nature human behaviour

ARTICLES

https://doi.org/10.1038/s41562-021-01084-x

Check for updates

Fresh teams are associated with original and multidisciplinary research

An Zeng[1], Ying Fan[1], Zengru Di[1], Yougui Wang[1] and Shlomo Havlin[2] ✉

Teamwork is one of the most prominent features in modern science. It is now well understood that team size is an important factor that affects the creativity of the team. However, the crucial question of how the character of research studies is related to the freshness of a team remains unclear. Here, we quantify the team freshness according to the absence of prior collaboration among team members. Our results suggest that papers produced by fresher teams are associated with greater originality and a greater multidisciplinary impact. These effects are even stronger in larger teams. Furthermore, we find that freshness defined by new team members in a paper is a more effective indicator of research originality and multidisciplinarity compared with freshness defined by new collaboration relationships among team members. Finally, we show that the career freshness of team members is also positively correlated with the originality and multidisciplinarity of produced papers.

In contrast to about a century ago, when individual scientists had an important role in scientific discoveries, teamwork is becoming increasingly common in recent modern science[1,2]. Indeed, it has been found that the fraction of scientific papers that were written by teams and the mean size of teams increased during the last century, indicating a notable shift in favour of teamwork[3–5]. Specifically, it has been shown that the mean team size of research papers increased from 1.9 to 3.5 authors per paper from 1955 to 2000 (ref. [3]). Moreover, the team size distribution has been found to change fundamentally from a simple Poisson distribution to a power-law-shape distribution[6]. These phenomena are attributed to the combination effect of the increasing scale, complexity and costs of big science[7–9].

Various models have been developed to better understand team formation in scientific research. A team is defined as the coauthors of a paper, and many classical studies focused on studying the collaboration features of individual scientists to understand team formation[10]. Related studies are numerous, mainly aiming to reveal the topological features, such as community structure and assortative mixing in collaboration networks[11,12], and model the evolution of collaboration networks and author–paper bipartite networks[13,14]. In recent years, attention has shifted to directly understanding the team-assembly mechanisms. For example, a recent study found that research teams include both small stable 'core' teams and large dynamically changing 'extended' teams[6]. The shift of team size distribution from Poisson to power law has been explained by the fast tendency towards extended teams[6]. Another study investigated how the mechanisms by which creative teams self-assemble determine the structure of collaboration networks, and observed a second-order phase transition of the giant component in the collaboration networks[4].

In many studies, the citations of papers have been used to measure the impact of the paper[15]. By comparing papers of multiple authors to papers that have a single author, a strong signal favouring teamwork has been detected[3]. The distribution of workload across team members was shown to largely affect the performance of teams[16]. It was also found that a greater number of authors and countries in a paper is associated with higher citation rates when examining the influence of international research teams on citation outcomes[17]. In a recent study, the authors used an index called disruption to measure the originality of a paper[18]. Interestingly, they found that small teams tend to disrupt science and technology with original ideas and opportunities, whereas larger teams tend to develop existing ones. This finding highlights the vital role that small teams have in expanding the frontiers of knowledge.

To date, various factors—such as team size[3,18], workload distribution[16], number of involved countries[17,19], universities[20,21] and disciplines[22,23]—have been found to substantially affect the outcome impact of teamwork. However, the role of team freshness in advancing science has rarely been studied. A research team may consist of some researchers who have not previously worked together, resulting in some freshness of the team. By contrast, the authors of a paper may have already published a number of papers together, therefore working as an old team. In team-formation models, the tendency of individual teams to select new team members has been found to decrease the giant component of the whole collaboration network and this tendency varies in papers of different journals[4]. Furthermore, less seniority of team members in their careers can be also regarded as a kind of freshness of a team. Using the data of Nobel-prize-winning papers, it was found that the performance of scientists typically peaks in their middle age and the average age of the peak has significantly increased during the past century[24–26]. A recent study pointed out that the most influential work is distributed randomly within a scientist's sequence of publications[27]. Note that most of these studies focused on the career of individual scientists. However, how the team freshness and career freshness of team members are related to the performance of teams in advancing science remains unclear.

In this Article, we address the association between team freshness and the originality and multidisciplinarity of the produced work by systematically investigating the prior collaboration relationships between team members. The freshness of a team is defined according to the fraction of team members that have not collaborated previously with other team members (Fig. 1). We found that papers produced by fresher teams have significantly higher originality and more multidisciplinary impact compared with papers produced by older teams. We found that the effect is even more prominent in larger teams. Our results suggest that freshness defined by new team

[1]School of Systems Science, Beijing Normal University, Beijing, China. [2]Department of Physics, Bar-Ilan University, Ramat-Gan, Israel.
✉e-mail: havlin@ophir.ph.biu.ac.il

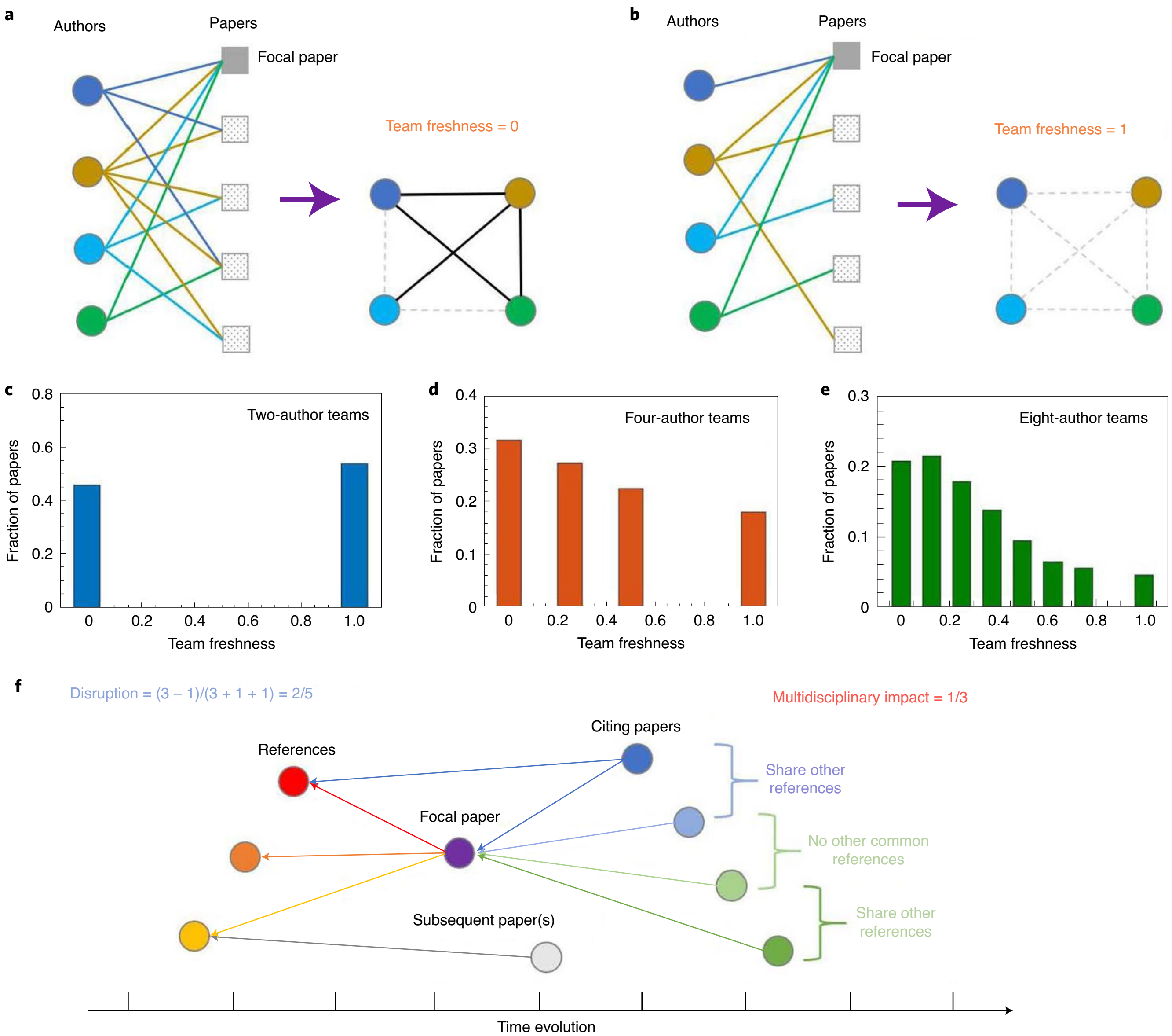

Fig. 1 | Illustration of the freshness of teams, disruption of papers and multidisciplinary impact of papers. a,b, Example of freshness = 0 (**a**) and freshness = 1 (**b**). The four authors (circles) in the toy bipartite network are the authors of the focal paper (filled square). The other papers (dotted squares) are the papers published by the four authors before the focal paper. The collaboration network of these authors before the focal paper can be constructed. The solid and dashed lines represent the existing and missing links, respectively. The fraction of nodes with zero collaboration links in the prior network is defined as the freshness of the team in the focal paper. Accordingly, the team of the focal paper in **a** has a freshness of 0 and the team of the focal paper in **b** has a freshness of 1. In our analysis, we studied all 482,566 papers published during the years 1893–2010 by the APS. **c–e**, The freshness distribution for 130,470 two-author papers (**c**), 51,391 four-author papers (**d**) and 6,965 eight-author papers (**e**). One immediate observation was that completely fresh teams are less common in larger teams. **f**, Demonstration of calculating the disruption[18,28] and multidisciplinary impact in a citation network. The citation network consists of a focal paper, its references (outgoing links) and its citing papers (incoming links). The disruption aims to measure the originality of a paper. To calculate the disruption of the focal paper, one should first calculate the difference between the number of its citing papers that do not cite its references and the number of its citing papers that cite its references. The disruption is obtained by dividing this difference by the number of all of the citing papers plus the number of subsequent papers of the focal paper that do not cite it but do cite its references. In the example, disruption of the focal paper is $(3-1)/(3+1+1)=2/5$. The disruption varies between −1 and 1; a larger disruption corresponds to a higher originality. The multidisciplinary impact aims to measure the diversity of the areas that a paper influences. Here we define it as the fraction of temporal adjacent citing papers that share no references apart from the focal paper. In the example, the focal paper has four citing papers, resulting in three adjacent pairs in time. Among these three pairs, one pair shares no other common references apart from the focal paper, and two pairs share other references apart from the focal paper. Thus, the multidisciplinary impact of the focal paper is 1/3. This index varies between 0 and 1, corresponding to a narrow and a diverse impact in disciplines, respectively.

members is more correlated with the originality and impact diversity of the resultant papers than freshness defined by new collaboration relationships among team members. Finally, we also studied the career freshness of team members and found that younger teams are associated with higher originality and a higher impact diversity of the produced studies.

Results

We started by defining the freshness of a team of a paper as the fraction of team members who have not collaborated with any of other team members before they coauthored this paper. According to this definition, the freshness varies between 0 and 1, corresponding to fully old teams and fully fresh teams, respectively. This definition can be easily calculated by constructing a collaboration network that represents all prior collaboration relationships among the team members of the considered paper. The freshness of a team can be obtained directly by computing the fraction of nodes with zero degree in this collaboration network. The definition is illustrated in Fig. 1, with two networks describing examples of freshness 0 and 1, respectively.

Here we analysed the scientific publication data of the American Physical Society (APS) journals, containing 482,566 papers, ranging from the year 1893 to 2010. Evaluating the freshness of a team requires the knowledge of the prior papers that each team member published. We therefore needed to assign each paper in the dataset correctly to its real authors. For this, we used the disambiguated author name data provided in ref. [27] to assign each paper to its authors, which resulted in 236,884 distinct scientists and 482,566 papers. Furthermore, we also examined three additional datasets from computer science, chemistry and multidisciplinary research (a description of the data is provided in the Methods). As their results are similar to those on the APS data, the analyses in the main text are based on the APS data, while the results of the other three datasets are presented in Supplementary Figs. 19–24. In Figs. 1c–e, we show the distribution of team freshness for two-author papers, four-author papers and eight-author papers, respectively. The results are consistent with the intuition that fresh teams are more common in small teams than in large teams. The teams with a freshness of 1 make up 54% of all two-author teams, while the fraction with a freshness of 1 is only 4.6% for eight-author teams.

To evaluate the role of a paper in advancing science, we considered a recently developed index[18,28], disruption (D), to measure the originality of a paper, and propose here a measure for multidisciplinary impact (M) to evaluate the diversity of disciplines that a paper influences (see Fig. 1 for an illustration and the Methods for more details on both measures). D varies between −1 and 1. A larger disruption of a paper reflects that more of the paper's citing papers cite it but none of its references, corresponding to higher originality. The multidisciplinary impact, M, of a paper is defined as the fraction of its temporal adjacent citing papers that share no other references apart from the focal paper. The multidisciplinary impact varies between 0 and 1, corresponding to narrow and diverse impacts in different disciplines, respectively.

The first question we examined here is whether and how team freshness is correlated with the originality and impact diversity of the produced work. To answer this, we show D and M of papers as a function of different team freshness (Fig. 2). Figure 2 contains results for two-author papers, four-author papers and eight-author papers. The results for all cases exhibit a consistent increasing trend of both D and M with increasing team freshness (the two-tailed Pearson correlation tests between these two metrics and team freshness, with a summary of the coefficients, P values and 95% confidence intervals, are provided in Supplementary Table 1). To examine the significance of the trend, we compared the distributions of the bootstrap disruption and bootstrap multidisciplinarity of papers with a team freshness of 0 and 1 (Methods). The results are presented in the inset of Fig. 2. A remarkable difference in the distributions can be observed between papers with a team freshness of 0 and 1. To further support the significance of the trend, we tested differences in the distribution of disruption between all different team freshness using two-tailed Kolmogorov–Smirnov tests. The results suggest that, even with small difference in team freshness, the difference in originality (disruption) is significant (Supplementary Fig. 1a–c); most P values were lower than 0.01. Similar results were obtained for the multidisciplinary impact trend using the two-tailed Kolmogorov–Smirnov test (Supplementary Fig. 1d–f).

One possible concern regarding the observed trend in Fig. 2 is whether the disruption and multidisciplinary impact detects the same property of a paper, such that the increasing trend of one index with team freshness is highly related to the increase of the other index. To test this, we first studied the relationship between these two indexes in Supplementary Fig. 2. We found that (1) both indexes, D and M, are only very weakly correlated with citations; and (2) the Pearson correlation coefficient between disruption and multidisciplinary impact is 0.320 ± 0.003 ($P < 0.001$, two-tailed Pearson correlation test), indicating some correlation. We next examined how independent both indexes are from each other. To this end, we studied the relationship between disruption and team freshness when controlling for the multidisciplinary impact. We analysed papers with multidisciplinary impacts of $M \approx 0.3$, $M \approx 0.5$ and $M \approx 0.7$ and found that, even for fixed M, the disruption of these papers still increases with team freshness (Supplementary Fig. 3). Similarly, we fixed disruption to $D \approx -0.1$, $D \approx 0$ and $D \approx 0.1$ and found that, even for fixed D, the multidisciplinary impact still increases with team freshness. These results suggest that the disruption and multidisciplinary impacts truly represent quite distinct properties.

The team members in fresh teams, according to our definition, do not have prior collaboration with any other in the team. The team freshness might actually be related to some degree to the prior productivity and career age of team members. If team members have fewer previous papers, the formed team is more likely to be a fresh team. It is therefore important to test whether the observed trend with freshness in Fig. 2 can simply be explained by the prior productivity and career age of team members. To remove this effect, we studied the dependence of disruption and multidisciplinary impact on freshness when controlling for the team member productivity and career age, respectively (Supplementary Figs. 4 and 5). The results suggest that the increasing trend of disruption and multidisciplinary impact with team freshness is preserved, indicating that the effect of team freshness is not simply caused by team member productivity or career age. As the APS data ranges from year 1893 to 2010, we also examined the mean disruption (originality) and mean multidisciplinary impact of papers in different years. We show that both indexes decrease with time, yet fresh teams constantly have higher originality and multidisciplinarity compared with old teams (Supplementary Fig. 6).

We performed several additional analyses to further support our findings. We first controlled for the citation of papers by repeating our analyses in papers with similar number of citations c. We compared two groups of papers: high-cited papers ($c \geq 30$) and low-cited papers ($3 \leq c \leq 5$) (Supplementary Fig. 7). We found that team freshness is positively correlated with disruption/multidisciplinarity in both high-cited and low-cited cited papers. Another concern is that the team freshness might be inflated by the one-shot authors (that is, authors who have just a single paper in the dataset). Although there are around 43% authors in APS who have a single paper, interestingly, we found that these one-shot authors represent only 15.7% papers in APS, as multiple one-shot authors often cluster in the same paper. Nevertheless, we further supported our findings by performing the same analyses in papers without any one-shot authors (Supplementary Fig. 8).

Next, we examined whether the effect of team freshness can simply be explained by the prior scientific distance between team members. We quantify the distance d_{ij} between scientist i and scientist j in the scientific space by their dissimilarity in research interests. For each scientist i, we constructed a set Γ_i, recording all of the references in his/her papers, representing the research literature he/she is interested in. The distance d_{ij} between scientist

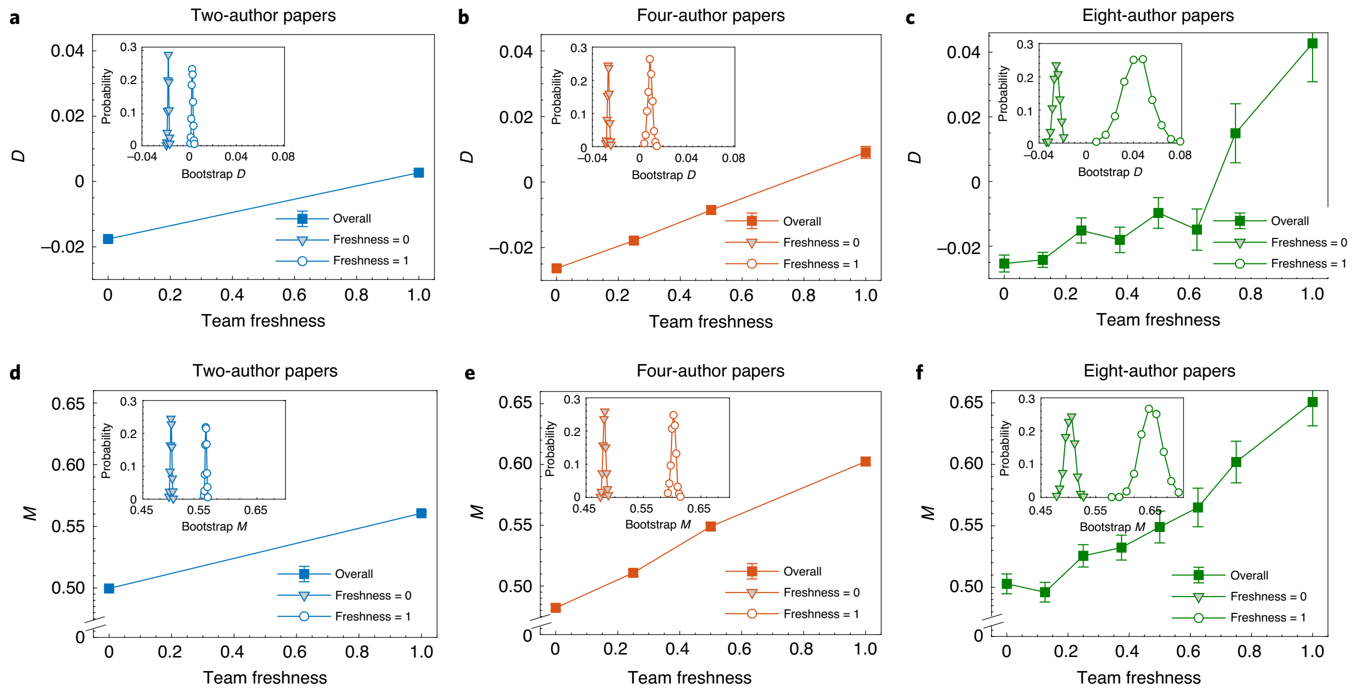

Fig. 2 | Fresh teams create more original and multidisciplinary research. a–f, The dependence of the disruption (originality) D (**a–c**) and multidisciplinarity M (**d–f**) of papers on the team freshness for 130,470 two-author papers (**a**,**d**), 51,391 four-author papers (**b**,**e**) and 6,965 eight-author papers (**c**,**f**). Data are mean ± s.e.m. Some error bars are not visible because they are smaller than the size of the symbols. The results suggest that both originality and multidisciplinarity significantly increase with team freshness. Insets: the distributions of 1,000 realizations of bootstrap disruption or bootstrap multidisciplinarity. A remarkable difference (that is, high significance) can be observed between the distributions of D of papers with a team freshness of 0 and 1. Two-tailed Kolmogorov–Smirnov tests of the distribution difference in D or M between papers with a team freshness of 0 and 1 all yeild $P < 0.001$ (other freshness values are provided in Supplementary Fig. 1). To further support the observed increasing trend, we directly performed two-tailed Pearson correlation tests between D or M and team freshness (coefficients, P values and 95% confidence intervals are provided in Supplementary Table 1). We obtained significant positive correlation coefficients; all P values from the Pearson correlation tests were lower than 0.001.

i and scientist j can therefore be calculated as their Jaccard dissimilarity $d_{ij} = 1 - |\Gamma_i \cap \Gamma_j| / |\Gamma_i \cup \Gamma_j|$ where $|.|$ is the size of the set. For each paper, we calculated the mean distance d_{ij} between team members using the data before they coauthored this paper. We found that the mean distance d_{ij} is indeed positively correlated with team freshness (Pearson correlation coefficients with 95% confidence intervals, 0.429 ± 0.005 (two-author papers), 0.401 ± 0.010 (four-author papers), 0.380 ± 0.033 (eight-author papers); all P values in the two-tailed Pearson correlation tests are lower than 0.001; Supplementary Fig. 9). We next analysed whether team freshness captures additional information beyond scientific distance between team members. To this end, we controlled for the mean distance between team members and again studied the relationship between team freshness and disruption/multidisciplinarity (Supplementary Fig. 10). Although less often, the scientists who are close in scientific space form fresh teams. Our results suggest that the papers published by those fresh teams are associated with greater disruption and multidisciplinarity (Supplementary Fig. 10).

We further support our findings by testing several variations of evaluating team freshness. First, we considered a generalization of the definition of team freshness by taking into account the prior strength of collaboration ties. The freshness of a team of a paper is defined as the fraction of team members who have collaborated fewer than m papers with any of other team members before they coauthored this paper. When $m = 1$, it returns to the original definition of team freshness. The case $m > 1$ corresponds to a looser definition of team freshness, whereby scientists who coauthored fewer than m papers with each other are still regarded as a fresh team the next time they work together. We found that $m > 1$ does not change the positive correlation between team freshness and disruption/multidisciplinarity (Supplementary Fig. 11).

Although the APS dataset records a large part of the publications of individual physicists and their collaboration relationships[29], some of physicists' collaborations are published outside the APS journals. We therefore examined the possibility of capturing collaborations outside the APS dataset in a statistical sense using link prediction algorithms. To this end, we used the APS dataset as the training set and used the Resource Allocation link prediction algorithm[30] to predict and add missing collaboration relationships to the data. After adding 10% or 50% predicted links to the collaboration networks, we recalculated the team freshness of each paper. In these cases, a team is fully fresh only if the team members have no prior collaboration in the APS data and no added missing links connecting them. We found that adding the predicted links did not change the increasing trend of disruption and multidisciplinarity with team freshness (Supplementary Fig. 12).

Team size has been found to be an important factor in affecting the disruption of a paper, that is, the disruption (originality) decreases with team size[18]. It is therefore natural to examine how team freshness is related to disruption and multidisciplinarity in teams of different sizes. To this end, we analysed the mean disruption D and multidisciplinary impact M as a function of the team size of papers published by old teams (freshness = 0) and fresh teams (freshness = 1) (Fig. 3). Indeed, the overall disruption D as well as old teams D tend to decrease with team size, supporting the finding of Wu et al[18]. However, interestingly, we found that the disruption (originality) D of the papers published by fresh teams tends to increase with team size. The significance of this increasing trend is

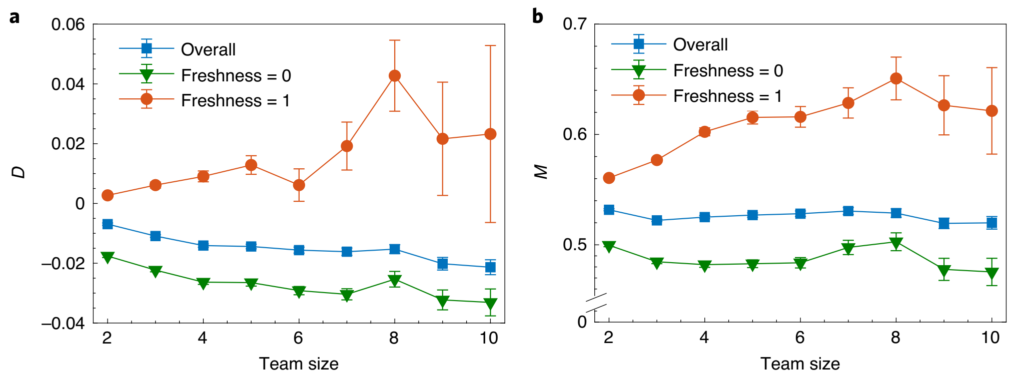

Fig. 3 | The difference between fresh and old teams is amplified in larger teams. a, The mean disruption D of papers of different team sizes (overall), showing a decreasing trend, as reported previously[18]. There are 342,809 papers with team sizes of 2–10 people. For each team size, we also studied the mean disruption D of papers published by old teams (freshness = 0) and fresh teams (freshness = 1). Unexpectedly, in contrast to the overall papers, for papers with freshness = 1, D increases with team size. This suggests that the overall decreasing disruption is due to the dominant non-fresh teams. **b**, The mean multidisciplinary impact M of papers of different team sizes (overall). For each team size, we also studied the mean multidisciplinary impact M of papers published by old teams (freshness = 0) and fresh teams (freshness=1). Similarly to the results for D, the difference in M between fresh and old teams is amplified in larger teams. Data are mean ± s.e.m. Some error bars are not visible because they are smaller than the size of the symbols. The P values of the two-tailed Kolmogorov–Smirnov test of the difference in D or M distribution between papers of different team sizes are provided in Supplementary Fig. 13. We also performed two-tailed Pearson correlation tests between D or M and team size for fresh and old teams (the coefficients, P values and 95% confidence intervals are provided in Supplementary Table 2). The correlation coefficients are consistent with the observed trends; the significance of the correlations is indicated by small P values, which were determined using the Pearson correlation test (most $P < 0.001$).

supported by two-tailed Kolmogorov–Smirnov tests of the disruption distribution of different team sizes shown in Supplementary Fig. 13. We also performed the two-tailed Pearson correlation test between D or M and team size for fresh and old teams (the coefficients, P values and the 95% confidence intervals are provided in Supplementary Table 2). These results suggest that papers published by large fresh teams are associated with higher disruption compared with those published by small fresh teams. Similar increasing trends were observed when we examined the relationship between multidisciplinary impact and the team size of fresh teams. Comparing the difference between fresh teams and old teams in D and M, we also found that the advantage of fresh teams is more prominent in larger teams.

In the above analysis, we defined the team freshness as the fraction of new team members in a paper. It was evaluated by calculating the fraction of nodes with no link to others in the collaboration network that represents the prior collaboration relationships of the team members (Fig. 1a,b). However, an alternative way to define the freshness of a team is to measure the number of new collaboration relationships (new links) created by the team. This can be regarded as a link freshness, which could be of interest. This link freshness can be easily calculated by the fraction of missing links in the collaboration network that represents the prior collaboration relationships of the team members (for example, 2/6 = 1/3 dashed links in Fig. 1a). To distinguish between these two types of freshness, we refer to them as node freshness f_n (new collaborators) and link freshness f_l (new collaborations) according to their calculations in the collaboration networks. An interesting question here is which types of freshness (node or link) are more-strongly correlated with the originality and impact diversity of the produced papers. To test and answer this question, we show a scatter plot of link freshness versus node freshness, with circle size and colour representing the mean disruption of the corresponding papers (Fig. 4a). Given a certain node freshness, higher link freshness is very little or even not associated with a higher disruption. This observation can be quantitatively supported using the two-tailed Pearson correlation test between link freshness and disruption for each given node freshness (the coefficients, P values and 95% confidence intervals are provided in Supplementary Table 3). One can also see in the insets of Fig. 4a that the Pearson correlations between link freshness and disruption are very weak and even, in some cases, negative—that is, at the level of noise. The results of other team sizes are provided in Supplementary Fig. 14.

To better estimate the role of link freshness, we designed a combined freshness measure f_m as a weighted linear combination of node freshness and link freshness, with a tunable parameter controlling the relative weights of the two types of freshness (Methods). We next computed the Pearson correlation between the combined freshness f_m and the disruption D. By tuning the relative weights of the two type of freshness, we found that the maximum correlation achieved with f_m (Fig. 4b) is not significantly higher than the correlation between f_n and D (further support is provided in Supplementary Fig. 15). These results suggest that incorporating link freshness does not provide notable additional information for predicting originality. In Fig. 4c,d, we performed a similar analysis for the multidisciplinary impact and found similar results.

We next considered another type of freshness of teams that we call here career freshness of team members. The career freshness of a team member can be measured by his/her career age, namely the number of years since he/she published their first paper. A shorter career age indicates a fresher scientist. The basic statistics of the mean career age of team members is shown (Supplementary Fig. 16). We further examined whether the career freshness of team members is related to the originality and impact diversity of their produced papers. In Fig. 5, we show the dependence of the mean disruption (originality) D and multidisciplinary impact M on the mean career age of the team members of a paper. Surprisingly, we observed a decreasing trend in both cases (for the two-tailed Pearson correlation test between these two metrics and mean career age, a summary of the coefficients, P values and 95% confidence intervals is provided in Supplementary Table 4). The decreasing trend is still present when we fix the team freshness (as low freshness 0 and high freshness 1 in Fig. 5). Note that a similar trend was observed when we used the mean productivity of team members to define the freshness of their careers (Supplementary Fig. 17). These results suggest that research produced by early-career team members, that

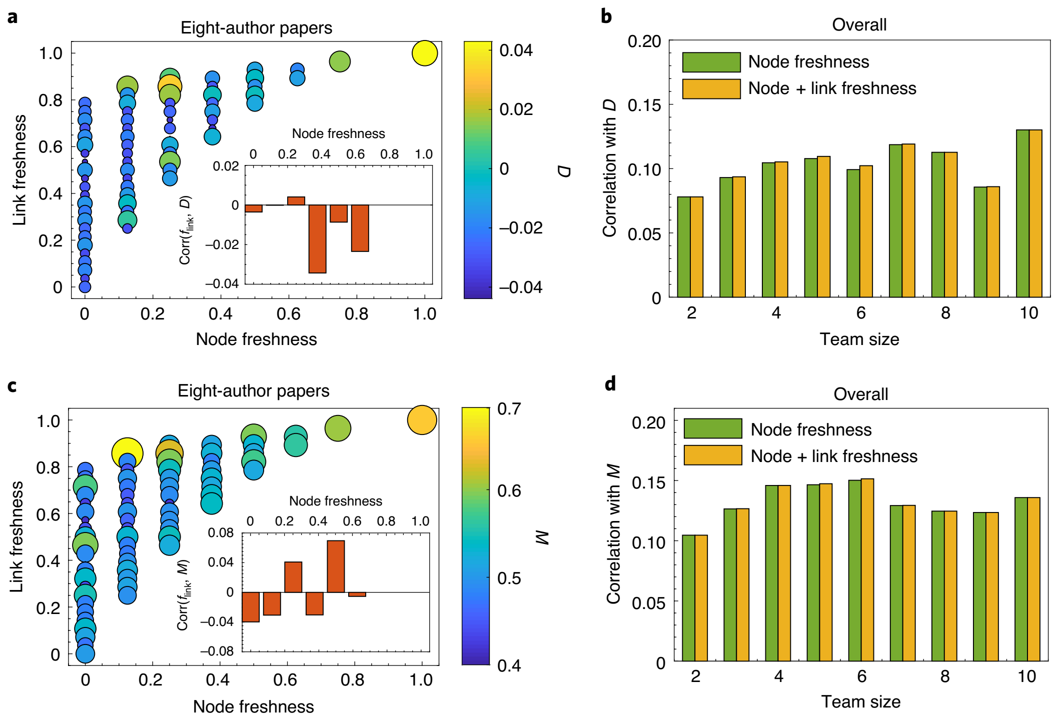

Fig. 4 | Team freshness defined by new team members and new collaboration relationships. We defined the node freshness of a paper as the fraction of new team members and defined the link freshness as the fraction of new collaboration relationships in the team. **a**, Link freshness versus node freshness for 6,965 eight-author papers. The circle size and the colour represent the mean originality (disruption) of the corresponding papers (similar results for other team sizes are provided in Supplementary Fig. 14). Given a certain node freshness of a paper, it is shown that higher link freshness is not associated with a higher disruption. This finding was supported by directly performing the two-tailed Pearson correlation test between link freshness and disruption for each node freshness. Inset: the coefficient of the two-tailed Pearson correlation (Corr) test between link freshness (f_{link}) and disruption (D) for each node freshness. The coefficient values, P values and 95% confidence intervals are provided in Supplementary Table 3. The large P values ($P > 0.1$) in the Pearson correlation tests indicate the lack of significance of the correlation between link freshness and disruption given a node freshness. **b**, The Pearson correlation of node freshness and originality (disruption) for papers of different team sizes (342,809 papers with team sizes of 2–10). For comparison, we calculated the maximum Pearson correlation when we consider team freshness as a weighted linear combination of node and link freshness. The results suggest that incorporating link freshness does not bring notable additional information for predicting disruption. **c**, Link freshness versus node freshness for eight-author papers. The circle size and colour represent the mean multidisciplinary impact (similar results for other team sizes are provided in Supplementary Fig. 14). Inset: the coefficient of the two-tailed Pearson correlation (Corr) test between link freshness (f_{link}) and multidisciplinarity (M) for each node freshness. The coefficients, P values and 95% confidence intervals are provided in Supplementary Table 3. **d**, The Pearson correlation of node freshness and multidisciplinary impact for papers of different team sizes. We also show the maximum Pearson correlation of the weighted linear combination of both node and link freshness.

is, higher career freshness, are associated with a more original and diverse impact. There could be many reasons for this surprising phenomenon. One possible explanation for this is that researchers in earlier stages of their career are less likely to be trapped by concepts and general beliefs that are common in the scientific field, resulting in higher originality in their work.

In the literature, the tendency of individual teams to select new team members has been found to be related to the impact factor of the journals of the published studies[4]. Thus, the question we ask here is how the team freshness is related to the citations impact of papers. To this end, we analysed how the team freshness is related to the number of citations that a paper will receive (Supplementary Fig. 18). To be able to compare papers from different years, we calculated the citations received by a paper within 10 years of publication (c_{10})[27]. We show that papers produced by fresh teams tend to have a smaller c_{10} compared with old teams (Supplementary Fig. 18), which is consistent with previous findings[4]. It has been shown previously[31] that the impact of a paper (number of citations) is positively correlated with the cumulated reputation of the authors (measured by their productivity). To remove this effect, we considered only papers published by teams with similar team member productivity. After controlling for this factor, we found that papers with a different team freshness do not exhibit a significant difference in c_{10}. Thus, our results suggest that the difference in the number of citations received by fresh teams and old teams can be attributed to the productivity of team members instead of team freshness. Note that we have also shown that the increasing trend of disruption and multidisciplinarity with team freshness is independent of team members' productivity (Supplementary Fig. 6).

Discussion

Despite intensive efforts in understanding team formation mechanism and the effect of team size on creativity, little is known about how the prior relationships between team members are related to the originality and impact diversity of the produced papers. Here, we define the freshness of a team according to the fraction of team members without prior collaboration with other team members. We found that papers of fresher teams are associated with higher originality. Furthermore, the impact of the papers produced by fresher teams was found to be more diverse, influencing multiple research areas. These two effects were found to be more prominent in larger teams. We also found that new team members have a stronger

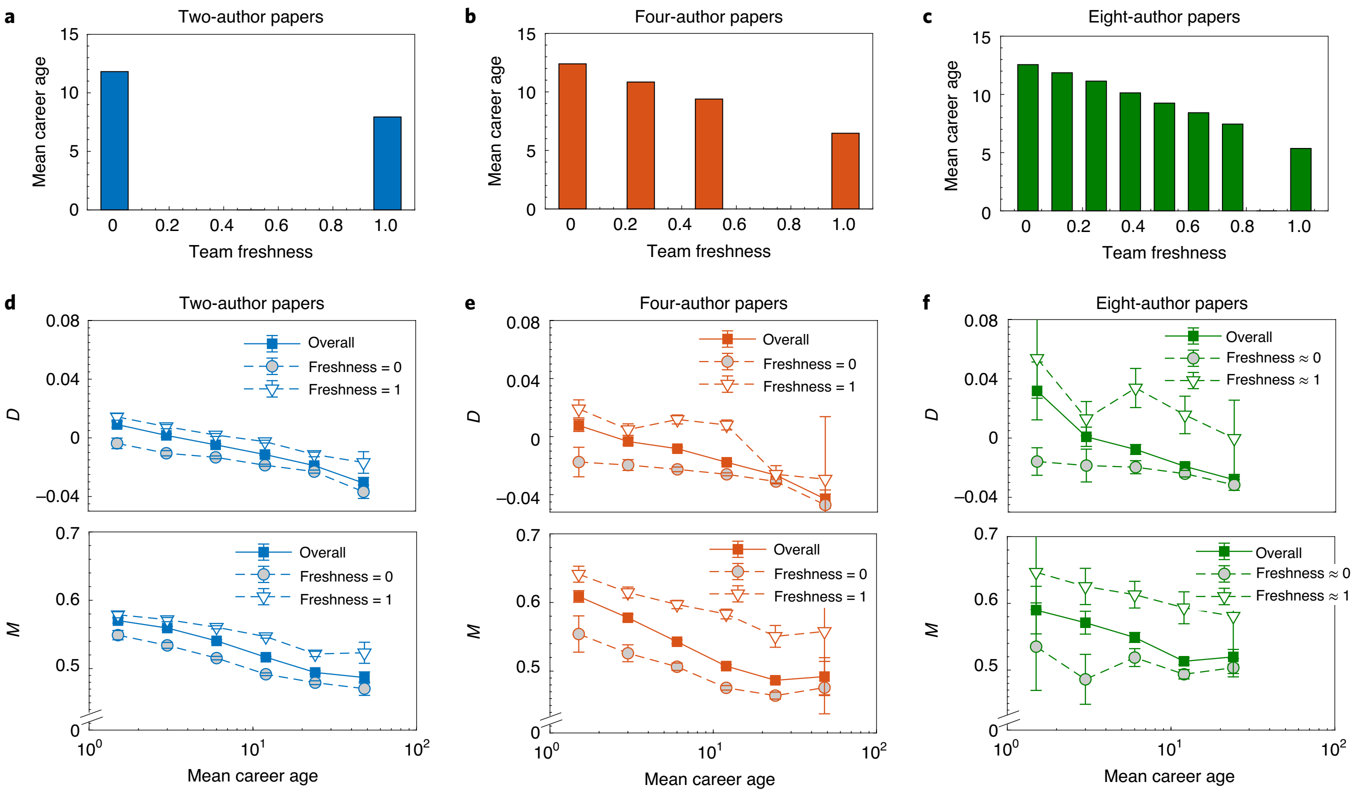

Fig. 5 | Freshness of team member's careers. a–c, The mean career age of team members in team sizes of two (**a**), four (**b**) and eight (**c**) members. The career age of a team member is defined as the number of years after he/she publishes the first paper. The results suggest that scientists in fresh teams tend to have a lower career age than those in old teams. **d–f**, The dependence of the mean disruption (originality) *D* (top) and multidisciplinarity *M* (bottom) on team members' mean career age in 130,470 two-author papers (**d**), 51,391 four-author papers (**e**) and 6,965 eight-author papers (**f**). To remove the effect of team freshness, we also show that the curves for old teams (freshness = 0) and fresh teams (freshness = 1) behave similarly, decreasing with mean career age. For better statistics in eight-author papers, we took freshness ≤ 0.25 as freshness ≈ 0 and freshness ≥ 0.85 as freshness ≈ 1. The results suggest that papers published by younger teams are associated with higher originality and multidisciplinarity. Data are mean ± s.e.m. Some error bars are not visible because they are smaller than the size of the symbols. To further support the observed decreasing trend, we directly performed two-tailed Pearson correlation tests between *D* or *M* and mean career age and obtained negative correlation coefficients (the coefficients, *P* values and 95% confidence intervals are provided in Supplementary Table 4). The significance of the negative correlation is indicated by the small *P* values, which were determined using Pearson correlation tests (most $P < 0.001$).

correlation than new collaboration relationships with the originality and impact diversity of the resultant studies. Although the researchers in fresh teams have a substantially smaller number of published papers compared with old teams, the productivity of team members is shown to not be a relevant factor that affects the increasing trend of originality and impact diversity with team freshness. Finally, we found that researchers in fresher careers have higher original and multidisciplinary papers. This could be because they are less likely to be trapped in conventional concepts and beliefs in the field, and tend to produce more original and multidisciplinary research.

Our research supports that originality decreases with team size, as reported previously[18], and reveals a possible origin of this discovery. The decreasing trend of originality with team size could be explained by lower freshness in larger teams (Fig. 1c–e). Indeed, for fully fresh teams, both the originality and multidisciplinary impact increased significantly with team size (Fig. 3). Furthermore, we note that our research has several limitations. The question of what comes first (that is, what causes what), the fresh team or the novel and original ideas, is not answered in our research. Our results mainly discovered correlations between freshness and originality, and between freshness and multidisciplinary impact, but not causality. For example, it is possible that one (or more) author(s) had a novel idea or a novel problem and created a suitable new team to study it. Another limitation is that our research defines freshness solely on the basis of the prior coauthorship relationships between team members. There are actually many other types of relationships between scientists, such as online and offline social connections, as well as coparticipation in conferences, research projects and patent invention. Thus, our research reveals only a specific dimension of team freshness.

One of the main findings in this paper is that the papers of fresh teams are associated with higher originality and a more-diverse impact. On the basis of the current work, several promising research extensions can be performed. A straightforward one is to investigate the performance of fresh teams in other activities, such as research funding applications, software development and patent invention. Another interesting research direction would be to study the mechanisms that drive the formation of fresh teams. Finally, we note that scientific collaboration is a complex phenomenon, the outcomes of which are driven by multiple factors. Apart from the team freshness, the freshness of the topic that the team studies is also a critical factor in determining the quality of produced papers[32,33]. Therefore, identifying the inter-relationships between team freshness and topic freshness would be an interesting topic for future study.

Methods

Data. Here we analysed the publication data of all journals of the APS. The data contain 482,566 papers, ranging from the year 1893 to 2010. For the sake of author name disambiguation, we used the author name dataset provided by Sinatra et al.[27], which was obtained using a comprehensive disambiguation process in the APS data. Eventually, a total number of 236,884 distinct authors were matched. Another set of data that we analysed (Supplementary Information) is the computer science data obtained by extracting the profiles of scientists from online web databases[34]. The data contain 1,712,433 authors and 2,092,356 papers ranging from 1948 to 2014. The author names in these data were already disambiguated. The third dataset we analysed is the publication data of the *Journal of the American Chemical Society*. The data contain 59,913 papers, ranging from 1997 to 2017. We performed the same name disambiguation process as described in ref. [27] and obtained 162,016 distinct authors. The fourth dataset contains all papers in five representative multidisciplinary journals: *Nature*, *Science*, *Proceedings of the National Academy of Sciences*, *Nature Communications* and *Science Advances*. The dataset consists of 633,808 papers and 1,077,399 authors, ranging from the year 1869 to 2020. The author names in these data are already disambiguated. The data were downloaded freely from Microsoft Academic Graph[35].

Disruption. The disruption index was originally designed to identify destabilization and consolidation in patented inventions[28]. In a recent article, it was extended to measure the originality of scientific papers[18]. The disruption D varies between −1 and 1. $D=1$ for a paper indicates that all of the paper's citing papers cite it but not any of its references. In this case, the paper is considered to disrupt science with new ideas and opportunities, corresponding to higher originality. If a paper has $D=-1$, all of its citing papers not only cite it but also cite at least one of its references. In this case, the paper is devoted to further developing existing findings and ideas. The calculation of the disruption is illustrated in Fig. 1.

Multidisciplinary impact. Here, we proposed a simple index, called the multidisciplinary impact M, to measure the diversity of the disciplines that a paper influences. Different from the various existing indexes that rely on external information, such as disciplinary categories[36,37], our method is based solely on the citation relationships. We define the multidisciplinary impact of a paper as the probability of two successive citing papers from different disciplines. It can be easily obtained by calculating the fraction of temporal adjacent citing papers that share no references apart from the focal paper. The multidisciplinary impact M varies between 0 and 1, corresponding to narrow and diverse impact in disciplines, respectively. The calculation of the multidisciplinary impact is illustrated in Fig. 1. Similar to the disruption index, the multidisciplinary impact of a paper is only very weakly correlated with its citations (Supplementary Fig. 2).

Bootstrap disruption and bootstrap multidisciplinary impact. We compared the distributions of bootstrap disruption of papers with a team freshness of 0 and 1 (Fig. 2, insets). The bootstrap disruption was obtained by random sampling of papers' disruption such that each paper's disruption has an equal chance to be selected and can be selected over and over again. The distributions were obtained by performing 1,000 realizations of bootstrap disruption. The bootstrap multidisciplinary impact in the insets of Fig. 2 was obtained similarly.

Combined freshness measure. The node freshness of a team f_n is defined as the fraction of nodes with no link to others in the collaboration network that represents the prior collaboration relationships of the team members. The link freshness of a team f_l is defined as the fraction of missing links in the collaboration network representing the prior collaboration relationships of the team members. Denoting the combined freshness measure as f_m, it is computed as $f_m=\lambda f_n+(1-\lambda)f_l$, where λ is a tunable parameter between 0 and 1. In Fig. 4b,d, we show the maximal Pearson correlations that can be achieved by adjusting λ.

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

Data availability

The APS data can be downloaded at https://journals.aps.org/datasets. The computer science data can be downloaded at https://www.aminer.cn/aminernetwork. The multidisciplinary data were download from https://docs.microsoft.com/en-us/academic-services/graph. Other related, relevant data are available from the corresponding author on reasonable request.

Code availability

Computational codes for data processing and analysis are available from the corresponding author on request.

Received: 12 July 2020; Accepted: 26 February 2021;
Published online: 05 April 2021

References

1. Fortunato, S. et al. Science of science. *Science* **359**, eaao0185 (2018).
2. Zeng, A. et al. The science of science: from the perspective of complex systems. *Phys. Rep.* **714-715**, 1–73 (2017).
3. Wuchty, S., Jones, B. F. & Uzzi, B. The increasing dominance of teams in production of knowledge. *Science* **316**, 1036–1039 (2007).
4. Guimera, R., Uzzi, B., Spiro, J. & Amaral, L. Team assembly mechanisms determine collaboration network structure and team performance. *Science* **308**, 697–702 (2005).
5. Leahey, E. et al. From sole investigator to team scientist: trends in the practice and study of research collaboration. *Annu. Rev. Sociol.* **42**, 81–100 (2016).
6. Milojevic, S. Principles of scientific research team formation and evolution. *Proc. Natl Acad. Sci. USA* **111**, 3984–3989 (2014).
7. Hunter, L. & Leahey, E. Collaborative research in sociology: trends and contributing factors. *Am. Sociol.* **39**, 290–306 (2008).
8. Xie, Y. 'Undemocracy': inequalities in science. *Science* **344**, 809–810 (2014).
9. Falk-Krzesinski, H. J. et al. Mapping a research agenda for the science of team science. *Res. Eval.* **20**, 145–158 (2011).
10. Barabasi, A. et al. Evolution of the social network of scientific collaborations. *Phys. A* **311**, 590–614 (2002).
11. Newman, M. E. J. The structure of scientific collaboration networks. *Proc. Natl Acad. Sci. USA* **98**, 404–409 (2001).
12. Petersen, A. M. Quantifying the impact of weak, strong, and super ties in scientific careers. *Proc. Natl Acad. Sci. USA* **112**, E4671–E4680 (2015).
13. Li, M. et al. Evolving model of weighted networks inspired by scientific collaboration networks. *Phys. A* **375**, 355–364 (2007).
14. Borner, K., Maru, J. T. & Goldstone, R. L. The simultaneous evolution of author and paper networks. *Proc. Natl Acad. Sci. USA* **101**, 5266–5273 (2004).
15. Redner, S. How popular is your paper? An empirical study of the citation distribution. *Eur. Phys. J. B* **4**, 131–134 (1998).
16. Klug, M. & Bagrow, J. P. Understanding the group dynamics and success of teams. *R. Soc. Open Sci.* **3**, 160007 (2016).
17. Hsiehchen, D., Espinoza, M. & Hsieh, A. Multinational teams and diseconomies of scale in collaborative research. *Sci. Adv.* **1**, e1500211 (2015).
18. Wu, L., Wang, D. & Evans, J. A. Large teams develop and small teams disrupt science and technology. *Nature* **566**, 378–382 (2019).
19. Coccia, M. & Wang, L. Evolution and convergence of the patterns of international scientific collaboration. *Proc. Natl Acad. Sci. USA* **113**, 2057–2061 (2016).
20. Jones, B. F., Wuchty, S. & Uzzi, B. Multi-university research teams: shifting impact, geography, and stratification in science. *Science* **322**, 1259–1262 (2008).
21. Gazni, A., Sugimoto, C. R. & Didegah, F. Mapping world scientific collaboration: authors, institutions, and countries. *J. Am. Soc. Inf. Sci. Technol.* **63**, 323–335 (2012).
22. Van Noorden, R. et al. Interdisciplinary research by the numbers. *Nature* **525**, 306–307 (2015).
23. Uzzi, B., Mukherjee, S., Stringer, M. & Jones, B. Atypical combinations and scientific impact. *Science* **342**, 468–472 (2013).
24. Stephan, P. E. & Levin, S. G. Age and the Nobel Prize revisited. *Scientometrics* **28**, 387–399 (1993).
25. Jones, B. F. & Weinberg, B. A. Age dynamics in scientific creativity. *Proc. Natl Acad. Sci. USA* **108**, 18910–18914 (2011).
26. Jones, B. F., Reedy, E. J. & Weinberg, B. A. *Age and Scientific Genius* (Wiley-Blackwell, 2014).
27. Sinatra, R., Wang, D., Deville, P., Song, C. & Barabasi, A.-L. Quantifying the evolution of individual scientific impact. *Science* **354**, aaf5239 (2016).
28. Funk, R. J. & Owen-Smith, J. A dynamic network measure of technological change. *Manag. Sci.* **63**, 791–817 (2017).
29. Sinatra, R., Deville, P., Szell, M., Wang, D. & Barabasi, A.-L. A century of physics. *Nat. Phys.* **11**, 791–796 (2015).
30. Zhou, T., Lu, L. & Zhang, Y.-C. Predicting missing links via local information. *Eur. Phys. J. B* **71**, 623–630 (2009).
31. Petersen, A. M. et al. Reputation and impact in academic careers. *Proc. Natl Acad. Sci. USA* **111**, 15316–15321 (2014).
32. Zeng, A. et al. Increasing trend of scientists to switch between topics. *Nat. Commun.* **10**, 3439 (2019).
33. Jia, T., Wang, D. & Szymanski, B. K. Quantifying patterns of research-interest evolution. *Nat. Hum. Behav.* **1**, 0078 (2017).
34. Tang, J. et al. ArnetMiner: extraction and mining of academic social networks. In *Proc. Fourteenth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (SIGKDD'2008)* (eds Li, Y., Liu, B. & Sarawagi, S.) 990–998 (Association for Computing Machinery, 2008).
35. Sinha, A. et al. An overview of Microsoft Academic Service (MA) and applications. In *Proc. 24th International Conference on World Wide Web (WWW '15 Companion)* (eds Gangemi, A., Leonardi, S. & Panconesi, A.) 243–246 (Association for Computing Machinery, 2015).
36. Stirling, A. A general framework for analysing diversity in science, technology and society. *J. R. Soc. Interface* **4**, 707–719 (2007).
37. Porter, A. & Rafols, I. Is science becoming more interdisciplinary? Measuring and mapping six research fields over time. *Scientometrics* **81**, 719–745 (2009).

Acknowledgements
This work is supported by the National Natural Science Foundation of China under Grant (71843005 and 71731002). S.H. thanks the Israel Science Foundation and the NSF-BSF for financial support. The funders had no role in study design, data collection and analysis, decision to publish or preparation of the manuscript.

Author contributions
A.Z. and S.H. designed the research. A.Z. performed the experiments. Y.F., Z.D. and Y.W. contributed analytical tools. A.Z. and S.H. analysed the data. All of the authors wrote the manuscript.

Competing interests
The authors declare no competing interests.

Additional information
Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41562-021-01084-x.

Correspondence and requests for materials should be addressed to S.H.

Peer review information *Nature Human Behaviour* thanks Filipi Nascimento Silva and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Peer reviewer reports are available.

nature research

Corresponding author(s): Shlomo Havlin

Last updated by author(s): Jan 26, 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 *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 *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	We did not use any open source codes for data collection in this study.
Data analysis	We did not use any open source codes for analysis in this study.

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 APS data can be downloaded via https://journals.aps.org/datasets. The computer science data can be downloaded via https://www.aminer.cn/aminernetwork. The multi-disciplinary data was download from https://docs.microsoft.com/en-us/academic-services/graph. Other related, relevant data are available from the corresponding author upon reasonable request.

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	We study the originality and multi-disciplinary impact of the papers published by fresh and old teams.
Research sample	We analyze in the paper the publication data from all journals of American Physical Society (APS). The computer science data analyzed in the supplementary materials was obtained by extracting scientists' profiles from on-line Web databases. The Chemistry data analyzed in the supplementary materials was the publication data of Journal of the American Chemical Society. The multi-disciplinary data analyzed in the supplementary materials contains all papers in five representative multi-disciplinary journals including Nature, Science, Proceedings of the National Academy of Sciences (PNAS), Nature Communications and Science Advances.
Sampling strategy	As our study focuses on papers published by teams, we take into account the papers with number of authors from 2 to 10.
Data collection	The American Physical Society data was downloaded via https://journals.aps.org/datasets, and the computer science data was downloaded via https://www.aminer.cn/aminernetwork. The data of Journal of the American Chemical Society was downloaded from http://apps.webofknowledge.com. The multi-disciplinary data was download from https://docs.microsoft.com/en-us/academic-services/graph.
Timing	The American Physical Society data ranges from year 1893 to year 2010. The computer science data ranges from year 1948 to year 2014. The data of Journal of the American Chemical Society ranges from 1997 to 2017. The data of multi-disciplinary journals ranges from 1869 to 2020.
Data exclusions	No data were excluded from the analysis.
Non-participation	No dropout participates.
Randomization	Participates 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