral category are local, while 36.4% of antis are local. Interestingly, climate change is the most popular topic among the magnified Illness communities (88.9%). Supplementary Fig. 1E magnifies pro groups and reveals that only 24.1% of communities are local, and there is a good mix of topics. Similarly, Supplementary Fig. 1F zooms in on parenting communities, where a variety of topics are also mixed. In Supplementary Fig. 1G, portions of conspiracy groups are magnified, and surprisingly, the topics of abortion and election appear to be very popular in this cluster, of which 50% are local. Nearby GMO groups also have a high percentage of local groups (53.8%). While 25.2% of communities are local in the entire dataset, 31.7% of communities are local in the network of only content producers in Supplementary Fig. 1B. COVID-19 and climate change are the most discussed topics among pros, at 75.9% and 60.2%, respectively, while mpox, abortions, and elections are discussed by 42.6%, 28.7%, and 25.9%, respectively. Examining the network in this way can provide insight into how the topics of engagement relate to the community itself and its position in the system and suggest that local communities drive potentially divisive discourse in unexpected ways.

Even though Facebook has removed some large and active extreme communities (nodes), operationally the ecology remains largely unchanged because of a hidden self-repair effect. [84] shows there was negligible impact on the subnetwork of top 20 nodes ranked by highest betweenness centrality, which is a measure of a node's (community's) ability to act as an efficient conduit of content. Red nodes (anti-vaccination communities) are the most prominent and highest ranked within these top 20: while many were removed by Facebook, there was a system rewiring such that other red nodes (anti-vaccination communities) took their place. This means that there is a self-repairing core 'mesh' of extreme anti-vaccination communities within the full ecology that have—and can continue to—share and distribute extreme content not only with each other but also with the other nodes in the full ecology, including the millions of users in mainstream neutral communities with whom they are connected, e.g., parenting communities.



Supplementary Figure 18: As in Fig. 2 in the main paper, the system in June 2025 is here filtered to only display nodes that are currently live, but may not discuss one of the five topics (COVID-19, mpox, abortions, elections, and climate change; non-discourse engaged nodes are represented with white circles). The nodes are now color-coded according to the topics they cover.

<u>Section 7:</u> Chi-square test for topic-glocality. Top 20 nodes by betweenness centrality. Statistical results for fan count, topic count, and betweenness centrality scores.

Chi-square test for topic-glocality.

Here we set out to answer the question of whether there is a relationship between the 5 different topics of study and the geographic local/global nature (i.e., geographic glocality) of the communities discussing them. To do this, we employed the chi-square test where:

H0: There is no association between the topics and geographic glocality, i.e., topic is independent of geographic glocality. The phenomenon in the population under statistical test is considered absent;

H1: There is some association between the topics and geographic glocality, i.e., topic is dependent upon geographic glocality.

In this way we can determine if there is a topic--geographic-glocality relationship. In this case, the row variable is geographic glocality, related to whether a community specifies a location in its username, or if the username has a more "global" focus, and column variables are the different combinations of topics, e.g., "COVID-19", "COVID-19 and mpox", etc. The result was statistically significant at the 0.0025 significance level, $\chi 2$ (28, N = 767) = 69.6, p = .0000214. To investigate the degree to which this association is present, we calculated the effect size, Cramér's V, which was .302, and can be considered moderate. Of course, Cramér's V, can be a biased estimator as it increases with the number of cells, thus after applying a bias correction, Cramér's V = .234. This effect size can be considered small, but it is important to remember that oftentimes effects are likely to be small due to the subtlety of the issues ¬involved—in this case, it's reasonable to assume that there may be a trivial or negligible effect between some of the topic combinations and geographic glocality, and a large effect between others, and .2 is not so small as to be negligible. It is important to keep in mind that all possible combinations of the 5 topics results in 31 different column variables, and the results of the chi-square test might be different had the column variables only been the 5 topics, had all these topics been mutually exclusive categories, with no overlap. As it is, as each observation is not categories to one and only one category, it is not possible to perform the chi-square test to determine if there is an association between only the 5 topics and glocality.

Using the results of the chi-squared test, we can compare the observed values to the expected values: for communities that only discussed 1 topic, for example, 25.21%, 35.14%, 17.65%, 38.71%, and 24.32% of communities discussing COVID-19, mpox, abortion, elections, and climate change, respectively, are local. This is in contrast with the expected percentage, which for all is 31.9%. Thus, while one's expectation that discussions surrounding elections would be highly localized, one might also expect the same of abortion, due to laws, providers, etc. We see however that abortion discussions are instead rather global and deviate the most from the expected value. We also see a difference between values for COVID-19 and mpox, whereas one might have expected them to be closer as this dataset was initially gathered pre-coronavirus pandemic regarding vaccine. One might have further expected COVID-19 discussions to be more localized, due to countries having different isolation, masking, and vaccination requirements, however this is not the case.

Supplementary Table 4. Top 20 nodes by betweenness centrality

Glocality	Fan Count	Vax Label	Vax Subcategory	COVID-19	Mpox	Abortion	Elections	Climate Change	Topic Count	Betweeness Centrality
global	15198	Anti Vaxxination	Anti Vaxxination						0	0.069378
local	23479	Anti Vaxxination	Anti Vaxxination	TRUE	TRUE	TRUE	TRUE	TRUE	5	0.027405
global	194013	Pro Vaccination	Pro Vaccination	TRUE	TRUE		TRUE	TRUE	4	0.023503
global	1520145	Pro Vaccination	Pro Vaccination	TRUE	TRUE	TRUE	TRUE	TRUE	5	0.022053
local	4378	Anti Vaxxination	Anti Vaxxination	TRUE	TRUE		TRUE		3	0.070424
local	4819	Anti Vaxxination	Anti Vaxxination	TRUE	TRUE	TRUE	TRUE		4	0.02587
local	10063	Anti Vaxxination	Anti Vaxxination	TRUE	TRUE		TRUE	TRUE	4	0.021685
global	9809	Anti Vaxxination	Anti Vaxxination	TRUE					1	0.020014
global	12745	Pro Vaccination	Pro Vaccination	TRUE	TRUE	TRUE	TRUE	TRUE	5	0.056072
global	231768	Pro Vaccination	Pro Vaccination	TRUE	TRUE	TRUE	TRUE	TRUE	5	0.021941
global	35447	Anti Vaxxination	Anti Vaxxination						0	0.04317
global	5762	Pro Vaccination	Pro Vaccination						0	0.025182
global	4700	Pro Vaccination	Pro Vaccination	TRUE	TRUE	TRUE	TRUE	TRUE	5	0.018195
global	1832	Anti Vaxxination	Anti Vaxxination			TRUE			1	0.019351
global	150738	Neutral	Neutral, Parent		TRUE		TRUE	TRUE	3	0.025439
global	53241	Neutral	Neutral, Parent						0	0.018388
global	38243	Neutral	Neutral, Movement						0	0.036284
global	2583	Anti Vaxxination	Anti Vaxxination	TRUE	TRUE	TRUE	TRUE	TRUE	5	0.03458
global	1620	Neutral	Neutral, Illness						0	0.031894
global	428	Pro Vaccination	Pro Vaccination					TRUE	1	0.0372

Statistical results for fan count, topic count, and betweenness centrality scores.

Two different statistical tests were performed using the R functions \`cor.test\` and \`aov\`. \`cor.test\` is the test for association/correlation between paired samples, using one of Pearson's product moment correlation coefficient [85]. \`aov\` is the test for fitting an analysis of variance model by a call to \`lm\` for each stratum [86]. \`cor.test\` is thus performs a correlation test to evaluate the association between two or more variables, whereas \`aov\` performs a one-way analysis of variance (ANOVA, also known as one-factor ANOVA), and is an extension of independent two-samples t-test for comparing means in a situation where there are more than two groups.

The results for the statistical tests are below:

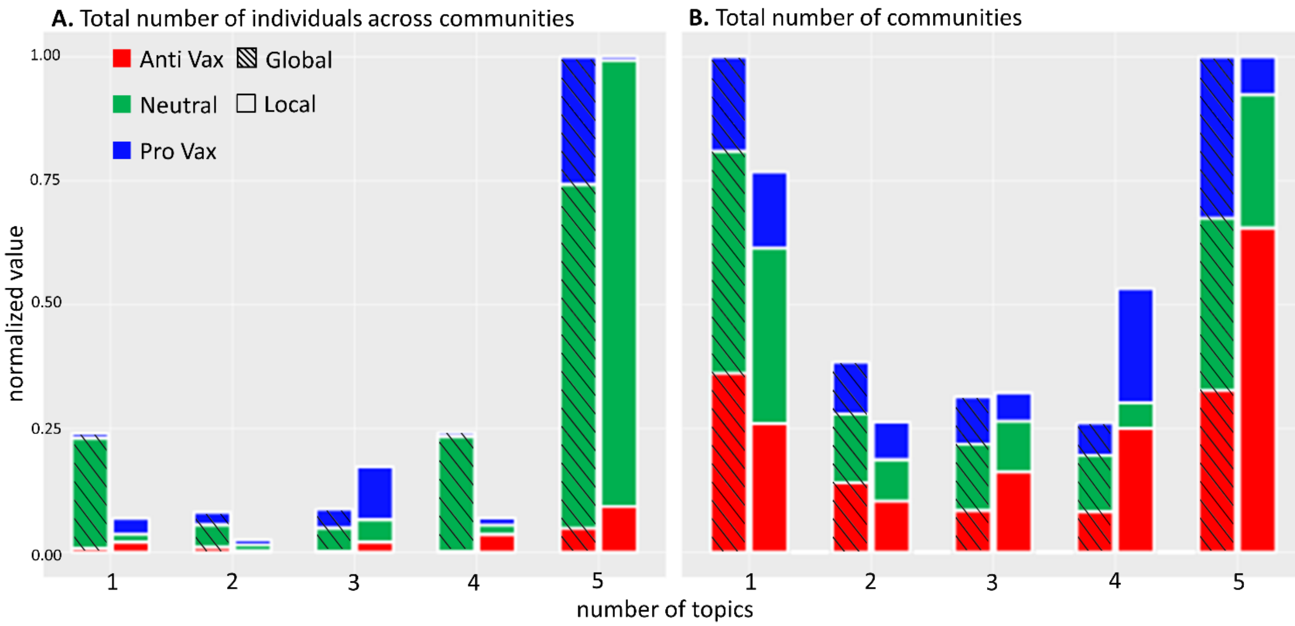
Supplementary Table 5. Pearson's product-moment correlation results

data	t	df	p-value	H1	95 percent CI	Sample estimates
fan count, betweenness centrality	0.08	1097	0.9	true correlation is not equal to 0	-0.0567, 0.0616	cor: 0.00249
topic count, betweenness centrality	4	1097	6e-05	true correlation is not equal to 0	0.0623, 0.1789	cor: 0.121
fan count, topic count	4	1097	6e-05	true correlation is not equal to 0	0.0625, 0.1791	cor: 0.121

Supplementary Table 6. one-factor ANOVA results

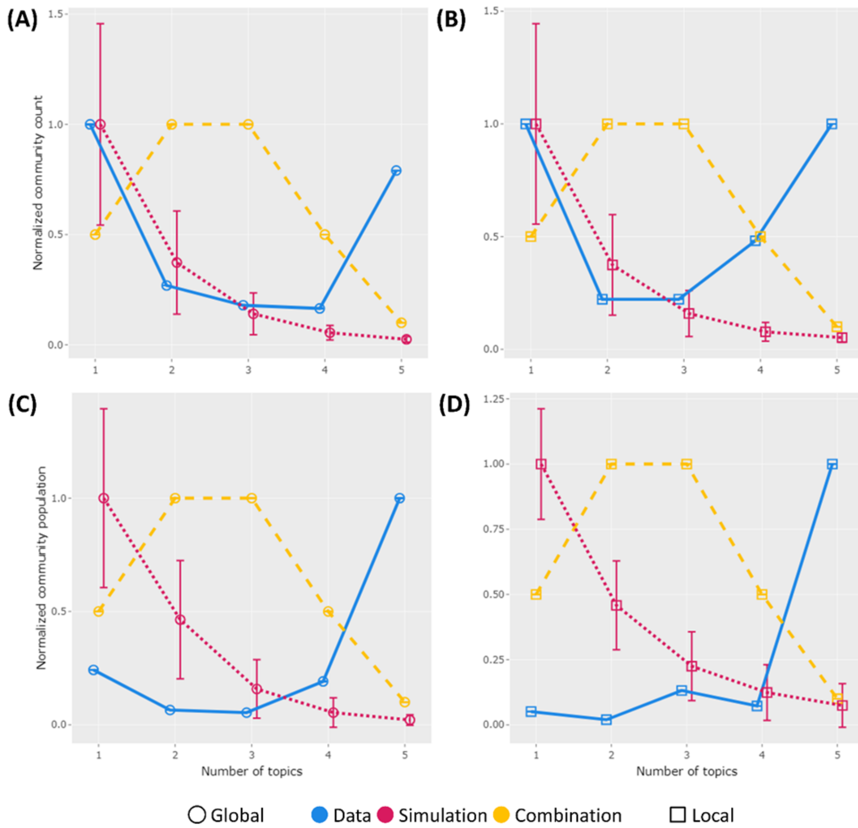
dependent variable	factor variable	residuals	df	sum sq	mean sq	F value	Pr (>F)
betweenness centrality	topic count		5	0.00053	1.07e-04	4.14	0.00099
		residuals	1093	0.02819	2.58e-05		
fan count	topic count		5	9.43e+12	1.89e+12	4.53	0.00043
		residuals	1093	4.55e+14	4.16e+11		

<u>Section 8:</u> Relationship between the number of topics and glocality. Topic heatmaps.



Supplementary Figure 19. The numbers of topics discussed by (A) individuals, and (B) communities. Categorized by their geographic scale and stance in the vaccine debate. Global communities are shaded. Results normalized by the maximum value. SI Sect. 9 confirms that these patterns do not arise by chance.

To explore the correlation between vaccine category, geographic glocality, and topic glocality, we also employ another approach: to categorize topic glocality into various subsets, i.e., we explore the various subsets of topics that exist, from having just one topic to all five. There are five different subsets for one element, ten for two, ten for three, five for four, and one for five. For instance, there is only one way to talk about all five topics, but five ways to discuss only one of the five topics. We started by determining the number of communities discussing 1 to 5 topics, and then examined the vaccine and geographic breakdown of these communities. Once we identified the communities and their topic discussions, we could then tally the number of individuals across communities, allowing us to understand the distribution of individuals in this system. However, it is worth noting that the dynamics of communities versus individuals can be vastly different, as some communities have tens of members while others have millions, and this is reflected in the data presented in Figure 2. In particular, Supplementary Fig. 26B, which focuses on the number of communities, shows a general "U" shape with regards to geographic glocality and vaccine category breakdowns. Interestingly, we found that local anti-communities were proportionally more likely to discuss all five topics, which may suggest a link between geographic locality and anti-vaccination sentiment. However, when examining Supplementary Fig. 26A, we observe a complete reversal of this trend. This is because in our dataset, neutral communities are extremely popular, with millions of followers. Thus, although local anti-communities may be actively involved in all 5 topics, the individuals in global and local neutral communities are heavily embedded in 5-topic discussions. In addition, we notice that the overwhelming majority of individuals across communities are involved in 5-topic discussions, unlike the community breakdown shown in Supplementary Fig. 26B. Furthermore, while there are few individuals in global pro-communities contributing to 5-topic discussions, the same cannot be said for local pro-communities.

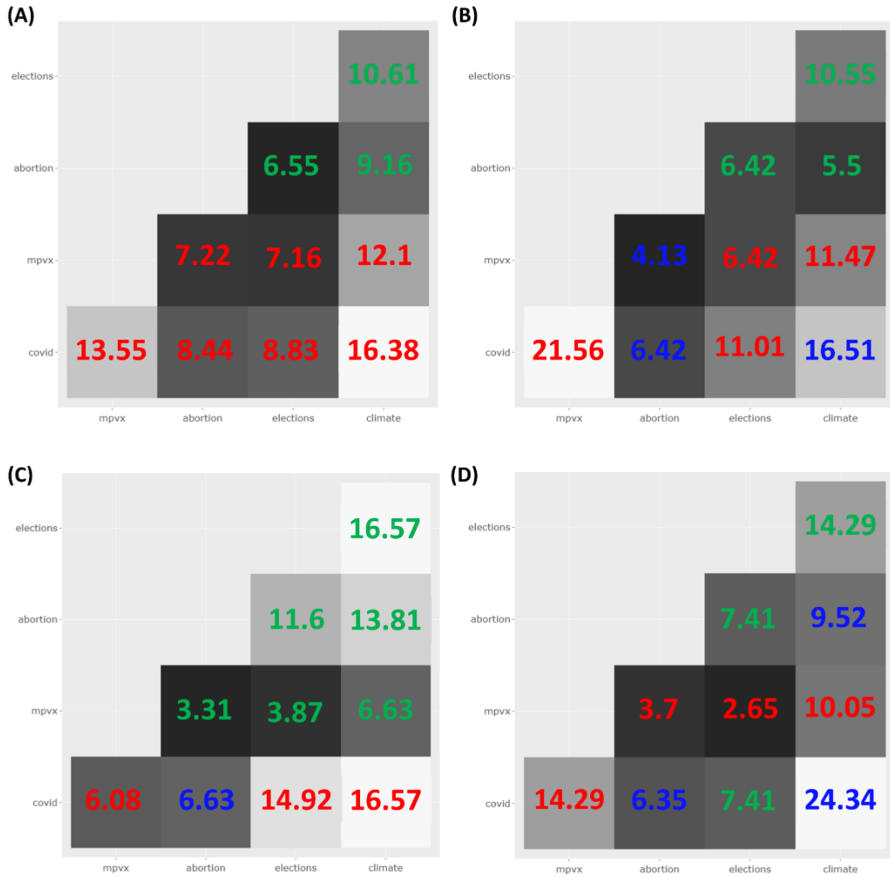


Supplementary Figure 20. Glocality of communities and populations compared to estimates and simulation. Blue solid line for data, magenta dotted line for simulation, and yellow dashed line for estimate. Plots on left are for global data whereas the right are for local data. Simulation is based on the idea that the number of communities discussing x topics is based on the vaccine type of the communities, and has a decreasing, concave up shape in every case. Estimate is based on the idea that the number of communities discussing x number of topics is based on combinations. For example, for communities discussing only one topic out of 5 possible topics, we would expect C(5,1) communities posting about only one topic, or 0.5 when normalized through division by maximum. (A) Comparison of global communities (solid blue) discussing x number of topics (ranging from only 1 to all 5) to simulated (dotted magenta) and estimated (dashed yellow) values. The actual data curve has a "U" shape and the simulation deviates from data at 4 topics. (B) Comparison of local communities (solid blue) discussing x number of topics (ranging from only 1 to all 5) to simulated (dotted magenta) and estimated (dashed yellow) values. The actual data curve has a "U" shape and the simulation deviates from data at 4 topics. In contrast to the global data, the estimate only predicts the normalized number of communities for 4 topics. (C) Comparison of global populations (solid blue) discussing x number of topics (ranging from only 1 to all 5) to simulated (dotted magenta) and estimated (dashed yellow) values. The "U" shape has a lower left end compared to the right end for the actual data curve and the simulation predicts only the normalized population value for 3 topics. (D) Comparison of local populations (solid blue) discussing x number of topics (ranging from only 1 to all 5) to simulated (dotted magenta) and estimated (dashed yellow) values. The actual data curve has an increasing, concave up shape and how simulation predicts only the normalized population value at 3 and 4 topics.

One might have imagined that as a community tackles more topics, the number of potential flashpoints for internal disagreements increases, which in turn would result in fewer communities engaging in a greater number of topics. Hence the curves in Supplementary Fig. 26 would all be expected to decay with an increasing number of topics—yet the opposite happens in Supplementary Fig. 26. This may be because social media's ability to connect like-minded individuals who share similar beliefs isolates them from those with differing viewpoints [3] hence leading to a general "distrust" that transcends topics. This can result in an increasing number of communities engaging in an increasing number of topics.

Using combinatorics, we might explore if Supplementary Fig. 2 can be explained mathematically by saying that there are C(5,n) ways to engage with n topics, where C(5,n) is the number of ways to choose a sample n of topics from the 5 topics in this study. Hence, out of 31 communities, one might expect there to be 5/31.31=5 who worry about 1 topic, 10 who worry about 2, 10 who worry about 3, 5 who worry about 4, and 1 who worries about 5. Since our dataset is larger than 31 communities, these values can be scaled accordingly. In order to compare the expectations of our simulation and combinatorics with the actual data, we can refer to Supplementary Fig. 27, which shows the data (solid blue) broken down into global and local communities (top row) and their populations (bottom row) on the left and right side, respectively. The simulation results are represented by a dotted magenta line, while the dashed yellow line represents our expectations based on combinatorics. To make the datasets comparable, all of them have been normalized by their maximum value. The simulation assumes that communities discussing topics are a random sample based on their vaccine categories, with a community's vaccination status being the primary factor in determining its topic engagement. However, as shown in Supplementary Fig. 27, none of these expectations match the actual data except in specific circumstances.

In Figs. S20A–B, the simulation predicts engagement in 1, 2, and 3 topics, but deviates greatly for 4 and 5 topics. The only case where the simulation matches the data is for 4 topics in Supplementary Fig. 20B. Differences between global and local trends for both communities and community populations are subtle, but for Supplementary Fig. 20A vs Supplementary Fig. 20B, 4 topics deviate more quickly and dramatically for local communities. Proportionately, more local communities discuss 5 topics than global. Both data curves have a "U" shape. In Fig. S20C, the simulation only predicts 3 topics, while in Supplementary Fig. 20D, it predicts 3 and 4 topics. Global and local trends in community populations differ, with Supplementary Fig. 20C having a "U" shape and Supplementary Fig. 20D having an almost increasing, concave up shape. The largest populations engage in all 5 topics. These "U" shapes in Figs. S20 might be explained by our expectations: engagement in up to 3 topics decreases with increasing potential for disagreement, but after this point, the number of communities involved sharply increases. Hence, we are left with the following as a candidate explanation: it is because social media's ability to connect like-minded individuals who share similar beliefs, isolates them from those with differing viewpoints [3] hence leading to a general "distrust" that transcends topics. This can result in an increasing number of communities engaging in an increasing number of topics.



Supplementary Figure 21. Heatmaps showing prevalence of cross-topic discussions. Symmetric squares are left empty since "Topic A and Topic B" is equivalent to "Topic B and Topic A". Colored tiles indicate magnitude of communities discussing the paired topics, with black being the lowest and white the highest. Text color denotes the dominant vaccine stance within those discussions. (A) 1 May 2022 - 17 October 2022: Most common were COVID-19-climate change, COVID-19-mpox, and mpox-climate change. No cross-topics were pro-dominated. (B) 12-26 May around first U.S. mpox case on 19 May: Highest interest in COVID-19-mpox and COVID-19-climate change. Unlike overall trends, some COVID-19 pairings were not anti-dominated, while abortion was split between pro/neutral stances. (C) 17 June - 1 July around Roe v. Wade reversal on 24 June: Spikes in abortion-elections and abortion-climate change. Climate change pairings dominated. Only COVID-19-abortion was pro-dominated. (D) 3-17 October before student loan forgiveness rollout: Continued interest in COVID-19-climate change and COVID-19-mpox. COVID-19-abortion and elections-climate change also pro-dominated in this period.

To examine how communities discussed combinations of topics, we generated heatmaps at three key time points: (1) 12-26 May around the first U.S. mpox case; (2) 17 June-1 July around the reversal of Roe v. Wade; and (3) 3-17 October before U.S. student loan forgiveness applications opened, a salient election issue. This figure shows that over the previous full study period (1 May - 17 October) [50], COVID-19 and climate change interactions were common, with anti-communities prevailing in disease/virus conversations. However, narrowed snapshots revealed fluctuations. Around the first mpox case, more pro-communities discussed “COVID-19 and climate change”. In early October as well, this pairing was pro-dominated. While anti-communities led most mpox discussions in May, “mpox and abortion” tilts toward pro-communities. “COVID-19 and abortion” also skewed pro during the 2-week windows, unlike the overall 6-month trend. Thus, while anti/neutral communities generally dominated cross-topic discourse, pros occasionally controlled specific conversations over short timescales. Leveraging these breakthroughs may increase pro influence on the broader system.

<u>Section 9:</u> Details regarding the softening simulation in Figure 4.

The agent-based simulation presented in Figure 4 in the main paper represents a Monte Carlo analysis designed to answer a fundamental question: if we repeatedly form small neighborhood-based deliberation groups within the Facebook vaccine network and allow mixed-opinion groups to engage in persuasive dialogue, how quickly can systematic opinion moderation be achieved across the entire network? The simulation framework models the empirically observed softening effects from pilot studies [87,88] at network scale, tracking both the temporal dynamics of opinion change and the structural properties that enable large-scale transformation.

Simulation architecture and parameters

Each complete simulation run consists of 1,000 independent Monte Carlo trials to ensure statistical robustness. For each trial, we initialize several tracking arrays to capture different aspects of the conversion process: the number of mixed-opinion groups formed in each run, the initial counts of red (anti-vaccine) and blue (pro-vaccine) nodes requiring conversion, the number of time steps needed until complete network neutralization, the time points at which each opinion group drops below 50% of its original size, and the complete trajectory of opinion counts throughout each run.

The core behavioral parameter governing opinion change is the per-time-step conversion probability, derived from empirical pilot study results showing that approximately one-third of participants in mixed-opinion groups moderate their extreme views. This probability applies equally to both red and blue nodes, reflecting the empirical finding that the likelihood of opinion change depends on participation in diverse dialogue rather than initial stance. Each simulation trial uses independent random number generation with fixed seeds to ensure reproducibility while maintaining statistical independence across runs.

Network partitioning and group formation

The simulation begins by partitioning the entire Facebook vaccine network into small neighborhood-based deliberation groups. Starting with the complete list of network nodes, the algorithm randomly selects an unused node as a "pivot" for each new group. For each pivot, the algorithm identifies all nodes within graph distance 2 (meaning nodes that are either directly connected to the pivot or connected through one intermediate node), creating a local neighborhood that reflects the existing social fabric of the network rather than arbitrary groupings.

From this neighborhood, the algorithm randomly selects up to 4 additional nodes to join the pivot, creating deliberation groups of 1-5 members that preserve the network's natural clustering structure. This process continues iteratively, with each selected node marked as "used" to prevent double-assignment, until all nodes in the network have been assigned to exactly one group. The result is a complete partitioning of the network into small clusters, each centered on a randomly chosen pivot but constrained by the network's existing topology.

Critically, the algorithm stores only the opinion colors (red, blue, or green) of nodes within each group rather than their specific identities, focusing the analysis on opinion dynamics rather than individual node properties. This abstraction enables the simulation to model opinion change processes while maintaining computational efficiency across the large-scale network.

Mixed-group identification and debate mechanics

Following network partitioning, the simulation identifies which groups contain multiple opinion types, as only these "mixed" groups can facilitate cross-perspective dialogue and subsequent opinion change. Groups containing only red nodes, only blue nodes, or only green nodes are excluded from the conversion process, reflecting the empirical finding that homogeneous groups reinforce existing views rather than promoting moderation.

The remaining mixed-opinion groups constitute the "battleground" where opinion change occurs. The simulation records both the absolute number of such groups and their proportion relative to total groups formed, providing insights into how network structure affects the availability of cross-perspective dialogue opportunities. This step reflects the core insight that network topology determines not just who talks to whom, but which conversations have the potential to generate opinion change.

Temporal dynamics and conversion process

With mixed groups identified, the simulation tracks the initial counts of red and blue nodes participating in these groups, establishing baseline measures for subsequent conversion tracking. The temporal evolution then proceeds through discrete time steps, with each step representing one round of deliberation across all mixed groups simultaneously.

During each time step, every red and blue node still participating in mixed groups faces an independent probability trial for conversion to neutral (green) status. This probability trial follows a Bernourial distribution, where success indicates opinion moderation and failure indicates persistence of the current view. The independence assumption reflects empirical findings that individual conversion decisions, while influenced by group dynamics, ultimately depend on personal psychological processes that vary across participants.

The simulation continues this time-step process until both red and blue node counts reach zero, indicating complete network neutralization. At each step, the algorithm records the current counts of unconverted nodes, enabling detailed analysis of conversion trajectories and identification of critical transition points in the opinion change process.

Statistical tracking and output generation

Throughout each simulation run, the algorithm maintains comprehensive statistics to capture different aspects of the conversion process. For temporal analysis, it records the exact time step at which complete neutralization occurs, providing measures of intervention duration. For milestone tracking, it identifies the first time step at which each opinion group (red and blue) drops below 50% of its original size, indicating the pace of opinion change for different starting positions.

Most importantly, the simulation maintains complete time-series data for red and blue node counts throughout each run, enabling detailed analysis of conversion trajectories and identification of characteristic patterns in opinion change dynamics. This granular data supports both the aggregate statistics presented in Figure 4B in the main paper and more detailed analyses of conversion patterns across different network structures and parameter settings.

Methodological validation and robustness

The 1,000-trial Monte Carlo approach ensures that results reflect genuine network properties rather than artifacts of particular random groupings or conversion sequences. Each trial uses independent random number generation while maintaining identical behavioral parameters, allowing statistical confidence intervals to capture the range of outcomes possible under different random realizations of the group formation process.

The simulation framework validates its approach through several consistency checks: ensuring complete network partitioning (every node assigned to exactly one group), confirming that only mixed groups contribute to opinion change (homogeneous groups remain static), and verifying that conversion probabilities remain constant across time steps and node types (maintaining consistency with empirical findings).

This methodology enables systematic exploration of how network structure interacts with local dialogue processes to generate large-scale opinion change, providing quantitative predictions for the effectiveness of network engineering approaches while remaining grounded in empirically-observed behavioral parameters from pilot studies.

Simulation results and network transformation metrics

The Monte Carlo simulation results demonstrate consistent and robust opinion moderation effects across the Facebook vaccine network. Table 7 presents the key statistical outcomes from 1,000 independent simulation runs, while Table 8 illustrates the network composition changes achievable through different intervention strategies.

Supplementary Table 7. Statistical summary of network engineering simulation (1,000 runs)

Metric	Mean	Standard Deviation
Red conversion probability	7.03%	-
Blue conversion probability	7.03%	-
Mixed groups formed	76.96	4.33
Red nodes converted	87.55	8.73
Blue nodes converted	88.02	7.59
Time steps to full neutralization	37.52	2.68
Time steps to 50% red conversion	11.08	1.52
Time steps to 50% blue conversion	511.15	1.49

Supplementary Table 8. Network composition transformation (representative simulation)

Scenario	Neutral (%)	Anti-vaccination (%)	Pro-vaccination (%)
Original network	47.9	34.2	17.9
Average case (both convert)	65.4	26.5	8.1
Better case (anti-only convert)	55.6	26.5	17.9

The simulation results reveal several important patterns in network-scale opinion change. The conversion probabilities of approximately 7% per time step for both red and blue nodes reflect the empirical finding that roughly one-third of participants moderate their views in mixed-opinion deliberations, distributed across multiple rounds of engagement. The formation of mixed groups averages nearly 77 per simulation, indicating that the network's structure provides substantial opportunities for cross-perspective dialogue despite its polarized organization.

The temporal dynamics show remarkably consistent conversion patterns, with 50% conversion milestones achieved for both opinion groups around 11 time steps, while complete network neutralization requires approximately 38 time steps. This suggests that network engineering interventions could achieve systematic opinion moderation within weeks if implemented at scale. The network composition transformations demonstrate significant shifts toward neutrality, with the proportion of neutral nodes increasing from 48% to either 65% (average case) or 56% (targeted intervention), representing substantial moderation of the overall ecosystem while preserving the network's essential connectivity and social fabric.

References

1. Johnson NF, Velásquez N, Restrepo NJ, Leahy R, Gabriel N, El Oud S, Zheng M, Manrique P, Wuchty S, Lupu Y. The online competition between pro- and anti-vaccination views. *Nature* (2020) **582**:230–233. doi:10.1038/s41586-020-2281-1

2. Johnson NF, Zheng M, Vorobyeva Y, Gabriel A, Qi H, Velasquez N, Manrique P, Johnson D, Restrepo E, Song C, et al. New online ecology of adversarial aggregates: ISIS and beyond. *Science* (2016) **352**:1459–1463. doi:10.1126/science.aaf0675

3. Johnson NF, Leahy R, Restrepo NJ, Velasquez N, Zheng M, Manrique P, Devkota P, Wuchty S. Hidden resilience and adaptive dynamics of the global online hate ecology. *Nature* (2019) **573**:261–265. doi:10.1038/s41586-019-1494-7

4. Centola D, Becker J, Brackbill D, Baronchelli A. Experimental evidence for tipping points in social convention. *Science* (2018) **360**:1116–1119. doi:10.1126/science.aas8827

5. Sear RF, Leahy R, Restrepo NJ, Lupu Y, Johnson NF. Machine Learning Language Models: Achilles Heel for Social Media Platforms and a Possible Solution. *AAIML* (2021) **01**:191–202. doi:10.54364/AAIML.2021.1112

6. Constantino AK. Health officials confirm first U.S. case of monkeypox virus this year in Massachusetts. *CNBC* (2022) Available at: https://www.cnbc.com/2022/05/19/monkeypox-virus-case-confirmed-in-massachusetts.html [Accessed March 14, 2023]

7. Liptak A. In 6-to-3 Ruling, Supreme Court Ends Nearly 50 Years of Abortion Rights. *The New York Times* (2022) Available at: https://www.nytimes.com/2022/06/24/us/roe-wade-overturned-supreme-court.html [Accessed March 14, 2023]

8. OLCA. The President Signs H.R. 5376, the "Inflation Reduction Act of 2022." Available at: https://www.ssa.gov/legislation/legis_bulletin_081622.html [Accessed March 14, 2023]

9. 2022 Midterm Election Calendar - 270toWin. *270toWin.com* Available at: https://www.270towin.com/2022-election-calendar/ [Accessed March 14, 2023]

10. Cosenza V. World Map of Social Networks. *Vincos - il blog di Vincenzo Cosenza* Available at: https://vincos.it/world-map-of-social-networks/ [Accessed March 14, 2023]

11. Lampe C, Vitak J, Gray R, Ellison N. Perceptions of facebook's value as an information source. in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* CHI '12. (New York, NY, USA: Association for Computing Machinery), 3195–3204. doi:10.1145/2207676.2208739

12. Silver A, Matthews L. The use of Facebook for information seeking, decision support, and self-organization following a significant disaster. *Information, Communication & Society* (2017) **20**:1680–1697. doi:10.1080/1369118X.2016.1253762

13. Moon RY, Mathews A, Oden R, Carlin R. Mothers' Perceptions of the Internet and Social Media as Sources of Parenting and Health Information: Qualitative Study. *Journal of Medical Internet Research* (2019) **21**:e14289. doi:10.2196/14289

14. Pearson H. How to speak to a vaccine sceptic: research reveals what works. *Nature* (2025) **642**:289–291. doi:10.1038/d41586-025-01771-z

15. United Nations Children's Fund. *The State of the World's Children 2023: For every child, vaccination*. Florence, Italy: UNICEF Innocenti – Global Office of Research and Foresight (2023).

16. Ingram T, Sidhu S. New data indicates declining confidence in childhood vaccines of up to 44 percentage points in some countries during the COVID-19 pandemic. *UNICEF* (2023) Available at: https://www.unicef.org/rosa/press-releases/new-data-indicates-declining-confidence-childhood-vaccines-44-percentage-points-some [Accessed July 15, 2025]

17. Vashist K, Yankey D, Elam-Evans LD, Mu Y, Valier MR, Pingali C, Hill HA, Santibanez TA, Singleton JA. Changes in vaccine hesitancy among parents of children aged 6 months – 17 Years, National Immunization Surveys, 2019–2022. *Vaccine* (2024) **42**:125989. doi:10.1016/j.vaccine.2024.05.037

18. Gardner S. The 2025 measles outbreak is largest in 3 decades. *POLITICO* (2025) Available at: https://www.politico.com/news/2025/07/09/the-2025-measles-outbreak-is-largest-in-three-decades-00444186 [Accessed July 15, 2025]

19. U.S. Measles Cases Hit Highest Level Since Declared Eliminated in 2000 | International Vaccine Access Center. (2025) Available at: https://publichealth.jhu.edu/ivac/2025/us-measles-cases-hit-highest-level-since-declared-eliminated-in-2000 [Accessed July 15, 2025]

20. Yousif N. How Canada became the centre of a measles outbreak in North America. *BBC News* (2025) Available at: https://www.bbc.com/news/articles/c4g8d39gdr0o [Accessed July 22, 2025]

21. Ledford H. 154 million lives and counting: 5 charts reveal the power of vaccines. *Nature* (2025) doi:10.1038/d41586-025-00862-1

22. Heidt A, Ledford H. Vaccine sceptic RFK Jr is now a powerful force in US science: what will he do? *Nature* (2025) doi:10.1038/d41586-025-00439-y

23. Jones N, Witze A, Tollefson J, Kozlov M. What Trump 2.0 means for science: the likely winners and losers. *Nature* (2025) **637**:532–535. doi:10.1038/d41586-025-00052-z

24. Levine D, Douglas L. Exclusive: US CDC plans study into vaccines and autism, sources say. *Reuters* (2025) Available at: https://www.reuters.com/business/healthcare-pharmaceuticals/us-cdc-plans-study-into-vaccines-autism-sources-say-2025-03-07/ [Accessed July 15, 2025]

25. Woolf S, Rosenthal J. RFK Jr. Is Systematically Undermining Vaccine Science and Endangering Health. *Center for American Progress* (2025) Available at:

https://www.americanprogress.org/article/rfk-jr-is-systematically-undermining-vaccine-science-and-endangering-health/ [Accessed July 15, 2025]

26. Nunn CJB, Stolberg SG. Kennedy Turns to a Discredited Vaccine Skeptic for Autism Study. *The New York Times* (2025) Available at: https://www.nytimes.com/2025/03/27/health/rfk-jr-autism-vaccines.html [Accessed July 15, 2025]

27. Myers SL. Trump Administration Cancels Scores of Grants to Study Online Misinformation. *The New York Times* (2025) Available at: https://www.nytimes.com/2025/05/15/business/trump-online-misinformation-grants.html [Accessed July 15, 2025]

28. NIH cuts research grants on vaccine hesitancy—with mRNA vaccine studies the next target | CIDRAP. (2025) Available at: https://www.cidrap.umn.edu/anti-science/nih-cuts-research-grants-vaccine-hesitancy-mrna-vaccine-studies-next-target [Accessed July 15, 2025]

29. NIH Cuts Funding for Studies on Vaccine Hesitancy and Acceptance. *Infectious Disease Advisor* (2025) Available at: https://www.infectiousdiseaseadvisor.com/news/nih-cuts-funding-for-studies-on-vaccine-hesitancy-and-acceptance/ [Accessed July 15, 2025]

30. Nobel Prize Summit. *NobelPrize.org* Available at: https://www.nobelprize.org/events/nobel-prize-summit/2023/ [Accessed March 28, 2023]

31. NewsDirect. TRUST SCIENCE PLEDGE CALLS FOR PUBLIC TO ENGAGE IN SCIENTIFIC LITERACY. *News Direct* Available at: https://newsdirect.com/news/trust-science-pledge-calls-for-public-to-engage-in-scientific-literacy-737528151 [Accessed March 14, 2023]

32. Meyers C. American Physical Society Takes On Scientific Misinformation. Available at: http://aps.org/publications/apsnews/202203/misinformation.cfm [Accessed March 14, 2023]

33. Jilani Z. AAAS 2022 Annual Meeting: How to Tackle Mis- and Dis-information | American Association for the Advancement of Science (AAAS). Available at: https://www.aaas.org/news/aaas-2022-annual-meeting-how-tackle-mis-and-dis-information [Accessed March 28, 2023]

34. Beyond disinformation – EU responses to the threat of foreign information manipulation. (2023). Available at: https://www.youtube.com/watch?v=YJf2pZGe36Q [Accessed March 28, 2023]

35. Tharoor I. Analysis | 2024 brings wave of elections with global democracy on the ballot. *The Washington Post* (2024) Available at: https://www.washingtonpost.com/world/2024/01/03/2024-elections-us-eu-portugal-austria-bangladesh-india-south-africa-indonesia-mexico-taiwan/ [Accessed July 15, 2025]

36. Rajvanshi A, Serhan Y. A Make-or-Break Year for Democracy Worldwide. *TIME* (2024) Available at: https://time.com/6551743/2024-elections-democracy-trump-putin/ [Accessed July 15, 2025]

37. Berger C, Freihse C. Super election year 2024: A super year for disinformation? *Upgrade Democracy* (2024) Available at: https://upgradedemocracy.de/en/superwahljahr-2024-superjahr-desinformation/ [Accessed July 15, 2025]

38. Check RF. Fact Check: Fake CNN headline says all Democrats need latest COVID shot to vote. *Reuters* (2024) Available at: https://www.reuters.com/fact-check/fake-cnn-headline-says-all-democrats-need-latest-covid-shot-vote-2024-09-16/ [Accessed July 15, 2025]

39. Caulfield T. Politics and vaccine misinformation: A horrifyingly bad mix. *Healthy Debate* (2024) Available at: https://healthydebate.ca/2024/09/topic/politics-vaccine-misinformation/ [Accessed July 15, 2025]

40. Dunai M. From vaccines to votes: Russian ad agency influences Europeans. *Financial Times* (2024) Available at: https://www.ft.com/content/a0a0c382-6253-494d-8cd6-13862c740c12 [Accessed July 15, 2025]

41. Pew Research Center. Americans' Views of Government's Role: Persistent Divisions and Areas of Agreement. (2024). Available at: https://www.pewresearch.org/politics/2024/06/24/americans-trust-in-federal-government-and-attitudes-toward-it/ [Accessed July 15, 2025]

42. Edelman. 2023 Edelman Trust Barometer Special Report: Trust & Health. Available at: https://www.edelman.com/trust/2023/trust-barometer [Accessed July 16, 2025]

43. Bhattacharya S, Singh A. Unravelling the infodemic: a systematic review of misinformation dynamics during the COVID-19 pandemic. *Front Commun* (2025) **10**: doi:10.3389/fcomm.2025.1560936

44. IQVIA Digital. Navigating Social Media in Healthcare Advertising: Key Trends for 2025. Available at: https://www.iqviadigital.com/resources/blogs/navigating-social-media-in-healthcare-advertising-key-trends-for-2025 [Accessed July 16, 2025]

45. St. Aubin C, Liedke J. Social Media and News Fact Sheet. *Pew Research Center* (2024) Available at: https://www.pewresearch.org/journalism/fact-sheet/social-media-and-news-fact-sheet/ [Accessed July 16, 2025]

46. Deb T. Social Media in Healthcare Statistics 2025 By Data, Insights, Engagement. *Market.us Media* (2025) Available at: https://media.market.us/social-media-in-healthcare-statistics/ [Accessed July 16, 2025]

47. World Health Organization. Infodemic. Available at: https://www.who.int/health-topics/infodemic#tab=tab_1 [Accessed July 16, 2025]

48. Meta. Keeping Our Services Stable and Reliable During the COVID-19 Outbreak. *Meta* (2020) Available at: https://about.fb.com/news/2020/03/keeping-our-apps-stable-during-covid-19/ [Accessed March 14, 2023]

49. Illari L, Restrepo NJ, Johnson NF. Complexity of the online distrust ecosystem and its evolution. *Front Complex Syst* (2024) **1**: doi:10.3389/fcpxs.2023.1329857

50. Illari L, Restrepo NJ, Johnson NF. Rise of post-pandemic resilience across the distrust ecosystem. *Sci Rep* (2023) **13**:15640. doi:10.1038/s41598-023-42893-6

51. Illari L, Restrepo NJ, Johnson NF. Losing the battle over best-science guidance early in a crisis: COVID-19 and beyond. *Science Advances* (2022) **8**:eabo8017. doi:10.1126/sciadv.abo8017

52. Restrepo NJ, Illari L, Leahy R, Sear RF, Lupu Y, Johnson NF. How Social Media Machinery Pulled Mainstream Parenting Communities Closer to Extremes and Their Misinformation During Covid-19. *IEEE Access* (2022) **10**:2330–2344. doi:10.1109/ACCESS.2021.3138982

53. Rosen G. Community Standards Enforcement Report, Q3 2020. Facebook. Inc. (2020). Available at: https://about.fb.com/news/2020/11/community-standards-enforcement-report-nov-2020/ [Accessed July 22, 2025]

54. Jin K-X. Keeping People Safe and Informed About the Coronavirus. *Meta Newsroom* (2020) Available at: https://about.fb.com/news/2020/12/coronavirus/ [Accessed July 22, 2025]

55. Bickert M. How We're Taking Action Against Vaccine Misinformation Superspreaders. *Meta Newsroom* (2021) Available at: https://about.fb.com/news/2021/08/taking-action-against-vaccine-misinformation-superspreaders/ [Accessed July 22, 2025]

56. Facebook Help Center. COVID-19 and Vaccine Policy Updates and Protections | Facebook Help Center. *Facebook* Available at: https://www.facebook.com/help/697825587576762/ [Accessed July 22, 2025]

57. Wardle C. Getting Ahead of Misinformation. *Democracy Journal* (2023) Available at: https://democracyjournal.org/magazine/68/getting-ahead-of-misinformation/ [Accessed March 28, 2023]

58. RAND Corporation. What's Being Done to Fight Disinformation Online. (2019) Available at: https://www.rand.org/research/projects/truth-decay/fighting-disinformation.html [Accessed March 28, 2023]

59. The WHO. Managing the COVID-19 infodemic: Promoting healthy behaviours and mitigating the harm from misinformation and disinformation. Available at: https://www.who.int/news/item/23-09-2020-managing-the-covid-19-infodemic-promoting-healthy-behaviours-and-mitigating-the-harm-from-misinformation-and-disinformation [Accessed March 14, 2023]

60. Government to relaunch 'Don't Feed the Beast' campaign to tackle Covid-19 misinformation – Society of Editors. Available at: https://www.societyofeditors.org/soe_news/government-to-relaunch-dont-feed-the-beast-campaign-to-tackle-covid-19-misinformation/ [Accessed March 14, 2023]

61. Get the facts on coronavirus. *Full Fact* Available at: https://fullfact.org/health/coronavirus/ [Accessed March 14, 2023]

62. COVID-19 Information Center | Meta. *COVID-19 Information Center* Available at: https://about.meta.com/covid-19-information-center [Accessed March 14, 2023]

63. Wone A. VERIFIED: UN launches new global initiative to combat misinformation. *Africa Renewal* (2020) Available at: https://www.un.org/africarenewal/news/coronavirus/covid-19-united-nations-launches-global-initiative-combat-misinformation [Accessed March 14, 2023]

64. Klein S, Hostetter M. Restoring trust in public health. *Commonwealth Fund* (2023) doi:10.26099/j7jk-j805

65. The Rockefeller Foundation. New USD10 Million Project Launched To Combat the Growing Mis- and Disinformation Crisis in Public Health. *The Rockefeller Foundation* Available at: https://www.rockefellerfoundation.org/news/new-usd10-million-project-launched-to-combat-the-growing-mis-and-disinformation-crisis-in-public-health/ [Accessed March 14, 2023]

66. Mosenkis J, Dukuly D. Mercury Project to Boost Covid-19 Vaccination Rates and Counter Public Health Mis- and Disinformation in 17 Countries Worldwide. *The Rockefeller Foundation* Available at: https://www.rockefellerfoundation.org/news/mercury-project-to-boost-covid-19-vaccination-rates-and-counter-public-health-mis-and-disinformation-in-17-countries-worldwide/ [Accessed July 22, 2025]

67. The National Academies. Navigating Infodemics and Building Trust during Public Health Emergencies A Workshop. Available at: https://www.nationalacademies.org/event/04-10-2023/navigating-infodemics-and-building-trust-during-public-health-emergencies-a-workshop [Accessed March 28, 2023]

68. Tiffany E. Taylor, Holly Rhodes, Leticia Garcilazo Green, Lauren Ryan. Understanding and Addressing Misinformation About Science A Public Workshop. National Academy of Sciences Building 2101 Constitution Ave NW Washington DC 20418 USA: The National Academies (2023). Available at: https://www.nationalacademies.org/event/04-19-2023/understanding-and-addressing-misinformation-about-science-a-public-workshop [Accessed March 28, 2023]

69. Mirza S, Begum L, Niu L, Pardo S, Abouzied A, Papotti P, Pöpper C. Tactics, Threats & Targets: Modeling Disinformation and its Mitigation. in *Proceedings 2023 Network and Distributed System Security Symposium* (San Diego, CA, USA: Internet Society). doi:10.14722/ndss.2023.23657

70. Roozenbeek J, van der Linden S, Goldberg B, Rathje S, Lewandowsky S. Psychological inoculation improves resilience against misinformation on social media. *Science Advances* (2022) **8**:eabo6254. doi:10.1126/sciadv.abo6254

71. van der Linden S, Leiserowitz A, Rosenthal S, Maibach E. Inoculating the Public against Misinformation about Climate Change. *Global Challenges* (2017) **1**:1600008. doi:10.1002/gch2.201600008

72. Smith R, Cubbon S, Wardle C. Under the surface: Covid-19 vaccine narratives, misinformation and data deficits on social media. First Draft (2020). Available at: https://firstdraftnews.org/long-form-article/under-the-surface-covid-19-vaccine-narratives-misinformation-and-data-deficits-on-social-media/ [Accessed March 28, 2023]

73. Calleja N, AbdAllah A, Abad N, Ahmed N, Albarracin D, Altieri E, Anoko JN, Arcos R, Azlan AA, Bayer J, et al. A Public Health Research Agenda for Managing Infodemics: Methods and

Results of the First WHO Infodemiology Conference. *JMIR Infodemiology* (2021) **1**:e30979. doi:10.2196/30979

74. Lazer DMJ, Baum MA, Benkler Y, Berinsky AJ, Greenhill KM, Menczer F, Metzger MJ, Nyhan B, Pennycook G, Rothschild D, et al. The science of fake news. *Science* (2018) **359**:1094–1096. doi:10.1126/science.aao2998

75. Lewandowsky, Stephan, Cook, John, Ecker, Ullrich, Albarracín, Dolores, Amazeen, Michelle A, Kendeou, Panayiota, Lombardi, Doug, Newman, Eryn J, Pennycook, Gordon, Porter, Ethan, et al. Debunking Handbook 2020. George Mason University (2020). Available at: https://www.climatechangecommunication.org/wp-content/uploads/2020/10/DebunkingHandbook2020.pdf [Accessed March 28, 2023]

76. Green Y, Gully A, Roth Y, Roy A, Tucker JA, Wanless A. Evidence-Based Misinformation Interventions: Challenges and Opportunities for Measurement and Collaboration. Carnegie Endowment for International Peace (2023). Available at: https://carnegieendowment.org/2023/01/09/evidence-based-misinformation-interventions-challenges-and-opportunities-for-measurement-and-collaboration-pub-88661 [Accessed August 10, 2023]

77. Price G. Social Science Research Council (SSRC) Launches a $10 Million Project to Combat the Growing Mis- and Disinformation Crisis in Public Health. *Library Journal infoDOCKET* (2021) Available at: https://www.infodocket.com/2021/11/16/social-science-research-council-ssrc-launches-a-10-million-project-to-combat-the-growing-mis-and-disinformation-crisis-in-public-health/ [Accessed July 22, 2025]

78. Understanding and Addressing Misinformation among Populations that Experience Health Disparities (R01 - Clinical Trials Optional). *NIH Grants & Funding* (2022) Available at: https://grants.nih.gov/grants/guide/rfa-files/RFA-MD-22-008.html [Accessed July 22, 2025]

79. RESCINDED - Notice of Special Interest (NOSI): Research to Address Vaccine Hesitancy, Uptake, and Implementation among Populations that Experience Health Disparities. *NIH Grants & Funding* (2021) Available at: https://grants.nih.gov/grants/guide/notice-files/NOT-MD-22-006.html [Accessed July 22, 2025]

80. Audit of CDC's COVID-19 "Vaccinate with Confidence" Strategy. (2021) Available at: https://oig.hhs.gov/reports-and-publications/workplan/summary/wp-summary-0000718.asp [Accessed July 22, 2025]

81. Delgado C. "It's Up to You" COVID-19 Vaccine Campaign Aims to Increase Vaccine Uptake. *Verywell Health* (2021) Available at: https://www.verywellhealth.com/its-up-to-you-vaccine-ad-campaign-5115183 [Accessed July 22, 2025]

82. CDC COVID-19 Response Vaccine Task Force. Building Confidence in COVID-19 Vaccines Among Your Patients: Tips for the Healthcare Team. (2021) Available at: https://stacks.cdc.gov/view/cdc/142817/cdc_142817_DS1.pdf

83. CDC. COVID-19 Vaccination Field Guide: 12 Strategies for Your Community. (2024). Available at: https://archive.cdc.gov/www_cdc_gov/vaccines/covid-19/downloads/vaccination-strategies.pdf

84. Illari L, Johnson NF. Network resilience in the face of deplatforming: the online anti-vaccination movement and COVID-19. (2022) Available at: https://researchshowcase2022-gwu.ipostersessions.com/?s=7F-4B-0F-52-FE-EB-CA-63-04-D1-99-44-89-00-52-E6

85. Best DJ, Roberts DE. Algorithm AS 89: The Upper Tail Probabilities of Spearman's Rho. *Applied Statistics* (1975) **24**:377. doi:10.2307/2347111

86. aov function - RDocumentation. Available at: https://www.rdocumentation.org/packages/stats/versions/3.6.2/topics/aov [Accessed October 27, 2023]

87. Restrepo EM, Moreno M, Illari L, Johnson NF. Softening online extremes organically and at scale. (2022) doi:10.48550/arXiv.2207.10063

88. Restrepo EM, Moreno M, Illari L, Johnson NF. Softening Online Extremes Using Network Engineering. *IEEE Access* (2024) **12**:55768–55783. doi:10.1109/ACCESS.2024.3388851