Science AAAS

Multi-University Research Teams: Shifting Impact, Geography, and Stratification in Science

Benjamin F. Jones, *et al.*

Science **322**, 1259 (2008);
DOI: 10.1126/science.1158357

15. B. P. Lazzaro, *Curr. Opin. Microbiol.* **11**, 284 (2008).
16. Y. Moret, M. T. Siva-Jothy, *Proc. R. Soc. Lond. B. Biol. Sci.* **270**, 2475 (2003).
17. B. M. Sadd, P. Schmid-Hempel, *Curr. Biol.* **16**, 1206 (2006).
18. L. N. Pham, M. S. Dionne, M. Shirasu-Hiza, D. S. Schneider, *PLoS Pathog.* **3**, e26 (2007).
19. A. J. Needham, M. Kibart, H. Crossley, P. W. Ingham, S. J. Foster, *Microbiology* **150**, 2347 (2004).
20. J. Garcia-Lara, A. J. Needham, S. J. Foster, *FEMS Immunol. Med. Microbiol.* **43**, 311 (2005).
21. N. Matova, K. V. Anderson, *Proc. Natl. Acad. Sci. U.S.A.* **103**, 16424 (2006).
22. P. E. Dunn, *Annu. Rev. Entomol.* **31**, 321 (1986).
23. G. G. Perron, M. Zasloff, G. Bell, *Proc. R. Soc. Lond. B. Biol. Sci.* **273**, 251 (2006).
24. We thank R. Clare, A. Dobson, S. Webster, and A. Williams for laboratory assistance. P. Brakefield, S. Foster, and M. Osborn made valuable comments that improved the manuscript. This research was funded by The Leverhulme Trust, and E.R.H. was also supported by The Wellcome Trust.

Supporting Online Material
www.sciencemag.org/cgi/content/full/322/5905/1257/DC1
Materials and Methods
Fig. S1
References

28 August 2008; accepted 17 October 2008
10.1126/science.1165265

Multi-University Research Teams: Shifting Impact, Geography, and Stratification in Science

Benjamin F. Jones,[1,2]* Stefan Wuchty,[3]*† Brian Uzzi[1,3,4]*‡

This paper demonstrates that teamwork in science increasingly spans university boundaries, a dramatic shift in knowledge production that generalizes across virtually all fields of science, engineering, and social science. Moreover, elite universities play a dominant role in this shift. By examining 4.2 million papers published over three decades, we found that multi-university collaborations (i) are the fastest growing type of authorship structure, (ii) produce the highest-impact papers when they include a top-tier university, and (iii) are increasingly stratified by in-group university rank. Despite the rising frequency of research that crosses university boundaries, the intensification of social stratification in multi-university collaborations suggests a concentration of the production of scientific knowledge in fewer rather than more centers of high-impact science.

In May 1845, Samuel F. B. Morse telegraphed the first electronic message, "What hath God wrought?" from Washington, DC, to Baltimore and declared an end to "the tyranny of distance." Yet, in the 150 years since Morse's breakthrough, the production of science has had a reputation for geographic localization. Early 20th-century German universities singularly led in chemistry and physics, creating the first commercial dyes and nuclear and rocket programs (*1*). Silicon Valley became a renowned incubator for excellence in technology (*2*), and the University of Chicago has been a persistent center for Nobel Prize winners in economics.

Nonetheless, recent observers have suggested a weakening link between location and scientific research—a "death of distance," in popular coinage (*3*). In this view, technology is inevitably removing the last barriers of distance, widening access to geographically distant collaborators with potential implications for the location, research quality, and social stratification of science (*4*, *5*). Researchers report, for example, a modest rise in collaborations between research institutions in a limited sample of fields (*6*–*8*), although other authors conclude that face-to-face contact and the "30-feet collaboration rule" nonetheless continue to encourage collaboration at home universities (*9*–*11*).

Although these studies suggest a possible trend toward collaboration between authors at different universities, the generalizability of the shift remains unclear, as does the association with research impact and the geographic and social structure of collaboration. With an eye toward understanding the role that multi-university collaborations play in the production of science, we examined (i) trends in multi-university collaborations across the full spectrum of academic fields, (ii) the association of these changes with research impact, and (iii) the role of elite and nonelite universities in these trends. Our sample focuses on a set of 662 major U.S. universities and includes 4.2 million research papers in the Web of Science database, covering 172 fields of science and engineering (SE), 54 fields of social sciences (SS), and 27 fields of arts and humanities (AH) from 1975 to 2005 [the supporting online material (SOM) further defines the sample design, and table S1 lists the universities].

Our analysis indicates a remarkable and nearly universal rise since 1975 in the frequency of collaborations between authors located at different universities. Figure 1 shows the share of research papers published by solo authors, collaborators at the same university, and collaborators between schools. In SE, between-school collaborations were rare in 1975, which suggests that this type of collaboration is a relatively modern phenomenon (Fig. 1A). Moreover, among all three authorship arrangements, between-school collaboration is the fastest- and only steadily growing segment, quadrupling its share in SE between 1975 and 2005 to 32.8%. Figure 1B shows that SS have experienced similar trends; the share of SS papers written in multi-university collaborations rose even more rapidly over the 30-year period to a peak share of 34.4%. These upward trends appear even stronger upon further including publications with collaborators outside the sample of major U.S. universities (fig. S1).

Although the rate of increase in multi-university collaborations is essentially constant over our entire period, we found an unusual jump of 3.4% in SE, which is large ($P < 0.0001$) compared to the typical year-on-year increase of 0.8%. Although this single event occurred in 1998, a period when the Internet and other information and computing technologies were spreading widely, multi-university collaborations in SE continued at the pre-1998 rate of growth thereafter. SS show little unusual acceleration in this period. These findings suggest that, although communication technologies may be generally important, multi-university collaborations were largely driven by factors that predated recent communication technologies.

Table 1 demonstrates the generality of these patterns across individual fields. It indicates that 98% (168 of 172) of the SE subfields increased their share of between-school collaborations. In SS, 100% (54 of 54) of the subfields increased their share of research written by coauthors at multiple schools. AH showed little disposition for teamwork in general (Fig. 1C); a more modest majority of fields (67%, 18 of 27) demonstrated a rise in multi-school collaboration. Our analyses further show that the trend toward multi-university collaboration in SE and SS is present for teams of any size (fig. S2) (*12*).

Although the incidence of between-school collaboration has grown rapidly, the average distance between collaborators has risen only slightly (fig. S3). In 1975, the mean distance between collaborators in SE was 750 miles, whereas it was 800 miles in 2005 (*13*). The mean distance between collaborators in SS increased from 725 to 800 miles over the same period. Median distances rose from 510 to 560 miles (SE) and 530 to 580 miles (SS). Spatially, this suggests that the death of distance in U.S. science is not primarily

[1]Kellogg School of Management, Northwestern University, Evanston, IL 60208, USA. [2]National Bureau of Economic Research, Cambridge, MA 02138, USA. [3]Northwestern Institute on Complexity (NICO), Northwestern University, Evanston, IL 60208, USA. [4]Haas School of Business, University of California at Berkeley, Berkeley, CA 94720, USA.

*These authors contributed equally to this work.
†Present address: National Institutes of Health, Bethesda, MD 20982, USA.
‡To whom correspondence should be addressed. E-mail: uzzi@northwestern.edu

driven by an increase in the distance between long-range collaborators but by an increase in the frequency of collaborations both near and far. It is not the length of a scientist's reach that has changed but rather the incidence of reaching across university boundaries.

To analyze and compare the citation impact of between-school collaborations and within-school collaborations, we estimated the probability that papers written with these two different authorship structures receive above-average citations. We focused on papers published from 1995 to 2005 and on between-school collaborations with two university affiliations, using regression analysis to account for subfield, team size, and publication year differences (*13*).

Figure 2 presents three primary findings. In both SE and SS, between-school collaborations have a citation-impact advantage over within-school collaborations. Averaging over all schools in our sample (Fig. 2A), within-school collaborations have a baseline probability of producing papers with above-average citations of 32.7% in SE and 34.1% in SS. For between-school collaborations, the marginal probability is 2.9% higher in SE ($P < 0.0001$) or 8.8% above the baseline rate (0.029/0.327). In SS, the marginal gains of between-school collaborations are larger at 5.8% ($P < 0.0001$) or 16.7% above the baseline rate (0.058/0.341).

In Fig. 2B, we further disaggregated the between-school collaborations based on the ranks of the collaborating schools. Schools were ranked by the total number of citations received by the papers published at the school in the corresponding period. We considered only within-school publications in this ranking method so that the school rankings are independent of each university's performance in between-school collaborations. Tier I schools are those ranked in the top 5%, tier II are in the top 6 to 10%, tier III are in the top 11 to 20%, and tier IV are the remainder. Other ranking methods were considered below as robustness checks (*14*). We took within-school collaborations as the baseline case for each tier and compared the citation impact of multi-university collaborations with that baseline. These results suggest that collaborations between tier I schools show a substantial impact advantage over tier I within-school collaborations. For example, a team of two authors from an elite school such as Harvard tends to produce lower-impact papers than a team of two authors with one at Harvard and the other at Stanford. This finding may challenge claims that within-school collaboration provides decisive advantages by lowering communication and monitoring costs, conveying tacit information, and enabling impromptu interaction. Specifically, in SE, circa 1995 to 2005, a collaboration between tier I institutions is 6.19% more likely ($P < 0.0001$) to be of high impact than a tier I within-school collaboration, which is a 17% increase in the baseline rate (0.062/0.372). In SS fields, there is an even larger 11.7 percentage-point increase ($P < 0.0001$) in the probability of being a high-impact paper, which is a 24% increase in the baseline rate (0.117/0.406). We also observe that the advantage of intra-tier partnerships exists for tier II and tier III partnerships and declines as the rank of schools decreases. For partnerships between tier IV schools, the advantage disappears in SS and is slightly negative for SE fields, which suggests that the advantage found in partnerships between similar-quality schools is dominated by, and most pronounced among, elite institutions.

Table S2 considers cross-tier collaborations and shows that their citation impact tends to lie between the impact of the upper and lower tiers' within-school collaborations. The between-school impact tends to lie closer to the upper-tier baseline than the lower-tier baseline. This finding suggests that cross-tier between-school collaboration may follow a "strongest-partner" rather than a "weakest-partner" model, with tier I pulling up tier IV more than tier IV pulls down tier I. Tables S3 and S4 substantiate all the results for alternative school-ranking methods (*14*), and fig. S5 presents the raw citation impact data, unconditional on field or team size, and shows similar findings.

Further exploring the special role of elite universities in the production of multi-school research, Table 2 decomposes the intensity of multi-university collaboration by university. Considering papers with two university affiliations, the table cells above the diagonal show the actual share of multi-university collaborations for 2001 to 2005 for each tier pairing, which indicates that elite schools dominate multi-university collaborations in the 2001–2005 period. Just 5% of schools (tier I) hold places in 59.7% of multi-university collaborations in SE (tier I row sum, above the diagonal). In SS, the corresponding fraction is 56.2%. Conversely, only 18.4% of SE multi-university collaborations were formed solely among authors from tier III and IV universities, even though these universities are 90% of the schools in the sample.

We further examined the intensity of stratification in multi-university partnerships by comparing actual collaboration rates between tier pairings with a baseline model of random matching across tiers. Defining P_j as the fraction of multi-university papers that include tier j, the probability that a multi-university paper includes tiers j and k under random matching is $2P_jP_k$ if $j \neq k$ and P_jP_k if $j = k$. The cells of Table 2, below the diagonals, show the ratio of the actual frequency of a given tier-pairing to the expected frequency under random matching. Ratios >1 indicate a greater propensity for tier pairings than expected by random matching and ratios <1 indicate a lesser propensity for tier pairings.

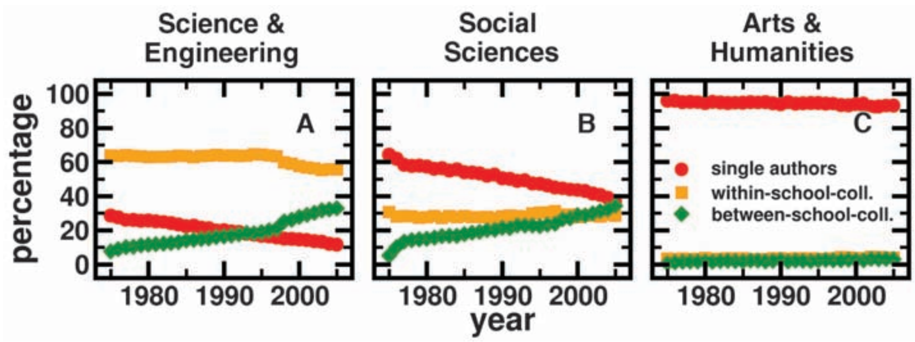

Fig. 1. The rise in multi-university collaboration. By comparing the incidence of papers produced by different authorship structures, we see that the share of multi-university collaborations strongly increases from 1975 to 2005. This rise is especially strong in SE (**A**) and SS (**B**), whereas it appears weakly in AH (**C**), in which collaboration of any kind is rare. The share of single-university collaborations remains roughly constant with time, whereas the share of solo-authored papers strongly declines in SE and SS.

Table 1. Patterns by subfield. For each subfield, we calculated the fractions of papers written as within-school collaborations and between-school collaborations in 1975–1979 and 2001–2005. Comparing the incidence of these authorship structures between the early and late periods, we determined the number of subfields (*N*) and the fraction of subfields (**%**) for which the share of each type of collaboration is increasing.

		Increasing collaborations			
		Within-school		Between schools	
Fields	Total subfields (*N*)	Subfields (*N*)	Subfields (%)	Subfields (*N*)	Subfields (%)
SE	172	114	66.3	168	97.7
SS	54	47	87.0	54	100.0
AH	27	16	59.3	18	66.7

We found that tier I–tier I collaborations in SE were 14% more common ($P < 0.001$) than expected under random matching, whereas tier I–tier IV collaborations were 19% less common ($P < 0.001$) than expected. In SS, this inequality is even stronger, for which tier I–tier I collaborations were 27% more common than expected ($P < 0.001$), and tier I–tier IV collaborations were 32% less common ($P < 0.001$). Meanwhile, both SE and SS show that collaborations limited to lower-tier schools were substantially more common than expected. Lower-tier schools did reach across university boundaries, but they tended to interact within their own tier, echoing the in-group status-matching behavior of elite schools. However, recalling the results in Fig. 2, the tendency to favor in-group matching was an advantage for tier I schools but a disadvantage for those of tier IV. Other ranking methods (*14*) demonstrated the robustness of these stratification tendencies (tables S6 and S7).

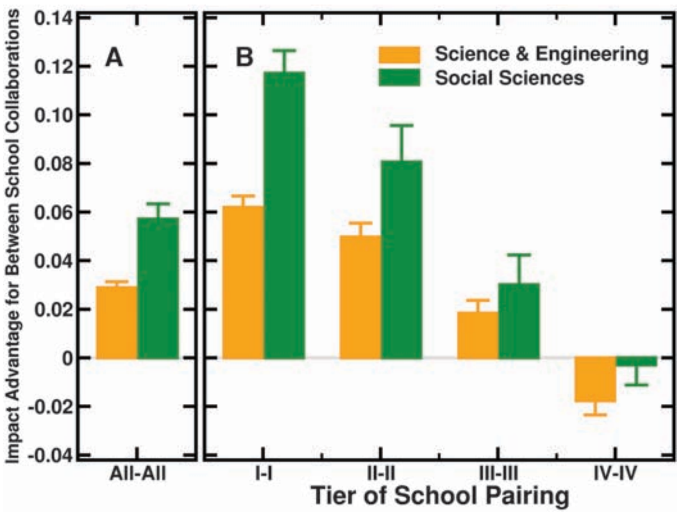

Fig. 2. The impact advantage of between-school collaborations. The citation impact of between- versus within-school collaborations is compared. Impact is measured as the probability that a paper receives above-average citations. For within-school publications, the baseline probabilities of a high-impact paper are 0.327, 0.372, 0.327, 0.295, and 0.235 in SE and 0.341, 0.406, 0.366, 0.303, and 0.229 in SS for all schools, tier I, tier II, tier III, and tier IV, respectively. The vertical axis is the probability that a between-school collaboration is high-impact minus the probability that a within-school collaboration is high-impact, while using regression to net out differences by subfield, team size, and year. (**A**) The marginal advantages when all schools are pooled. (**B**) The marginal advantage for schools in the same tier. Each bar represents a separate regression; all marginal advantages are significant ($P < 0.001$), except for tier IV–tier IV pairings in SS, with 95% confidence intervals represented by T bars. Data cover the 1995 to 2005 period.

Table 2. Who collaborates with whom? For teams that span two different institutions, we calculated the frequency of pairings by school ranks for the period 2001–2005. For each school tier pairing, we show the actual percentage of two-school collaborations above the diagonal. Below the diagonal, we show the ratio of the actual frequency to the expected frequency under random matching. A ratio greater than 1 indicates that such tier pairings are more common than expected, and a ratio less than 1 indicates that such tier pairings are less common than expected. All quantities are statistically significant with $P < 0.001$ unless marked by a †. Statistical significance was determined by bootstrapping.

Science and Engineering

Tier	I	II	III	IV
I	16.6% / 1.14	17.8% / 0.98	15.5% / 0.91	9.8% / 0.81
II		4.9% / 0.90	9.9% / 0.94	7.1% / 0.97
III			5.6% / 1.03	8.4% / 1.19
IV				4.4% / 1.89

Social Sciences

Tier	I	II	III	IV
I	16.2% / 1.27	15.0% / 1.10	13.7% / 0.86	11.2% / 0.68
II		4.1% / 0.97†	9.2% / 1.05	7.9% / 0.91
III			5.1% / 0.95	10.9% / 1.14
IV				6.7% / 1.35

P / R

P: Frequency of tier combination (in %)
R: Ratio of frequency to expected frequency

The SOM further considers the evolution of social and geographic stratification over time. Table S5 demonstrates that tier I–tier I collaborations rose with time as compared with the expected rate, whereas tier I–tier IV collaborations fell. Thus elite schools have been increasingly likely to capture the impact advantages of their collaborations, whereas tier IV schools appear to lose ground in accessing top-tier collaborations. Finally, figs. S6 to S9 illustrate the intersection of social and geographic distance for the Boston, Chicago, Los Angeles, New York, North Carolina, and San Francisco research centers. Top schools in these regions increasingly collaborated with geographically distant top schools, whereas their other partnerships declined. Partnership choice increasingly appears to be based on who the collaborators are rather than where they are, with an emphasis on in-group status matching.

The shift toward cross-institutional collaboration suggests a fundamental transformation in the production of high-impact science. Prior work established that teamwork has become ubiquitous across scientific fields and supplanted solo-authored work in the production of the highest-impact research (*15*, *16*). We demonstrate that the rising collaboration in science is increasingly composed of collaborations that span university boundaries. Moreover, elite universities play a dominant role in this shift. We found that multi-university partnerships (i) are the fastest growing type of authorship structure, (ii) produce the highest-impact papers when they include a top-tier university, and (iii) are increasingly stratified by in-group university rank. Thus, although geographic distance is of decreasing importance, social distance is of increasing importance in research collaborations. Elite universities are more intensely interdependent, playing a higher-impact and increasingly visible role in SE and SS.

Consistent with some claims, we observed that rapid technology advances in the 1990s may have modestly accelerated between-school collaboration, but our results more generally indicate a smooth growth in multi-school collaboration that substantially predated this period. These findings suggest a more complex relationship between social relations and technology in science, with collaboration technologies and social networks potentially being an endogenous outcome of the burgeoning interest in research partnerships (*8*, *17*–*19*). Others have argued that increasing specialization drives collaboration (*20*), which, coupled with the limited capacity of academic departments to encapsulate more than a fraction of a field's specializations, may promote institutionally boundaryless collaborations in the search for high-impact science. However, whereas the greater geographic interconnectedness of universities would appear to make geography less important, the corresponding intensification of social stratification in multi-university collaboration tends to embed the production of outstanding scientific knowledge in fewer rather than more centers of high-impact science. The dominant role of elite universities suggests several ideas for future research, including scale and resource advantages, social networks, journal leadership, and other factors such as matching based on status or quality, that promote a widening rather than a narrowing of stratification in science through the vehicle of multi-university partnerships.

References and Notes

1. J. P. Murmann, *Knowledge and Competitive Advantage* (Cambridge Univ. Press, Cambridge, 2003).
2. A. L. Saxenian, *Regional Advantage: Culture and Competition in Silicon Valley and Route 128* (Harvard Univ. Press, Cambridge, MA, 1994).
3. F. Cairncross, *The Death of Distance* (Harvard Univ. Press, Cambridge, MA, 1997).
4. S. Teasley, S. Wolinsky, *Science* **292**, 2254 (2001).
5. A. Finholt, M. Gary, Olson. *Psychol. Sci.* **8**, 28 (1997).
6. F. Havemann, M. Heinz, H. Kretschmer, *J. Biomed. Discov. Collab.* **1**, 6 (2006).
7. T. S. Rosenblat, M. M. Mobius, *Q. J. Econ.* **119**, 971 (2004).
8. A. Agrawal, A. Goldfarb, *Am. Econ. Rev.* **98**, 1578 (2008).
9. G. M. Olson, J. S. Olson, *Hum. Comp. Interact.* **15**, 139 (2000).
10. E. Bradner, G. Mark, in *Proceedings of the 2002 ACM Conference on Computer Supported Cooperative Work* (ACM, New York, 2002), pp. 226–235.
11. J. Cummings, S. Keisler, *Res. Policy*, in press.
12. We also found (fig. S2) that social scientists have a substantially higher propensity to engage in multi-university collaboration for any given team size. For example, a team of size 2 in SS operated across universities nearly 40% of the time in the year 2000, whereas the corresponding propensity in SE was 18%, a 2-to-1 ratio that holds approximately over time. This difference may follow from conditions such as the relative capital intensity of lab sciences, which particularly favor colocation in SE. The relative tendency toward multi-university collaboration in SS is offset by the greater incidence of large teams in SE and the fact that larger teams are more likely to be assembled from multiple universities. Hence, the absolute fraction of multi-university work turns out to be similar in both SE and SS (Fig. 1).
13. We focused on cases with two university affiliations, which allowed us to categorize multi-university collaborations into explicit tier pairings, as seen in Fig. 2. Two-university collaborations represent over 90% of multi-university collaborations (fig. S4). The regressions are linear probability models (ordinary least squares), in which an indicator for whether the paper is above the mean impact in its field and year is regressed on indicator variables for authorship structure (solo-authored, within-school collaboration, or between-school collaboration), publication year, team size, and subfield, with standard errors clustered by subfield.
14. To determine the ranking of each school, we pooled all papers that have been published entirely by a given school in each year by using only within-school collaborations and papers of single authors. We then counted the number of citations those papers received and ranked the schools according to their total citation count. This ranking method is equivalently described as the average citation impact (a quality measure) times the number of publications (a quantity measure). Between-school papers were omitted from the measure of school ranking to ensure that the papers used to assign tiers were independent of the papers used to assess between-school impact. Alternative measures are average citation impact, as a pure quality measure, and the university Hirsch index (*21*), which emphasizes very-high-impact papers. A university's Hirsch index for a given publication period is the largest number *H* such that *H* of its papers received at least *H* citations. As noted in the text and detailed in tables S2, S4, S6, and S7, the patterns of impact and stratification are consistent across these alternative ranking approaches. Table S8 lists the top 50 universities under different ranking methods.
15. S. Wuchty, B. Jones, B. Uzzi, *Science* **316**, 1036 (2007).
16. S. Wuchty, B. Jones, B. Uzzi, *Science* **317**, 1496 (2007).
17. V. Gurbaxani, *Commun. ACM* **33**, 65 (1990).
18. R. K. Merton, *On Social Structure and Science* (Univ. Chicago Press, Chicago, IL, 1996).
19. R. Guimera, B. Uzzi, J. Spiro, L. A. N. Amaral, *Science* **308**, 697 (2005).
20. B. F. Jones, *Rev. Econ. Studies*, in press.
21. J. E. Hirsch, *Proc. Natl. Acad. Sci. U.S.A.* **102**, 16569 (2005).

22. We thank K. Murnighan, K. Williams Phillips, S. Stern, S. S. Sanchez, and two anonymous referees for helpful comments. The Northwestern Institute on Complex Systems provided financial support.

Supporting Online Material

www.sciencemag.org/cgi/content/full/1158357/DC1
Materials and Methods
Figs. S1 to S9
Tables S1 to S8

27 March 2008; accepted 25 September 2008
Published online 9 October 2008;
10.1126/science.1158357
Include this information when citing this paper.