PNAS NEXUS

*PNAS Nexus*, 2025, **4**, pgae522
**https://doi.org/10.1093/pnasnexus/pgae522**
Advance access publication 21 November 2024
**Research Report**

# Face-to-face or face-to-screen: A quantitative comparison of conferences modalities

Emma R. Zajdela[a,b,c,*,1], Kimberly Huynh[d], Andrew L. Feig[d], Richard J. Wiener[d] and Daniel M. Abrams[a,e,f,*]

[a]Department of Engineering Sciences and Applied Mathematics, Northwestern University, Evanston, IL 60208, USA
[b]Department of Ecology and Evolutionary Biology, Princeton University, Princeton, NJ 08544, USA
[c]Santa Fe Institute, Santa Fe, NM 87501, USA
[d]Research Corporation for Science Advancement, 4703 East Camp Lowell Drive, Suite 201, Tucson, AZ 85712, USA
[e]Northwestern Institute for Complex Systems, Northwestern University, Evanston, IL 60208, USA
[f]Department of Physics and Astronomy, Northwestern University, Evanston, IL 60208, USA
*To whom correspondence should be addressed: Email: emma.zajdela@princeton.edu (E.R.Z.); Email: dmabrams@northwestern.edu (D.M.A.)
[1]Present address: Department of Ecology and Evolutionary Biology, Princeton University, NJ 08544, USA
**Edited By** Noshir Contractor

## Abstract

The COVID-19 pandemic forced a societal shift from in-person to virtual activities, including scientific conferences. As society navigates a "new normal," the question arises as to the advantages and disadvantages of these alternative modalities. We introduce two new comprehensive datasets enabling direct comparison between virtual and in-person conferences: the first, from a series of nine small conferences, encompasses over 12,000 pairs of potential scientific collaborators across five virtual and four in-person meetings on a range of scientific topics; the expressed goal of these conferences is to create novel collaborations. The second dataset, from a series of three large physics conferences, encompasses >250,000 possible pairs of scientific collaborators. Our study provides quantitative insight into benefits and drawbacks of virtual and in-person conferences for team formation, community building, and engagement. We demonstrate the causal role of formal interaction on team formation across both modalities. Our findings show that formal interaction impacted team formation significantly more in virtual settings, while informal interaction played a larger role at in-person conferences as compared with virtual. We show that a nonlinear memory model for predicting team formation based on interaction outperforms seven alternative models. The model suggests that prior knowledge and interaction time contribute to catalyzing collaborations in both settings. Our results underscore the critical responsibility of organizers for optimizing professional interactions, whether virtual or in-person.

**Keywords:** virtual conferences, team science, scientific collaboration, network connectivity, sociophysics

### Significance Statement

Scientific collaborations are needed to solve many challenges facing society. Conferences play a crucial role in team formation by connecting scientists who may otherwise never meet. We introduce two new comprehensive datasets which enable quantitative comparison between virtual and in-person conferences. We show that formal interaction at conferences is causally linked to team formation and is similarly effective in both modalities, whereas informal interaction plays a significantly larger role connecting scientists at in-person conferences. We develop a mechanistic mathematical model that predicts which pairs of scientists are likely to form those collaborations. Our work shows how conference structure can determine which teams form and thus influence the direction of scientific inquiry, in both in-person and virtual modalities.

## Introduction

Conferences play a crucial role in the scientific community, serving as a platform for networking, collaboration, and knowledge dissemination. According to a 2014 study, 16% of collaborators who do not live in the same city meet at conferences (1). However, the direct costs of conferences amount to tens of billions of US dollars each year (2, 3) and entail well-documented negative environmental impacts (4–6) and other inequities (7, 8). With the advent of virtual conferencing technology and the recent pandemic-induced shift to virtual environments across many facets of society (9), virtual conferences have become increasingly common. Research has shown that virtual conferences are less effective at facilitating social interactions compared with in-person events, but there is evidence that virtual communications can

**Competing Interest:** Three of the authors have worked at the Research Corporation for Science Advancement which provided data analyzed in the study.
**Received:** March 7, 2024. **Accepted:** October 8, 2024

reduce biases (10–12). Virtual conferences may offer certain benefits, such as increased diversity, equity and inclusion (8, 13) through increased accessibility, reduced carbon footprint (6), and lower costs. However, concerns linger about their impact on scientific collaborations, productivity, and innovation (14, 15). These concerns have prompted discussions about how scientific conferences should be held even when it is safe to convene them in person (16). Some posit that they should continue to be held virtually rather than in person (3), others that hybrid features should be included (17, 18), while yet others argue that a lack of in-person interaction will cause a significant damper on scientific productivity and innovation (14, 19).

During the COVID-19 pandemic, quarantine, and safety considerations required changing the modality of many conferences from in-person to virtual. Drawing on data from multiple virtual and in-person conferences held over a 6 year time span, this study provides quantitative insight into the potential benefits and drawbacks of each conference format in terms of three criteria: team formation, engagement, and community building. The first data source analyzed comes from a series of conferences dubbed "Scialogs" that span a wide range of scientific fields; the data captures detailed information on over 12,000 pairs of participants, including their demographics, preconference and postconference awareness of one another, assigned discussion sessions, and formation of new collaborations. The second dataset encompasses over 250,000 pairs of participants who spoke in sessions at the American Physical Society (APS) March Meetings and their publication records.[a]

The Scialog conferences are highly interdisciplinary, involving around 50 participants with the explicit goal of fostering community and generating new collaborations. At each conference, teams of participants write proposals for seed funding. In contrast, the APS conferences are large-scale events primarily targeted towards physicists with broader objectives that include disseminating new research, connecting participants, and facilitating collaborations. The diverse participant makeups and goals of the Scialog and APS conferences indicate that our results are potentially applicable to various conference types.

By shedding light on the differences between in-person and virtual scientific conferences, this study aims to inform the scientific community about strengths and weaknesses of these formats in a postpandemic world.

# Results

## Data set

The first dataset comprises information from nine conferences—five virtual and four in-person—held between 2015 and 2021 and involving a total of 573 participants. Research Corporation for Science Advancement (RCSA) organized these conferences as part of its Scialog ("Science Dialog") program, and the conference themes spanned a variety of scientific fields. For each conference, RCSA invited approximately 50 early career scientists (*"fellows"*) and several senior scientists (*"facilitators"*), recognized as world-leading researchers in the area of focus, to attend. Scialog conferences have an interactive structure with the expressed goals of fostering dialogue, networking, and the formation of new collaborations to initiate novel seed projects based on highly innovative ideas that emerge at the conference (20).

Near the end of each conference, fellows self-assemble into teams of two to four members to write collaborative research proposals. Team members may not have previously collaborated; we treat co-authoring a proposal as formation of a new collaboration. A fellow may participate in a maximum of two teams, with different team members on each team. The number of proposals submitted at any given conference varies from 20 to 35, from which 5–10 are selected for funding. Our study collected comprehensive data encompassing participants' levels of preconference and postconference familiarity with one another, session co-attendance, and team composition following the conclusion of each conference. See Methods section for the scientific topics, acronym definitions, and descriptive statistics of the conferences.

The shift from an in-person to a virtual setting in 2020 due to the COVID-19 pandemic provided the grounds for a direct comparison between conference modalities[b] as the organizers attempted to recreate the in-person conference structure in the virtual environment to the extent possible. Total formal interaction time (e.g. breakout sessions and keynote talks) was almost identical for in-person and virtual conferences, while total possible informal interaction time (e.g. meals, social activities) was higher at in-person conferences. See Methods section for a comparison of virtual and in-person conference structures.

## Effect of interaction on team formation

We examine the impact of formal interaction on the formation of new scientific teams, and whether that is consistent across conference modalities. Specifically, we focus on the role of formal interaction in small to medium sized breakout discussion groups in promoting team formation. Our analysis reveals that small groups of three to four individuals, tasked with conceiving potential research projects, are especially effective for this purpose, with a multiplier of 7 compared with participants who did not co-attend the groups. To understand this effect, we developed a team formation model that accounts for both interaction during the conference and prior network awareness, incorporating the effect of memory beyond formal interaction during the meeting. Previously, we demonstrated the effectiveness of this model for in-person conferences (21), and here we extend our findings to show that the same model holds for virtual conferences.

We define total scaled interaction time for a given pair $I_{\text{tot}}$ as proportional to the time spent in a session and inversely proportional to its size (21)—see section "Defining interaction" for more detail.

As shown in Fig. 1, collaborators interacted significantly more than noncollaborators at all conferences. That is, the typical pair that formed a collaboration interacted far more than the typical pair that did not end up collaborating. Thus, the immediate suggestion is that more interaction leads to more likely collaboration, even without further statistical investigation.

## Causal effect

Although there may have been confounding factors due to the pandemic, we uncover evidence that formal interaction time has a causal effect on collaboration in both modalities.

### Indirect evidence

We define *interaction ratio* (*IR*) as the ratio of total scaled interaction for collaborators (blue/left bar in Fig. 1) to total scaled interaction for noncollaborators (red/right bar in Fig. 1). We interpret *interaction ratio* as the degree to which the probability of forming a collaboration was impacted by formal interaction. When comparing across conferences, since the structure of formal interaction is nearly identical, a higher interaction ratio means that the pairs

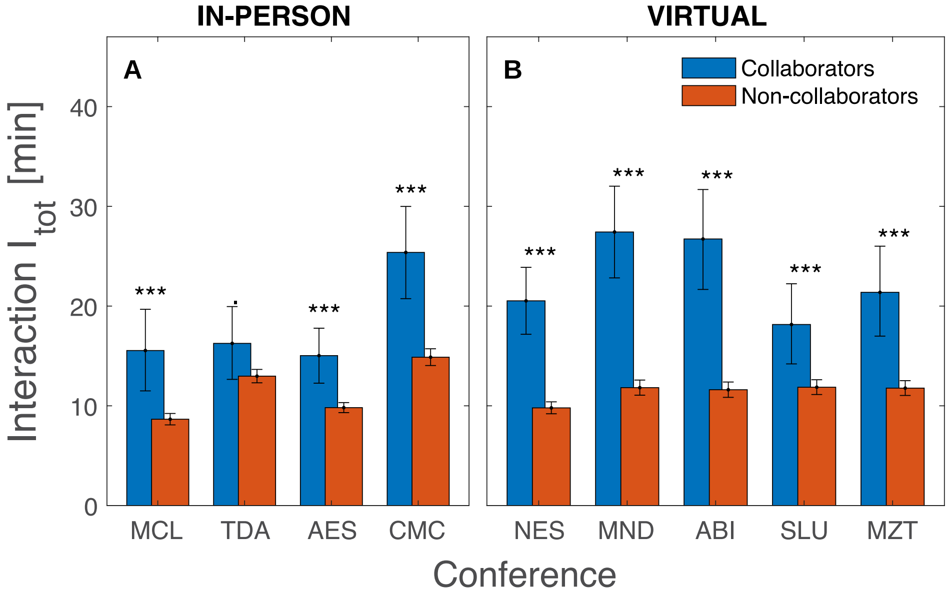


**Fig. 1.** Interaction and collaboration. Attendees with greater formal interaction were significantly more likely to collaborate, and that effect was more pronounced at virtual meetings. Grouped bars are shown for each in-person A) and virtual B) conference analyzed. Left (blue) and right (red) paired bars show bootstrap estimates for mean total scaled interaction time $I_{tot}$ [minutes] for collaborators and noncollaborators, respectively. Error bars show mean values of the bootstrapped data with 95% CI. *P*-values of the Mann–Whitney *U* test: MCL, $6.0 \times 10^{-4}$; TDA, $6.3 \times 10^{-2}$; AES, $1.4 \times 10^{-4}$; CMC, $7.7 \times 10^{-6}$; NES, $5.8 \times 10^{-12}$; MND, $1.7 \times 10^{-11}$; ABI, $1.8 \times 10^{-9}$; SLU, $2.1 \times 10^{-4}$; MZT $2.9 \times 10^{-6}$.

who formed collaborations were more predictable from interaction alone.

As an example, at the MCL conference (in-person), collaborators spent an average of 16 scaled minutes interacting, whereas noncollaborators spent an average of only 9, an interaction ratio of 1.7. At the NES conference (virtual), the relevant numbers are 21 min and 10, an interaction ratio of 2.1. In both cases, those pairs who ended up collaborating were disproportionately the ones who had spent more time interacting.

### *Direct evidence*

We also find direct evidence that interaction has a causal effect on collaboration formation across both conference modalities. For many conferences, alternative schedules were retained that could have been selected but were not.[c] These formed the basis of tests performed on 2,500 plausible counter-factual schedules. We consider pairs of fellows who collaborated at the actual conference. Then, we compute their mean total scaled interaction at (i) the actual conference (blue/left bar in Fig. 2), and average over all counter-factual conferences where (ii) the mini session assignments were identical to the actual conference (red/second to left bar in Fig. 2), (iii) breakout session assignments were identical to the actual conference (yellow/second to right bar in Fig. 2), and (iv) different session assignments (gray/right bar in Fig. 2). See Methods section for details.

The only counter-factual cases close to the value of the actual conference correspond to scenarios sharing the same exact mini session assignments but with variations in the larger breakout session assignments (in Fig. 2, blue/left and red/second to left bars). Using this method, we are able to infer small group assignments knowing only which pairs ultimately collaborated. These results strongly suggest a causal connection between intense interaction in a small-group setting and team formation. There appears to be no significant effect of co-attending a larger breakout session on collaboration formation in 5 of 7 conferences, and even for those at which it was significant, the effect size is small to moderate. Therefore, at both in-person and virtual conferences, *assigning participants to mini sessions has a strong causal impact on which teams ultimately form.*

## Effect of co-attending a small group session

In addition to showing that interaction has a statistically significant effect on collaboration probability, we also wish to determine the size of the effect. To evaluate that, we restrict our data to pairs without prior knowledge and use bootstrap statistics to estimate the odds of collaboration for pairs who co-attended one mini session $P_{\text{Collab}}^{(\text{mini session})}$ and for those who did not co-attend any mini session, but could have in one of the 2,500 counter-factual scenarios $P_{\text{Collab}}^{(\text{no-mini})}$. We compute how much co-attending a mini session multiplied the chance of a pair collaborating (see column *Mltp.* in Table 1). The effect of co-attending a mini session increased the odds of collaborating at both virtual and in-person conferences, with no notable difference between the two modalities. On average, co-attending a mini session increased the chances of a pair collaborating from about 1.7% to about 11%, an increase by a factor of 7.[d]

## Modality comparison

In evaluating the indirect evidence for a causal relationship between interaction and team formation, we assessed interaction ratios for each of nine Scialog conferences. Results differed significantly based on conference modality: numbers were 30% higher on average for virtual (IR=2.17) modalities as compared with in-person (IR=1.67) conferences. This suggests that formal interaction has a greater impact on team formation in the virtual modality than in the in-person modality.

## Mathematical model

We developed a mathematical model for estimating the probability of any given pair of participants forming a new collaboration at a conference, based on their level of interaction throughout the

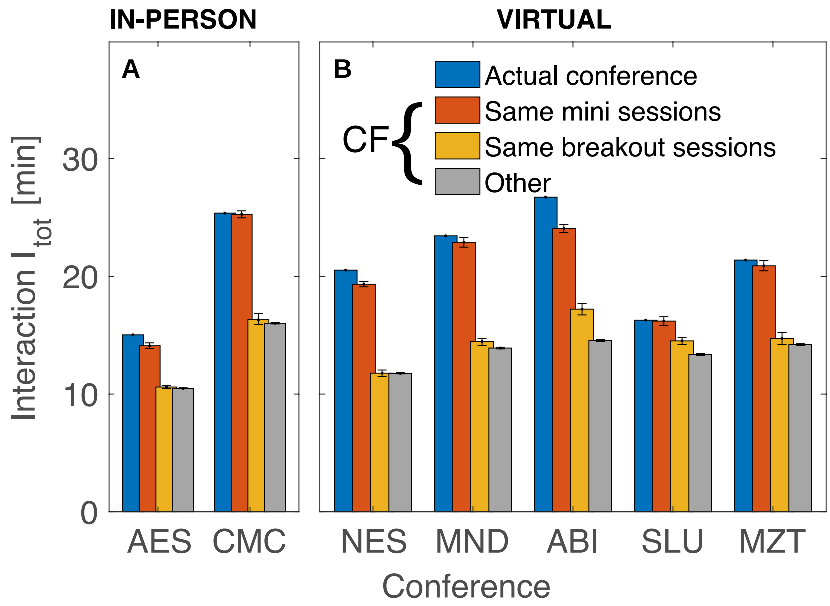


**Fig. 2.** Average actual vs. counterfactual interaction for pairs of actual collaborators. Panels show bootstrap estimates for mean total scaled interaction of pairs who collaborated at the actual conference, but with differing session assignments in counterfactual scenarios. Actual conference: blue/left bars, same mini sessions: red/second-to-left bars, same breakout sessions: dark yellow/second-to-right bars, other counterfactual scenarios: gray/right bars. Error bars show mean values of the bootstrapped data with 95% CI.

**Table 1.** Effect size of co-attending a mini session.

| Conf. | Mod. | $N^{\text{pairs}}_{K_0=0}$ | $P^{(\text{no-mini})}_{\text{Collab}}$ | $P^{(\text{mini session})}_{\text{Collab}}$ | Mltp. |
|---|---|---|---|---|---|
| NES | V | 1,348 | 0.019 [0.012, 0.028] | 0.14 [0.10, 0.19] | 7.6 |
| MND | V | 1,170 | 0.015 [0.0080, 0.025] | 0.13 [0.092, 0.18] | 9.0 |
| ABI | V | 1,233 | 0.016 [0.0079, 0.026] | 0.11 [0.071, 0.16] | 6.9 |
| SLU | V | 880 | 0.015 [0.0067, 0.029] | 0.042 [0.019, 0.075] | 2.8 |
| MZT | V | 1,295 | 0.021 [0.012, 0.034] | 0.077 [0.046, 0.12] | 3.7 |
| AES | IP | 896 | 0.015 [0.0078 0.025] | 0.10 [0.068 0.15] | 7.1 |
| CMC | IP | 984 | 0.017 [0.0085 0.028] | 0.15 [0.10 0.21] | 8.7 |

Rows with modality (Mod.) V correspond to virtual conferences and those with IP correspond to in-person conferences.
$N^{\text{pairs}}_{K_0=0}$: Num. pairs with no prior knowledge of one another.
$P^{(\text{mini session})}_{\text{Collab}}$: Prob. for pairs who co-attended one mini session.
$P^{(\text{no-mini})}_{\text{Collab}}$: Prob. for pairs who did not co-attend any mini session, but could have in one of the 2,500 counter-factual scenarios.
Mltp.: Multiplier (ratio of previous two columns).
Numbers in brackets are the 95% CI.

**Table 2.** Model selection.

| Conf. | Best AIC | Next Best AIC | Next Best Model | Relative Likelihood |
|---|---|---|---|---|
| NES | 558.71 | 561.99 | $aK_0 + bI_{\text{tot}} + c$ | 0.19 |
| MND | 364.43 | 384.29 | Linear | 4.9e-05 |
| ABI | 367.72 | 378.07 | Threshold | 0.0057 |
| SLU | 371.12 | 369.98 | Memory | 0.57 |
| MZT | 354.32 | 372.52 | Threshold | 1.2e-4 |
| MCL | 375.88 | 385.62 | $aK_0 + bI_{\text{tot}} + c$ | 0.0077 |
| TDA | 337.02 | 340.79 | $aK_0 + b$ | 0.15 |
| AES | 526.62 | 541.64 | $aK_0 + bI_{\text{tot}} + c$ | 0.00055 |
| CMC | 407.96 | 430.53 | $aI_{\text{tot}} + b$ | 1.3e-05 |

Note: the nonlinear memory model is the best performing model for all Scialogs except SLU, for which it was the next best model. In the case of SLU, $aK_0 + bI_{\text{tot}} + c$ performed the best.

duration of the conference and their prior knowledge of each other before the conference. Initially, a linear model was developed based on the assumptions that: (i) probability of collaborating $P(t)$ increases (at maximum rate S) as interaction $I$ increases, (ii) the smaller the group of participants the pairwise interaction took place in, the more intense the interaction, and (iii) the probability decays (at maximum rate $W$) when interaction ceases. Defining $I_{\max}$ as the maximum pairwise interaction intensity, we constructed the following ordinary differential equation (ODE) model (1):

$$\frac{dP}{dt} = \underbrace{S\frac{I}{I_{\max}}(1-P)}_{\text{strengthening}} - \underbrace{WP\left(1-\frac{I}{I_{\max}}\right)}_{\text{weakening}}. \tag{1}$$

This linear model has a major limitation, namely, that the probability $P(t)$ decays exponentially to zero when interaction has ceased. To capture human behavior more realistically, we developed a nonlinear "memory" model which incorporates the assumption that if people have sufficiently interacted, they will remember one another. We also modified the models to account for prior knowledge $K_0$ that participants have of one another prior to attending a conference. See the Methods section for a description of how $K_0$ values were determined and incorporated into interaction, (21, 22) for details of the model construction and (23) for an interactive version of the models.

We compare the nonlinear memory model to seven other candidate models: two models based on randomness, three models with a linear combination of total scaled interaction time and prior knowledge, a threshold model where pairs had one higher probability if they interacted over a certain threshold and one lower probability otherwise, and the linear model. Note that several of these models can be seen as limits of the nonlinear memory model with certain parameter choices. We compute the maximum log likelihood of the data with respect to the model and the Aikike Information Critera (AIC) of each of the models with best-fit parameters, which penalizes the addition of new parameters. We then compute the relative likelihood of the next best model compared with the best model. Results of the model selection can be found in Table 2. The nonlinear memory model outperforms all other models in all but one case (in which the nonlinear memory model still exhibits a higher likelihood).[e] These findings demonstrate the robustness of the model in capturing the essential mechanisms underlying team formation arising from interaction among individuals. They also suggest that prior knowledge and interaction time are similarly indicative of the probability of forming a new collaboration in both modalities.

## Community building

Another goal of conferences is to provide a venue for networking and to create connections between people who otherwise would not have met. To assess how virtual versus in-person conferences impact networking, we compute a quantity that we will refer to as "connecting efficiency:"

Definition 1 Connecting efficiency.

Given a network before and after a certain point in time (the "event"), the connecting efficiency of that event, denoted CE, is the fraction of initially missing ties $m^{(\text{before})}$ that became connected:

$$\mathrm{CE} = \frac{m^{(\text{before})} - m^{(\text{after})}}{m^{(\text{before})}}. \tag{2}$$

Note that connecting efficiency is proportional to the change in network density before / after the conference, rescaled by a factor of $(1 - \text{density}^{(\text{before})})$ to allow for cross-comparison among events with different initial density.

To quantify an event's effectiveness at increasing acquaintanceship among attendees (or, equivalently, reducing the number of pairs of attendees who do not know one another), we compute the connecting efficiency at the level of acquaintanceship, where a link between a pair is present if at least one participant is aware of the other. In our case, we define pairs as acquainted if a participant reported they knew the other at the level of awareness or higher for the relevant question in preconference or postconference surveys. We compute the connecting efficiency of four in-person and five virtual conferences using bootstrap resampling with $10^4$ repetitions and find $\mathrm{CE}_{\text{in-person}} = 0.40$ (95% CI [0.38 0.42], $n = 3{,}134$ pairs) and in the virtual case: $\mathrm{CE}_{\text{virtual}} = 0.22$ (95% CI [0.21 0.24], $n = 2{,}841$ pairs), treating each conference modality as a single group. *This suggests that in-person conferences are much more effective at connecting attendees* (the confidence intervals of the distributions generated by bootstrap resampling connecting efficiency do not overlap at the $\alpha = 0.001$ level). At in-person conferences, 40% of initially disconnected pairs became connected, but only 22% at virtual conferences. See Fig. 3 for a graphical representation of how connections formed at conferences.

## Engagement

As a measure of how engaged in the conference participants were, we compute the *team participation rate* (TPR): the fraction of attendees who participated in at least one proposal team. We find that $\mathrm{TPR}_{\text{in-person}} = 0.88$ (90% CI [0.840 0.915], $n = 211$ participants) and $\mathrm{TPR}_{\text{virtual}} = 0.80$ (90% CI [0.756 0.836], $n = 275$ participants). Thus, team participation rate was only weakly significantly higher at in-person than virtual conferences (confidence intervals generated from bootstrap resampling with $10^4$ draws do not overlap at the 90% level, data aggregated for four in-person and five virtual conferences), with 8% more attendees joining a team at in-person compared with virtual conferences.

We also compute the average number of teams per person for those who participated in at least one team.[f] We find that the median for in-person conferences is 1.47 and at virtual conferences 1.48, with CI overlapping even at the 90% level. This result suggests that, although the team participation rate is lower at virtual conferences, the team-forming behavior for those who are engaged is similar.

## What about other types of conferences?

To determine whether the results demonstrated through the Scialog dataset generalize to other types of conferences, we compiled a new dataset for a series of conferences at the opposite end of the spectrum in terms of size: the American Physical Society (APS) March Meeting conferences. These annual conferences typically have more than 10, 000 total presenters who are divided among over 800 sessions, each consisting of approximately one dozen speakers. The APS March Meeting took place exclusively in-person prior to the COVID-19 pandemic and was held fully virtually and synchronously in 2021. Our dataset contains the names of the speakers in each March Meeting session, as well as their publication records from the arXiv preprint repository (24). We determine which speakers who co-presented in a session later coauthored a preprint for the first time after the conference (authors who co-published prior to the conference were omitted), as an indicator of a new collaboration being formed.

To evaluate the impact of the in-person to virtual transition for both the Scialog conferences and the APS conferences, we examine the percentage of all possible new pairwise collaborations that actually formed. Table 3 shows, first of all, that the small Scialog meetings were more than two orders of magnitude more effective at generating new collaborations. This difference is likely attributable to many factors, among them the more intensive small-group interactions among participants at Scialogs, the opportunity to apply for seed funding from the Scialog conference organizers, and the different indicators used to identify a new collaboration—in the case of Scialog meetings coauthoring a proposal and in the case of APS meetings coauthoring a preprint for the first time. In fact, nearly as many collaborations are generated from among the ∼50 Scialog attendees as from the ∼10,000 APS attendees.[g]

To the question of virtual vs. in-person, though, Table 3 shows that the mean percentage of possible collaborations that formed slightly increases in the virtual modality for both the small Scialogs and the large APS conferences (refer to Supplementary Materials for more details about these calculations). Although the Scialogs and APS March Meetings are vastly different types of conferences, the finding that virtual conferences are at least as effective for both types of conferences at generating new collaborations suggests that results from the small Scialog conferences may be widely applicable.

# Discussion

Before performing this research, we anticipated (based on our own anecdotal experiences with distraction and screen fatigue) that virtual conferences would be less effective at generating novel collaborations than in-person conferences. Surprisingly, the results of our analysis imply the opposite: based on data from six in-person and six virtual conferences, formal interaction at virtual conferences is at least as effective, if not more so, but comes at the cost of community building and engagement. The fact that the percentage of possible collaborations that formed *increased*

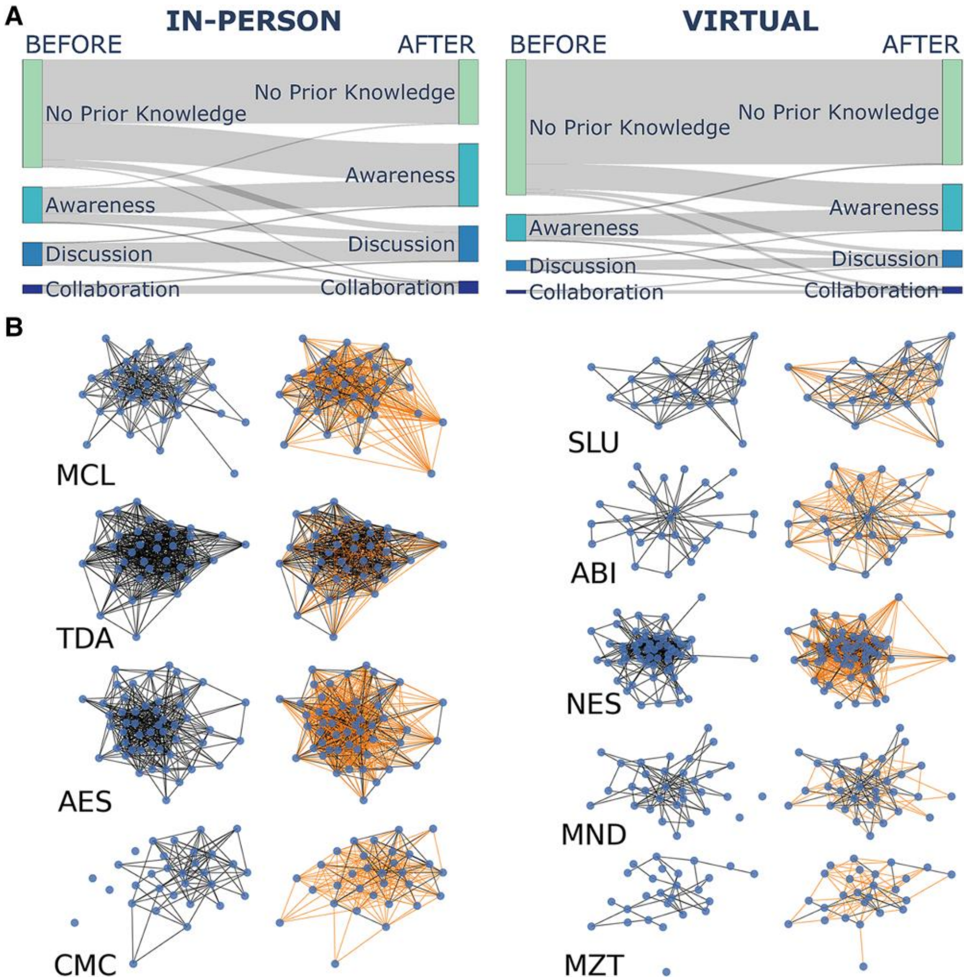


**Fig. 3.** Connecting efficiency. A) Sankey plots showing how participants' knowledge of one another shifted after first-year conferences. Data from preconference surveys, aggregated for four in-person conferences (MCL, TDA, AES, and CMC) and five virtual conferences (SLU, ABI, NES, MND, and MZT). B) Social network representation before and after conferences. Here nodes represent participants who attended both the first and second conference in the series; a link between two nodes is present if the pair of nodes it is connecting reported knowing each other at the level of awareness or higher. Orange links are new connections formed at the conference.

**Table 3.** New collaborations by modality.

| Modality | Sessions | New collabs | Total pairs | Percent new collabs (%) | |
|---|---|---|---|---|---|
| Scialog IP | 78.5 | 42.0 | 837.3 | 5.0 | [4.3 5.8] |
| Scialog V | 90.8 | 41.0 | 736.0 | 5.6 | [4.9 6.3] |
| APS IP | 882.0 | 75.0 | 254,733.0 | 0.029 | [0.025, 0.034] |
| APS V | 999 | 84 | 192,454 | 0.044 | [0.035, 0.053] |

Percentage of all possible pairs (pairs who co-attended any session) forming a new collaboration. Note: Each row is averaged over all meetings of the given modality. IP, In-person; V, Virtual. Numbers in brackets are 95% CI. See Supplementary Material for detailed breakdown.

in the virtual modality for both large APS and small Scialog conferences suggests that the lessons learned from our detailed analysis of Scialog conferences are applicable to other types of conferences with vastly different sizes and participant makeups. As such, our research extends prior findings from literature on face-to-face versus virtual interaction by providing quantitative measures of the differences between in-person and virtual interactions in a real-world setting.

The importance of existing relationships and familiarity in research teams has been previously highlighted (25, 26). Our research supports these findings, with the nonlinear memory model which incorporates prior knowledge outperforming all others tested. Several questions arise about the role of familiarity among participants at conferences: (i) Given that connecting efficiency is halved at virtual compared with in-person conferences, what are the effects on the formation of future collaborations? (ii) We have only considered the initial initiative in the Scialog series, and novel collaborations at Scialog and APS conferences. How are relationships between participants developed and sustained over time, and what are the implications for collaboration and innovation?

The nonlinear memory model—conceptually simple, only requiring two inputs (scaled interaction time and prior knowledge) alongside the incorporation of a memory effect—appears to capture the essence of how collaborations are formed at conferences, both in the case of virtual and in-person conferences. Our model assumes that collaborations between any two participants are equally likely to form, despite the fact that social network degrees are heavily skewed and that reducing the network to a pairwise analysis loses relevant information about the structure of the network (27). A more detailed model could capture more of the known skewing and other properties of social networks, like triadic closure and homophily effects. Despite these approximations, model selection based on AIC shows that the nonlinear memory model performs well at predicting who is likely to form a collaboration, suggesting that the influence of interaction is important and sufficient to overcome biases.

This model is particularly powerful when combined with the result that interaction has a causal effect on collaboration, because conference organizers are effectively able to control interaction by assigning participants to sessions. For the in-person conference modality, research indicates long-lasting effects of interaction on scientific output and new collaborations, as measured by co-publications and the diffusion of scientific ideas (28). This shows that organizers have a responsibility to think carefully about the participants they are assigning to sessions, as their decisions can have effects on scientific progress well beyond the duration of the conference itself.

Although participants may believe they are independently selecting their collaborators, the reality is that their choices are constrained by the opportunities afforded by the conference format, so their sense of control is to some extent illusory (29). This freedom is further reduced at virtual conferences, where there are fewer opportunities for serendipitous encounters outside of assigned sessions. The disproportionate impact of formal interaction at virtual conferences, though, comes at the cost of community engagement and connecting efficiency. Given the importance of "weak ties" (30) for generating novel opportunities and research directions, virtual conference organizers should incorporate design elements to encourage informal interaction, including through the use of innovative technologies (31, 32).

The disproportionate impact of formal interaction at virtual conferences can be leveraged by policymakers, university administrators, or other academic communities to better promote and orchestrate new collaborations in this setting. Virtual platforms could facilitate multidisciplinary collaborations by assigning to groups experts from various fields who might not typically meet at in-person events. Early-career researchers may also receive more opportunities to network, present their work, and receive feedback from a global audience. This can be particularly beneficial in early career stages, where building a professional network and gaining visibility are critical.

## Limitations

Like any research based on real-world data our work has limitations. We cannot control for all the changes that occurred in the world during the pandemic (e.g. changes in willingness or availability to attend even virtual conferences), though we believe confining our Scialog analysis to events that occurred during the conferences reduces the impact of exogenous confounding effects. This approach cannot be used with the APS March Meeting data, which means that possible confounding effects from the pandemic (e.g. access to facilities or technology) may play a larger role.

We also note that the comparative analysis of new collaborations formed at APS March Meetings (as measured by arXiv publications by co-presenters) was performed longitudinally over time. There is likely a growth in the total number of preprints across time, so there is a possible confounding factor of changing publication rates per capita.

We examine a set of small conferences and a set of very large conferences in the United States, but these of course do not cover the huge span of other conference types with varying sizes, goals, attendance criteria, or participant backgrounds. The Scialog conferences are not representative of all types of conferences, as their goal is explicitly to generate new collaborations and they provide an incentive in the form of grants for participants to form collaborations, contrary to the APS March Meetings which have a broader set of goals targeted towards disseminating new research, connecting participants, and forming collaborations. Furthermore, although the disciplinary backgrounds of participants at Scialog conferences are varied (with scientists from fields including chemistry, biology, physics, computer science, veterinary science, neurophysiology, engineering, earth and planetary science, astronomy, microbiology, and biochemistry), the APS March Meetings primarily convene scientists from the physics community. We leave further exploration of other types of conferences for future work.

We hope that future work can further explain and improve on the success of the model for team formation that we have provided.

## Conclusion

The COVID-19 pandemic has catalyzed rapid and widespread behavioral shifts in society, notably the adoption of virtual meeting technologies across various domains encompassing healthcare, education, and professional settings (9). As societal activities return to in-person formats, the debate surrounding the merits of face-to-face versus face-to-screen interactions continues to unfold across various sectors. In professional spheres, organizations grapple with the decision to mandate in-office attendance or embrace fully remote workforces (33). The landscape of scientific conferences reflects a spectrum of approaches, ranging from exclusively in-person to exclusively virtual to hybrid models with synchronous and/or asynchronous virtual components.

The competing benefits and drawbacks of virtual and in-person conferences will need to be balanced moving forward. To make informed decisions, data collection and analysis should be given high priority, with the goal of answering the question: what is optimal conference design in the 21st century?

We believe that our investigation offers a glimpse at the answer. The uniquely detailed Scialog dataset allows for a level of analysis not achievable for other conferences, but, as demonstrated by the APS March meeting analysis, appears to have application far beyond its specific context. This study represents an important first step in exploring the impact of virtual and in-person interactions on the dynamics and outcomes of collaborative endeavors, and underscores the crucial role of conference organizers. However, further research is needed to comprehensively understand and optimize virtual—and/or hybrid—meetings across different settings, with the ultimate goal of enhancing collaboration, knowledge dissemination, and innovation.

## Methods

More details about the purpose and organizing principles of Scialog conferences are available in reference (20). Table 4 shows descriptive statistics for each of the conferences analyzed.

Key conference structures include:

- Whole group formal interaction time (e.g. keynote talks)
- Medium-sized breakout sessions of 8–12 fellows moderated by one or two senior scientists (each session lasts 1 h–1 h 15 min)
- Mini breakout sessions of 3–4 fellows (each session lasts 30–45 min)

**Table 4.** Descriptive Scialog conference Statistics.

| Conf. | Topic | Type | Date | Participants | Fellows | Fraction returning | Pairs of fellows | Proposals Submitted | Proposals funded | Mean pairwise prior knowledge | Pairs who collaborated (%) |
|---|---|---|---|---|---|---|---|---|---|---|---|
| NES | Negative Emissions Science | Virtual | 2020 | 69 | 60 | 0.92 | 1,770 | 32 | 8 | 0.55 | 4.3 |
| MND | Microbiome, Neurobiology, and Disease | Virtual | 2021 | 61 | 53 | 0.60 | 1,378 | 23 | 6 | 0.25 | 3.6 |
| ABI | Advancing BioImaging | Virtual | 2021 | 63 | 54 | 0.48 | 1,431 | 28 | 10 | 0.20 | 3.1 |
| SLU | Signatures of Life in the Universe | Virtual | 2021 | 63 | 54 | 0.39 | 1,431 | 26 | 8 | 1.02 | 3.2 |
| MZT | Mitigating Zoonotic Threats | Virtual | 2021 | 63 | 54 | 0.48 | 1,431 | 24 | 8 | 0.43 | 4.6 |
| MCL | Molecules Come to Life | In-person | 2015 | 64 | 52 | 0.75 | 1,326 | 20 | 5 | 0.91 | 3.5 |
| TDA | Time Domain Astrophysics | In-person | 2015 | 59 | 49 | 0.78 | 1,176 | 30 | 6 | 2.1 | 4.3 |
| AES | Advanced Energy Storage | In-person | 2017 | 71 | 60 | 0.83 | 1,170 | 35 | 6 | 0.76 | 6.3 |
| CMC | Chemical Machinery of the Cell | In-person | 2018 | 60 | 50 | 0.62 | 1,225 | 24 | 8 | 0.43 | 4.6 |

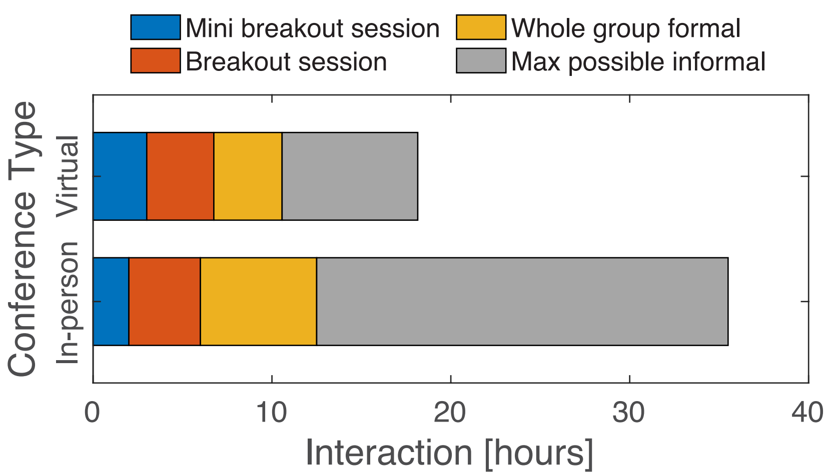


**Fig. 4.** Total interaction time in hours for representative programs at virtual and in-person Scialog conferences.

- Informal interaction time (in-person: meals, receptions; virtual: social mixer hosted on the Gather platform)[h]

Refer to Supplementary Materials for detailed schedules of representative virtual and an in-person conferences.

Participants were assigned to breakout sessions and mini breakout sessions using an algorithm taking into account their prior knowledge of each other, scientific interests, and background. This algorithm produced numerous possible conference schedules with nearly equivalent priority; these formed the basis of our counter-factual analysis to assess causality.

Figure 4 presents a summary of cumulative interaction time in hours at representative virtual and in-person conferences, categorized by interaction type. Total time was computed from the beginning of the conference until the formation of teams by fellows. Notably, participants spent comparable amounts of time in mini breakout sessions, breakout sessions, and formal interaction in whole-group settings at virtual and in-person conferences. The primary difference in conference schedules is total potential time available for informal interaction, which was about 7 h 30 min at virtual conferences compared with 23 h at their in-person counterparts, despite efforts by organizers to recreate informal interaction opportunities in the virtual setting (through use of the Gather platform for social mixers and other activities).

Organizers attempted to replicate the in-person conference experience as closely as possible in the virtual setting, despite less total possible informal interaction time at virtual conferences.

This affords us a remarkable opportunity to investigate the impact of virtual environments on team formation and social network evolution.

## Defining interaction

For each pair, total scaled interaction $I_{tot}$ is defined as the sum of scaled interaction time across all sessions. Scaled interaction during a co-attended session was taken to be proportional to the time one participant spent listening to the other, under the unrealistic but convenient assumption that all participants spoke equally. Thus, for a given pair of participants co-attending a session of time $T_k$ with $N_k$ participants, we assumed:

$$I_k^{\text{session}} \propto \frac{T_k - T_k/N_k}{N_k - 1} \propto \frac{T_k}{N_k} \tag{3}$$

where the numerator is the total time spent listening to others and the denominator is the number of people to listen to. When the pair are in different sessions, $I_k^{\text{session}} = 0$. Normalizing so that when $N_k = 2$, $I_k^{\text{session}} = T_k$, we find: $I_k^{\text{session}} = \frac{2T_k}{N_k}$. Then with $k$ the index of the session and $m$ the number of sessions (here, $6 \le m \le 8$),

$$I_{\text{tot}} = \sum_{k=1}^{m} I_k^{\text{session}} = \sum_{k=1}^{m} \begin{cases} \frac{2T_k}{N_k} & \text{during co-attended session} \\ 0 & \text{else} \end{cases}. \tag{4}$$

The pairwise interaction intensity profile $I(t)$ was constructed in a similar fashion. To get instantaneous interaction intensity, we divided by the session length $T_k$ and chose units such that a 2-person session would have maximum intensity $I_{\max}$. We assumed a minimum interaction term (corresponding to informal interaction between sessions) $I_{\min} = 2/N_{\text{tot}}$ that depends on the total number of participants $N_{\text{tot}}$. We then added a term proportional to the prior knowledge $K_0$, which was measured in a preconference survey: each attendee rated familiarity with each other attendee on a 0–3 scale[i] and $K_0 \in [0, 6]$ was defined as the sum of reciprocal pair ratings.[j] Eq. (5) summarizes the interaction function as it was implemented:

$$I(t) = \frac{I_{\max}}{6a + 1} \left( aK_0 + \begin{cases} \frac{2}{N} & \text{in co-attended sessions} \\ 0 & \text{in non co-attended sessions} \\ \frac{2}{N_{\text{tot}}} & \text{outside of session times} \end{cases} \right). \tag{5}$$

Figure 3(a) from reference (21) shows an example interaction function, with $T = 0$ corresponding to 1 h before the start of the first topical discussion session and (23) shows an interactive version of scaled interaction as group assignments change.

## Causal analysis through counterfactual conference schedules

We define counterfactual conference schedules as agendas that could have been selected for the conference but were not. These exist because conference agendas were generated using an optimization algorithm (simulated annealing) that attempts to balance many competing factors: e.g. group diversity on many fronts (disciplines, research methodologies, gender, interests in discussion topics), minimizing preconference knowledge of one another among group members, maximizing mixing of group members across multiple groups. Organizers typically generated 500–1,000 potential larger breakout group assignments and 500–1,000 minigroup assignments, then chose the 50 best of each. Typically, those top 50 group assignments (for each group type) varied little in their optimization scores. In total, there were therefore 2,500 ($50 \times 50$) counterfactual conference schedules considered. Note that these counterfactual data were retained for all the virtual conferences, but only for two of the in-person conferences we report on.

For each of the two in-person and five virtual conferences used to analyze causality, we identify the pairs of fellows who formed a collaboration (i.e. submitted a proposal together) and label them with an index $i = 1..N_{\text{pairs}}$. We compute the total amount of scaled interaction time for each pair of collaborators at the actual conference $I_i^A$. We average this interaction over all pairs of collaborators, which gives us the mean total scaled interaction time spent by pairs of collaborators at the actual conference, denoted: $I^A = N_{\text{pairs}}^{-1} \sum_{i=1}^{N_{\text{pairs}}} I_i^A$ (blue/left bar in Fig. 2).

For each of the 2,500 counter-factual schedules indexed $j = 1 \ldots 2{,}500$, we compute the total amount of scaled interaction time for each pair $i$ of *actual* collaborators $I_i^{CF_j}$. We average this interaction over all pairs of actual collaborators and over all 2,500 counter-factual schedules to obtain the average over all counter-factual conferences of the mean total scaled interaction time spent by pairs of collaborators denoted $I^{CF} = (2{,}500 N_{\text{pairs}})^{-1} \sum_{j=1}^{2{,}500} \sum_{i=1}^{N_{\text{pairs}}} I_i^{CF_j}$. For each of the two in-person and five virtual conferences, we considered the counter-factual scenarios where participants had the same mini session assignments as in the actual conference (red/second to left bar in Fig. 2), same breakout session assignments as in the actual conference (yellow/second to right bar in Fig. 2), and different session assignments (gray/right bar in Fig. 2).

Northwestern IRB reviewed and approved this study under IRB ID STU00213499. It was deemed an exempt study. Participants were notified of the ongoing research and provided the opportunity to opt out. All data have been anonymized for privacy of the participants.

## Notes

[a] We present the number of possible pairs in each dataset as a metric for the dataset size and to emphasize the dyad as our unit of analysis, however, we note that social networks are not random.

[b] We note that this does not constitute a perfect natural experiment as there may be confounding variables associated with the pandemic.

[c] Counter-factual scheduling data were retained for two in-person and five virtual conferences—see Methods section for details.

[d] This would be referred to as the "risk ratio" in the context of medical studies.

[e] The conference where the nonlinear memory model was second-best (SLU) was unusual in that many participants knew each other prior to the conference, so $K_0$ values were particularly high. That prior knowledge may have played a larger role than interaction compared with other conferences. The conference with the highest average prior knowledge of all conferences is TDA. Although the memory model outperformed all null models for TDA, the second-best model was the one which considered only prior knowledge and its relative likelihood was closer to the memory model than for other in-person conferences. The topics of both SLU and TDA were related to astrophysics, which may point to disciplinary specificities of the community attending these conferences.

[f] Participants were limited to a maximum of two teams.

[g] We emphasize again that the definition of a collaboration is quite different, though even among Scialog attendees who did not receive any funding, co-authorship of a proposal did lead to later publication in roughly 30% of cases.

[h] Gather is a video-conferencing platform that facilitates organic and unstructured interactions through the use of customizable 2D environments and avatars, replicating the experience of being in a physical space (https://www.gather.town/about).

[i] 0 = unfamiliar, 1 = aware, 2 = had discussions, 3 = collaborated

[j] For simplicity in the model exposition, we considered $K_0 = 0$. When $K_0 > 0$, the model parameter $I_{max}$ needs to be rescaled by the new maximum possible interaction, $6a + 1$ (from a hypothetical 2-person session for a pair with maximal $K_0 = 6$).

## Acknowledgments

E.R.Z. acknowledges the valuable contributions of Maher Said for thoughtful insights, Elena Graetz for discussions on statistical methods, and Guy Amichay for comments on the manuscript.

## Supplementary Material

Supplementary material is available at *PNAS Nexus* online.

## Funding

The authors acknowledge the United States Department of Agriculture NACA 58-3022-0-005 and the Research Corporation for Science Advancement for providing data and assistance. E.R.Z. thanks the National Science Foundation Graduate Research Fellowship Program DGE-184216 and the Buffett Institute for Global Affairs' Global Impacts Graduate Fellowship program for financial support.

## Author Contributions

E.R.Z.: conceptualization, formal analysis, funding acquisition, methodology, software, visualization, writing—original draft; K.H.:. data curation, software; A.L.F.: conceptualization, funding acquisition, investigation, resources, supervision, writing—review & editing; R.J.W.: conceptualization, funding acquisition, investigation, resources, supervision, writing—review & editing; D.M.A.: conceptualization, formal analysis, funding acquisition, methodology, supervision, visualization, writing—original draft.

## Data Availability

The datasets generated by the research and/or analyzed during the current study will be made available upon publication in the Northwestern Arch repository, https://doi.org/10.21985/n2-xckb-qd60.

## References

1 Freeman RB, Ganguli I, Murciano-Goroff R. 2014, Aug. *Why and Wherefore of Increased Scientific Collaboration*. University of Chicago Press. p. 17–48.
2 Klöwer M, Hopkins D, Allen M, Higham J. 2020. An analysis of ways to decarbonize conference travel after COVID-19. *Nature*. 583(7816):356–359.
3 Sarabipour S, Khan A, Seah YFS. 2021. Changing scientific meetings for the better. *Nat Hum Behav*. 5:296–300.
4 Burtscher L, *et al.* 2020, Sep. The carbon footprint of large astronomy meetings. *Nat Astron*. 4(9):823–825.
5 Holden MH, *et al.* 2017, Sep. Academic conferences urgently need environmental policies. *Nat Ecol Evol*. 1(9):1211–1212.
6 Tao Y, Steckel D, Klemeš JJ, You F. 2021, Dec. Trend towards virtual and hybrid conferences may be an effective climate change mitigation strategy. *Nat Commun*. 12(1):7324.
7 Niner HJ, Johri S, Meyer J, Wassermann SN. 2020, Sep. The pandemic push: can COVID-19 reinvent conferences to models rooted in sustainability, equitability and inclusion? *Socioecol Pract Res*. 2(3):253–256.
8 Skiles M, *et al.* 2022. Conference demographics and footprint changed by virtual platforms. *Nat Sustain*. 5(2):149–156.
9 Vargo D, Zhu L, Benwell B, Yan Z. 2021. Digital technology use during COVID-19 pandemic: a rapid review. *Hum Behav Emerg Technol*. 3(1):13–24.
10 Denstadli JM, Julsrud TE, Hjorthol RJ. 2012. Videoconferencing as a mode of communication: a comparative study of the use of videoconferencing and face-to-face meetings. *J Bus Tech Commun*. 26(1):65–91.
11 Staples DS, Zhao L. 2006. The effects of cultural diversity in virtual teams versus face-to-face teams. *Group Decis Negot*. 15:389–406.
12 Weisband S, Atwater L. 1999. Evaluating self and others in electronic and face-to-face groups. *J Appl Psychol*. 84(4):632–639.
13 Wu J, *et al.* 2022. Virtual meetings promise to eliminate geographical and administrative barriers and increase accessibility, diversity and inclusivity. *Nat Biotechnol*. 40(1):133–137.
14 Brucks MS, Levav J. 2022. Virtual communication curbs creative idea generation. *Nature*. 605(7908):108–112.
15 Lin Y, Frey CB, Wu L. 2023, Nov. Remote collaboration fuses fewer breakthrough ideas. *Nature*. 623(7989):987–991.
16 Remmel A. 2021. Scientists want virtual meetings to stay after the COVID pandemic. *Nature*. 591(7849):185–186.
17 Margolis A, Balmer JT, Zimmerman A, López-Arredondo A. 2020. The extended congress: reimagining scientific meetings after the COVID-19 pandemic. *MedEdPublish*. 9(1):128.
18 Porpiglia F, *et al.* 2020. Traditional and virtual congress meetings during the COVID-19 pandemic and the post-COVID-19 era: is it time to change the paradigm? *Eur Urol*. 78(3):301–303.
19 Subramanya SH, Lama B, Acharya KP. 2020. Impact of COVID-19 pandemic on the scientific community. *Qatar Med J*. 2020(1):21. https://doi.org/10.5339/qmj.2020.21.
20 Wiener RJ, Ronco S. 2019. Scialog: the catalysis of convergence. *ACS Energy Lett*. 4(5):1020–1024.
21 Zajdela ER, *et al.* 2022. Dynamics of social interaction: modeling the genesis of scientific collaboration. *Phys Rev Res*. 4(4):L042001.
22 Zajdela E. 2023. Mathematical modeling of complex systems with applications to scientific collaboration at conferences and chimera states for coupled oscillators. (Doctoral dissertation, Northwestern University).
23 Zajdela E. 2021. Exploring the dynamics of team formation at conferences. https://mathemmatician.shinyapps.io/DynamicsOfConferences/ [accessed 2023 Mar 15].
24 Ginsparg P. 2021. Lessons from arXiv's 30 years of information sharing. *Nat Rev Phys*. 3(9):602–603.
25 Atzmueller M, Lemmerich F. 2018. Homophily at academic conferences. In: Companion Proceedings of the The Web Conference 2018. Republic and Canton of Geneva: International World Wide Web Conferences Steering Committee. p. 109–110.
26 Lungeanu A, Huang Y, Contractor NS. 2014. Understanding the assembly of interdisciplinary teams and its impact on performance. *J Informetr*. 8(1):59–70.
27 Contractor NS, Wasserman S, Faust K. 2006. Testing multitheoretical, multilevel hypotheses about organizational networks: an analytic framework and empirical example. *Acad Manage Rev*. 31(3):681–703.
28 Teplitskiy M, Park S, Thompson N, Karger D. 2022. The effect of in-person conferences on the diffusion of ideas. arXiv,

arXiv:2209.01175. https://doi.org/10.48550/arXiv.2209.01175, preprint: not peer reviewed.

29 Langer EJ. 1975. The illusion of control. *J Pers Soc Psychol*. 32(2): 311–328.

30 Granovetter MS. 1973. The strength of weak ties. *Am J Sociol*. 78(6): 1360–1380.

31 Correa J, Hartline J, Immorlica N. 2022. BLOG@CACM: a semester virtual institute: https://cacm.acm.org/blogcacm/a-semester-virtual-institute [accessed 2023 Mar 29].

32 Park S, Song J, Karger DR, Malone TW. 2024. Who2chat: a social networking system for academic researchers in virtual social hours enabling coordinating, overcoming barriers and social signaling. *Proc ACM Hum Comput Interact*. 8(CSCW1): 1–34.

33 Tahlyan D, *et al.* 2024. In-person, hybrid or remote? Employers' perspectives on the future of work post-pandemic. *Transp Res Part A Policy Pract*. 190:104273. https://doi.org/10.1016/j.tra.2024.104273.