nature sustainability

ANALYSIS

https://doi.org/10.1038/s41893-021-00823-2

Check for updates

# Conference demographics and footprint changed by virtual platforms

Matthew Skiles[1], Euijin Yang[1], Orad Reshef[2], Diego Robalino Muñoz[1], Diana Cintron[3], Mary Laura Lind[4], Alexander Rush[5], Patricia Perez Calleja[6], Robert Nerenberg[6], Andrea Armani[7]✉, Kasey M. Faust[1]✉ and Manish Kumar[1]✉

**Conferences disseminate research, grow professional networks and train employees. Unfortunately, they also contribute to climate change and present barriers to achieving a socially sustainable work environment. Here, we analyse the recent impact of transforming in-person conferences into virtual conferences on improving diversity, equity and inclusion in science and engineering conferences. Factors including cost, gender, career stage and geographic location were evaluated. Virtual conferences demonstrated a clearly discernable and, in some cases, orders of magnitude improvement across nearly all metrics. On the basis of participant survey results, this improvement may be attributed to a combination of reduced financial and personal-life burdens. However, despite this clear impact, further development of virtual networking features and poster sessions is necessary to achieve widespread adoption and acceptance of this new format.**

Conferences fulfil a range of needs by facilitating dissemination of ideas, initiating collaborative relationships and providing education, training and career opportunities. Traditional in-person conferences (IPCs) have filled this role for centuries[1] and these events cut across all sectors: academia, industry and government. However, this format has been criticized as outdated and detrimental to the environment[2–4]. More recently, emerging evidence is also connecting this modality to social sustainability issues as well, notably poor retention of a diverse workforce. In this context, the two dominant contributors are the intrinsic power-imbalance in the workplace and an imbalance in home-life responsibilities[5,6].

Over the past two decades, the creation and sustainment of a diverse, equitable and inclusive (DEI) work environment in the scientific and engineering community has not kept pace with many other fields. In part, this can be attributed to career expectations revolving around conference travel and participation. Participation in conferences can be cost prohibitive for many, as the cumulative expenses can be thousands of US dollars per person. International travel creates additional barriers[7] which are exacerbated by the frequent changes in document requirements and lengthy delays in obtaining visas. These financial and documentation barriers can also dissuade scientists that have difficulty securing funding to cover conference costs such as students, postdoctoral researchers or scientists from historically under-represented institutions. These factors can also exclude participants from countries that do not have very high research activity, such as nations that are not in the top ten research countries as defined by the Nature Index (NI; ref. [8]), NI > 10.

However, even for those researchers who are able to travel, the time away from home necessitated by work-related travel is intrinsically exclusionary to care-givers, who are primarily women[3,7,9]. Yet, given how important conference attendance is to career advancement, this community is frequently faced with the decision of choosing between work and family. Lastly, despite conference organizers' attempts to solve accessibility concerns of the disabled community, many conferences still fall short of providing an equitable experience.

The recent surge in virtual events is forcing the scientific community to re-evaluate its long-held position against virtual conferences (VCs). The initial anecdotal evidence indicated that VCs enabled a more diverse population to participate. But a quantitative analysis of the impact on DEI challenges has yet to be performed. Such analysis is critical to make decisions regarding the format of future events, potentially resulting in a paradigm shift in the field. Here, we evaluate several metrics, including cost, carbon footprint, impact of conference format and attendee demographics. We collected historical data from three IPCs based in the United States, of varying sizes and disciplines within science, technology, engineering and mathematics (STEM). These results were compared to the same three conferences after they transitioned to a VC format in 2020. These scientific conferences were among the early conferences to transition online in response to the COVID-19 pandemic and were chosen to investigate the impact of an abrupt transition from historically IPCs to a new virtual format.

The historically IPCs-turned-VCs analysed here are the annual International Conferences on Learning Representations (ICLR), the American Astronomical Society (AAS) and the North American Membrane Society (NAMS) conferences. Also analysed here are several conference series that were originally designed for the VC environment, including the Photonics Online Meetups (POM 1, January 2020; POM 2, June 2020) and the International Water Association (IWA) Biofilms conference. These conferences span

[1]Department of Civil, Architectural and Environmental Engineering, The University of Texas at Austin, Austin, TX, USA. [2]Department of Physics, University of Ottawa, Ottawa, Ontario, Canada. [3]Department of Chemical Engineering, The University of Texas at Austin, Austin, TX, USA. [4]School for Engineering of Matter, Transport and Energy, Arizona State University, Tempe, AZ, USA. [5]Department of Computer Science, Cornell University, Ithaca, New York, NY, USA. [6]Department of Civil and Environmental Engineering and Earth Sciences, University of Notre Dame, Notre Dame, IN, USA. [7]Department of Chemical Engineering and Materials Science, University of Southern California, Los Angeles, CA, USA. ✉e-mail: armani@usc.edu; faustk@utexas.edu; manish.kumar@utexas.edu

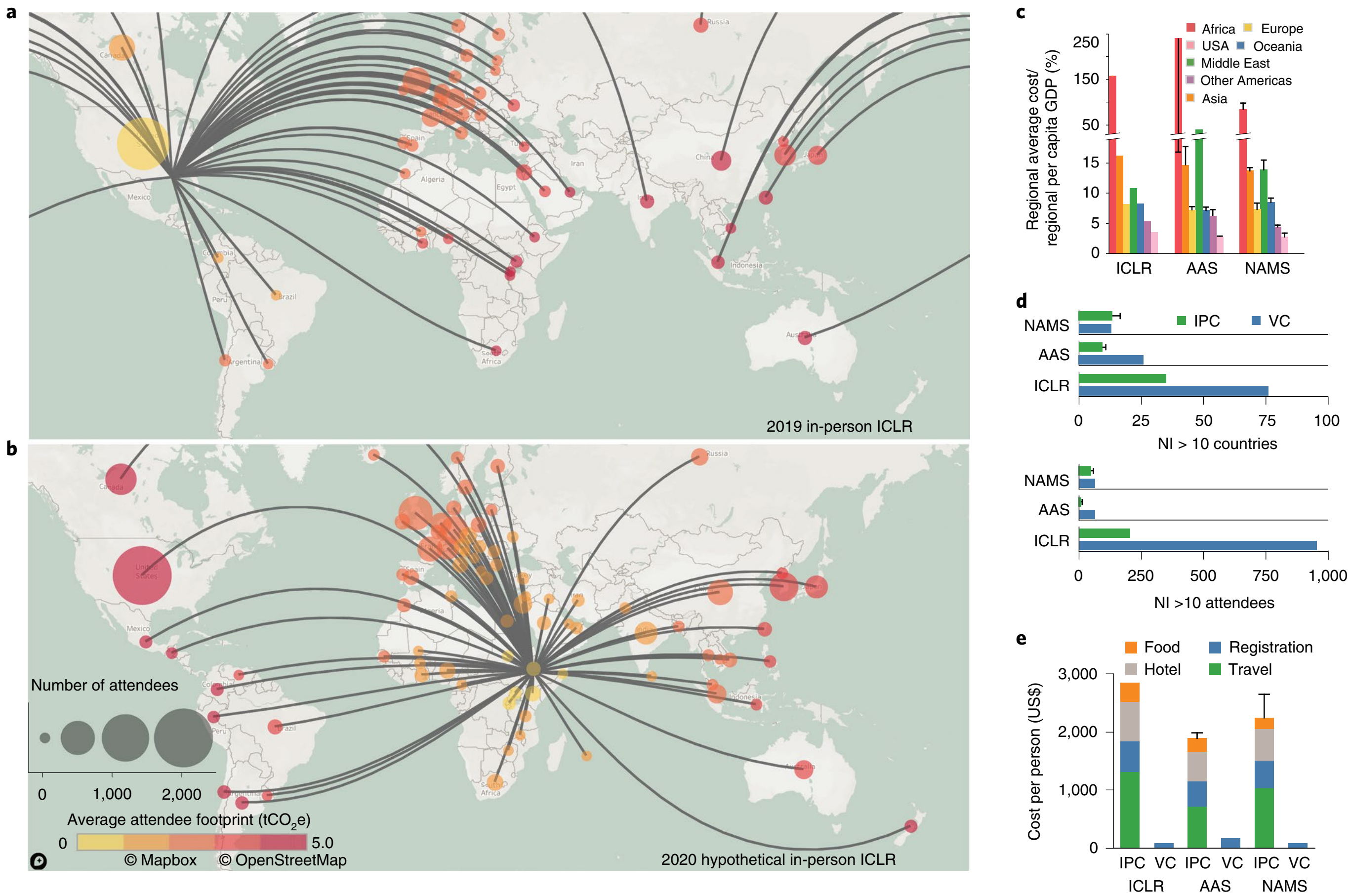


**Fig. 1 | VCs increase overall attendance and geographical diversity while reducing costs.** **a**, The delegation for the 2019 ICLR IPC located in the United States was global but concentrated in the United States ($n = 2,584$). **b**, The delegation for the 2020 ICLR, which was originally scheduled to occur in Ethiopia but transitioned online, was larger ($n = 4,980$) and more geographically diverse. **c**, Regional average cost of attendance to IPCs as a percentage of attendee country's GDP per capita for ICLR ($n = 2$), AAS ($n = 4$) and NAMS ($n = 4$) conferences was higher for African participants and very low for US participants. Error bars are not included for AAS Middle East because $n < 3$. **d**, The delegations for 2020 ICLR ($n = 1$), AAS ($n = 1$) and NAMS ($n = 1$) VCs generally represented more countries that were not in the top ten research countries as defined by NI > 10 (ref. [8]) and included a higher number of attendees from those countries compared to the average delegations from IPCs. **e**, Average registration, food, hotel and travel costs for a single attendee to past ICLR ($n = 2$), AAS ($n = 4$) and NAMS ($n = 4$) IPCs totalled thousands of US$, compared to <US$200 for 2020 ICLR ($n = 1$), AAS ($n = 1$) and NAMS ($n = 1$) VCs. Error bars are the propagated uncertainty for food, registration, hotel and travel costs. Error bars in all panels are defined as standard deviation and are not included for ICLR IPC data because $n < 3$. Credit: **a**,**b**, © Mapbox and © OpenStreetMap contributors.

five fields of science and engineering and range from small- to large-scale events. All have international audiences.

We focused our analysis on the environmental, social and economic costs of VCs versus IPCs and accompanying demographic impacts (global participation), participation from women, early career researchers and scientists from under-represented institutions. We also assessed the challenges and benefits of the VC format.

## Results

**Demographic impact.** The elimination of the travel and cost burdens realized with the VC format resulted in a large increase in attendance at all events (Fig. 1). The increase in attendance was particularly pronounced for international attendees. We propose that this trend may be related to the substantial decrease in costs as compared to IPCs as described below.

The cost of attending IPCs for international attendees was dominated by airfare (Fig. 1 and Supplementary Tables 1 and 2). When compared to US attendees, the average researcher from Africa, Asia, Europe, the Middle East and Oceania paid between 90% and 210% more to attend NAMS IPCs (Supplementary Table 3). When placed in financial context, the cost of attendance for scientists from Africa to past ICLR (2018–2019), AAS (2016–2019) and NAMS (2015–2019) IPCs was on average between 80% and 250% of their country's annual per capita gross domestic product (GDP), compared to ~3% of per capita GDP for US participants (Fig. 1c and Supplementary Table 4). Cost of attendance for participants from Asia to past ICLR (2018–2019), AAS (2016–2019) and NAMS (2015–2019) IPCs was ~15% of their country's per capita GDP (Fig. 1c and Supplementary Table 4). However, it is important to note that many conferences not included in this analysis have registration fees in excess of $700. For these events, registration fees can begin to compete with airfare as an important financial consideration.

The 2020 ICLR, AAS and NAMS VC delegations were more geographically diverse, probably due to the elimination of these travel and registration costs as seen from responses to our surveys (Supplementary Information). Notably, the audiences were ~40–120% larger than the historical average for IPCs (Supplementary Table 5). Attendance by scientists from NI > 10 countries increased from the historical average at ICLR, AAS and NAMS IPCs to the 2020 ICLR, AAS and NAMS VCs (Fig. 1d and

Supplementary Table 6). The increased representation was more comparable to delegations seen at conferences originally designed for the virtual environment; specifically, 31–38% of attendees at the POM 1, POM 2 and IWA VCs were from NI > 10 countries (Supplementary Fig. 1).

In this context, the environmental impact of international conferences can also be considered. In a collection of decarbonization pathways designed to limit global warming to 1.5 °C with a small overshoot, the median global per capita carbon budget for the entire year of 2030 was 3.26 t of $CO_2$ equivalents ($CO_2$e) (ref. [10]). The carbon footprint for a single international attendee to the 2019 ICLR, AAS or NAMS IPCs approached this value. Conversely, the cumulative footprints of the >7,000 attendees to 2020 ICLR, AAS and NAMS VCs (1.07 t$CO_2$e) was comparable to the average footprint of a single attendee (combined average of domestic and international) to one of the analysed 2019 IPCs as shown in Supplementary Fig. 2 and Supplementary Table 7 and discussed further in the Supplementary Information.

**Participation of women.** The VC format also eliminated travel burdens that can act as a barrier to attendance for certain sociodemographic groups. This impact is probably reflected in changes in the gender make up of VC delegations (Supplementary Table 8) and supported by survey responses to a follow-up survey sent separately to men and women attendees of NAMS 2020 (Supplementary Table 9). Attendance by women increased between 60% and 260% at ICLR, AAS and NAMS VCs compared to the IPC baselines (Fig. 2 and Supplementary Table 10). On average, women made up larger fractions of the conference delegations at 2020 VCs as compared to IPCs (Fig. 2g and Supplementary Table 11). The increase in the number of female attendees is particularly notable considering that women make up smaller portions of STEM fields compared to men. For example, women comprise only 33–34% of STEM researchers in the countries that made up the delegations for historical ICLR, AAS and NAMS IPCs (Supplementary Tables 12, 13 and 14). Survey responses confirmed that the elimination of the travel requirement realized with VCs partially explain trends in attendance by gender. For example, about half (47%) of the 2020 NAMS VC survey respondents that did not plan on attending the originally scheduled 2020 NAMS IPC indicated that the primary reason for attending the VC was convenience (Supplementary Fig. 3).

Abstract submittals to the 2020 NAMS conference from before and after the decision to switch from an in-person to a virtual format also indicated an increase in interest and participation from female researchers for the VC. Approximately a quarter (26%) of abstracts submitted to the 2020 NAMS IPC were from female researchers, which was aligned with historical average attendance by women to 2015–2019 NAMS IPCs (Supplementary Fig. 4). After it was announced that the 2020 NAMS conference would be held online, 37% of submitted abstracts came from female scientists (Supplementary Fig. 4). The 2020 ICLR VC also saw an increase in attendance from gender queer and transexual scientists. On average, 2018–2019 ICLR IPCs were attended by one gender queer scientist and zero transgender scientists. The 2020 ICLR VC was attended by eight gender queer scientists and two transgender scientists (Fig. 2a). However, it should be noted that this increase in reported attendance by LGBTQ scientists could be the result of an increased willingness to identify as LGBTQ.

**Participation of students and postdoctoral researchers.** High costs characteristic of IPCs can also be exclusionary to certain sociodemographic groups that may face challenges securing funding for travel, such as students and postdoctoral researchers. Cost of attendance to historical NAMS IPCs was on average US$1,612 for students and US$2,142 for postdoctoral researchers. The shift to a virtual environment resulted in a substantial growth in this population of attendees (Fig. 3a–c and Supplementary Table 15). Additionally, on average, for all conferences evaluated, the VC delegations had higher proportions of students (29–42%) and postdoctoral researchers (5–11%) compared to historical IPCs (Fig. 3d and Supplementary Tables 16 and 17). Additionally, the audiences of conferences designed for the virtual environment (POM 1, POM 2 and IWA) were all comprised of >45% students and postdoctoral scholars, demonstrating the impact that virtual events can have on the careers of emerging scholars (Supplementary Fig. 5). The AAS conference surprisingly did not show much change in conference composition as seen from surveys (32% completion) (Supplementary Fig. 6). The role of cost on attendance was evident in survey responses, as 33% of respondents to NAMS surveys indicated that they were not planning on attending the scheduled 2020 NAMS IPC before the decision to move online (Supplementary Fig. 7). Of the respondents that were not planning on attending, 34% indicated that cost was the primary motivation for attending the 2020 NAMS VC (Supplementary Fig. 3).

**Participation of historically under-represented institutions.** A unique and particularly challenging subset of researchers to engage are those from primarily undergraduate institutions (PUIs) and high research activity (R2) universities (as distinct from the very high research activity category, R1). Attendance from both groups increased at VCs. At the 2020 NAMS VC, attendance by researchers from PUIs and R2 universities increased from the IPC baseline by 157% and 45%, respectively. Similarly, attendance at the 2020 AAS VC from PUIs and R2 universities increased by 72% and 106%, respectively (Fig. 3e and Supplementary Table 18). Increasing participation of researchers from these historically excluded institutions will enhance their educational experiences and provide more research opportunities. Additionally, attending technical events will provide students with mentoring and networking opportunities, potentially increasing the likelihood that they pursue graduate degrees.

**Effect of time zones and conference format.** While VCs may eliminate many barriers to participation, the impact on international attendances seen in this work was strongly dependent on the VC format (Supplementary Fig. 8 and Supplementary Table 19) with the primary variations being synchronous, asynchronous or blended (both options available) content delivery. The 2020 NAMS VC was organized around synchronous live talks. Consequently, attendance from regions where the conference was held during normal work hours was higher than in other regions. As a result, attendance from Europe and the Middle East increased by 102% and 76%, respectively, when compared to the 2015–2019 NAMS IPC average. Conversely, for Asia, where the 2020 NAMS VC was held around or past midnight local time, attendance decreased by 62%. In the case of the 2020 AAS VC which was also synchronous, attendance increased for all regions compared to AAS IPCs (60–700% increase) and the largest percentage increases came from Europe, Oceania and other Americas. Therefore, the dependence on working hours was not universally observed. However, it is important to note that some regions had very small participant numbers which could influence the analysis.

The 2020 ICLR VC was asynchronous, with only a few live events and most talks prerecorded and released for consumption at the attendee's leisure. A live question-and-answer session was held for each keynote speaker after the video had been available for some time, thus affording the opportunity to interact with the speaker. As a result of this format, attendance at the 2020 ICLR VC increased for all regions (by 57–1,700%), when compared to the 2018–2019 ICLR IPC average. Additionally, unlike the AAS and NAMS conferences, >50 people attended the 2020 ICLR VC from every region in the world, increasing confidence in the analysis. On the basis

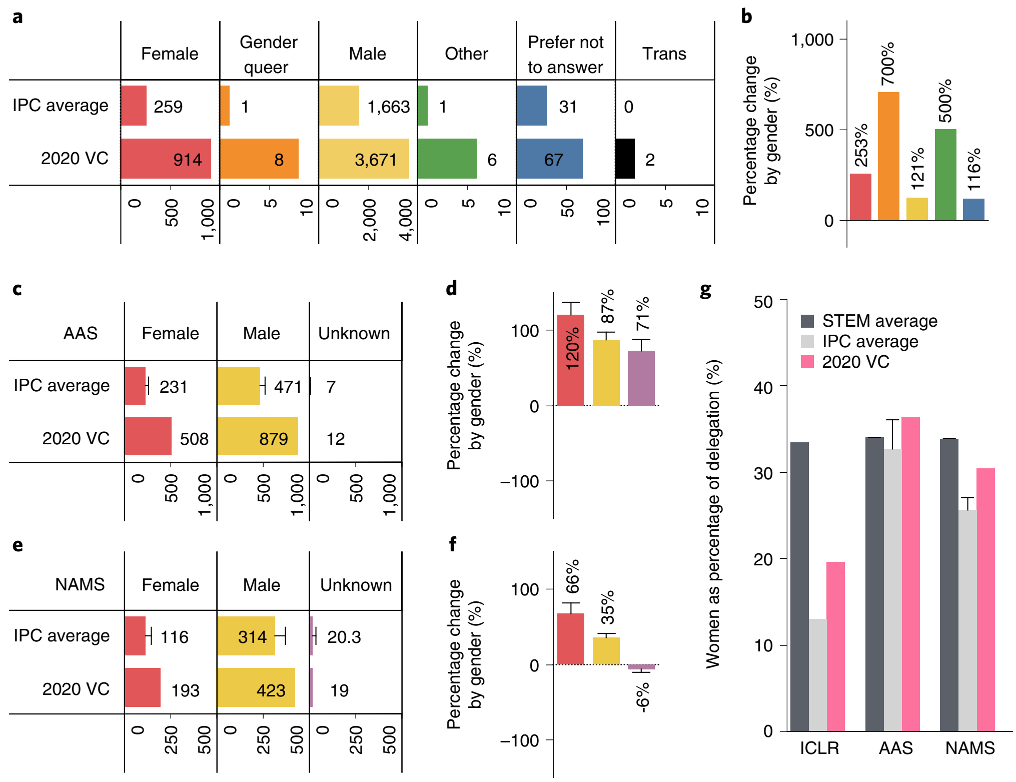

**Fig. 2 | VCs increase gender diversity. a**, The 2020 ICLR VC ($n=1$) was attended by more scientists of all genders compared to the 2018–2019 ICLR IPCs ($n=2$). **b**, A positive percentage change in attendance for all genders was observed between the 2018–2019 ICLR IPCs and the 2020 ICLR VC, with the highest percentage increase in attendance observed for gender queer scientists and scientists that identified as a gender that was not included in the survey. **c**, The 2020 AAS VC ($n=1$) was attended by more male and female scientists compared to the 2016–2019 AAS IPCs ($n=4$). **d**, A positive percentage change in attendance for males and females was observed between the 2016–2019 AAS IPCs and the 2020 AAS VC, with a larger percentage increase for female scientists. **e**, The 2020 NAMS VC ($n=1$) was attended by more male and female scientists compared to the 2015–2019 NAMS IPCs ($n=4$). **f**, A positive percentage change in attendance for males and females was observed between the 2015–2019 NAMS IPCs and the 2020 NAMS VC, with a larger percentage increase for female scientists. **g**, The female fractions of the delegations at the ICLR ($n=1$), AAS ($n=1$) and NAMS ($n=1$) VCs were larger than at historical ICLR ($n=2$), AAS ($n=4$) and NAMS ($n=4$) IPCs and were more comparable to the delegation-specific STEM average, with female fractions in STEM calculated as a weighted average of females in STEM for the origin countries of conference attendees[26,27]. For **a** and **b**: female, red; gender queer, orange; male, yellow; other, green; prefer not to answer, blue; trans, black. For **c**–**f**: female, red; male, yellow; unknown, purple. For **g**: STEM average, dark grey; IPC average, light grey; 2020 VCs, magenta. Error bars in all panels are defined as standard deviation and are not included for ICLR IPC data because $n<3$.

of these results, it is clear that to take full advantage of the virtual format and to make these events effective at disseminating science, it is necessary to offer content asynchronously or using a blended format. A similar blended approach was used by the IWA VC. At IWA, prerecorded presentations were released at a specified time and presenters were available to answer questions during and after the video presentation.

**Initial attendee perceptions of virtual conferences.** The VC format, in general, was well received by attendees and helped to shift negative perceptions to more positive views towards this format. No major alterations in the type of content presented was observed between IPCs and VCs, as discussed in the Supplementary Information. Attendees to 2020 VCs indicated via preconference surveys that they were initially sceptical about the efficacy of VC components but overall felt that the format could possibly improve IPCs in some ways. When asked what they foresaw as the biggest challenge with the virtual format, networking and social interaction was the most common response for NAMS surveys (42% of respondents) and POM 2 surveys (25% of respondents) (Supplementary Fig. 9). Aversion to engaging with the virtual format was lowest among students, as indicated by the fact that only 25% of graduate students and no undergraduate students who submitted abstracts to the 2020 NAMS IPC elected to withdraw from the conference once it was moved online. Conversely, 37% of industry personnel and 39% of postdoctoral researchers who applied to the 2020 NAMS IPC elected not to attend the 2020 NAMS VC (Supplementary Fig. 11). NAMS survey respondents indicated that they were looking forward to some aspects of the virtual format, particularly the opportunity to seamlessly transition between sessions and quickly access the internet to research unfamiliar concepts that arose during the conference.

Part of the success realized by VCs is related to the wide range of currently available virtual environments for hosting oral sessions. Oral sessions at analysed conferences were either livestreamed via webinar (synchronous format) (Supplementary Fig. 12) or prerecorded and released at a specified time (asynchronous format). They were popular among attendees, with 43% of NAMS survey respondents and 74% of POM 2 survey respondents indicating that they preferred the virtual format for oral sessions over the in-person format (Supplementary Figs. 13 and 14). Some presentations and question-and-answer sessions were recorded and made available indefinitely, eliciting persistent viewing after the conference ended. The ICLR platform drew 652,087 total page views during the

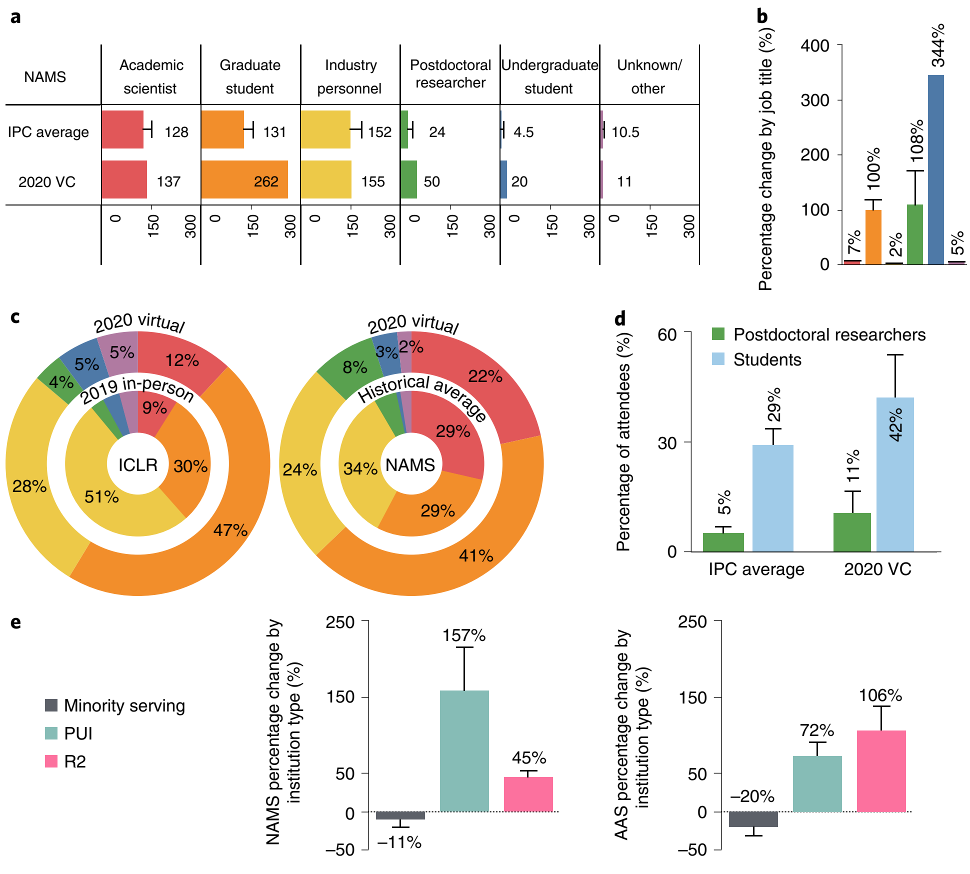


**Fig. 3 | VCs increase participation by early career scientists (students and postdoctoral researchers) and from non-research-intensive institutions.** **a**, The 2020 NAMS VC ($n=1$) was attended by substantially more students and postdoctoral researchers than the 2015–2019 NAMS IPCs ($n=4$), while attendance by other job types remained fairly constant. **b**, A positive percentage change in attendance for all categories was observed between 2015–2019 NAMS IPCs and the 2020 NAMS VC and percentage increase in attendance by students and postdoctoral researchers was very high. Error bar for undergraduate students is too large to be included. **c**, Students and postdoctoral researchers made up a larger percentage and industry personnel and academic scientists represented smaller fractions of both the 2020 ICLR VC delegation ($n=1$) compared to the 2019 ICLR IPC ($n=1$) and the 2020 NAMS VC delegation ($n=1$) compared to the 2015–2019 NAMS IPCs ($n=4$). **d**, On average, postdoctoral researchers and students made up smaller fractions of the delegations at historical IPCs (total $n=6$: ICLR ($n=1$), AAS ($n=1$) and NAMS ($n=4$)) compared to the fractions they represented at analysed 2020 VCs (total $n=6$: ICLR ($n=1$), AAS ($n=1$), NAMS ($n=1$), POM ($n=2$) and IWA ($n=1$)). **e**, A positive percentage change in attendance by persons from PUIs and R2 universities was observed at the 2020 NAMS ($n=1$) and AAS ($n=1$) VCs compared to the 2015–2019 NAMS IPCs ($n=4$) and 2016–2019 AAS IPCs ($n=4$), while attendance from minority serving institutions decreased but this is probably a result of small sample sizes (attendees from minority serving institutions <10). Error bars in all panels are defined as standard deviation.

scheduled conference days and then views increased again by 74% (481,092 additional views) in the 3 months following the conference, indicating an increase in exposure for presenters and sponsors compared to the in-person format (Supplementary Fig. 15).

Analysed VCs had poster authors publish their posters via Twitter, using a web-based iPoster sharing platform, or by uploading a 5-minute prerecorded presentation to the conference website. The poster presentations had high view counts (NAMS iPosters had on average 142 views) (Supplementary Fig. 16) but presenters could not tell how many attendees were viewing their posters and features for communicating with poster viewers were not effective. In contrast, Twitter-based poster sessions are increasing in frequency and allow asynchronous communication. However, Twitter is not available in every country, limiting access. Consequently, virtual posters were less popular, with 85% of NAMS survey respondents and 43% of POM 2 survey respondents indicating that they preferred in-person poster sessions to virtual poster sessions (Supplementary Figs. 16 and 14).

Analysed VCs attempted to facilitate networking by using a variety of social media, messaging, video chat and virtual reality features (Supplementary Table 20). However, survey respondents indicated that the interactions felt inauthentic and contrived. As a result, 75% of POM 2 survey respondents and 96% of NAMS survey respondents indicated that they preferred in-person networking to virtual networking (Supplementary Figs. 13 and 14). In response to this feedback, VCs that occurred later in 2020 and in early 2021 took advantage of improvements in virtual networking technology. These features included robust central chat and discussion board features, as well as Gather.town, an app that allowed participants to navigate a virtual room with an avatar and video chat with other avatars in close proximity. The January 2021 POM used Gather.town to hold a virtual job fair and poster session among other networking events. Gather.town was also used at the 2020 IWA VC and was popular with attendees, as all 56 survey respondents indicated that they would like the Gather.town Interactive Lounge feature to be included in future IWA VCs.

## Discussion

Our findings reveal that VCs reduce the environmental impact of conferences, the financial burden and the social cost. In the VC format, researchers are much more likely to be able to overcome economic and travel-related barriers that are intrinsic to IPCs and that ultimately discourage participation from institutions and countries with limited resources, women, disabled scientists and early career researchers and practitioners (for example, students and postdoctoral researchers). These factors are discussed further in the Supplementary Information. Thus, virtual formats can provide an excellent avenue to address DEI challenges stemming from barriers to participation and representation at IPCs and other professional events. However, despite these clear benefits, the difficulties networking in a virtual environment are routinely emphasized as a limitation.

Of the survey respondents, 75% for POM 2 and 96% for NAMS indicated that they preferred in-person networking to virtual networking (Supplementary Figs. 13 and 14). Analysed VCs experimented with incorporating social media and organizing virtual breakout rooms to facilitate networking with some success. However, survey respondents indicated that the interactions felt inauthentic and contrived. Therefore, while virtual networking technology has improved considerably, there is substantial need for further development of these features as well as research into their efficacy.

One approach to overcome this challenge and increase in-person interaction without increasing cost or travel was piloted during POM 1 by creating locally organized viewing and networking sites (POM-hubs). This 'conference within a conference' approach allowed for reduced cost and travel, increased local and regional networking and created an international conference. Notably, approximately half of the POM 1 attendees participated in the conference from a local hub-site[1]. This hybrid hub approach pioneered by POM 1 (ref. [1]) is a promising solution to this challenge that warrants further study. A hybrid format could allow communities to realize many of the advantages identified by this analysis of COVID VCs, while still offering the option of a traditional IPC experience. It would be ideal for postpandemic conferences to use the knowledge gained on the benefits of expanding inclusion using virtual tools. The resultant conferences could facilitate networking and effective dissemination of scientific knowledge to diverse audiences in an environmentally sustainable manner, moving toward more equitable environments and opportunities. Innovative VC strategies and platforms used to administer oral and poster sessions and virtual networking are further discussed in the Supplementary Information along with additional discussion on organization recommendations.

Our study is characterized by one important limitation. While nearly all interactions made the abrupt shift from in-person to virtual, our analysis is focused on STEM subjects. In some ways, the demographic and financial sensitivities of this population are distinct from other academic communities or an industry or government audience. However, they do share several similarities, particularly for global industry consortiums. Notably, all groups are sensitive to international politics and visa policies, fluctuations in currencies and the financial markets, and gender inequities. However, the attendees at scientific events tend to be highly educated (BSc degree or higher in a STEM field) and speak English as a primary or secondary language. These limits do not adversely affect our conclusions, as we are focusing on STEM. However, to extend our conclusions outside of higher education and STEM fields specifically, a broader population analysis should be performed with appropriate benchmarking. Such an analysis will require engaging conference organizers in other areas including humanities, commerce, business as well as related industry, non-profits and government organizations.

In addition to extending the analysis outside of STEM, the present research findings motivate several new areas of investigation. A few examples include: (1) developing strategies for improving virtual networking; (2) role of organization type on the impact of travel (small versus large business, domestic versus global); (3) policy development by technical/scientific societies, funding agencies and universities; and (4) longitudinal study tracking travel and career progression. These topics are discussed further in the Supplementary Information. In this context, we consider the present conclusions to be an important step in understanding the positive impact of VCs, paving the way for future policy decisions and reducing DEI challenges in the workplace.

## Methods

**Data.** Registration, digital platform and survey data were collected from three IPCs-turned-VCs and are presented in Supplementary Table 21. The three analysed IPCs-turned-VCs include the annual ICLR (~2,300 historical average attendees), the AAS summer meeting (~700 historical average attendees) and the NAMS annual conference (~450 historical average attendees). Complementing this are data from the POMs (~1,000 attendees) and the IWA (~350 attendees), conference series that were specifically designed for the virtual ecosystem. These conferences represent varying fields and community sizes and allow for comparisons across a range of STEM backgrounds. Data for IPCs-turned-VCs were collected for 2020 VCs and for historical IPCs. POM and IWA data provided a control for an always VC, while the baseline data for historically IPCs allowed for the elimination of effects from other variables, facilitating direct analysis of the impact that virtual components had on conference performance.

Specific data collected include registration and abstract information, spanning information such as the number and type of participants (for example, students and industry personnel), geographic participation, institution and gender. For IPCs-turned-VCs, these data were collected for registrations accrued before and after moving online. Carbon footprint and cost of attendance were estimated on the basis of attendee work locations and conference destinations. Descriptive statistics[11] and thematic mapping[12] were applied to understand changing sociodemographics realized in the shift to a virtual format. Additional data collected on webinar attendance and virtual platform activity were used to assess the efficacy with which the VCs distributed content to attendees. Qualitative data were collected by asking participants to fill out polls as well as pre- and postconference surveys designed to interrogate the participant experience and field suggestions for improvement. Surveys were also used to investigate the impact of travel burden and cost barriers for female versus male NAMS attendees. Survey questions included multiple-choice and open-ended questions about specific conference components and the participant experience. The surveys were produced by the authors for the conferences that they organized. Survey and polling questions underwent IRB review receiving and exempt status (protocol 2020-05-0026) at The University of Texas at Austin.

Sociodemographic data were provided by conference organizers and filled in as necessary. Attendee countries were manually categorized by region for analysis. Job-type data (that is, graduate student, industry personnel) were provided by conference organizers via registration or survey data. Data that included specific job titles (that is, operations director, research scientist) for attendees were categorized manually by job type. Gender data were provided by organizers for some conferences via voluntary surveys. Gender data for the NAMS conference were manually assigned on the basis of author familiarity with the participants and through internet search of attendee names. The Gender API[13] was also used to assign gender to attendee names for NAMS and AAS conference attendees. Due to confidence in the accuracy of manually assigned names for NAMS attendees, discrepancies in the genders assigned to NAMS attendees by the manual process and the Gender API indicated that the Gender API was less accurate than the manual process (Supplementary Table 8). Consequently, the Gender API was only applied to assign gender to AAS participants. Attendee academic institutions were manually categorized according to databases of institution types. Minority serving institutions were defined according to the 2007 US Department of Education database[14]. High research institutions (R2) were defined as any institution that was included in the 2018 Carnegie classification of R2 universities[15]. PUI were defined as any university that awarded 20 or fewer PhD degrees in NSF-supported fields during the combined previous two academic years[16] as reported by the US National Science Foundation (NSF) records on PhD degrees for major science and engineering fields awarded by universities during 2017 and 2018[17,18]. Non-research-intensive countries were defined as countries that were not in the top ten countries for scientific research as defined by the NI that measured top countries in terms of contributions to papers published in 82 leading journals during 2019 (NI > 10) (ref. [8]).

**Travel distance.** Attendee travel distance, carbon footprint and cost were calculated via python scripts using attendee origin location data provided by conference organizers. NAMS and AAS registrant origin locations were provided by organizers via registration data as a list of attendees with attendee-specific locations. If location for an attendee was not included, origin location was determined via internet search of the attendee name. ICLR and POM registrant origin location data was provided by conference organizers and comprised a list of countries in attendance and the number of attendees from each country. While the sample size of data for single ICLR conferences varied by data type (origin country, gender and job title), origin country was the largest dataset for all ICLR conferences and was thus assumed to be the true size of the conference delegations.

Conference city and attendee origin coordinates were determined by querying the Google Maps API[19] with the location names. If a city-specific attendee origin was not recognized by the API, the attendee origin was set to the attendee's origin country name. Google Maps API queries of only country name return coordinates for the geographical centre of the country. Travel distance between attendee origin and conference location were calculated as the great circle distance (great_circle python package).

**Carbon footprint of attendance.** The carbon footprint of conference attendees was calculated for all IPCs-turned-VCs as the cumulative emissions associated with the flight and hotel stay. The air travel carbon footprint was calculated according to the methodology for the myclimate air travel emissions calculator[20]. The myclimate calculator computes air travel footprint by adding 95 km to the great circle distance to account for flightpath inefficiencies and calculating GHG emissions associated with the fuel burn and life-cycle footprint of the airplane and associated aviation infrastructure. The GHG emissions are then converted to $CO_2$e. It was assumed that all conference attendees flew economy class. If city-specific attendee origin data were available and the attendee was local ($\leq$100 km from the conference city) it was assumed that the attendee did not fly to the conference city and their travel $CO_2$e was set to 0. If registrant origin coordinates were not found, the attendee travel distance and travel footprint were set to the average for that conference.

The carbon footprint per night for the attendee hotel stay was determined using the Hotel Carbon Measurement Initiative (HCMI) rooms footprint per occupied room from the Hotel Sustainability Benchmarking Tool published by the Cornell Center of Hospitality Research[21]. The tool provides city-specific and country-specific footprint data. If data were not available in the Hotel Sustainability Benchmarking Tool for the conference city, then the footprint per night was set to the country average in the tool. If no data were available for the country in which the conference was held, the footprint was set to the value that was closest to the conference location geographically. Student hotel footprint calculations were adjusted to assume shared hotel rooms, that is footprint per night was divided by two. If attendee-specific job title (student versus non-student) information was not available, percentage of students as defined by the voluntary survey data was multiplied by the number of attendees from each country to estimate the number of students from each country. When computing total hotel footprint, it was assumed that attendees stayed for all but one night of the conference (so, for a four-day conference, nightly hotel footprint was multiplied by three). If the attendee was local, the hotel footprint was set to 0. If the attendee origin was not near the conference city and their job title (student versus non-student) was not known, the attendee hotel footprint was set to the conference average.

**Cost of attendance.** Cost of attendance for individual attendees was computed for historically IPCs-turned-VCs by calculating their cost of travel based on air travel distance and summing with the estimated cost of the hotel, food and conference registration fees. Travel cost was calculated as the one-way air travel distance multiplied by the cost distance for air travel defined in ref. [22] and doubled to represent the cost of a round-trip flight. If the registrant was local, their travel cost was set to 0. If the registrant origin was not known, the travel cost was set to the average conference travel distance and converted to cost using ref. [22]. To account for a potential overestimate of travel cost, a sensitivity analysis where the one-way flight cost is multiplied by 1.5 instead of 2 was conducted and is presented in Supplementary Table 1.

NAMS hotel cost was taken from NAMS records. 2020 ICLR hotel cost was set to the average of hotel options provided by the ICLR website. For 2018–2019 ICLR and all AAS conferences, the cost of US hotels was set to the US General Services Administration lodging maximum per diem for the conference city. For 2018–2019 ICLR the cost of all hotels outside of the United States was set to the US State Department lodging maximum per diem for the conference city. Nightly hotel costs were divided by two for students to assume shared rooms. If attendee-specific job title (student versus non-student) information was not available, percentage of students as defined by the voluntary survey data was multiplied by the number of attendees from each country to estimate the number of students from each country. ICLR 2020 student hotel cost data were taken from 'double room rate' and ICLR 2020 non-student hotel cost data were taken from the 'single room rate' cost on the ICLR website. When computing total hotel cost, it was assumed that attendees stayed for all but one night of the conference (so, for a four-day conference, nightly hotel cost was multiplied by three). If the attendee was local, the hotel cost was set to 0. If the attendee was not local but their job title (student versus non-student) was not known, the hotel cost was set to the conference average.

Food cost for conferences held in US cities was taken from US General Services Administration city-specific per diem rates for breakfast, lunch and dinner. For NAMS, one dinner is subtracted from the total cost to account for the banquet dinner provided by NAMS. Food cost for conference cities outside of the United States was taken from US State Department city-specific meals and incidental expenses per diem. Attendees were assumed to stay for all but one night of the conference. If the attendee was local, food cost was set to 0. If the attendee origin was not known, the food cost was set to the conference average.

Registration costs for historical NAMS IPCs was set to the recorded registration fee per registrant. Fees for the sponsor and exhibitor registration types, where sponsors made their contributions via the registration fee, at historical NAMS conferences were set to conference average of that year (these registration types are excluded from the average).

Hypothetical registration fees for a 2020 NAMS IPC were assigned to attendees to the 2020 NAMS VC. The 2020 NAMS attendees with registrant type 'student' were assigned a hypothetical 2020 NAMS IPC registration fee equal to the average fee for students at the 2015–2019 NAMS IPCs (average based on title category, with 'unknown/other' title category excluded from the average). The 2020 NAMS VC attendees with registrant type 'professional/academic' were assigned a registration fee equal to the average fee for non-students at the 2015–2019 NAMS IPCs (average based on title category, 'unknown/other' excluded).

Student and non-student registration fees for 2018–2019 ICLR IPCs were set to early registration fees from the conference website. The registration fees for the 2020 ICLR VC were set to the 2018–2019 ICLR IPC average fees. As attendee-specific job title (student versus non-student) information was not available, percentage of students as defined by the voluntary survey data was multiplied by the number of attendees from each country to estimate the number of students from each country (total student registration fees by country = percentage of students from job title data × total attendees from country × student registration fee).

Registration fees for 2016–2019 AAS IPCs were set to the early registration fees for 'full member/educator/international affiliate', 'graduate student member', 'undergraduate student member', 'emeritus member' and 'amateur affiliate' from the 2020 winter meeting website. As attendee-specific job title information was not available, percentage for attendee job title as defined by the voluntary survey data was multiplied by the number of attendees to estimate the number of each job type in attendance. The total registration fee for each conference was calculated accordingly. The total registration fees were then divided by the number of attendees and the average registration fee was assigned to each registrant.

VC registration fees for ICLR and NAMS were set to US$50 for students and US$100 for non-students. VC registration fees for AAS were set to the full meeting fees for 'full member/LAD member', 'graduate student', 'undergraduate student/high school student', 'emeritus member' and 'amateur affiliate' from the 2020 VC website.

**World map figures.** Attendee origin coordinates and conference city coordinates were converted to great circle distance paths and saved in .kml files using the lxml and geographiclib.geodesic python packages. World maps were plotted using Tableau and MapBox.

**Global annual per capita carbon budget for 2030 and 2050.** Median global carbon budget calculated in terms of Kyoto GHG as $CO_2$e for 2030 and 2050 were taken from a set of decarbonization pathways as outlined in the IPCC report on mitigation pathways compatible with 1.5 °C in the context of sustainable development[10]. The global carbon budget was divided by the medium variant of global population projections for 2030 and 2050 produced by the United Nations Department of Economic and Social Affairs[23].

**Car travel footprint.** Car travel footprint per mile was taken from US Environmental Protection Agency (EPA) estimates for average passenger vehicles[24].

**Virtual conference carbon footprint.** VC footprints were estimated on the basis of emissions for YouTube video streaming multiplied by the projected duration of conference webinar and video streaming by attendees[4].

**Regional average cost/regional per capita GDP.** Country-specific GDP per capita was defined as the 2019 GDP per capita in the attendee country's national currency converted to US$ and divided by the total country population as calculated in the World Economic Outlook Database[25]. Total representative GDP per capita for conference attendees from each region was calculated as the sum of GDP per capita for all the countries in each region multiplied by the number of conference attendees from each country in the region. Total cost of attendance for each region was calculated as the sum of the cost of attendance for all the participants from each region. The regional average cost divided by the regional per capita GDP was calculated by dividing the total cost of attendance for all the attendees from each region by the total representative GDP for the attendees from each region.

**Gender make up of STEM researchers from conference attendee's countries.** Country-specific percentage of women data are taken from 'female researchers as a percentage of total researchers (full-time equivalents)—natural sciences and engineering (subtotal)' published as ref. [26] with the exception of the United States which is not included in that dataset. US percentage of women is derived from women as a percentage of MSc and PhD graduates employed in science and engineering occupations[27]. Overall percentage of women in STEM for the countries represented in the conference delegations was calculated with percentage values from each country represented at the conference, weighted by the number of attendees from each country.

**Reporting Summary.** Further information on research design is available in the Nature Research Reporting Summary linked to this article.

## Data availability

The data that support the plots within this paper and other findings of this study have been deposited on Github[28] (https://doi.org/10.5281/zenodo.5567764). Source data are provided with this paper.

## Code availability

The custom code used to process and analyse the data for this study has been deposited on Github[28] (https://doi.org/10.5281/zenodo.5567764).

Received: 11 November 2020; Accepted: 28 October 2021;
Published online: 9 December 2021

## References

1. Reshef, O. et al. How to organize an online conference. *Nat. Rev. Mater.* **5**, 253–256 (2020).
2. Yakar, D. & Kwee, T. C. Carbon footprint of the RSNA annual meeting. *Eur. J. Radiol.* **125**, 5 (2020).
3. Parker, M. & Weik, E. Free spirits? The academic on the aeroplane. *Manag. Learn.* **45**, 167–181 (2014).
4. Klöwer, M., Hopkins, D., Allen, M. & Higham, J. An analysis of ways to decarbonize conference travel after COVID-19. *Nature* **583**, 356–359 (2020).
5. Hewlett, S. A. et al. The Athena factor: reversing the brain drain in science, engineering, and technology. *Harv. Bus. Rev. Res. Rep.* **10094**, 1–100 (2008).
6. Simard, C., Henderson, A., Gilmartin, S., Schiebinger, L. & Whitney, T. *Climbing the Technical Ladder: Obstacles and Solutions for Mid-Level Women in Technology* (Michelle R. Clayman Institute for Gender Research, 2008).
7. Urry, J. Social networks, mobile lives and social inequalities. *J. Transp. Geogr.* **21**, 24–30 (2012).
8. *Nature Index 2020 Annual Tables* (Springer Nature, 2020); https://www.natureindex.com/annual-tables/2020
9. Cohen, S. A. & Gossling, S. A darker side of hypermobility. *Environ. Plan. A* **47**, 1661–1679 (2015).
10. Rogelj, J. et al. in *IPCC Special Report on Global Warming of 1.5 °C* (eds Masson-Delmotte, V. et al.) Ch. 2 (WMO, 2018).
11. Fisher, M. J. & Marshall, A. P. Understanding descriptive statistics. *Aust. Crit. Care* **22**, 93–97 (2009).
12. Schaab, G., Adams, S. & Coetzee, S. Conveying map finesse: thematic map making essentials for today's university students. *J. Geogr. High. Educ.* https://doi.org/10.1080/03098265.2020.1850656 (2020)
13. *Gender API* (gender-api.com, accessed 30 October 2020); https://gender-api.com
14. *Accredited Postsecondary Minority Institutions* (US Department of Education, accessed 30 October 2020); https://www2.ed.gov/about/offices/list/ocr/edlite-minorityinst-list-tab.html
15. *The Carnegie Classification of Institutions of Higher Education* (Indiana Univ. School of Education, accessed 30 October 2020); https://carnegieclassifications.iu.edu/lookup/srp.php?clq=%7B%22basic2005_ids%22%3A%2216%22%7D&start_page=standard.php&backurl=standard.php&limit=0,50
16. *Facilitating Research at Primarily Undergraduate Institutions: Research in Undergraduate Institutions (RUI) and Research Opportunity Awards (ROA)* (United States National Science Foundation, accessed 30 October 2020); https://www.nsf.gov/pubs/2014/nsf14579/nsf14579.htm
17. *Doctorates Awarded, by State or Location, Broad Field of Study, and Sex of Doctorate Recipients: 2017* (United States National Science Foundation, 2018).
18. *Doctorates Awarded, by State or Location, Broad Field of Study, and Sex of Doctorate Recipients: 2018* (United States National Science Foundation, 2019).
19. *Google Maps v3 API* (Google Inc., accessed 30 October 2020); https://developers.google.com/maps/documentation/geocoding/overview
20. *The myclimate Flight Emission Calculator* (Foundation myclimate, accessed 30 October 2020); https://www.myclimate.org/
21. Chong, H. G. & Ricaurte, E. E. *Hotel Sustainability Benchmarking Tool 2015: Energy, Water, and Carbon* (Cornell Hospitality Reports, Cornell Univ., 2015).
22. Dudas, G., Boros, L., Pal, V. & Pernyesz, P. Mapping cost distance using air traffic data. *J. Maps* **12**, 695–700 (2016).
23. *World Population Prospects 2019* (United Nations, 2019); https://population.un.org/wpp/Publications/
24. *Greenhouse Gases Equivalencies Calculator - Calculations and References* (United States Environmental Protection Agency, 2020); https://www.epa.gov/energy/greenhouse-gases-equivalencies-calculator-calculations-and-references
25. *World Economic Outlook Database* (International Monetary Fund, accessed 30 October 2020); https://www.imf.org/external/pubs/ft/weo/2017/02/weodata/index.aspx
26. *Research and Development* (UNESCO Institute for Statistics, 2020).
27. *Employed Scientists and Engineers, by Occupation, Highest Degree Level, and Sex: 2017* (National Science Foundation, accessed 30 October 2020); https://ncses.nsf.gov/pubs/nsf19304/data
28. Skiles, M. *Virtual-Conferences-Project* (Zenodo, 2021); https://doi.org/10.5281/zenodo.5567764

## Acknowledgements

We acknowledge G. Ragusa's contribution to designing the pre- and postsurveys used for the POM 1 and 2 conferences. The work in this paper was supported by the NSF (award no. CBET 2029219 and CBET 1946392). We gratefully acknowledge access to data and consultation support provided by NAMS, ICLR, POM and IWA meeting leadership. K. Marvel's input on AAS data and trends and his role in providing access to AAS meeting data is also acknowledged. This material is based on work supported by the NSF Graduate Research Fellowship Program under grant no. DGE-1610403. Any opinions, findings and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the NSF.

## Author contributions

M.K., K.M.F. and M.S. conceived the idea. M.K., M.S., D.R.M. and D.C. collected data. M.S. and E.Y. analysed data. O.R., M.L.L, P.P.C, R.N., A.R. and A.A. provided access to data and provided insights on data. M.K., K.M.F., A.A. and M.S. wrote the manuscript.

## Competing interests

M.K. and M.L.L. were organizers of NAMS. A.R. was an organizer of ICLR. P.P.C. and R.N. were organizers of IWA. A.A. and O.R. were organizers of POM. All other authors have no competing interests.

## Additional information

**Supplementary information** The online version contains supplementary material available at https://doi.org/10.1038/s41893-021-00823-2.

**Correspondence and requests for materials** should be addressed to Andrea Armani, Kasey M. Faust or Manish Kumar.

**Peer review information** *Nature Sustainability* thanks Meagan Mauter and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

nature research

Corresponding author(s): Andrea Armani, Kasey Faust, Manish Kumar
Last updated by author(s): Oct 13, 2021

# Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see our Editorial Policies and the Editorial Policy Checklist.

## Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

| n/a | Confirmed | |
|---|---|---|
| ☐ | ☒ | The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement |
| ☐ | ☒ | A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly |
| ☒ | ☐ | The statistical test(s) used AND whether they are one- or two-sided<br>*Only common tests should be described solely by name; describe more complex techniques in the Methods section.* |
| ☒ | ☐ | A description of all covariates tested |
| ☐ | ☒ | A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons |
| ☒ | ☐ | A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals) |
| ☒ | ☐ | For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted<br>*Give P values as exact values whenever suitable.* |
| ☒ | ☐ | For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings |
| ☒ | ☐ | For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes |
| ☒ | ☐ | Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated |

*Our web collection on statistics for biologists contains articles on many of the points above.*

## Software and code

Policy information about availability of computer code

| | |
|---|---|
| Data collection | Usage data for an open source conference-hosting platform called mini-conf were collected via google analytics. Participation data for webinar sessions were collected with the commercial platforms Zoom and Webex. Data on attendee interaction with virtual posters were collected with the commercial iPoster platform. Survey data were collected using the commercial platforms Qualtrics and SurveyMonkey. Coordinates for attendee origin locations were collected using the Google Maps API. |
| Data analysis | Custom code was written in Python 3 to analyze collected data. The code will be made available on Github. |

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.

## Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:
- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

The data that support the plots within this paper and other findings of this study have been deposited on Github (DOI: doi.org/10.5281/zenodo.5567764).

# Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☐ Life sciences ☒ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

# Behavioural & social sciences study design

All studies must disclose on these points even when the disclosure is negative.

| | |
|---|---|
| Study description | The performance of in-person and virtual scientific conferences was analyzed along with the impact of the conference format on the environment and attendee socio-demographics. Informing this analysis was quantitative data on attendee socio-demographics, cost and carbon footprint of conference attendance, as well as qualitative data from surveys distributed to conference participants. |
| Research sample | The research sample for the project was the delegations at three historically in-person turned virtual conferences: The International Conference on Learning Representations (ICLR), The North American Membrane Society Conference (NAMS), and the American Astronomical Society Conference (AAS). For each of the three conferences, data was available for multiple historical in-person years and for a 2020 virtual conference. The sample also included data from two always virtual conferences (The Photonics Online Meetup and The International Water Association Biofilms conference). Data were collected on the job type, gender, place of origin, and institution of the conference attendees. The sample includes conferences that represent varying fields, allowing for comparison across different STEM backgrounds. |
| Sampling strategy | Sample sizes were dictated by the number of attendees at each conference and the number of respondents to distributed surveys. Historical average attendance to the three conventionally in-person turned virtual conferences was ~450 historical average attendees (NAMS), ~700 historical average attendees (AAS), and ~2300 historical average attendees (ICLR). The wide size range of the analyzed events allows for comparison across difference conference sizes. |
| Data collection | Data was collected from registration information provided by attendees, responses to virtual surveys, and through virtual platforms. |
| Timing | Data was collected between January 2020 and July 2021. |
| Data exclusions | Some registration fee data for the NAMS conference was excluded as it comprised of a donation to the organization and was not a true representation of the cost of attendance to NAMS. |
| Non-participation | No participants dropped out or declined participation. |
| Randomization | Participants were not allocated into experimental groups. |

# Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

## Materials & experimental systems

| n/a | Involved in the study | |
|---|---|---|
| ☐ | ☐ | Antibodies |
| ☐ | ☐ | Eukaryotic cell lines |
| ☐ | ☐ | Palaeontology and archaeology |
| ☐ | ☐ | Animals and other organisms |
| ☐ | ☒ | Human research participants |
| ☐ | ☐ | Clinical data |
| ☐ | ☐ | Dual use research of concern |

## Methods

| n/a | Involved in the study | |
|---|---|---|
| ☐ | ☐ | ChIP-seq |
| ☐ | ☐ | Flow cytometry |
| ☐ | ☐ | MRI-based neuroimaging |

## Antibodies

| | |
|---|---|
| Antibodies used | *Describe all antibodies used in the study; as applicable, provide supplier name, catalog number, clone name, and lot number.* |
| Validation | *Describe the validation of each primary antibody for the species and application, noting any validation statements on the manufacturer's website, relevant citations, antibody profiles in online databases, or data provided in the manuscript.* |

## Eukaryotic cell lines

Policy information about cell lines

| | |
|---|---|
| Cell line source(s) | *State the source of each cell line used.* |
| Authentication | *Describe the authentication procedures for each cell line used OR declare that none of the cell lines used were authenticated.* |
| Mycoplasma contamination | *Confirm that all cell lines tested negative for mycoplasma contamination OR describe the results of the testing for mycoplasma contamination OR declare that the cell lines were not tested for mycoplasma contamination.* |
| Commonly misidentified lines (See ICLAC register) | *Name any commonly misidentified cell lines used in the study and provide a rationale for their use.* |

## Palaeontology and Archaeology

| | |
|---|---|
| Specimen provenance | *Provide provenance information for specimens and describe permits that were obtained for the work (including the name of the issuing authority, the date of issue, and any identifying information).* |
| Specimen deposition | *Indicate where the specimens have been deposited to permit free access by other researchers.* |
| Dating methods | *If new dates are provided, describe how they were obtained (e.g. collection, storage, sample pretreatment and measurement), where they were obtained (i.e. lab name), the calibration program and the protocol for quality assurance OR state that no new dates are provided.* |

☐ Tick this box to confirm that the raw and calibrated dates are available in the paper or in Supplementary Information.

| | |
|---|---|
| Ethics oversight | *Identify the organization(s) that approved or provided guidance on the study protocol, OR state that no ethical approval or guidance was required and explain why not.* |

Note that full information on the approval of the study protocol must also be provided in the manuscript.

## Animals and other organisms

Policy information about studies involving animals; ARRIVE guidelines recommended for reporting animal research

| | |
|---|---|
| Laboratory animals | *For laboratory animals, report species, strain, sex and age OR state that the study did not involve laboratory animals.* |
| Wild animals | *Provide details on animals observed in or captured in the field; report species, sex and age where possible. Describe how animals were caught and transported and what happened to captive animals after the study (if killed, explain why and describe method; if released, say where and when) OR state that the study did not involve wild animals.* |
| Field-collected samples | *For laboratory work with field-collected samples, describe all relevant parameters such as housing, maintenance, temperature, photoperiod and end-of-experiment protocol OR state that the study did not involve samples collected from the field.* |
| Ethics oversight | *Identify the organization(s) that approved or provided guidance on the study protocol, OR state that no ethical approval or guidance was required and explain why not.* |

Note that full information on the approval of the study protocol must also be provided in the manuscript.

## Human research participants

Policy information about studies involving human research participants

| | |
|---|---|
| Population characteristics | See above. |
| Recruitment | Participants that attended the selected conferences were included in the study. As surveys were voluntary, survey data only represents conference attendees who elected to participate. |
| Ethics oversight | The Institutional Review Board (IRB) at the University of Texas at Austin. |

Note that full information on the approval of the study protocol must also be provided in the manuscript.

## Clinical data

Policy information about clinical studies

All manuscripts should comply with the ICMJE guidelines for publication of clinical research and a completed CONSORT checklist must be included with all submissions.

| | |
|---|---|
| Clinical trial registration | *Provide the trial registration number from ClinicalTrials.gov or an equivalent agency.* |
| Study protocol | *Note where the full trial protocol can be accessed OR if not available, explain why.* |

| | |
|---|---|
| Data collection | *Describe the settings and locales of data collection, noting the time periods of recruitment and data collection.* |
| Outcomes | *Describe how you pre-defined primary and secondary outcome measures and how you assessed these measures.* |

## Dual use research of concern

Policy information about dual use research of concern

### Hazards

Could the accidental, deliberate or reckless misuse of agents or technologies generated in the work, or the application of information presented in the manuscript, pose a threat to:

| No | Yes | |
|---|---|---|
| ☐ | ☐ | Public health |
| ☐ | ☐ | National security |
| ☐ | ☐ | Crops and/or livestock |
| ☐ | ☐ | Ecosystems |
| ☐ | ☐ | Any other significant area |

### Experiments of concern

Does the work involve any of these experiments of concern:

| No | Yes | |
|---|---|---|
| ☐ | ☐ | Demonstrate how to render a vaccine ineffective |
| ☐ | ☐ | Confer resistance to therapeutically useful antibiotics or antiviral agents |
| ☐ | ☐ | Enhance the virulence of a pathogen or render a nonpathogen virulent |
| ☐ | ☐ | Increase transmissibility of a pathogen |
| ☐ | ☐ | Alter the host range of a pathogen |
| ☐ | ☐ | Enable evasion of diagnostic/detection modalities |
| ☐ | ☐ | Enable the weaponization of a biological agent or toxin |
| ☐ | ☐ | Any other potentially harmful combination of experiments and agents |

## ChIP-seq

### Data deposition

☐ Confirm that both raw and final processed data have been deposited in a public database such as GEO.

☐ Confirm that you have deposited or provided access to graph files (e.g. BED files) for the called peaks.

| | |
|---|---|
| Data access links<br>*May remain private before publication.* | *For "Initial submission" or "Revised version" documents, provide reviewer access links. For your "Final submission" document, provide a link to the deposited data.* |
| Files in database submission | *Provide a list of all files available in the database submission.* |
| Genome browser session<br>(e.g. UCSC) | *Provide a link to an anonymized genome browser session for "Initial submission" and "Revised version" documents only, to enable peer review. Write "no longer applicable" for "Final submission" documents.* |

### Methodology

| | |
|---|---|
| Replicates | *Describe the experimental replicates, specifying number, type and replicate agreement.* |
| Sequencing depth | *Describe the sequencing depth for each experiment, providing the total number of reads, uniquely mapped reads, length of reads and whether they were paired- or single-end.* |
| Antibodies | *Describe the antibodies used for the ChIP-seq experiments; as applicable, provide supplier name, catalog number, clone name, and lot number.* |
| Peak calling parameters | *Specify the command line program and parameters used for read mapping and peak calling, including the ChIP, control and index files used.* |
| Data quality | *Describe the methods used to ensure data quality in full detail, including how many peaks are at FDR 5% and above 5-fold enrichment.* |

Software | *Describe the software used to collect and analyze the ChIP-seq data. For custom code that has been deposited into a community repository, provide accession details.*

## Flow Cytometry

### Plots

Confirm that:

☐ The axis labels state the marker and fluorochrome used (e.g. CD4-FITC).

☐ The axis scales are clearly visible. Include numbers along axes only for bottom left plot of group (a 'group' is an analysis of identical markers).

☐ All plots are contour plots with outliers or pseudocolor plots.

☐ A numerical value for number of cells or percentage (with statistics) is provided.

### Methodology

Sample preparation | *Describe the sample preparation, detailing the biological source of the cells and any tissue processing steps used.*

Instrument | *Identify the instrument used for data collection, specifying make and model number.*

Software | *Describe the software used to collect and analyze the flow cytometry data. For custom code that has been deposited into a community repository, provide accession details.*

Cell population abundance | *Describe the abundance of the relevant cell populations within post-sort fractions, providing details on the purity of the samples and how it was determined.*

Gating strategy | *Describe the gating strategy used for all relevant experiments, specifying the preliminary FSC/SSC gates of the starting cell population, indicating where boundaries between "positive" and "negative" staining cell populations are defined.*

☐ Tick this box to confirm that a figure exemplifying the gating strategy is provided in the Supplementary Information.

## Magnetic resonance imaging

### Experimental design

Design type | *Indicate task or resting state; event-related or block design.*

Design specifications | *Specify the number of blocks, trials or experimental units per session and/or subject, and specify the length of each trial or block (if trials are blocked) and interval between trials.*

Behavioral performance measures | *State number and/or type of variables recorded (e.g. correct button press, response time) and what statistics were used to establish that the subjects were performing the task as expected (e.g. mean, range, and/or standard deviation across subjects).*

### Acquisition

Imaging type(s) | *Specify: functional, structural, diffusion, perfusion.*

Field strength | *Specify in Tesla*

Sequence & imaging parameters | *Specify the pulse sequence type (gradient echo, spin echo, etc.), imaging type (EPI, spiral, etc.), field of view, matrix size, slice thickness, orientation and TE/TR/flip angle.*

Area of acquisition | *State whether a whole brain scan was used OR define the area of acquisition, describing how the region was determined.*

Diffusion MRI ☐ Used ☐ Not used

### Preprocessing

Preprocessing software | *Provide detail on software version and revision number and on specific parameters (model/functions, brain extraction, segmentation, smoothing kernel size, etc.).*

Normalization | *If data were normalized/standardized, describe the approach(es): specify linear or non-linear and define image types used for transformation OR indicate that data were not normalized and explain rationale for lack of normalization.*

Normalization template | *Describe the template used for normalization/transformation, specifying subject space or group standardized space (e.g. original Talairach, MNI305, ICBM152) OR indicate that the data were not normalized.*

Noise and artifact removal | *Describe your procedure(s) for artifact and structured noise removal, specifying motion parameters, tissue signals and physiological signals (heart rate, respiration).*

Volume censoring | *Define your software and/or method and criteria for volume censoring, and state the extent of such censoring.*

## Statistical modeling & inference

Model type and settings | *Specify type (mass univariate, multivariate, RSA, predictive, etc.) and describe essential details of the model at the first and second levels (e.g. fixed, random or mixed effects; drift or auto-correlation).*

Effect(s) tested | *Define precise effect in terms of the task or stimulus conditions instead of psychological concepts and indicate whether ANOVA or factorial designs were used.*

Specify type of analysis: ☐ Whole brain ☐ ROI-based ☐ Both

Statistic type for inference (See Eklund et al. 2016) | *Specify voxel-wise or cluster-wise and report all relevant parameters for cluster-wise methods.*

Correction | *Describe the type of correction and how it is obtained for multiple comparisons (e.g. FWE, FDR, permutation or Monte Carlo).*

## Models & analysis

| n/a | Involved in the study |
|---|---|
| ☐ | ☐ Functional and/or effective connectivity |
| ☐ | ☐ Graph analysis |
| ☐ | ☐ Multivariate modeling or predictive analysis |

Functional and/or effective connectivity | *Report the measures of dependence used and the model details (e.g. Pearson correlation, partial correlation, mutual information).*

Graph analysis | *Report the dependent variable and connectivity measure, specifying weighted graph or binarized graph, subject- or group-level, and the global and/or node summaries used (e.g. clustering coefficient, efficiency, etc.).*

Multivariate modeling and predictive analysis | *Specify independent variables, features extraction and dimension reduction, model, training and evaluation metrics.*