SCIENTIFIC PUBLISHING

10.1126/sciadv.1500211

Multinational teams and diseconomies of scale in collaborative research

David Hsiehchen,[1]* Magdalena Espinoza,[1] Antony Hsieh[2]

Collaborative research has become the mainstay in knowledge production across many domains of science and is widely promoted as a means of cultivating research quality, enhanced resource utilization, and high impact. An accurate appraisal of the value of collaborative research efforts is necessary to inform current funding and research policies. We reveal contemporary trends in collaborative research spanning multiple subject fields, with a particular focus on interactions between nations. We also examined citation outcomes of research teams and confirmed the accumulative benefits of having additional authors and unique countries involved. However, when per capita citation rates were analyzed to disambiguate the effects of authors and countries, decreasing returns in citations were noted with increasing authors among large research teams. In contrast, an increasing number of unique countries had a persistent additive citation effect. We also assessed the placement of foreign authors relative to the first author in paper bylines of biomedical research articles, which demonstrated a significant citation advantage of having an international presence in the second-to-last author position, possibly occupied by foreign primary co-investigators. Our analyses highlight the evolution and functional impact of team dynamics in research and suggest empirical strategies to evaluate team science.

INTRODUCTION

Scientific discoveries are customarily recognized as the feat of individuals or small teams. For example, conventional funding grants are typically awarded to a single investigator, and internationally renowned science prizes (for example, the Lasker and Nobel awards) are bestowed to no more than three persons per achievement. However, enlarging team sizes, increasing interdisciplinarity, and intensifying ties across institutional and geographic borders demonstrate how scientific research has evolved from a solitary enterprise to an expanding social movement (*1*–*5*). This paradigm shift of how knowledge is created has spurred the development of the "science of team science" with the goals of elucidating mechanisms and outcomes of research teams and defining the optimal circumstances for team-based inquiries to foster evidence-based research policies or practices (*6*).

Citations are popularly used among bibliometric studies to gauge the significance and usage of articles. In addition, citation-based metrics have been implicated in performance review, funding, and promotion, despite their problematic interpretation (*7*). Nonetheless, few other measures are as universal or as simple to calculate while preserving the same connotations. It was previously demonstrated that large multinational research teams were associated with increased citations or publication in high-impact journals (*8*–*13*). It has been reported that citations correlate with the physical distance between collaborators in the European Union and the number of country affiliations (*14*, *15*). These findings promote multicountry collaboration as a critical component of national portfolios due to its association with greater research impact. However, dissenting studies suggest a negligible or context-specific effect of multinational teams on citation rates in publications (*16*–*19*). This controversy likely stems from limitations in scope or reliance on small sample sizes of the aforementioned studies, resulting in fragmented perspectives. In addition, the quantitative relationship between research impact and the precise constitution of research teams remains ambiguous. Conspicuously, the relationship between the differential contribution of authors in multicountry collaborations and research impact has also not been investigated to date.

Herein, we perform a comprehensive analysis of multinational research spanning the last four decades and its citation impact. Across multiple disciplines, we discovered that large country teams were the fastest-growing modes of scholarship, with significant citation benefits being associated with increasing numbers of authors and national affiliations in papers. However, subset analyses revealed that citation advantages were dependent on team sizes, with additional author members but not necessarily country affiliations having a decreased or negative effect on citations among the largest team compositions. Furthermore, using a manually curated database designed to capture more qualitative aspects of international collaborations in papers, we reveal, as far as we are aware, the role of authorship position and the allocation of credit to foreign authors in citation outcomes. Our results highlight the differential effects of knowledge production strategies by demonstrating that greater personnel or human capital investments by foreign countries may lead to highly cited works. Our study also advocates the establishment of evidence-based measures to evaluate and guide team science, which has been hitherto widely promoted and directed in research policies largely based on anecdotal expectations.

RESULTS

Global trends in research team sizes

We analyzed about 24 million articles published over four decades indexed in the Thomson Reuters Web of Science (WOS) database, which covers the natural, social, and applied research disciplines. Our data set was restricted to articles and conference proceedings presenting original research. Although texts not pertaining to primary research constitute a considerable and growing fraction (nearly a quarter in recent years) of the citable literature, they were excluded in our investigation because they indirectly contribute to knowledge production and have dissimilar citation patterns (*15*, *20*, *21*). Here, international collaboration was

[1]Mount Auburn Hospital, Cambridge, MA 02138, USA. [2]Northwestern Memorial Hospital, Northwestern University, Chicago, IL 60611, USA.
*Corresponding author. E-mail: gbtwnow@gmail.com

defined as having more than one country among all author affiliations, and the number of national affiliations was calculated from the sum of unique country partners in a paper.

In examining global changes in research behavior, we found that the ratio of authors and countries to publications in all research fields has steadily risen over time with a marked surge apparent by the 2000s (Fig. 1A). Grouping publications by categories of author team sizes demonstrated the waning dominance of small team (composed of two to four authors) publications and a rapid decline in the proportion of single-authored works (Fig. 1B). This was contrasted by an increase in the fraction of papers attributed to medium team (composed of five to eight authors) and large team (composed of more than eight authors) papers. An analogous dissection of publications by categories of country team sizes showed a preponderance of single-country publications that has steadily diminished, an increase in the proportion of small country team (composed of two countries) research, and seemingly marginal changes in the proportion of medium (composed of three to four countries) and large country teams (composed of more than four countries) (Fig. 1B). Similar results were seen in the subset analyses of papers that had been sorted into technology and engineering, physical, life, social, and arts disciplines with some caveats (fig. S1). In particular, solo authors in the arts were the principal knowledge producers in the 1970s and have only experienced a slight relative decline since. Meanwhile, the dominant contribution of solo authors in the social sciences has ebbed with the emergence of principally small team size papers by the late 1990s. It is important to note that despite large shifts in the relative contribution of different team sizes to the aggregate body of knowledge, the absolute number of papers credited to all classes of author and country team sizes has actually increased over time. For example, whereas large author team output has grown from 667 to 87,525 papers between 1973 and 2009, single-authored works have also grown from 73,035 to 110,785 papers in the same time period. Similarly, notwithstanding contrasting trends in their relative publication output, large country teams and small country teams have increased publication production from 28 to 5507 papers and from 208,917 to 825,956 papers, respectively.

Notably, despite diminutive changes in the absolute research contribution of larger country teams, closer scrutiny of relative changes in publication patterns demonstrated that large country teams have exhibited near-exponential growth over time among all research disciplines except for the arts (Fig. 1C). They were also the fastest-growing collaborative structure, followed by medium and then small country relationships. This pattern was conserved but attenuated in recent years (fig. S2).

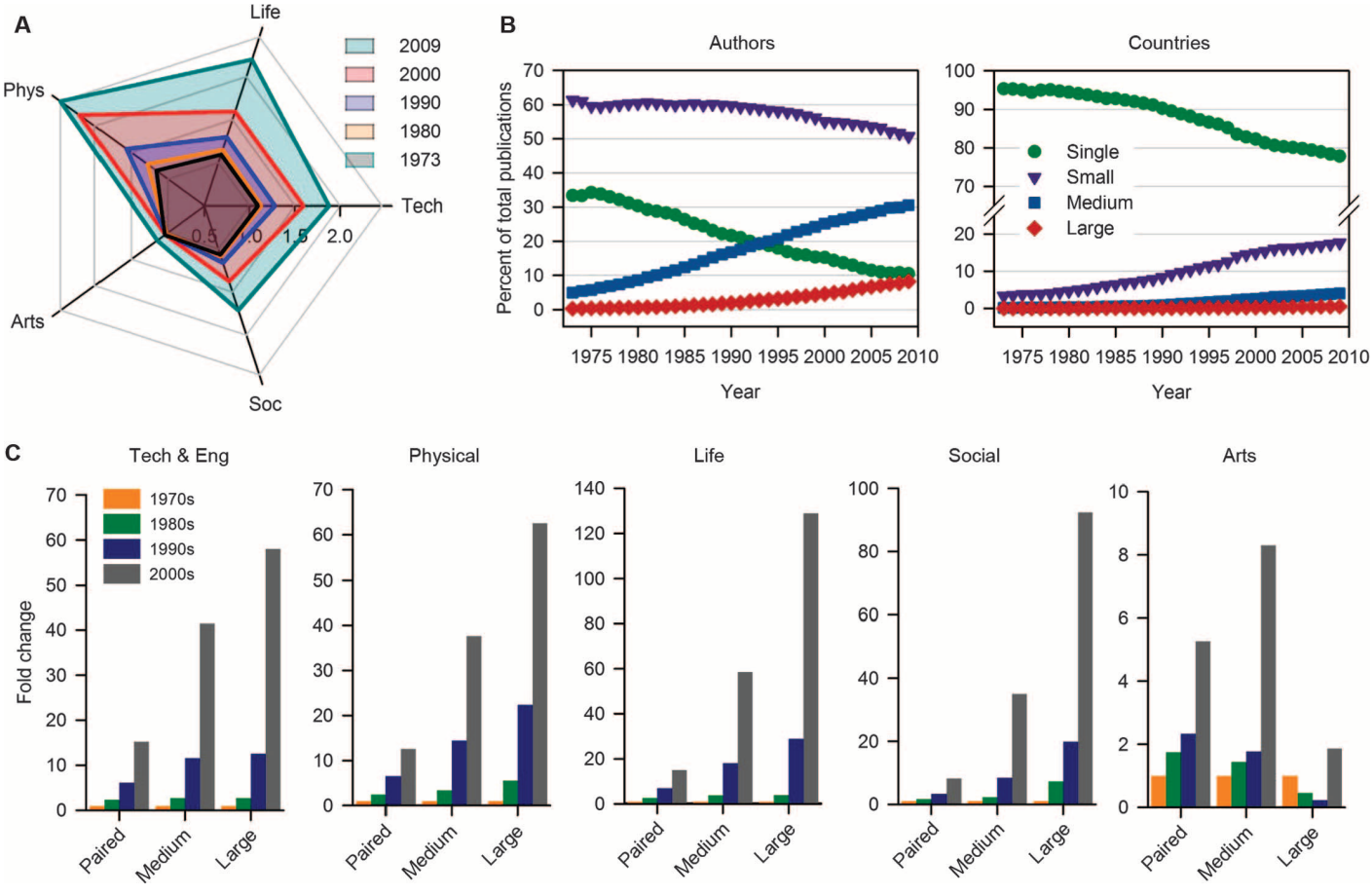

Fig. 1. Demographic shifts in knowledge producers. (**A**) The ratio of countries per publication was averaged for each time period and research discipline. Data were further normalized to data from 1973 to show relative changes over time. (**B**) Proportion of publications produced by different team sizes stratified by number of authors or countries. Single, small, medium, and large teams denote one author or country, two to four authors or two countries, five to eight authors or three to four countries, and more than nine authors or five countries, respectively. (**C**) Fold change in publication number by different country team sizes over different decades relative to 1970s data.

The differing behavior of arts publications, distinguished by the persistence of single-country papers and the rapid rise of paired and medium but not large country team sizes, likely pertains to field-specific norms of knowledge creation that remain to be defined. Notably, our study uses papers as proxies for research teams, and the marked increase in large country team papers may be secondary to their increased productivity. The fact that small country teams still constitute the great majority of all multinational research suggests that barriers to collaboration are least when it is with only one partner.

Citations and international collaborations

We next sought to compare the citation impact of papers involving international authors and papers exclusively involving domestic authors. Across all disciplines and conserved over time, a multinational presence was coupled with a decreased probability of not being cited and an increased probability of being among the most cited papers (Fig. 2, A to D). Differences in the probability of not being cited or being highly cited were not different over time. Congruent with this finding, the ratio of citations garnered by international collaborations versus domestic collaborations was always greater than 1 but largely static over the last few decades (Fig. 2E).

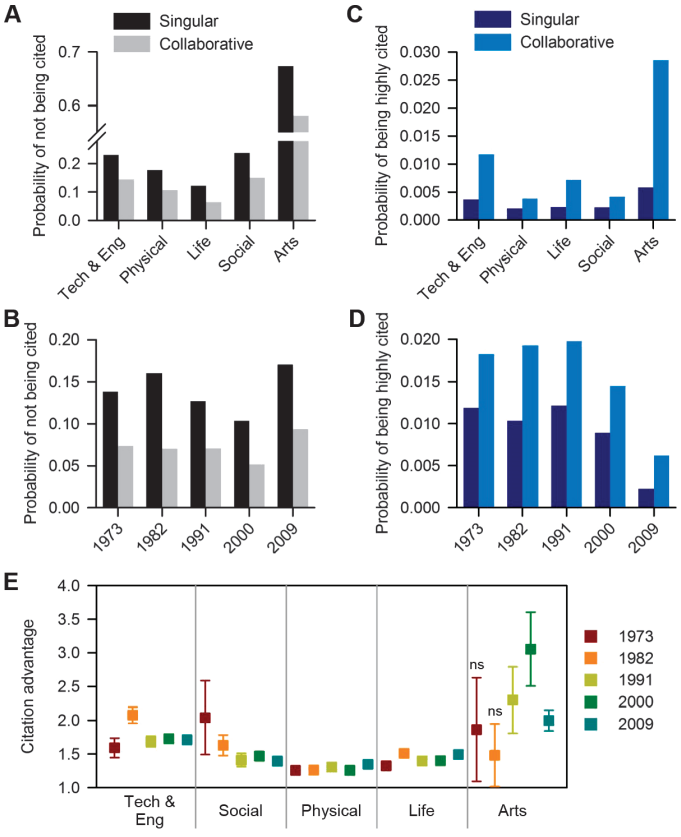

Fig. 2. Increased citations are associated with multicountry collaborations. (**A** and **B**) The probability of not being cited is decreased in collaborative papers compared to singular-nation papers, and this was conserved across subject areas (A) and time (B). (**C** and **D**) The increased probability of being highly cited (being in the top percent of all articles published in a given year by citations) was also conserved across subject areas (C) and time (D). All data were found to be significant ($P < 0.001$) using the χ^2 test. (**E**) The citation advantage ratio, defined as the mean citation of multinational papers divided by the mean citation of singular-nation papers, was calculated for the indicated years. Citations of multicountry and singular-nation papers were significantly different ($P < 0.05$) for all comparisons except for years 1973 and 1982 in the arts discipline. ns, not significant.

Larger author compositions were previously shown to be associated with greater citation rates (*1*). Thus, an increased number of authors in multicountry collaboratives could confound our previous assertion that multination affiliations in research are associated with a heightened citation impact. To disambiguate the influence of national affiliations from author numbers, we analyzed citation rates of all publications stratified by authorship composition that showed a dose-dependent relationship between the number of country partners and citations among small (composed of 2 to 4 authors), medium (composed of 5 to 9 authors), and large (composed of 9 to 20 authors) teams. However, the association between citations and national affiliations per paper became less apparent in publications authored by mega teams (composed of more than 20 authors) over the different time periods studied (fig. S3, A to D). An exception to these results was seen in data from 1973, where no papers could be classified into the mega team category, although ambiguous citation benefits were nonetheless noted in large teams (fig. S3E). As a robustness check, we also analyzed the top fraction of papers with the most authors from multiple years and showed an enduring lack of difference in citations between domestic and international collaborations (fig. S3F).

Declining citations per capita with increasing author team sizes

Given a dearth of papers written by large author teams within the arts and social science disciplines, we focused our subsequent citation analyses on papers pertaining to the life, physical, and technology and engineering sciences. When we examined the underlying distribution of citations before data binning, the disassociation between author numbers and citations was apparent when author team sizes grew beyond 20 members (mega teams) in all subject fields (Fig. 3A and fig. S4). These findings suggest a declining or detrimental effect of either author or country numbers on citations. This would be consistent with diseconomies of scale observed in many human endeavors with marginal costs increasing once firms or organizations surpass an optimal mass. We tested this hypothesis using generalized linear models (GLMs) to quantify the relationship between citations and researcher team size. Specifically, we used GLMs, a generalization of linear regressions that allows for skewed distributions of response variables, because of the non-normal distribution of citations. Whereas the contribution of an additional country or author was positive and comparable when analyzing all publications in recent years, subset analyses of top-ranked papers by team size, defined by being at least within the top 1% of all papers ranked by the number of authors, countries, or institutional affiliations, showed a reduced or even deleterious effect of additional authors across different years (fig. S5, A to C). The application of different criteria for larger teams did not change our results, suggesting that decreased citations per capita among top-ranked papers are not strictly dependent on a precise definition of team sizes. However, the median author count of the top 1% of papers ranked by any authorship attribute was about 20, indicating some redundancy in the composition of analyzed papers. The citation gain related to national affiliations was largely preserved or increased among large-team papers. An expanded analysis of more than 10 million papers published over a decade also showed a decreased or inhibitory effect of author numbers on citation

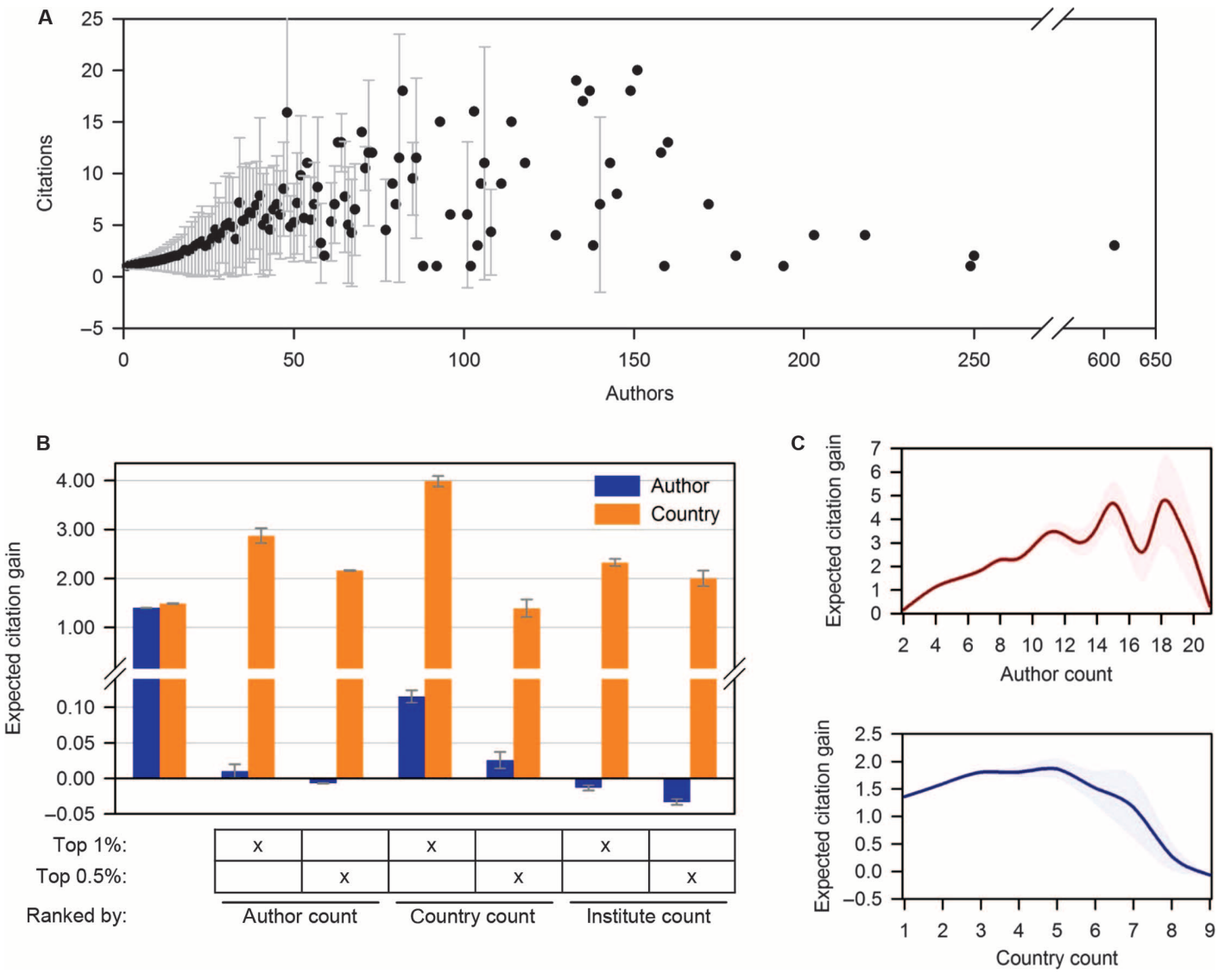

Fig. 3. Diseconomy of scale in citation benefits among papers produced by large teams. (**A**) Mean citations for different author team sizes. (**B**) The expected citation gain from additional authors or national affiliations was determined using a GLM of all publications (far left column) and the top 1 or 0.5% of papers (as indicated) after they were ranked by increasing author, country, or institute numbers. Regression coefficients were significant for all models ($P < 0.01$). (**C**) The citation benefit of additional national affiliations (top panel) and authors (bottom panel) was determined for specific team masses as defined by author and country counts, respectively. Lightly shaded areas surrounding lines depict 95% confidence intervals.

gains across different definitions of large team sizes after adjustment for year of publication and subject field content (Fig. 3B). Although the length of articles and number of references in a paper have been previously reported to correlate with citations, inclusion of the total number of pages or references in regression models did not change the effect or significance of our results (*22*, *23*).

To elucidate circumstances in which additional national affiliations or author count have provided the greatest citation benefit, we assessed citation gains from either an additional author or country for different author or country team sizes, respectively. Among all publications, there was a slow rise in citation benefits from national affiliations with increasing author team sizes, peaking at 18 individuals, before a sharp decline (Fig. 3C). The citation gains for additional authors also increased with growing country team sizes, peaking at five national members, before declining (Fig. 3C). These results allude to optimal team sizes in citation benefits and demonstrate both efficiency and inefficiency in scaling for multinational teams. However, unlike for author numbers, the citation benefits for country affiliations near but have yet to reach or fall below zero.

Citation impact of the number and order of international authors

We next sought to determine whether the number or organization of international authors (with respect to the first author's national affiliation) in multinational papers was also associated with citation outcomes. Because of the lack of authorship detail captured in our

initial WOS data set, we manually ascertained the domestic or foreign status relative to the first author and authorship position of all authors among nearly 4700 primary articles published across 21 biomedical journals that were selected to provide a range of impact factors. The impact factors of journals correlated well with the mean number of international authors in a paper and the average size of international teams, suggesting a numerical advantage in citations for collaborations involving foreign teams contributing a larger number of researchers (Fig. 4, A and B).

At the paper level, our biomedical literature panel demonstrated conserved citation benefits with increasing author numbers and increasing number of unique affiliated countries (Fig. 4C). In univariate analyses of international author numbers and average international author team sizes, there was a perceptible citation advantage with increasing mass of either authorship property (Fig. 4C). However, after adjustment for total author count, national affiliations, or both, only the citation effects stemming from the total number of international authors remained significant. Collectively, these findings suggest that citation benefits associated with multinational research teams may be dependent on both a diverse representation of state entities (that is, the number of unique affiliated countries) and a substantial investment of foreign human capital, although the value of average international team sizes is ambiguous.

In biomedical research, authorship position often follows the convention of author order being inversely related to an individual's contribution, with the exception of the last author positions, which are reserved for team leaders or principal investigators (PIs). Although the exact proportion of credit that can be attributed to an author position is dependent on social dynamics and team-specific practices, we considered authorship order as an approximate measure of an individual's involvement in the project. Among multicountry collaborations involving more than two authors, the presence of international authors in the second or third author position, likely signifying a substantial allocation of credit, had an insignificant impact on citations (Fig. 4D). Foreign authors in successive author positions also did not affect research impact. Conversely, even after adjustment for other authorship properties, the occupation of the second-to-last and last author positions by foreign contributors was associated with a citation advantage and disadvantage, respectively (Fig. 4D). We surmised that the second-to-last author position in international team papers likely denotes a significant contribution by a foreign PI in the provision of management, resources, skills, or creative input but not to the extent of the primary team leader or investigator. On the other hand, an international author as the last author, a position typically occupied by a corresponding author ultimately accountable for the study, likely represents a foreign individual who is principally responsible for the work or provided the greatest degree of oversight. The physical proximity between the first and last authors in papers has been previously linked to increased citations among papers from a single institute (*24*). Our results indicate similar citation benefits when the first and last authors are from the same

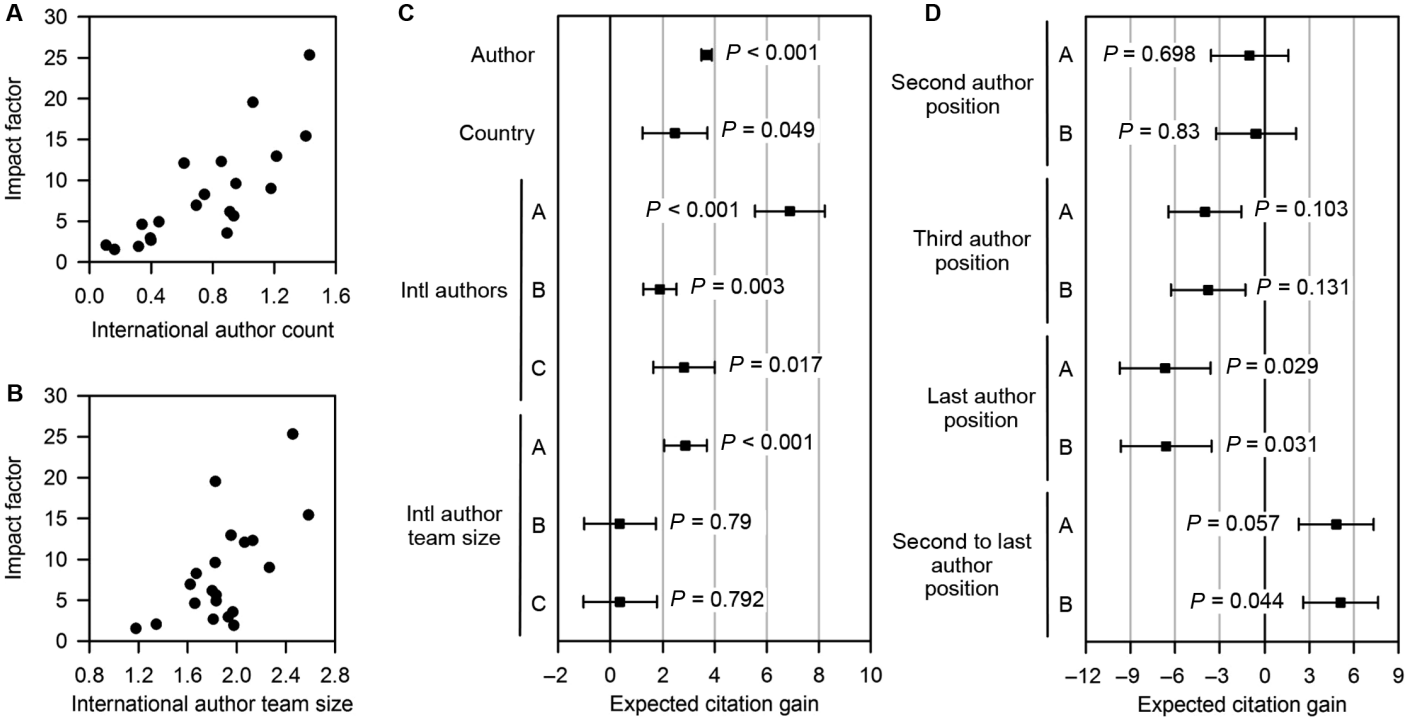

Fig. 4. Quantitative measures of multicountry collaborations and citations. (**A** and **B**) Journal impact factors correlate with average international (Intl) author counts (A) and international author team sizes (B). (**C**) Coefficients from regression analyses of author and country numbers of papers in the biomedical literature panel displayed in the top two rows demonstrate the conserved citation advantage of author and unique country numbers. The citation effects related to the number of international authors per paper and the average author team sizes of foreign authors are displayed in the indicated rows. Rows designated A display coefficients obtained from univariate models. Rows designated B and C display coefficients obtained from models adjusting for author and both author numbers and national affiliations, respectively. (**D**) Citation effects associated with different author positions occupied by foreign authors are displayed as indicated. Rows designated A and B display coefficients obtained from univariate analyses and adjusted models, respectively.

country, although in our data set the outcome did not necessitate that the first and last authors originate from the same institute.

Because the role of authors and their authorship order may follow different conventions in different disciplines, additional analyses are needed to determine whether our findings may be generalized to teams outside of biomedical research. Nonetheless, a particular social hierarchy or organization in multinational teams may be critical in sustaining efficient and effective research practices. Our data suggest that qualitative differences in the degree of participation by international team members, particularly of senior members rather than subordinate researchers, may influence the impact of multinational research, although this is countered by a detrimental effect if the first author is not proximally supervised.

DISCUSSION

Here, we uncovered several themes in authorship dynamics in research production. In particular, we highlight how team sizes are growing in part because of the increased participation of international authors and how a growing fraction of our knowledge stems from larger multinational teams. Although the increasingly collaborative nature of modern research is multifactorial, our findings demonstrate the waning role of geographic and political borders as barriers in knowledge production, with technological advances likely enhancing the flow of resources, personnel, and knowledge worldwide. The accelerating growth of publications originating from growing team sizes defined by either author numbers or national affiliations led us to question the functional impact of evolving research practices.

This study is distinct from past research, which only examined aggregate citation rates of publications rather than citations per capita. Our strategy allowed us to disambiguate the citation effects of different authorship traits and define exploitable strategies of maximizing research outcomes. In addition, per capita measurements enabled us to account for dynamic changes in research team structures. Although our results initially support a numerical advantage for authors and national affiliations on citations, this benefit was subsequently found to be attenuated, predominantly for the number of authors, in larger research teams. The mechanisms underlying the mutable citation effects of team sizes remain to be elucidated, although diseconomies of scale are omnipresent across social endeavors due to ineffective communication, redundant efforts, and increased bureaucracy. Thus, we hypothesize that the organization of researchers in social networks may contribute to increasing costs associated with scale. This can be illustrated by depicting researchers as individual nodes in a network, with links representing communication or information flow. The links between researchers depict the social infrastructure of research teams, akin to distribution networks hypothesized to underlie the scaling properties of biological and social dimensions (*25*–*27*). In particular, Bettencourt and colleagues showed that select urban indicators (including gross domestic product, patents, employment, and crimes) grow disproportionately with the population size of cities, which is governed by the spatial organization of social interactions in a physical network (such as a grid of streets) where persons, goods, or information may flow (*25*). For example, more populous cities are associated with a greater rate of innovation and wealth creation compared to less populous cities in part due the increased density of persons and a resultant greater likelihood of social interactions that may lead to creative or material outputs (*28*). Bettencourt further demonstrated that the costs of creating or maintaining links (such as the amount of time, resources, efforts, and even opportunity costs expended on relationships) in an infrastructure network may be analogous to dissipative processes in electrical circuits (*26*). Given that networks incur resistance (energy lost to enable the flow of information or innovation) per network segment, energy dissipation also grows disproportionately with innovation or wealth as city populations rise. The tension between social interactions and the costs of maintaining such contacts underlie the success of cities and research enterprises, both of which must balance social connectivity with infrastructure costs.

The question of why smaller teams are associated with lower link costs and higher citations per capita remains. It was recently reported that team distributions in astronomy depict two modes of team growth: the formation of a small core team followed by the accumulation of new members (*29*). On the basis of this concept, we propose that collaborative research teams originate from a small core group of researchers characterized by high interconnectedness (such as a clique). Assuming that the number of links a researcher may be involved in is bounded, that the costs of links are uniform, and that there is a maximum network cost that can be accrued, an increasing number of researchers may find themselves outside the core group and situated at the periphery of the research network as more researchers join the team to maintain the number of interactions among core team members. This may be due to some researchers being pushed to more ancillary roles or the fact that newly added or junior members operate in more isolated settings, which parallels the preserved overall distribution of interactions in an individual's social network even when there is turnover or added network members (*30*). Whatever the case, the density of links in the infrastructure network of large research teams declines with scale, suggesting a relative decline in opportunities for social interactions that may stifle innovation. Thus, the ideal structure of research teams may entail a degree of intimacy as well as a critical mass (such as a core team) for collective input to engender creativity or productivity. The reason why increasing national affiliations may have a persistent citation advantage may be ascribed to the ability of international authors [such as foreign primary investigators (PIs)] to bridge multiple core groups to maintain a compact social infrastructure (likely providing closer supervision of members by PIs or other members), reduce the impact of network costs by supplying additional financial, material, or intellectual assets, and supporting enhanced communication. Regarding the latter factor, effective communication likely includes mutually beneficial flow of information or innovation rather than only allowing for directed content flow in one direction, which may limit mechanisms for feedback or reciprocation. Such one-way communication may be found between peers because of limits in expertise or interest, social conflicts, and other barriers, although the same factors could also pertain to communication with PIs or leaders. Although these notions remain speculative, reconstructing social networks of domestic and international research teams of diverse sizes may help elucidate the exact mechanisms contributing to either the economies or diseconomies of scale in research and address remaining questions, such as the appropriate size of core teams for different disciplines.

Although the identity and roles of collaborators in large cooperatives were not examined in this study, it is likely that the quality of participation by individuals also plays a role in limiting the impact of some collaborations. For example, as research teams expand, the pool of the most qualified or available experts within a nation may diminish, potentially

necessitating the involvement of experts from other countries although the number of nations is also limited. Thus, for very large teams, collaborations may necessitate compromises in member recruitment during its formative stages. In addition, inequalities in scientific enterprises and practices, such as disproportionate rewards and esteem given to top performers and skewed allocation of funding among researchers, may promote the increasing stratification of collaborations, resulting in fewer rather than more projects of high impact (*4*, *31*). The declining citation gain of larger research teams is not likely due to specific forms of research given its occurrence across different subject disciplines.

The causes of an additive citation effect among large research teams, observed in papers generally authored by less than 20 authors, may stem from access to geographically or politically restricted resources, synergistic social interactions, and complementation of regional research infrastructure, which may heighten or cultivate research impact. The inclusion of additional authors may also lead to increased self-citations and can be viewed as being consistent with the incremental nature of science, although it may also be secondary to self-promotion. Nonetheless, recent studies restricted to specific subject fields indicate only a slight increase in self-citations with increased authors and a negligible role for international collaborations on self-citations in the overall research impact of papers, suggesting only a limited role for self-promotion in the citation advantage of multinational teams (*9*, *21*, *32*).

Although it appears that modern research has already crossed the threshold of efficiency in regards to author numbers, citation benefits related to international affiliations continue to persist. This could suggest that research teams saturated with authors may profit from incorporating additional international expertise, particularly from a different locale, to enhance the impact of their work. It would be of particular interest in future queries to determine whether team size influences are preserved in most or select country relationships. Crucially, our work not only promotes the value of diverse national representation in research but also substantial investments or more direct participation by foreign individuals or countries in collaborations. This is supported by the link between the number of international authors and conceivably the presence of foreign co-PIs with citations. Although we relied on authorship proxies to portray properties of collaborations, the recent advent of explicated authorship roles in papers may provide a promising avenue to apply more precise measurements of author contributions in research teams.

Although the broad application of citations remains under scrutiny, many citation-based indicators that are used to gauge researcher or journal performance, such as the impact factor, *h*-index, *g*-index, eigenfactor, and other derivatives, do not account for individual-level contributions (*33*–*38*). Thus, papers that are produced by either two or a hundred individuals may not be differently treated by most indexes although the degree of author participation may be drastically different between the two publications. Our analysis of per capita citation rates adds an extra dimension to the evaluation of a paper's scientific impact by attempting to account for potential resources invested, namely, the number of scientists. This strategy may also be valuable for measuring the relative impact of researchers and incentivizing authors to curtail practices leading to the inflation or misappropriation of authorship that may be refractory to publication policies (*39*–*41*). In addition, although more precise indicators remain to be developed, multiple lines of evidence, including results contained herein and the increasing incidence of joint authorship positions, support the incorporation of authorship order in evaluating the relative impact of researchers and their work (*7*, *42*–*45*).

A limitation to this study is our reliance on the WOS database. However, to date, there are few other databases that are as comprehensive or complete as WOS. In addition, the scale of our study may suggest that our results are likely robust despite our dependence on one bibliometric source. Nevertheless, given slight differences among citation databases, it would be of interest for future investigations to analyze other resources beyond WOS (*46*, *47*). Another limitation in our study was our focus on English language articles, due to the differing citation patterns of non-English articles (*48*). Although non-English articles only contribute to a small fraction (less than 5% of 2011 articles are non-English) of our body of knowledge, further studies are needed to determine whether our findings are generalizable to research teams in disciplines where non-English articles may predominate.

Our results do not suggest that there is less to gain for individuals who collaborate within large teams. Collaborations likely allow scientists to simultaneously partake in a larger number of projects and thus increase productivity. Consequently, it may be advantageous at the author level to participate in as many collaborations as possible, although this strategy may not be prudent for all individuals because only a substantial contribution on papers, such as the first or last author placement, may lead to career advancement (*7*). An assessment of time inputs by researchers in collaborations and their effect on research impact remains an important direction for future work. Our study also does not imply a lack of need for large collaboratives because unique or massive projects requiring tremendous resources or analytical power and integration of multiple disciplines, such as the sequencing of the human genome and the discovery of the Higgs boson, would necessitate large teams of experts (*49*–*51*). However, not all large-scale projects have led to the expected paradigm shift or breakthrough in knowledge, as exemplified by the recent fundamental discoveries in the molecular genetics of several cancers by investigator-led research teams rather than massive and widely celebrated consortiums (*52*). Thus, at the paper or project level, very large collaboratives do not necessarily yield the greatest return on investments, highlighting the potential costs of collaborations, which have yet to be fully characterized. The growing prominence of very large research teams despite their association with diminishing returns in terms of citation impact suggests the impracticalities of science being indiscriminately conducted at the expense of smaller and possibly more efficient teams. These results may be particularly relevant to the growing need for research accountability and cost-effective practices.

MATERIALS AND METHODS

Bibliometric records of primary research articles in all subject areas published from 1973 to 2009 were downloaded from the Thomson Reuters (WOS) database. To query primary research articles, WOS publication results were refined using Boolean operators to include only English language entries coded as "article" or "conference proceedings" document types without anonymous authors. Publications coded in WOS as "article" document types were distinguished from nonprimary research articles if they included more than 100 references and contained the words "review" or "overview" within the title. WOS category fields were used to categorize the articles into the most relevant subject areas: technology and engineering, physical science, life science, social science, and arts and humanities. Records with missing or unreliable data in any record field were excluded from this study. In total, 24,161,726 articles were analyzed in this study, with 2,838,004 papers in technology and engineering,

7,479,363 papers in the physical sciences, 12,199,030 papers in the life sciences, 1,225,686 papers in the social sciences, and 419,643 papers in the arts.

To extract the country of origin, unique text strings corresponding to a country's name were ascertained within the author address and reprint address field tags to ensure that only one country was recognized from each address. Data from countries that have changed their names or composition since 1970, such as the Union of Soviet Socialist Republics, Yugoslavia, Germany, and Czechoslovakia, were incorporated into the publication counts of modern sovereign states that are politically and geographically representative of the previous entity or succeeded the previous state identity.

To qualitatively assess the degree of participation by foreign countries in the papers, the authorship data were manually curated from the primary research articles published in 2009 across 21 biomedical journals: *Cell*, *Cancer Cell*, *Nature Cell Biology*, *The Journal of Clinical Investigation*, *PLOS Biology*, *Nature Structural & Molecular Biology*, *Genes & Development*, *The Journal of Cell Biology*, *The EMBO Journal*, *Cell Death and Differentiation*, *EMBO Reports*, *Journal of Cell Science*, *BMC Biology*, *The International Journal of Biochemistry & Cell Biology*, *Journal of Cellular Physiology*, *FEBS Letters*, *Journal of Cellular Biochemistry*, *BMC Cell Biology*, *Molecular Biology Reports*, *Molecular and Cellular Biochemistry*, and *Cell Biochemistry and Function*. Citation data were determined from WOS, and only citations accrued up to 24 months after publication for each paper were included in our data set. Data curation was performed by a single author (D.H.) with repeat independent collection of authorship and citation data for 60 randomly selected articles per journal performed by the remaining two authors (M.E. and A.H.). Discrepancies were typically found in less than 5% of publications for each journal and were resolved by consensus among all authors. For three journals (*Cell*, *Cancer Cell*, and *Journal of Cellular Biochemistry*) where discrepancies were initially found in more than 5% of publications during repeat curation, the repeat independent collection of authorship and citation data was performed for all papers.

Statistical analysis was performed using SPSS Statistics 20 (IBM). Differences in the citations of multicountry and singular-nation papers were assessed using the Mann-Whitney test.

SUPPLEMENTARY MATERIALS

Supplementary material for this article is available at http://advances.sciencemag.org/cgi/content/full/1/8/e1500211/DC1

Fig. S1. Demographic shifts in team size of knowledge producers.

Fig. S2. Recent changes in collaborative publication output across disciplines.

Fig. S3. After controlling for author team size, publications with an increasing number of national affiliations are associated with increasing citations, except in papers authored by very large author team sizes.

Fig. S4. The correlation between author count and citations diminishes for large teams across research disciplines.

Fig. S5. Diseconomy of scale in citation benefits in recent years.

REFERENCES AND NOTES

1. S. Wuchty, B. F. Jones, B. Uzzi, The increasing dominance of teams in production of knowledge. *Science* **316**, 1036–1039 (2007).
2. A. L. Porter, I. Rafols, Is science becoming more interdisciplinary? Measuring and mapping six research fields over time. *Scientometrics* **81**, 719–745 (2009).
3. R. K. Pan, S. Sinha, K. Kaski, J. Saramäki, The evolution of interdisciplinarity in physics research. *Sci. Rep.* **2**, 551 (2012).
4. B. F. Jones, S. Wuchty, B. Uzzi, Multi-university research teams: Shifting impact, geography, and stratification in science. *Science* **322**, 1259–1262 (2008).
5. R. J. W. Tijssen, L. Waltman, N. J. van Eck, Collaborations span 1,553 kilometres. *Nature* **473**, 154 (2011).
6. K. Börner, N. Contractor, H. J. Falk-Krzesinski, S. M. Fiore, K. L. Hall, J. Keyton, B. Spring, D. Stokols, W. Trochim, B. Uzzi, A multi-level systems perspective for the science of team science. *Sci. Transl. Med.* **2**, 49cm24 (2010).
7. D. van Dijk, O. Manor, L. B. Carey, Publication metrics and success on the academic job market. *Curr. Biol.* **24**, R516–R517 (2014).
8. Z. Chinchilla-Rodríguez, M. Benavent-Pérez, F. de Moya-Anegón, S. Miguel, International collaboration in medical research in Latin America and the Caribbean (2003–2007). *J. Am. Soc. Inf. Sci. Technol.* **63**, 2223–2238 (2012).
9. A. F. J. Van Raan, The influence of international collaboration on the impact of research results. *Scientometrics* **42**, 423–428 (1998).
10. W. Y. Low, K. H. Ng, M. A. Kabir, A. P. Koh, J. Sinnasamy, Trend and impact of international collaboration in clinical medicine papers published in Malaysia. *Scientometrics* **98**, 1521–1533 (2014).
11. F. Didegah, M. Thelwall, Which factors help authors produce the highest impact research? Collaboration, journal and document properties. *J. Informetr.* **7**, 861–873 (2013).
12. A. Inzelt, A. Schubert, M. Schubert, Incremental citation impact due to international co-authorship in Hungarian higher education institutions. *Scientometrics* **78**, 37–43 (2009).
13. B. S. Lancho-Barrantes, V. P. Guerrero-Bote, F. Moya-Anegón, Citation increments between collaborating countries. *Scientometrics* **94**, 817–831 (2013).
14. Ö. Nomaler, K. Frenken, G. Heimeriks, Do more distant collaborations have more citation impact? *J. Informetr.* **7**, 966–971 (2013).
15. R. K. Pan, K. Kaski, S. Fortunato, World citation and collaboration networks: Uncovering the role of geography in science. *Sci. Rep.* **2**, 902 (2012).
16. J. N. Parker, S. Allesina, C. J. Lortie, Characterizing a scientific elite (B): Publication and citation patterns of the most highly cited scientists in environmental science and ecology. *Scientometrics* **94**, 469–480 (2013).
17. M. Thelwall, P. Sud, No citation advantage for monograph-based collaborations? *J. Informetr.* **8**, 276–283 (2014).
18. J. M. Levitt, M. Thelwall, Does the higher citation of collaborative research differ from region to region? A case study of Economics. *Scientometrics* **85**, 171–183 (2010).
19. O. Persson, Are highly cited papers more international? *Scientometrics* **83**, 397–401 (2010).
20. O. Persson, W. Glänzel, R. Danell, Inflationary bibliometric values: The role of scientific collaboration and the need for relative indicators in evaluative studies. *Scientometrics* **60**, 421–432 (2004).
21. M. L. Wallace, V. Larivière, Y. Gingras, A small world of citations? The influence of collaboration networks on citation practices. *PLOS One* **7**, e33339 (2012).
22. M. E. Falagas, A. Zarkali, D. E. Karageorgopoulos, V. Bardakas, M. N. Mavros, The impact of article length on the number of future citations: A bibliometric analysis of general medicine journals. *PLOS One* **8**, e49476 (2013).
23. Z. Corbyn, An easy way to boost a paper's citations. *Nature* 10.1038/news.2010.406 (2010).
24. K. Lee, J. S. Brownstein, R. G. Mills, I. S. Kohane, Does collocation inform the impact of collaboration? *PLOS One* **5**, e14279 (2010).
25. L. M. A. Bettencourt, J. Lobo, D. Helbing, C. Kühnert, G. B. West, Growth, innovation, scaling, and the pace of life in cities. *Proc. Natl. Acad. Sci. U.S.A.* **104**, 7301–7306 (2007).
26. L. M. A. Bettencourt, The origins of scaling in cities. *Science* **340**, 1438–1441 (2013).
27. G. B. West, J. H. Brown, B. J. Enquist, A general model for the origin of allometric scaling laws in biology. *Science* **276**, 122–126 (1997).
28. W. Pan, G. Ghoshal, C. Krumme, M. Cebrian, A. Pentland, Urban characteristics attributable to density-driven tie formation. *Nat. Commun.* **4**, 1961 (2013).
29. S. Milojević, Principles of scientific research team formation and evolution. *Proc. Natl. Acad. Sci. U.S.A.* **111**, 3984–3989 (2014).
30. J. Saramäki, E. A. Leicht, E. López, S. G. B. Roberts, F. Reed-Tsochas, R. I. M. Dunbar, Persistence of social signatures in human communication. *Proc. Natl. Acad. Sci. U.S.A.* **111**, 942–947 (2014).
31. Y. Xie, "Undemocracy": Inequalities in science. *Science* **344**, 809–810 (2014).
32. R. Costas, T. N. van Leeuwen, M. Bordons, Self-citations at the meso and individual levels: Effects of different calculation methods. *Scientometrics* **82**, 517–537 (2010).
33. C. T. Bergstrom, J. D. West, M. A. Wiseman, The Eigenfactor™ Metrics. *J. Neurosci.* **28**, 11433–11434 (2008).
34. L. Bornmann, R. Mutz, S. E. Hug, H.-D. Daniel, A multilevel meta-analysis of studies reporting correlations between the *h* index and 37 different *h* index variants. *J. Informetr.* **5**, 346–359 (2011).
35. L. Egghe, Theory and practise of the *g*-index. *Scientometrics* **69**, 131–152 (2006).
36. J. E. Hirsch, An index to quantify an individual's scientific research output. *Proc. Natl. Acad. Sci. U.S.A.* **102**, 16569–16572 (2005).
37. S. Lehmann, A. D. Jackson, B. E. Lautrup, Measures for measures. *Nature* **444**, 1003–1004 (2006).

38. R. K. Pan, S. Fortunato, Author Impact Factor: Tracking the dynamics of individual scientific impact. *Sci. Rep.* **4**, 4880 (2014).
39. A. Flanagin, L. A. Carey, P. B. Fontanarosa, S. G. Phillips, B. P. Pace, G. D. Lundberg, D. Rennie, Prevalence of articles with honorary authors and ghost authors in peer-reviewed medical journals. *JAMA* **280**, 222–224 (1998).
40. J. S. Wislar, A. Flanagin, P. B. Fontanarosa, C. D. DeAngelis, Honorary and ghost authorship in high impact biomedical journals: A cross sectional survey. *BMJ* **343**, d6128 (2011).
41. T. Bates, A. Anić, M. Marušić, A. Marušić, Authorship criteria and disclosure of contributions: Comparison of 3 general medical journals with different author contribution forms. *JAMA* **292**, 86–88 (2004).
42. J. Gomez-Alonso, Author! Author! *JAMA* **292**, 1815–1816 (2004).
43. D. W. Shapiro, N. S. Wenger, M. F. Shapiro, The contributions of authors to multiauthored biomedical research papers. *JAMA* **271**, 438–442 (1994).
44. E. Akhabue, E. Lautenbach, "Equal" contributions and credit: An emerging trend in the characterization of authorship. *Ann. Epidemiol.* **20**, 868–871 (2010).
45. J. D. Wren, K. Z. Kozak, K. R. Johnson, S. J. Deakyne, L. M. Schilling, R. P. Dellavalle, The write position. A survey of perceived contributions to papers based on byline position and number of authors. *EMBO Rep.* **8**, 988–991 (2007).
46. M. E. Falagas, E. I. Pitsouni, G. A. Malietzis, G. Pappas, Comparison of PubMed, Scopus, Web of Science, and Google Scholar: Strengths and weaknesses. *FASEB J.* **22**, 338–342 (2008).
47. A. V. Kulkarni, B. Aziz, I. Shams, J. W. Busse, Comparisons of citations in Web of Science, Scopus, and Google Scholar for articles published in general medical journals. *JAMA* **302**, 1092–1096 (2009).
48. K. Sangwal, Some citation-related characteristics of scientific journals published in individual countries. *Scientometrics* **97**, 719–741 (2013).
49. E. S. Lander, L. M. Linton, B. Birren, C. Nusbaum, M. C. Zody, J. Baldwin, K. Devon, K. Dewar, M. Doyle, W. FitzHugh, R. Funke, D. Gage, K. Harris, A. Heaford, J. Howland, L. Kann, J. Lehoczky, R. LeVine, P. McEwan, K. McKernan, J. Meldrim, J. P. Mesirov, C. Miranda, W. Morris, J. Naylor, C. Raymond, M. Rosetti, R. Santos, A. Sheridan, C. Sougnez, N. Stange-Thomann, N. Stojanovic, A. Subramanian, D. Wyman, J. Rogers, J. Sulston, R. Ainscough, S. Beck, D. Bentley, J. Burton, C. Clee, N. Carter, A. Coulson, R. Deadman, P. Deloukas, A. Dunham, I. Dunham, R. Durbin, L. French, D. Grafham, S. Gregory, T. Hubbard, S. Humphray, A. Hunt, M. Jones, C. Lloyd, A. McMurray, L. Matthews, S. Mercer, S. Milne, J. C. Mullikin, A. Mungall, R. Plumb, M. Ross, R. Shownkeen, S. Sims, R. H. Waterston, R. K. Wilson, L. W. Hillier, J. D. McPherson, M. A. Marra, E. R. Mardis, L. A. Fulton, A. T. Chinwalla, K. H. Pepin, W. R. Gish, S. L. Chissoe, M. C. Wendl, K. D. Delehaunty, T. L. Miner, A. Delehaunty, J. B. Kramer, L. L. Cook, R. S. Fulton, D. L. Johnson, P. J. Minx, S. W. Clifton, T. Hawkins, E. Branscomb, P. Predki, P. Richardson, S. Wenning, T. Slezak, N. Doggett, J. F. Cheng, A. Olsen, S. Lucas, C. Elkin, E. Uberbacher, M. Frazier, R. A. Gibbs, D. M. Muzny, S. E. Scherer, J. B. Bouck, E. J. Sodergren, K. C. Worley, C. M. Rives, J. H. Gorrell, M. L. Metzker, S. L. Naylor, R. S. Kucherlapati, D. L. Nelson, G. M. Weinstock, Y. Sakaki, A. Fujiyama, M. Hattori, T. Yada, A. Toyoda, T. Itoh, C. Kawagoe, H. Watanabe, Y. Totoki, T. Taylor, J. Weissenbach, R. Heilig, W. Saurin, F. Artiguenave, P. Brottier, T. Bruls, E. Pelletier, C. Robert, P. Wincker, D. R. Smith, L. Doucette-Stamm, M. Rubenfield, K. Weinstock, H. M. Lee, J. Dubois, A. Rosenthal, M. Platzer, G. Nyakatura, S. Taudien, A. Rump, H. Yang, J. Yu, J. Wang, G. Huang, J. Gu, L. Hood, L. Rowen, A. Madan, S. Qin, R. W. Davis, N. A. Federspiel, A. P. Abola, M. J. Proctor, R. M. Myers, J. Schmutz, M. Dickson, J. Grimwood, D. R. Cox, M. V. Olson, R. Kaul, C. Raymond, N. Shimizu, K. Kawasaki, S. Minoshima, G. A. Evans, M. Athanasiou, R. Schultz, B. A. Roe, F. Chen, H. Pan, J. Ramser, H. Lehrach, R. Reinhardt, W. R. McCombie, M. de la Bastide, N. Dedhia, H. Blöcker, K. Hornischer, G. Nordsiek, R. Agarwala, L. Aravind, J. A. Bailey, A. Bateman, S. Batzoglou, E. Birney, P. Bork, D. G. Brown, C. B. Burge, L. Cerutti, H. C. Chen, D. Church, M. Clamp, R. R. Copley, T. Doerks, S. R. Eddy, E. E. Eichler, T. S. Furey, J. Galagan, J. G. Gilbert, C. Harmon, Y. Hayashizaki, D. Haussler, H. Hermjakob, K. Hokamp, W. Jang, L. S. Johnson, T. A. Jones, S. Kasif, A. Kaspryzk, S. Kennedy, W. J. Kent, P. Kitts, E. V. Koonin, I. Korf, D. Kulp, D. Lancet, T. M. Lowe, A. McLysaght, T. Mikkelsen, J. V. Moran, N. Mulder, V. J. Pollara, C. P. Ponting, G. Schuler, J. Schultz, G. Slater, A. F. Smit, E. Stupka, J. Szustakowski, D. Thierry-Mieg, J. Thierry-Mieg, L. Wagner, J. Wallis, R. Wheeler, A. Williams, Y. I. Wolf, K. H. Wolfe, S. P. Yang, R. F. Yeh, F. Collins, M. S. Guyer, J. Peterson, A. Felsenfeld, K. A. Wetterstrand, A. Patrinos, M. J. Morgan, P. de Jong, J. J. Catanese, K. Osoegawa, H. Shizuya, S. Choi, Y. J. Chen; International Human Genome Sequencing Consortium, Initial sequencing and analysis of the human genome. *Nature* **409**, 860–921 (2001).
50. S. Chatrchyan, V. Khachatryan, A. M. Sirunyan, A. Tumasyan, W. Adam, E. Aguilo, T. Bergauer, M. Dragicevic, J. Erö, C. Fabjan, M. Friedl, R. Frühwirth, V. M. Ghete, J. Hammer, M. Hoch, N. Hörmann, J. Hrubec, M. Jeitler, W. Kiesenhofer, V. Knünz, M. Krammer, I. Krätschmer, D. Liko, W. Majerotto, I. Mikulec, M. Pernicka, B. Rahbaran, C. Rohringer, H. Rohringer, R. Schöfbeck, J. Strauss, Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC. *Phys. Lett. B* **716**, 30–61 (2012).
51. ATLAS Collaboration, Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC. *Phys. Lett. B* **716**, 1–29 (2012).
52. M. B. Yaffe, The scientific drunk and the lamppost: Massive sequencing efforts in cancer discovery and treatment. *Sci. Signal.* **6**, pe13 (2013).

Funding: There were no specific sources of funding. **Author contributions:** D.H. designed the study, collected and analyzed the data, and wrote the manuscript. M.E. and A.H. participated in data collection and drafting of the manuscript. **Competing interests:** The authors declare that they have no competing interests. **Data and materials availability:** Our access to the WOS comes through a contract with Thomson Reuters that forbids redistribution of their database; researchers who desire the raw data on which to run our analytics can obtain it via a paid subscription to Thomson Reuters.

Submitted 14 February 2015
Accepted 26 July 2015
Published 18 September 2015
10.1126/sciadv.1500211

Citation: D. Hsiehchen, M. Espinoza, A. Hsieh, Multinational teams and diseconomies of scale in collaborative research. *Sci. Adv.* **1**, e1500211 (2015).

Science Advances

Multinational teams and diseconomies of scale in collaborative research

David Hsiehchen, Magdalena Espinoza and Antony Hsieh

Sci Adv **1** (8), e1500211.
DOI: 10.1126/sciadv.1500211