

Chapter 12

How Atypical Combinations of Scientific Ideas Are Related to Impact: The General Case and the Case of the Field of Geography

Satyam Mukherjee, Brian Uzzi, Benjamin F. Jones, and Michael Stringer

Scientific enterprises are increasingly concerned that research within narrow boundaries is unlikely to be the source of the most fruitful ideas (National Academy of Sciences, National Academy of Engineering, & Institute of Medicine of the National Academies, 2004). Models of creativity emphasize that innovation is spurred by original combinations that spark new insights (Becker, 1982; Guimera, Uzzi, Spiro, & L.A. Amaral, 2005; Jones, 2009; Jones, Wuchty, & Uzzi, 2008; Schilling, 2005; Schumpeter, 1939; Usher, 1929/1998; Uzzi & Spiro, 2005; Weitzman, 1998). Current interest in team science and how scientists search for ideas is premised in part on the idea that teams can span scientific specialties, effectively combining knowledge that prompts scientific breakthroughs (Evans & Foster, 2011; Falk-Krzesinski et al., 2010; Fiore, 2008; Stokols, Hall, Taylor, & Moser, 2008; Wuchty, Jones, & Uzzi, 2007).

The production and consumption of boundary-spanning ideas can also raise well-known challenges (Azoulay, Zivin, & Manso, 2011; Collins, 1998; Einstein, 1949; Fleming, 2001; Henderson & Clark, 1990; Schilling & Green, 2011). If, as

S. Mukherjee

Indian Institute of Management, Udaipur, Rajasthan, India

e-mail: satyam.mukherjee@gmail.com

B. Uzzi (✉)

Kellogg School of Management, Northwestern University,

2001 Sheridan Road, 60208 Evanston, IL, USA

e-mail: uzzi@northwestern.edu

B.F. Jones

Department of Management and Strategy, Kellogg School of Management, Northwestern University, 2001 Sheridan Road, Jacobs Center 371, 60208-2001 Evanston/Chicago, IL, USA

e-mail: bjones@kellogg.northwestern.edu

M. Stringer

Northwestern Institute on Complex Systems, Northwestern University,

600 Foster, 60208 Evanston, IL, USA

e-mail: mike.stringer@datascopeanalytics.com

Einstein (1949) believed, individual scientists inevitably become narrower in their expertise as the body of scientific knowledge expands, then reaching effectively across boundaries may be increasingly challenging (Jones, 2009), especially given the difficulty of searching unfamiliar domains (Fleming, 2001; Schilling & Green, 2011). Moreover, novel ideas can be difficult to absorb (Henderson & Clark, 1990) and communicate, leading scientists to intentionally display conventionality. In his *Principia*, Newton presented his laws of gravitation using accepted geometry rather than his newly developed calculus, despite the latter's importance in developing his insights (Whiteside, 1970). Similarly, Darwin devoted the first part of *On the Origin of Species* to conventional, well-accepted knowledge of the selective breeding of dogs, cattle, and birds. Given these tendencies, the balance between extending science with atypical combinations of knowledge while maintaining advantages of conventional domain-level thinking is critical to the link between innovativeness and impact. However, little is known about the composition of this balance or how scientists can achieve it. In this paper, our analysis of 17.9 million papers spanning all scientific fields suggests that science follows a nearly universal pattern: The highest-impact science is primarily grounded in exceptionally conventional combinations of prior work yet simultaneously features an intrusion of unusual combinations. Papers of this type were twice as likely to be highly cited works. Notably, novel combinations of prior work are rare, yet teams are 37.7% more likely than solo authors to insert novel combinations into familiar knowledge domains.

Data and Methods

Data

We examined 17.9 million scientific publications across 15,613 journals, constituting all research articles indexed in the Thomson Reuters Web of Science (WOS) database that was published between 1950 and 2000. According to each journal's subject area, the Institute for Scientific Research (ISI, a.k.a. Web of Science) currently defines three fields and constituent subfields: science and engineering (171 subfields), social sciences (54 subfields), and arts and humanities (27 subfields) with coverage for research publications in science and engineering since 1945, social sciences since 1956, and arts and humanities since 1975. For each paper, the WOS records the citations, number of authors, and citation links to other papers in the database.

Methods

We measured the relative conventionality and novelty of the prior work that a paper combines by examining the papers referenced in a paper's bibliography (Small, 1973; Stringer, Sales-Pardo, & Amaral, 2010). This section first provides an overview of our methodology, followed by an illustrative example and further details.

Overview

Our basic measurement question is to assess how common or novel any pairwise combination of prior work is. To determine this, we want to know both the (i) observed frequency of any given pairing of references in the WOS and (ii) the frequency of that pairing that would have occurred by chance. Comparing the observed frequency to the frequency expected by chance creates a normalized z -score measure for whether any given pairing appears novel or conventional.

To measure the *observed frequency* of any given pairing in the WOS, we took the following five steps:

- (1) Took the references listed in a given paper's bibliography.
- (2) Considered all pairwise combinations of the papers referenced in the bibliography of the paper.
- (3) For each pairwise combination, recorded the two journals that were paired.
- (4) Repeated steps (1–3) for every paper in the WOS.
- (5) Counted the aggregate, population-wide frequency of each journal pairing for all referenced pairs from a given publication year.

Figure 12.1 presents a stylized example for steps (1–3), showing for a given paper how pairs of references are counted from that paper's reference list. The algorithm repeats this counting process for every article in the WOS and aggregates the counts for each given publication year.

Our method counts specific journal pairings, using different journals as a proxy for different areas of knowledge. Journal-level analysis is well positioned to distinguish domains of knowledge while having precedence in the literature for being relatively transparent, interpretable, and computationally feasible (Bollen et al., 2006; Izkovitz et al., 2003; Small, 1973; Stringer et al., 2010).¹

Having determined the observed frequency of each journal pairing, we considered the frequency distribution for each journal pairing that would have occurred by chance. The null model randomly reassigns the citation links between papers. As further detailed below, the method uses a variation of the Markov Chain Monte Carlo (MCMC) algorithm to randomly switch co-citations between all 17.9 million papers into a synthetic network with 302 million citations (edges), the same number of papers and citations as the observed network. Note that this method preserved the detailed paper-level structure of the global citation network. The number of citations to and from each paper was preserved backward and forward in time.

Using this approach, we created 10 synthetic instances of the entire WOS, each with its own set of randomized citation links. For each instance of the WOS, we then repeated steps (1–5) above, calculating the frequency of each co-referenced journal pair. Looking at all 10 randomized cases of the WOS, we generated a distri-

¹Other operationalizations might consider lower resolution pairings using the ISI's 252 subfield categories, text-based combinations, or conceptualizations for measuring novelty beyond combinatorial pairs (Rosenkopf & McGrath, 2011).

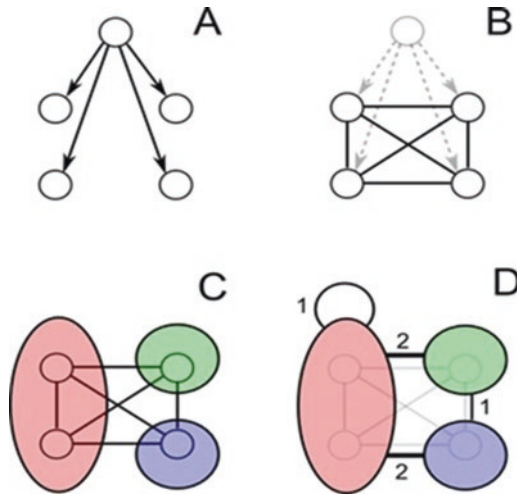


Fig. 12.1 Paper pairs and journal pairs. This figure presents a stylized example of how paper pairs and journal pairs are drawn from the network structure of citations. In *panel A*, the *circular nodes* represent papers; the directed links exist when the top paper cites the bottom four papers. In *panel B*, the *circular nodes* represent papers, and the undirected co-citation links between papers are shown in *black*. A co-citation exists between each pair of papers that occurs in the reference list of the focal paper. Here, there are 4 references and therefore 6 (i.e., 4 choose 2) co-citation links. In *panel C*, paper nodes are grouped by journal; the shaded ovals represent the three journals in which each of the cited papers is published. Finally, in *panel D*, the co-citation links between papers are mapped to the journal level, and the black links represent journal co-citations. Note that the total number of paper-to-paper co-citation links (6) is preserved at the journal co-citation level (From Uzzi et al. (2013b, p. 10). Copyright 2013 by *Science*. Adapted with permission from the authors and *Science*)

bution of frequencies for each journal pair. We could then evaluate the z -score for each observed journal pair relative to what was expected by chance:

$$z = (obs - exp) / \sigma$$

Where *obs* is the observed frequency of the journal pair in the actual WOS while *exp* is the mean and σ is the standard deviation of the number of journal pairs obtained from the 10 randomized simulations of the paper-to-paper citation network.

Finally, returning to categorizing a paper's prior work regarding novelty and conventionality, we could then assign a z -score to each of the journal pairs in that paper's reference list. Each paper thus had a distribution of journal pairings, where any given pairing could be more or less common compared to chance. To summarize the information in this distribution, we took two primary summary statistics:

- (i) The median z -score for that paper.
- (ii) The 10th percentile z -score for that paper.

The first measure is a summary statistic for the central tendency of the combinations of journals that a paper cites. The larger the median z -score for a paper, the more common the main mass of journal combinations in that paper compared to chance. The second measure is a summary statistic for the left tail of combinations of journals that a paper cites—journal pairings that are relatively unusual, compared to chance, among the set of journal pairings in that paper's reference list.

Illustrative Example of Methodology and Further Detail

To illustrate these procedures, consider the following example, based on a single paper in the field of geography.

1. Step 1. Take the references in a bibliography in a given paper. Consider the paper, "The Tropical Cyclone Hazard Over the South China Sea 1970–1989: Annual Spatial and Temporal Characteristics," which was published in *Applied Geography* in 1995. This paper has 22 references, of which 10 are known references (Fig. 12.2).
2. Step 2. Consider all pairwise combinations of the papers referenced in the bibliography of that paper. As can be seen in Fig. 12.2, pairwise paper combinations include, for example, (i) Deser et al. 1992 with Black 1990, (ii) Deser et al. 1992 with Thompson 1987, and (iii) Thompson 1987 with Black 1990. With 10 known references, we have 45 (i.e., 11 choose 2) pairwise paper combinations.
3. Step 3. Map the observed paper pairs into observed journal pairs. The 45 paper pairs are mapped into 45 journal pairs, where some journal pairs in this list appear multiple times. For example, *Nature* and *Monthly Weather Review* are paired twice.
4. Step 4. Repeat steps (1–3) for every paper in the WOS. The above steps, shown in a single article, are now repeated for every paper in the WOS. References to materials outside the WOS (for example, books) are not included.
5. Step 5. Count the frequency of each observed journal pairing for a given publication year, using the referenced works of every paper published that year in the WOS. Information from the sample paper above would be counted as part of the year 1995. Hence, we allow journal pair frequencies varying over time.

Having completed steps (1–5) for the observed papers in the WOS, we repeated them for each synthetic instance of the WOS, as created by the null model. Comparing the observed frequency of journal pairs under the real WOS with the frequency distribution that appears across instances of the null model, we computed a z -score for each journal pair. Continuing our illustrative example, the observed frequency, expected frequency, and z -score for several journal pairings that appear in the paper, "The Tropical Cyclone Hazard Over the South China Sea 1970–1989: Annual Spatial and Temporal Characteristics," are presented in Table 12.1. As Table 12.1 demonstrates (for a subsample of journal pairs), each published paper has a distribution of journal pairs, some of which are highly conventional (such as

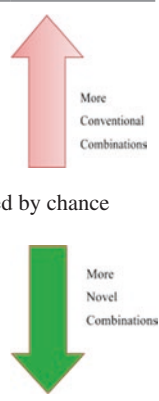
- Title: Long Term Variations in Western Tropical Pacific Cyclogenesis Associated With the Southern Oscillation
 Author(s): Harding, J.M.
 Source: *Nature*, 262, pp. 41–43, published 1976.
- Title: Network Autocorrelation in Transport Network and Flow Systems
 Author(s): Black, W. R.
 Source: *Geographical Analysis*, 24, pp. 207–222, published 1992.
- Title: Tropical Cyclone Activity in the Northwest Pacific in Relation to El Niño/Southern Oscillation Phenomenon
 Author(s): Chan, J. C. L.
Monthly Weather Review, 113, pp. 599–606, published 1985.
- Title: Large Scale Atmospheric Circulation Features of Warm and Cold Episodes in the Tropical Pacific
 Author(s): Deser, C., & Wallace, J. M.
 Source: *Journal of Climate*, 3, pp. 1254–1281, published 1990.
- Title: Rainfall at Tuvalu, Tokelau, and the Northern Cook Islands and Its Relationship to the Southern Oscillation
 Author(s): Thompson, C. S.
 Source: *New Zealand Journal of Geology and Geophysics*, 30, pp. 195–198, published 1987
- Title: Predictability of Interannual Variation of Australian Seasonal Tropical Cyclone Activity
 Author(s): Nicholas, N. N.
 Source: *Monthly Weather Review*, 113, pp. 1144–1149, published 1985.
- Title: El Niño and Tropical Cyclone Frequency in the Australian Region and Northwest Pacific
 Author(s): Dong, K.
 Source: *Australian Meteorological Magazine*, 36, pp. 219–225, published 1988.
- Title: El Niño/Southern Oscillation Modification to the Structure of the Monsoon and Tropical Cyclone Activity in the Australasian Region
 Author(s): Evans, J. L., & Allan, R. J.
 Source: *International Journal of Climatology*, 12, pp. 611–623, published 1992.
- Title: Temporal and Spatial Characteristics of Coastal Rainfall Anomalies in Papua New Guinea and Their Relationship to the Southern Oscillation
 Author(s): McGregor, G. R.
 Source: *International Journal of Climatology*, 12, pp. 449–468, published 1992.
- Title: El Niño: The Ocean–Atmospheric Connection
 Author(s): Rasmusson, E. M.
 Source: *Oceanus*, 27, pp. 5–12, published 1984.

Fig. 12.2 Reference list for example paper. The paper, “The Tropical Cyclone Hazard Over the South China Sea 1970–1989: Annual Spatial and Temporal Characteristics,” cites 10 different known references. From Uzzi et al. (2013b, p. 11). Copyright 2013 by *Science*. Adapted with permission from the authors and *Science*

Monthly Weather Review–Monthly Weather Review) while others are unusual compared to chance (such as *Nature–Monthly Weather Review*). Fig. 12.4a presents the distribution of z -scores for this illustrative paper and indicates the median z -score and the 10th percentile z -score in that paper’s distribution.

Table 12.1 Examples of journal pair frequencies for illustrative paper

| Journal pairs | Observed | Expected | Z-score |
|--|----------|----------|----------|
| <i>Mon. Weather Rev.</i> – <i>Mon. Weather Rev.</i> | 28,685 | 68.999 | 3252.658 |
| <i>Int. Journ. of Climatology</i> – <i>Int. Journ. of Climatology</i> | 402 | 0.1 | 1339.667 |
| <i>Nature</i> – <i>Journ. of Climate</i> | 2720 | 1392.016 | 47.344 |
| <i>Geogr. Anal.</i> – <i>Journ. of Clim.</i> | 3 | 2.199 | 0.743 |
| A z-score of zero means the actual journal pair frequency is the same as expected by chance pairings | | | |
| <i>N.Z. Journ. Geol. & Geophys.</i> – <i>Journ. of Climate</i> | 4 | 4.0 | 0 |
| <i>Nature</i> – <i>Aust. Meteorol. Mag.</i> | 89 | 95.494 | –0.915 |
| <i>Geogr. Anal.</i> – <i>Mon. Weather Rev.</i> | 2 | 6.192 | –1.436 |
| <i>Nature</i> – <i>Geogr. Anal.</i> | 30 | 138.889 | –8.78 |
| <i>Nature</i> – <i>Mon. Weather Rev.</i> | 2226 | 3779.287 | –24.554 |



Note. *Mon. Weather Rev.* Monthly Weather Review, *Int. Journ. of Climatology* International Journal of Climatology, *Geogr. Anal.* Geographical Analysis, *N.Z. Journ. Geol. & Geophys.* New Zealand Journal of Geology and Geophysics, *Aust. Meteorol. Mag.* Austrian Meteorology Magazine. From Uzzi et al. (2013b, p. 21). Copyright 2013 by *Science*. Adapted with permission from the authors and *Science*

Table 12.1 further shows the importance of normalizing the observed frequencies. For example, compare the pairings (1) *Nature* and *Journal of Climate* and (2) *Nature* and *Monthly Weather Review*. Both have similarly observed co-citation frequencies in the WOS: 2720 and 2226, respectively. However, compared to chance, the first pairing appears to have high conventionality while the second pairing seems to have high novelty.

Null Model Detail

The null model creates random synthetic instances of the WOS while incorporating realistic aspects of the data and its network structure. In particular, the null model incorporates two basic empirical facts about citation patterns:

- Citation distributions are skewed. Some papers and journals are cited far more often than other papers and journals and consequently are referenced more frequently in bibliographies.
- Citation counts are dynamic processes that vary by the journal (Stringer, Sales-Pardo, & Amaral, 2008), so that the rate at which papers accumulate citations is journal dependent.

Keeping these facts in mind, the null model preserves for each paper in the WOS the same number of references to past work, the same number of citations from

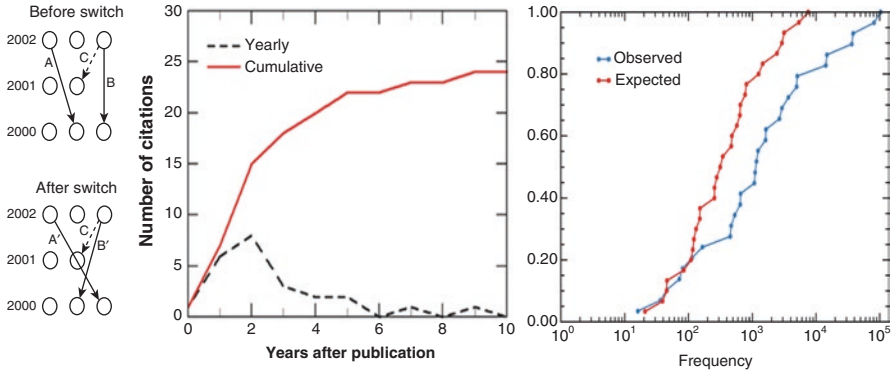


Fig. 12.3 Link switching in the null model and example distributions of observed and expected frequency of journal pairs. Citation links between papers are switched randomly but constrained to have the same origin year and target year. Thus in the *left panel*, switching links *A* and *B* are allowed, while switching links *A* and *C* are not allowed. The switching algorithm thus preserves for each paper its (i) number of references, (ii) citation count, (iii) citation accumulation dynamics, and (iv) the age distribution of referenced work. Performing QE switches converges to a random graph from the configuration model (Itzkovitz et al., 2003) where the number of and dynamics of citations are preserved, but the origin of the citations is randomized. Since each node is equally likely to be the originating node of any citation, given the constraints, we know a priori that no disciplines exist in this randomized citation network. The *middle panel* above demonstrates the citation history of a paper. The citation history of every paper is exactly preserved under our null model, ensuring that we control for both the variation in magnitude and dynamics of citation accumulation to papers. The *right panel* above further shows, for the example paper highlighted in Table 12.1, the frequency distribution for the observed journal pairings (*blue line*) and the frequency distribution for these journal pairings when averaged across instances of the null model (*red line*). From Uzzi et al. (2013b, p. 12). Copyright 2013 by *Science*. Reprinted with permission from the authors and *Science*

subsequent papers, and the same distribution of these citations over time (Fig. 12.3, left panel and middle panel). The right panel of Fig. 12.3 showed the distributions of observed frequency and expected frequency of journal papers for the example paper above.

Specifically, we used a variation of the Markov Chain Monte Carlo (MCMC) algorithm to construct randomized citation networks for all papers in the WOS database. The switching of endpoints of citation links was constrained to randomly chosen endpoints within the same class (Fig. 12.3), where the link classes are defined as having the same origin year and target year (Itzkovitz et al., 2003). One can think of each link class as a sub-graph of the global citation network, which can then be randomized in the usual way by performing $Q \cdot E$ switches, where E is the number of links in the subgraph. There is no proof for when the Markov Chain converges; however, it is suggested (Itzkovitz et al., 2003) to set Q at a safe value of 100. Since the citation network has 302 million edges, the scale of the computation is large, and we used a slightly less conservative value of $Q = 2 \log(E)$ to reduce computational burden. As can be noted in the original paper on the MCMC switching algorithm (Itzkovitz et al., 2003), this value of Q is well within the region where correlations with the original network cannot be detected.

Results

In this study, we examined 17.9 million research articles in the Web of Science (WOS) to see how prior work is combined. We present facts that inform (i) the extent to which scientific papers reference novel versus conventional combinations of prior work, (ii) the relative impact of papers based on the combinations they draw upon, and (iii) how (i) and (ii) are associated with collaboration.

We considered pairwise combinations of references in the bibliography of each paper (Small, 1973; Stringer et al., 2010). We counted the frequency of each citation pair across all papers published that year in the WOS and compared these observed frequencies to those expected by chance, using randomized citation networks. In the randomized citation networks, all citation links between all papers in the WOS were switched using a Monte Carlo algorithm. The switching algorithm preserves the total citation counts to and from each paper, and the distribution of these citations counts forward and backward in time to ensure that a paper (or journal) with n citations in the observed network will have n citations in the randomized network. For both the observed and the randomized paper-to-paper citation networks, we aggregated counts of paper pairs into their respective journal pairs to focus on domain-level combinations (Itzkovitz et al., 2003; Stringer et al., 2008, 2010). In the data, there were over 122 million potential journal pairs created by the 15,613 journals indexed in the WOS.

Comparing the observed frequency with the frequency distribution created with the randomized citation networks, we generated a z -score for each journal pair. This normalized measure describes whether any given pair appeared novel or conventional. Z -scores above zero indicate pairs that appeared more often in the observed data than expected by chance, indicating relatively common or “conventional” pairings. Z -scores below zero indicate pairs that appear less often in the observed WOS than expected by chance, indicating relatively atypical or “novel” pairings. For example, in the year 1995, the pairing *Nature* and *Journal of Climate* had a high z -score (47.344) indicating a conventional pairing, while *Nature* paired with *Monthly Weather Review* had a negative z -score (−24.554) indicating a pairing more unusual than chance.

The above method assigns each paper a distribution of journal pair z -scores based on the paper’s reference list (Fig. 12.4a). To characterize a paper’s tendency to draw together conventional and novel combinations of prior work, we took two summary statistics. First, to characterize the central tendency of a paper’s combinations, we considered the paper’s median z -score. The median allowed us to characterize conventionality in the paper’s main mass of combinations. Second, we considered the paper’s 10th percentile z -score. The left tail allows us to characterize the paper’s more unusual journal combinations where novelty may reside.

We found that papers typically relied on very high degrees of conventionality. Figure 12.4b presents the distribution of papers’ median z -scores for the WOS in the indicated decades. Considering that a z -score below zero represents a journal pair that occurs less often than expected by chance, the analysis of median z -scores sug-

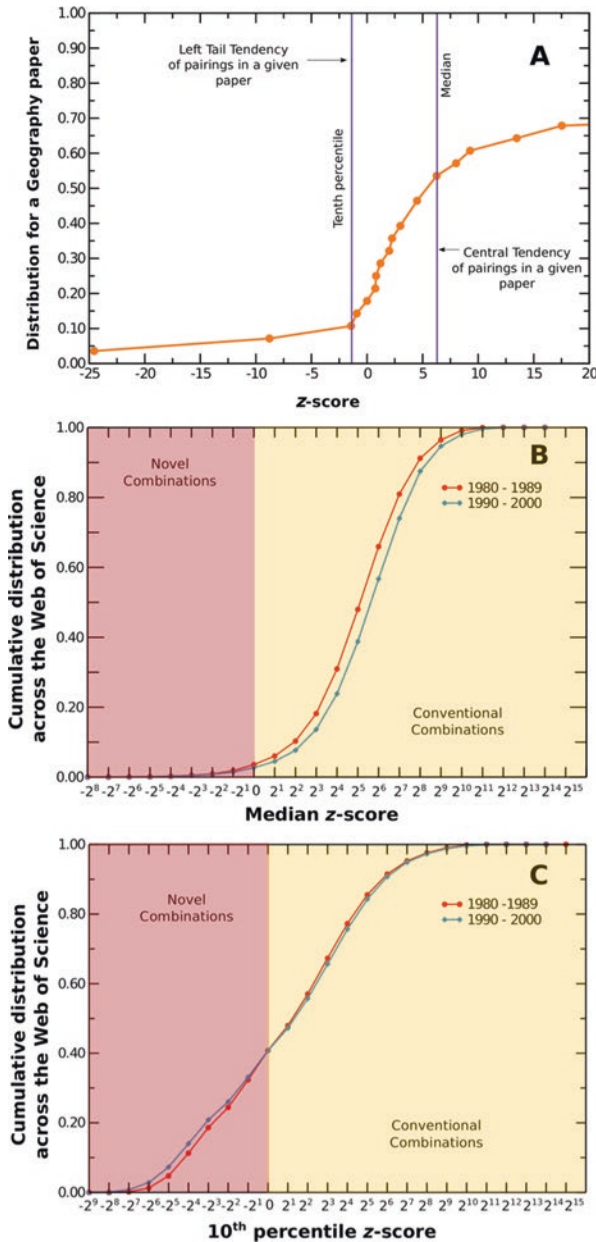


Fig. 12.4 Novelty and conventionality in science. For a sample paper, (a) shows the distribution of z-scores for that paper’s journal pairings. The z-score shows how common a journal pairing is compared to chance. For each paper we take two summary measures: its median z-score, capturing the paper’s central tendency in combining prior work, and the 10th percentile z-score, capturing the paper’s journal pairings that are relatively unusual. For the population of papers, we then consider these values across all papers in the WOS published in the 1980s or 1990s. (b) considers the

gests very high degrees of conventionality. Half the papers had median z -scores exceeding 69.0 in the 1980s and 99.5 in the 1990s. Moreover, papers with a median z -score below zero were rare. In the 1980s only 3.54% of papers had this feature, while in the 1990s the percentage fell to 2.67%, indicating a persistent and prominent tendency for high conventionality.

Focusing on each paper's left tail combinations, we found that even among the paper's relatively unusual journal combinations, the majority of papers did not feature atypical journal pairs. Figure 12.4c shows that 40.8% of the papers in the 1980s and 40.7% in the 1990s had a 10th percentile z -score below zero. Overall, by these measures, science typically relies on highly conventional combinations and rarely incorporates journal pairs that are uncommon compared to chance.

Our next finding indicates a powerful relationship between combinations of prior work and ensuing impact. Figure 12.5 presents the probability of a "hit" paper conditional on the combination of its referenced journal pairs. Hit papers are operationalized as those in the upper 5th percentile of citations received across the whole dataset, as measured by total citations through 8 years after publication. The vertical axis shows the probability of a hit paper conditional on a 2×2 categorization indicating the paper's (i) "median conventionality" (an indicator of whether the paper's median z -score is in the upper or lower half of all median z -scores) and (ii) "tail novelty" (an indicator of whether the paper's 10th percentile z -score is above or below zero).

Papers with "high median conventionality" and "high tail novelty" display a hit rate of 9.11 out of 100 papers, or nearly twice the background rate of 5 out of 100 papers. All other categories show significantly lower hit rates. Papers featuring high median conventionality but low tail novelty displayed hit rates of 5.82 out of 100 papers, while those featuring low median conventionality but high tail novelty display hit rates of 5.33 out of 100 papers. Finally, papers low on both dimensions have hit rates of just 2.05 out of 100.

Further analyses suggest universality of these relationships for scientific work across time and fields. In Fig. 12.6, we show that the results hold (a) over five decades of data recorded in the WOS from 1950 to 2000 and (b) using the upper 1st or 10th percentiles of citation impact. In Fig. 12.7, we define the cutoff for high and low tail novelty at different percentiles of a paper's z -score: The 1st, 5th, 20th, 30th, and 40th. Figure 12.7 shows that using the 1st, 5th, 10th, or 20th percentile captures significant positive associations between impact and tail novelty in the 1990s. Beyond the 30th percentile, the significant association between impact and tail nov-

Fig. 12.4 (continued) median z -scores and shows that the vast majority of papers display a high propensity for conventionality; in the 1980s and 1990s fewer than 4% of papers have median z -scores below 0 and more than 50% of papers have median z -scores above 64. (c) considers the 10th percentile z -scores, which further suggest a propensity for conventionality; only 41% of papers in the 1980s and 1990s have a 10th percentile z -score below 0. Overall, by these measures, science rarely draws on atypical pairings of prior work. From Uzzi et al. (2013a, p. 469). Copyright 2013 by *Science*. Adapted with permission from the authors and *Science*

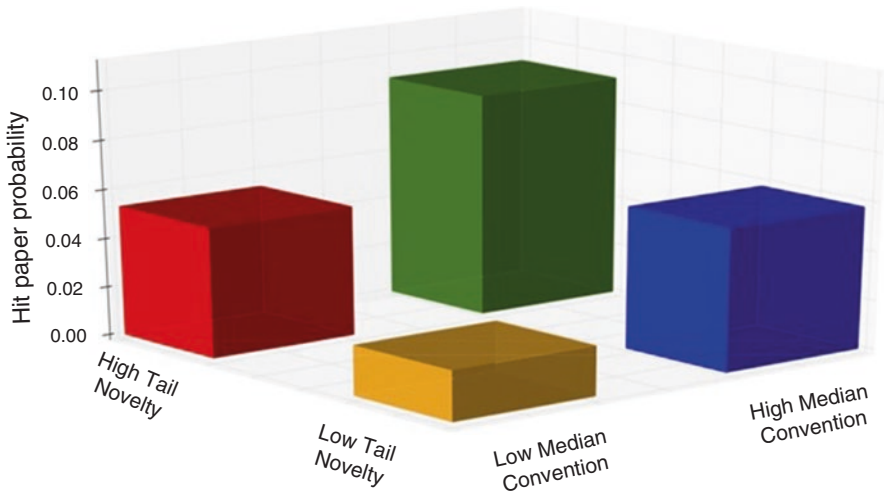


Fig. 12.5 The probability of a “hit” paper conditional on novelty and conventionality. Figure 12.5 presents the probability of a paper being in the top 5% of the citation distribution, conditional on two dimensions: whether a paper exhibits (1) high or low median conventionality and (2) high or low tail novelty, as defined in the text. Papers that combine high median conventionality and high tail novelty are hits in 9.11 out of 100 papers, a rate nearly double the background rate of 5%. Papers that are high on one dimension only—high median conventionality or high tail novelty but not both—have hit rates about half as large. Papers with low median conventionality and low tail novelty have hit rates of only 2.05 out of 100 papers. The sample includes all papers published in the WOS from 1990 to 2000. Figure 12.6 shows similar findings when considering (i) all other decades from 1950 to 2000; (ii) “hit” papers defined as the top 1% or 10% by citations, hinting at a universality of these relationships for scientific work. The difference in the hit probabilities for each category is statistically significant ($p < 0.00001$). The percentage of WOS papers in each category are; Green Bar (6.7%), Gold Bar (23%), Red Bar (26%), and Blue Bar (44%). From Uzzi et al. (2013a, p. 470). Copyright 2013 by *Science*. Reprinted with permission from the authors and *Science*

elty disappears. These patterns suggest that the concept of tail novelty is not sensitive to a single value and that beyond a precise focus on the 10th percentile the construct is related to impact so long as one continues to consider the left tail of the distribution.

Results by Subfields

The following analysis shows that the results presented in the main text for the whole of the WOS continued to appear quite broadly when examining patterns within individual subfields. By subfield, we presented (1) the tendency for tail novelty and median conventionality, and (2) the relationship between novelty, conventionality, and hit papers. We examined all 243 subfields that appeared in the WOS over the 1990s.

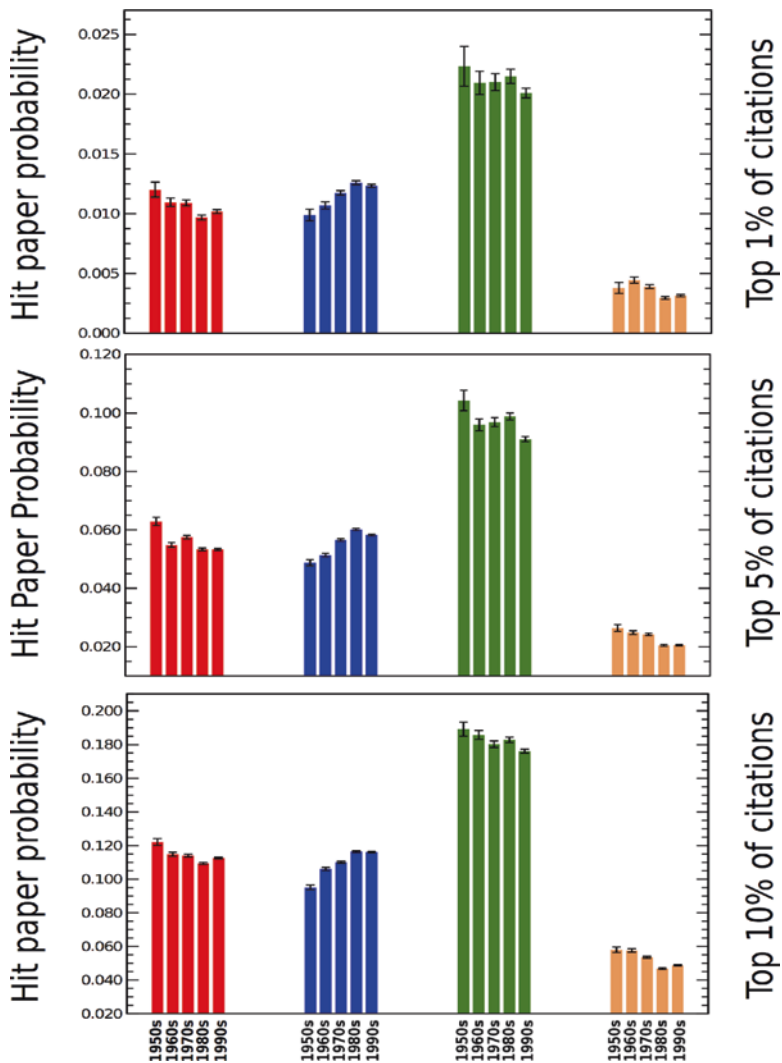


Fig. 12.6 Citation impact results generalize by decade and by definition of “hit” paper. This figure shows broadly consistent patterns both over time and by the definition of “hit” paper, suggesting a remarkably robust and strong empirical regularity between scientific impact and how prior work is combined. Specifically, the figure shows that high tail novelty combined with high median conventionality (*Green bars*) outperforms other categories in all decades from 1950 to 2000, regardless of whether a “hit” paper is defined as a top 1%, 5%, or 10% by citations received, and broadly shows hit rates that are approximately twice the background rate. By contrast, papers that feature neither high tail novelty nor high median conventionality (*Orange bars*) see hit rates at only half or less the background rate. From Uzzi et al. (2013b, p. 13). Copyright 2013 by *Science*. Reprinted with permission from the authors and *Science*

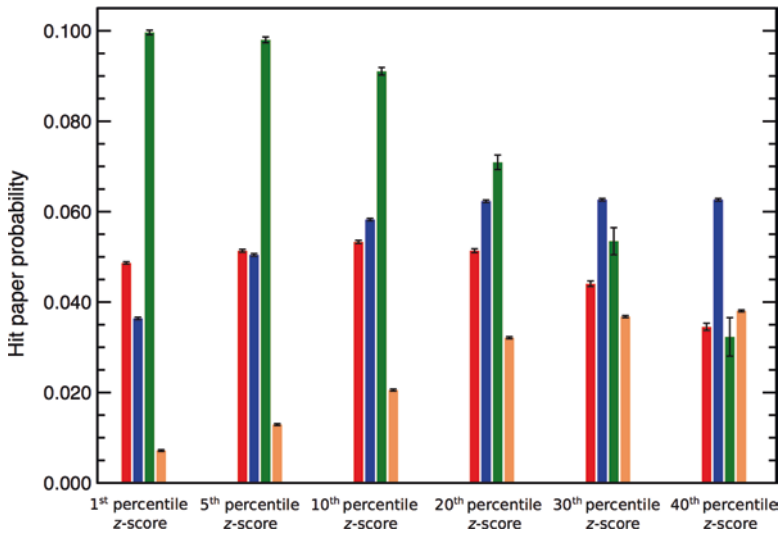


Fig. 12.7 Citation impact results generalize to broader definitions of left tail novelty. The figure presents the relationship between tail novelty and impact using alternative definitions of tail novelty. In each case, tail novelty is defined as an indicator for whether the e th percentile of a paper’s z-score distribution is less than zero. The x-axis indicates the value of e . It is seen that for $e \leq 20$, high tail novelty combined with a high median conventionality (Green bars) outperforms other categories. The results in the main text, which use the 10th percentile, thus extend broadly to other definitions of tail novelty so long as the measure emphasizes the paper’s left tail of combinations. From Uzzi et al. (2013b, p. 14). Copyright 2013 by *Science*. Reprinted with permission from the authors and *Science*

To examine any field-specific relationships between novelty, conventionality, and hit papers, we calculated the subfield-specific probabilities of a “hit” by the four categories used in Fig. 12.2 and defined in the text. We then ranked these four categories in each subfield, where 1 indicates the highest probability of a hit, 2 indicates the second highest probability of a hit and so on. Consistent with the main results, Table 12.2 shows that in 64.4% of fields, a paper’s likelihood of being a hit paper was greatest when combining prior work characterized by high tail novelty and high median conventionality. This category (Row 3 in Table 12.2) is ranked first or second in 86.3% of subfields. Notably, to the extent that this category is not dominant within a subfield, the category featuring a more general shift toward novelty (Row 1 in Table 12.2) appears prominently, suggesting that tail novelty is an especially generic feature of the highest-impact papers. Conversely, the category featuring low tail novelty and low median conventionality (Row 4 in Table 12.2) ranks lowest in 70.4% of subfields. Thus, novelty and conventionality are not opposing factors in the production of science; rather, papers with an injection of novelty into an otherwise exceptionally familiar mass of prior work are unusually likely to have high impact. Next, we focus on the effect of teams on novelty.

Role of Teams in Production of Knowledge

Collaboration is often claimed to produce more novel combinations of ideas (Falk-Krzesinski et al., 2010; Fiore, 2008; Stokols et al., 2008; Uzzi & Spiro, 2005; Wuchty et al., 2007), but the extent to which teams incorporate novel combinations across the universe of fields is unknown. Team-authored papers were more likely to show atypical combinations than single or pair-authored papers. Figure 12.8a shows that the distribution of 10th percentile z -scores shifted significantly leftward as the number of authors increased (*Kolmogorov-Smirnov* [KS] tests indicate solo vs. pair $p=0.016$, pair vs. team $p=0.001$, team vs. solo $p<0.001$). Papers written by one, two, three, or more authors showed high tail novelty in 36.1%, 39.8%, and 49.7% of cases, respectively, indicating that papers with three or more authors showed an increased frequency of high tail novelty over the solo-author rate by 37.7%.

Teams were neither more nor less likely than single authors or pairs of authors to display high median conventionality. Figure 12.8b indicates no significant statistical difference in the median z -score distributions (KS tests indicate solo vs. pair $p=0.768$, pair vs. team $p=0.417$, team vs. solo $p=0.164$). Teams thus achieve high tail novelty more often than solo authors, yet teams were not simply “more novel” but rather displayed a propensity to incorporate high tail novelty without giving up a central tendency for high conventionality.

Regression Methods

In our final analysis, we examined the interplay between citation, combination, and collaboration using regression methods (Fig. 12.9). Papers were binned into eleven equally sized categories of median conventionality. We used logistic regression to

Table 12.2 Novelty, conventionality, and citation impact by field

| | Rank | | | |
|---|-------|-------|-------|-------|
| | 1st | 2nd | 3rd | 4th |
| High tail novelty and low median conventionality | 20.3% | 44.5% | 28.7% | 6.5% |
| Low tail novelty and high median conventionality | 9.7% | 26.7% | 50.6% | 13.0% |
| High tail novelty and high median conventionality | 64.4% | 21.9% | 3.6% | 10.1% |
| Low tail novelty and low median conventionality | 5.7% | 6.9% | 17.0% | 70.4% |

Note. For each of 243 subfields indexed by the WOS in the 1990s, we ranked the categories of papers according to their probability of producing hit papers. Hit papers are defined as those in the upper 5% of citations received in that subfield. We focused on all papers published across all subfields in the 1990s. This analysis revealed that high tail novelty and high median conventionality were the highest impact papers in 64.4% of subfields and either first or second in 86.3% of fields. By contrast, low tail novelty and low median conventionality rank lowest or second lowest in 87.4% of fields. From Uzzi et al. (2013a, p. 22). Copyright 2013 by *Science*. Reprinted with permission by the authors and *Science*

predict the probability of hit papers in the 1990s and ran these regressions in a flexible manner that avoided imposing functional forms on the data. In particular, we first divided papers into subsamples based on their median conventionality (11 categories, from least to greatest median conventionality, as defined in the main text) and the number of authors (3 categories, for solo authors, two-author pairs, and three or more authors). This created 33 distinct subsamples. We then ran a separate regression for each subsample.

For a given subsample, a regression takes the form

$$\Pr(y_i) = f(\beta Tail_Novelty_i + \sum_f \gamma_f Field_{fi})$$

where $y_{ij} \in \{0,1\}$ is an indicator variable for a “hit” paper, and $Tail_Novelty_i \in \{0,1\}$ is an indicator variable for whether a paper’s 10th percentile z -score is below zero. The regression includes a full set of fixed effects for each of 243 subfields indexed by the WOS in the 1990s, where the indicator variables $Field_{fi} \in \{0,1\}$ are equal to 1 if the paper i is in field f . The inclusion of these fixed effects accounts for any mean differences in hit probabilities and tail novelty across subfields. We further restricted the sample to papers with at least ten known references, which ensured that each paper in the sample had many pairwise combinations of prior work.

Figure 12.9 establishes a large positive relationship between tail novelty and hit papers, which appears independently in each of the 33 subsamples. The regressions further establish that the probability of hit papers increases with median conventionality, peaking at approximately the 85th percentile of median conventionality. These strong empirical regularities extend to alternative analyses. Figure 12.10 reconsiders these regressions defining hit papers to be in the top 1% of citations received. The results for this higher threshold for a “hit” paper look extremely similar.

Novelty and Conventionality in Geography

In Fig. 12.11 we show the novel and conventional combinations in papers written in the field of geography. We do not observe any high degrees of conventionality when compared with fields like physics and economics. Figure 12.11a presents the distribution of papers’ median z -scores for papers published in geography in the indicated decades. Half the papers have median z -scores exceeding 16 in the 1980s in the 1990s. Papers with a median z -score below zero are rare; only 4% of papers displayed this feature in the 1980s and 1990s. Focusing on each paper’s left tail combinations, we found that 25% of papers in the 1980s and 1990s have a 10th percentile z -score below 0 (Fig. 12.11b). Overall, by these measures, we observed that geography papers rarely draw atypical knowledge from prior works.

We investigated the role of team authors in geography in production of knowledge. Our findings support the previous observations in Fig. 12.8: Each team size

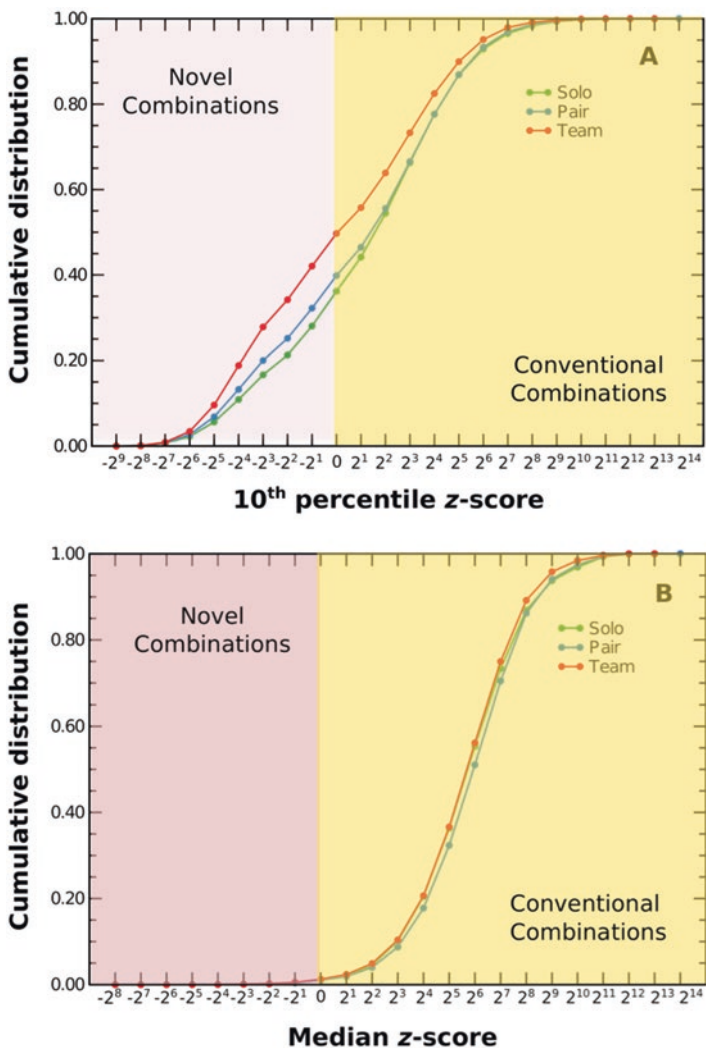


Fig. 12.8 Authorship structure, novelty, and conventionality. Team-authored papers are more likely to incorporate tail novelty but without sacrificing a central tendency for high conventionality. Papers introduce tail novelty (a 10th percentile z -score less than 0) in 36.2%, 39.9%, and 49.7% of cases for solo authors, dual authors, and three or more authors, respectively (a). Kolmogorov-Smirnov tests confirm the distributions of tail novelty are distinct (solo vs. pair $p=0.016$, pair vs. team $p=0.001$, team vs. solo $p<0.001$). By contrast, each team size shows similar distributions for median conventionality (b KS tests indicate no statistically significant differences). These findings suggest that a distinguishing feature of teamwork, and teams' exceptional impact, reflects a tendency to incorporate novelty. From Uzzi et al. (2013a, p. 470). Copyright 2013 by Science. Reprinted with permission from the authors and Science

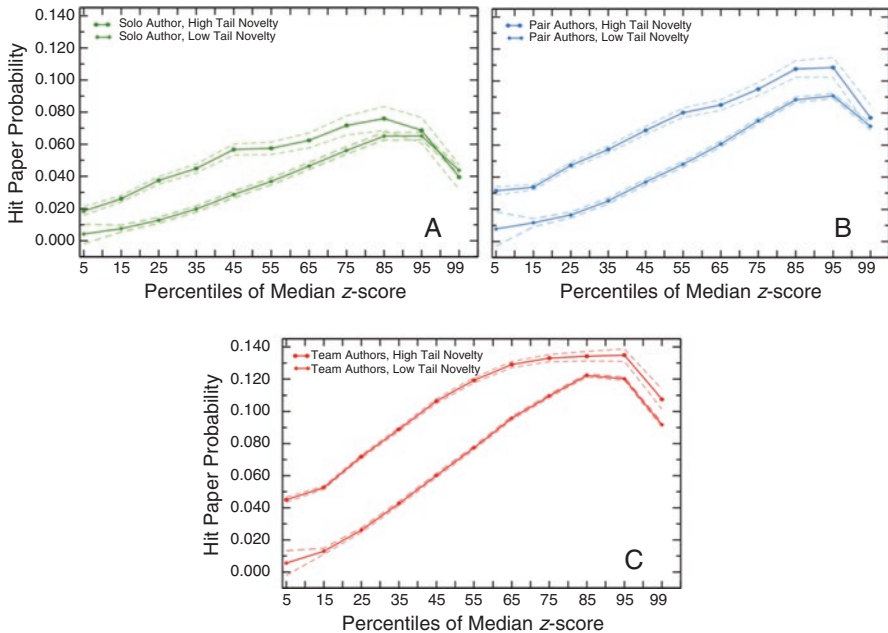


Fig. 12.9 Novel and conventional combinations in the production of science. The interplay between tail novelty, median conventionality, and hit paper probabilities show remarkable empirical regularities (a–c). First, high tail novelty papers have higher impact than low tail novelty papers at (i) any level of conventionality and (ii) regardless of authorship structure. Second, increasing median conventionality is associated with higher impact up to the 85–95th percentile of median conventionality after which the relationship reverses. Third, larger teams obtain higher impact given the right mix of tail novelty and median conventionality. Nonetheless, at low levels of median conventionality and tail novelty, even teams have low impact, further emphasizing the fundamental relationship between novelty, conventionality, and impact in science. From Uzzi et al. (2013a, p. 471). Copyright 2013 by *Science*. Reprinted with permission from the authors and *Science*

shows similar distributions for median conventionality (Fig. 12.11c) and papers with team authors show greater novelty than solo-authored or pair-authored papers (Fig. 12.11d).

We plot the average hit citations by considering the top 5% of highly cited papers written in the fields of geography, economics, and physics over time (Fig. 12.12). The average of hit citations for geography is significantly lower than that of economics and physics. For papers published between 1980 and 2000, half the papers in geography have median z-scores above 36. Thus papers in geography combine less conventionality when compared to physics or economics, where half the papers have median z-scores above 145. Moreover, for geography, conventionality doesn't increase with time when compared with physics and economics (Fig. 12.13). This indicates that for geography, mixing novel and conventional combinations does not result in high impact work.

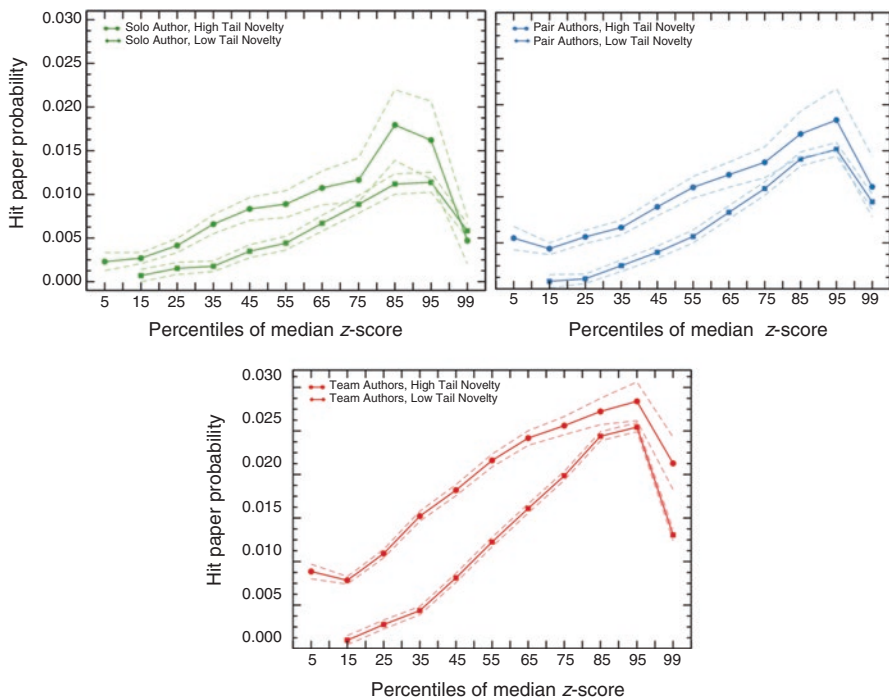
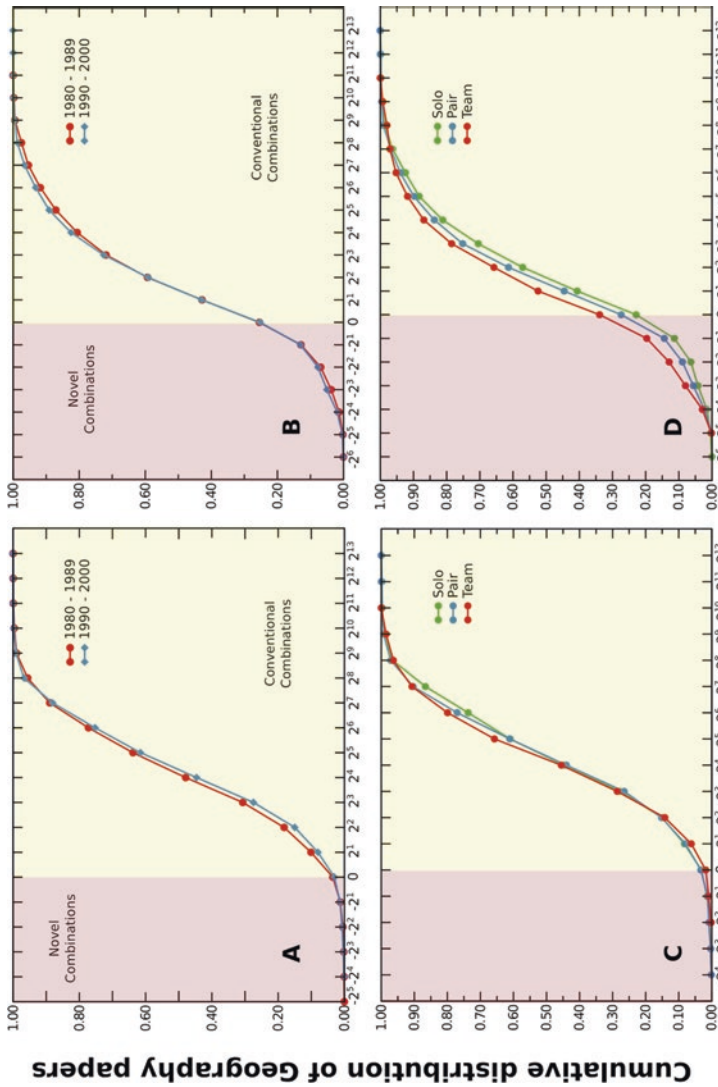


Fig. 12.10 Novelty, authorship and impact for top 1% of papers. This figure repeats Fig. 12.9 but defines hit papers as those that receive citations within 8 years of publication that are in the upper 1% of all papers published that year. From Uzzi et al. (2013b, p. 16). Copyright 2013 by *Science*. Adapted with permission from the authors and *Science*

Discussion and Conclusion

There were three primary findings. First, high tail novelty papers had higher impact than low tail novelty papers, an impact advantage that occurred at any level of conventionality and regardless of authorship structure. Second, peak impact occurs in the 85–95th percentile of median conventionality, an exceptionally high level. This peak and its position appeared irrespective of tail novelty/no tail novelty or authorship structure. These generic features suggest fundamental underlying rules relating combinations of prior work to the highest impact science.

Finally, Fig. 12.4 indicates that for virtually all possible mixes of tail novelty and median conventionality, larger teams were associated with higher impact. Thus, while teams incorporated the highest impact mixes more frequently (Fig. 12.3), teams also tended to obtain higher impact for any particular mix (Fig. 12.4). Nonetheless, despite this advantage in citations across virtually all fields of science (Wuchty et al., 2007), even teams had low impact at low levels of median conventionality and tail novelty.



10th percentile z-score

Median z-score

Fig. 12.11 Novel and conventional combinations in geography. The median z-scores shows that the vast majority of geography papers display a high propensity for conventionality: in the 1980s and 1990s fewer than 4% of papers have median z-scores below 0 and more than 50% of papers have median z-scores above 16 (a). The plot for 10th percentile z-scores, which further suggest a propensity for conventionality; ~ 25% of papers in the 1980s and 1990s have a 10th percentile z-score below 0 (b). From these two figures, we observe that geography papers rarely draw atypical knowledge from prior works. For geography papers published in the 1990s, each team size shows similar distributions for median conventionality (c). Papers published in the 1990s introduce tail novelty (a 10th percentile z-score < 0) in ~ 20%, 25%, and 35% of cases for solo authors, dual authors, and three or more authors, respectively (d) (Design by authors)

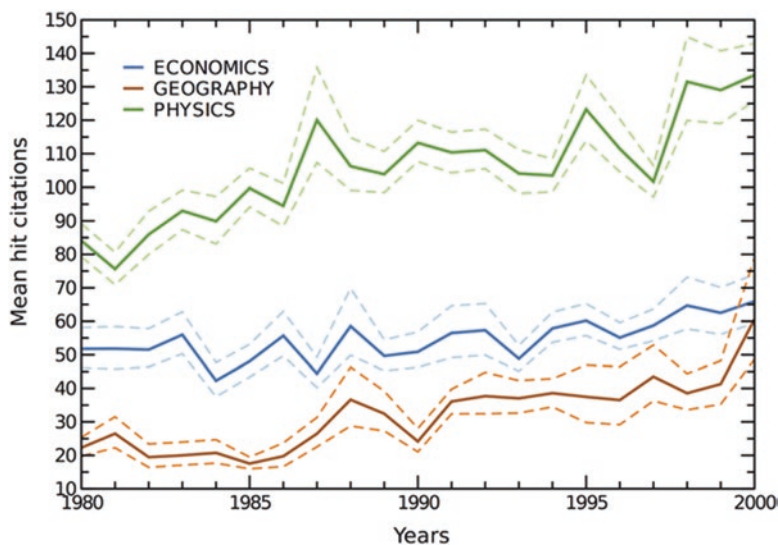


Fig. 12.12 Mean of hit citations with time for three fields. For each of the fields—geography, economics, and physics—we consider the “hit” papers. Hit papers are defined as those in the top 5% of citations. The plot shows the evolution of mean hit citations in time. For geography, the value of mean hit citations is much lower when compared to hit papers in physics and economics (Design by authors)

Our analysis of 17.9 million papers across all scientific fields suggests that the highest-impact science draws on primarily highly conventional combinations of prior work with an intrusion of combinations unlikely to have been joined before. These patterns suggest that novelty and conventionality are not factors in opposition; rather, papers that mix high tail novelty with high median conventionality have nearly twice the propensity to be unusually highly cited.

These findings have implications for theories about creativity and scientific progress. Combinations of existing material are centerpieces in theories of creativity, whether in the arts, the sciences, or commercial innovation (Becker, 1982; Collins, 1998; Guimera et al., 2005; Jones, 2009; Schilling & Phelps, 2007; Schumpeter, 1939; Usher, 1929/1998; Uzzi & Spiro, 2005; Weitzman, 1998). Across the sciences, the propensity for high impact work is sharply elevated when combinations of prior work are anchored in substantial conventionality while mixing in a left tail of combinations that are rarely seen together. In part, this pattern may reflect advantages to being within the mainstream of a research trajectory, where scientists are currently focused while being distinctive in one’s creativity. For example, as mentioned in the beginning of the chapter, Newton remained in the mainstream of traditional geometry and at the same time remained creative while communicating the laws of gravitation in *Principia*. Combinations of prior work also relate to “burden of knowledge” theory, which emphasizes the growing knowledge demands upon scientists (Einstein, 1949; Fleming, 2001; Jones, 2009). New articles indexed by the WOS now exceed 1.4 million per year across 251 fields, encouraging specialization

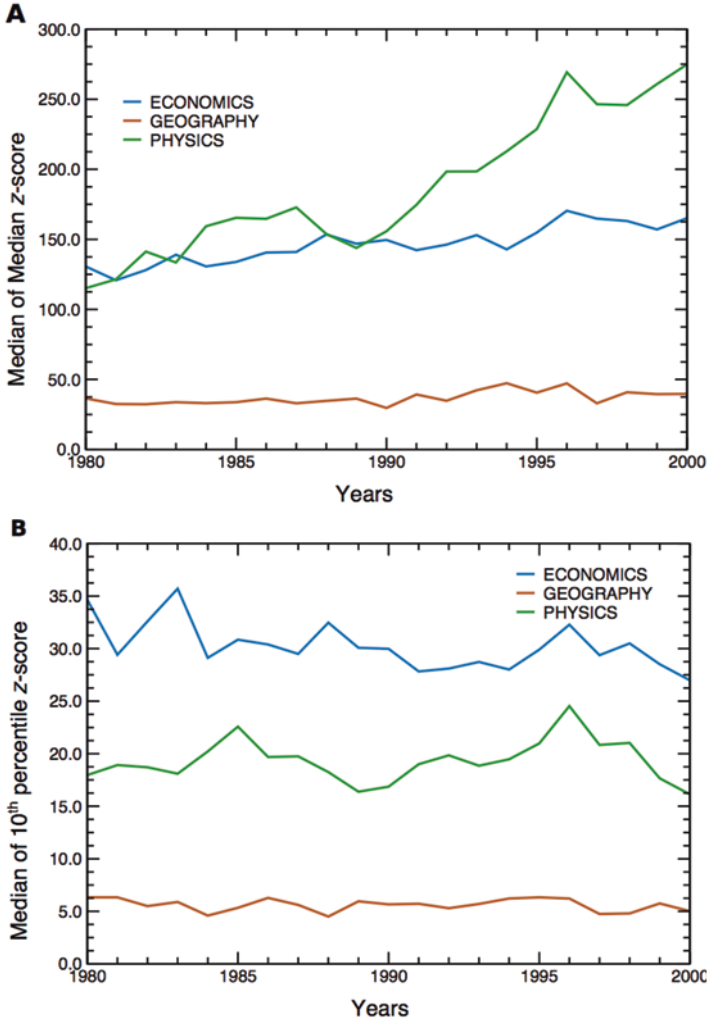


Fig. 12.13 Median of conventionality and novelty with time for three fields. (a) shows the median conventionality of papers published in geography exhibit no change with time. On the contrary, median conventionality of physics and economics papers increases with time, indicating a trend towards high conventionality for both these fields. Moreover, geography does not draw enough conventional knowledge as compared to physics or economics. In (b) we show the median novelty of geography, economics, and physics. While for physics and economics, novelty decreases with time, in the case of geography there is no such behavior—a trend we observed earlier in Fig. 12.11b (Design by authors)

and challenging scientists' capacity to comprehend new thinking across domains. The finding that teams preserve high conventionality yet introduce tail novelty suggests that teams help meet the challenge of the burden of knowledge by balancing domain-level depth with a capacity for atypical combinations.

Our methodology considered paper and journal pairings but can be applied at the level of disciplines, papers, or topics within papers, allowing the examination of combinations of prior work at different resolutions in future studies of creativity and scientific impact. Beyond science, links between novelty and conventionality in successful innovation also appear. E-books retain page-flipping graphics to remind the reader of physical books, and blue jeans were designed with a familiar watch pocket to look like conventional trousers. From this viewpoint, the balance between extending technology with atypical combinations of prior ideas while embedding them in conventional knowledge frames may be critical to human progress in many domains. Future research questions also arise from our findings. Science is dynamic, with research areas shifting and new fields arising. While we find that the regularities relating novelty, conventionality, and impact persist across time and fields, understanding how research trajectories shift and how new fields are born are questions that measures of novelty and convention may valuably inform. At root, our work suggests that creativity in science appears to be a nearly universal phenomenon of two extremes. At one extreme is conventionality, and at the other is novelty. Curiously, notable advances in science appear most closely linked not with efforts along one boundary or the other but with efforts that reach toward both frontiers.

Acknowledgements Sponsored by the Northwestern University Institute on Complex Systems (NICO), by the Army Research Laboratory under Cooperative Agreement Number W911NF-09-2-0053 and DARPA BAA-11-64, Social Media in Strategic Communication. The views and conclusions contained in this document are those of the authors and should not be interpreted as representing the official policies, either expressed or implied, of the Army Research Laboratory or the U.S. Government. 11B.

References

- Azoulay, P., Zivin, J. G., & Manso, G. (2011). Incentives and creativity: evidence from the Academic Life Science. Howard Hughes Medical Investigator Program. *The RAND Journal of Economics*, 42, 527–554. doi:[10.1111/j.1756-2171.2011.00140.x](https://doi.org/10.1111/j.1756-2171.2011.00140.x)
- Becker, H. S. (1982). *Art Worlds*. Berkeley: University of California Press.
- Bollen, J., Rodriguez, M. A., & van de Sompel, H. (2006). Journal status. *Scientometrics*, 69, 669–687. doi:[10.1007/s11192-006-0176-z](https://doi.org/10.1007/s11192-006-0176-z)
- Collins, R. (1998). *The sociology of philosophies: A global theory of intellectual change*. Cambridge, MA: Harvard University Press.
- Einstein, A. (1949). *The world as I see it*. Secaucus: Citadel Press.
- Evans, J., & Foster, J. (2011). Metaknowledge. *Science*, 331, 721–725. doi:[10.1126/science.1201765](https://doi.org/10.1126/science.1201765)

- Falk-Krzesinski, H. J., Börner, K., Contractor, N., Fiore, S. M., Hall, K. L., Keyton, J., & Uzzi, B. (2010). Advancing the science of team science. *CTS: Clinical and Translational Science Journal*, 3, 263–266. doi:[10.1111/j.1752-8062.2010.00223.x](https://doi.org/10.1111/j.1752-8062.2010.00223.x)
- Fiore, S. M. (2008). Interdisciplinarity as teamwork: How the science of teams can inform team science. *Small Group Research*, 39, 251–277. doi:[10.1177/1046496408317797](https://doi.org/10.1177/1046496408317797)
- Fleming, L. (2001). Recombinant uncertainty in technological search management. *Management Science*, 47, 117–132. doi:[10.1287/mnsc.47.1.117.10671](https://doi.org/10.1287/mnsc.47.1.117.10671)
- Guimera, R., Uzzi, B., Spiro, J., & Nunes Amaral, L. A. (2005). Team assembly mechanisms determine collaboration network structure and team performance. *Science*, 308, 697–702. doi:[10.1126/science.1106340](https://doi.org/10.1126/science.1106340)
- Henderson, R. M., & Clark, K. B. (1990). Architectural innovation: The reconfiguration of existing technologies and the failure of established firms. *Administrative Science Quarterly*, 35, 9–30. doi:[10.2307/2393549](https://doi.org/10.2307/2393549)
- Itzkovitz, S., Milo, R., Kashtan, N., Newman, M. E. J., & Alon, U. (2003). Subgraphs in random networks. *Physical Review E*, 68. doi:[10.1103/PhysRevE.68.026127](https://doi.org/10.1103/PhysRevE.68.026127)
- Jones, B. (2009). The burden of knowledge and the death of the renaissance man: Is innovation getting harder? *The Review of Economic Studies*, 76, 283–317. doi:[10.1111/j.1467-937X.2008.00531.x](https://doi.org/10.1111/j.1467-937X.2008.00531.x)
- Jones, B., Wuchty, S., & Uzzi, B. (2008). Multi-university research teams: Shifting impact, geography, and stratification in science. *Science*, 322, 1259–1262. doi:[10.1126/science.1158357](https://doi.org/10.1126/science.1158357)
- National Academy of Sciences, National Academy of Engineering, & Institute of Medicine of the National Academies (2004). *Facilitating interdisciplinary research*. Washington, DC: National Academies Press.
- Rosenkopf, L., & McGrath, P. (2011). Advancing the conceptualization and operationalization of novelty in organizational research. *Organization Science*, 22, 1297–1311. doi:[10.1287/orsc.1100.0637](https://doi.org/10.1287/orsc.1100.0637)
- Schilling, M. A., & Phelps, C. C. (2007). Interfirm collaboration networks: The impact of small world connectivity on firm innovation. *Management Science*, 53, 1113–1126.
- Schilling, M., & Green, E. (2011). Recombinant search and breakthrough idea generation: An analysis of high impact papers in the social sciences. *Research Policy*, 40, 1321–1331. doi:[10.1016/j.respol.2011.06.009](https://doi.org/10.1016/j.respol.2011.06.009)
- Schumpeter, J. (1939). *Business Cycles*. New York, McGraw-Hill.
- Small, H. (1973). Co-citation in scientific literature: A new measure of the relationship between two documents. *Journal of the American Society for information Science*, 24, 265–269. doi:[10.1002/asi.4630240406](https://doi.org/10.1002/asi.4630240406)
- Stokols, D., Hall, K. L., Taylor, B. K., & Moser, R. P. (2008). The science of team science: Overview of the field and introduction to the supplement. *American Journal of Preventive Medicine*, 35, 77–89. doi:[10.1016/j.amepre.2008.05.002](https://doi.org/10.1016/j.amepre.2008.05.002)
- Stringer, M. J., Sales-Pardo, M., & Amaral, L. A. (2008). Effectiveness of journal ranking schemes as a tool for locating information. *PLOS ONE*, 3. doi:[10.1371/journal.pone.00016](https://doi.org/10.1371/journal.pone.00016)
- Stringer, M. J., Sales-Pardo, M., & Amaral, L. A. (2010). Statistical validation of a global model for the distribution of the ultimate number of citations accrued by papers published in a scientific journal. *Journal of the American Society for Information Science and Technology*, 61, 1377–1385. doi:[10.1002/asi.21335](https://doi.org/10.1002/asi.21335)
- Usher, A. P. (1998). *A History of mechanical invention*. Cambridge, MA: Harvard University Press. (Original work published 1929)
- Uzzi, B., Mukherjee, S., Stringer, M., & Jones, B. (2013a). Atypical combinations and scientific impact. *Science*, 342, 468–472. doi:[10.1126/science.1240474](https://doi.org/10.1126/science.1240474)
- Uzzi, B., Mukherjee, S., Stringer, M., & Jones, B. (2013b). Supplementary materials for atypical combinations and scientific impact. *Science*, 342. Retrieved from www.sciencemag.org/content/342/6157/468/suppl/DC1
- Uzzi, B., & Spiro, J. (2005). Collaboration and creativity: The small world problem. *American Journal of Sociology*, 111, 447–504. doi:[10.1086/432782](https://doi.org/10.1086/432782)

- Weitzman, M. L. (1998). Recombinant growth. *The Quarterly Journal of Economics*, 113, 331–360. doi:[10.1162/003355398555595](https://doi.org/10.1162/003355398555595)
- Whiteside, D. T. (1970). The mathematical principles underlying Newton’s *principia mathematica*. *Journal for the History of Astronomy*, 1, 116–138. doi:[10.1177/002182867000100203](https://doi.org/10.1177/002182867000100203)
- Wuchty, S., Jones, B., & Uzzi, B. (2007). The increasing dominance of teams in the production of knowledge. *Science*, 316, 1036–1039. doi:[10.1126/science.1136099](https://doi.org/10.1126/science.1136099)

Open Access This chapter is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, duplication, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the work’s Creative Commons license, unless indicated otherwise in the credit line; if such material is not included in the work’s Creative Commons license and the respective action is not permitted by statutory regulation, users will need to obtain permission from the license holder to duplicate, adapt or reproduce the material.

