AFRL-AFOSR-VA-TR-2020-0007

Discovering the Extent of Estimable Prediction in Science and Technology

James Evans
UNIVERSITY OF CHICAGO THE

12/12/2019
Final Report

DISTRIBUTION A: Distribution approved for public release.

Air Force Research Laboratory
AF Office Of Scientific Research (AFOSR)/ RTB1
Arlington, Virginia 22203
Air Force Materiel Command

DISTRIBUTION A: Distribution approved for public release.

REPORT DOCUMENTATION PAGE	*Form Approved* *OMB No. 0704-0188*

1. REPORT DATE *(DD-MM-YYYY)*
01-10-2019

2. REPORT TYPE
Final Report

3. DATES COVERED *(From - To)*
August 2015-September 2019

4. TITLE AND SUBTITLE
Discovering the Extent of Estimable Prediction (DEEP) in Science and Technology

5a. CONTRACT NUMBER

5b. GRANT NUMBER
FA9550-15-1-0162

5c. PROGRAM ELEMENT NUMBER

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

6. AUTHOR(S)
James Evans & Dashun Wang

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
University of Chicago & Northwestern University (formerly Pennsylvania State University)

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
N/A

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
A / Approved for public release; distribution is unlimited.

13. SUPPLEMENTARY NOTES

14. ABSTRACT
We proposed to establish mathematical and empirical foundations regarding the nature and extent of quantifiable prediction in science and technology (S&T), the central question of science policy, a fundamental challenge for complex systems research, with the potential to dramatically improve the productivity and focus of science. Research based on this program yielded a wave of relevant discoveries published in Nature, Science, PNAS, Nature subfield journals, every major sociology outlet, and top venues in research policy, social, computer and information science. Moreover, we have drafted two forthcoming books from Cambridge University Press and Princeton University Press, and many more articles that will be published in the coming year. These works review the state of the art in science and technology prediction, but also probe and exceed those limits by predicting science & technology success and failure, career and team productivity and influence, the disruptiveness and popularity of novel idea and technology combinations, team and community conflict, and a host of indicators that predict future focus and impact. Moreover, this research generated new public data, and the development and calibration of new models that allowed us to push the limits of science

15. SUBJECT TERMS
science & technology prediction, science of science, science of team science

16. SECURITY CLASSIFICATION OF:			**17. LIMITATION OF ABSTRACT**	**18. NUMBER OF PAGES**	**19a. NAME OF RESPONSIBLE PERSON**
a. REPORT	**b. ABSTRACT**	**c. THIS PAGE**			**19b. TELEPHONE NUMBER** *(Include area code)*

Reset

Standard Form 298 (Rev. 8/98)
Prescribed by ANSI Std. Z39.18

INSTRUCTIONS FOR COMPLETING SF 298

1. REPORT DATE. Full publication date, including day, month, if available. Must cite at least the year and be Year 2000 compliant, e.g. 30-06-1998; xx-06-1998; xx-xx-1998.

2. REPORT TYPE. State the type of report, such as final, technical, interim, memorandum, master's thesis, progress, quarterly, research, special, group study, etc.

3. DATES COVERED. Indicate the time during which the work was performed and the report was written, e.g., Jun 1997 - Jun 1998; 1-10 Jun 1996; May - Nov 1998; Nov 1998.

4. TITLE. Enter title and subtitle with volume number and part number, if applicable. On classified documents, enter the title classification in parentheses.

5a. CONTRACT NUMBER. Enter all contract numbers as they appear in the report, e.g. F33615-86-C-5169.

5b. GRANT NUMBER. Enter all grant numbers as they appear in the report, e.g. AFOSR-82-1234.

5c. PROGRAM ELEMENT NUMBER. Enter all program element numbers as they appear in the report, e.g. 61101A.

5d. PROJECT NUMBER. Enter all project numbers as they appear in the report, e.g. 1F665702D1257; ILIR.

5e. TASK NUMBER. Enter all task numbers as they appear in the report, e.g. 05; RF0330201; T4112.

5f. WORK UNIT NUMBER. Enter all work unit numbers as they appear in the report, e.g. 001; AFAPL30480105.

6. AUTHOR(S). Enter name(s) of person(s) responsible for writing the report, performing the research, or credited with the content of the report. The form of entry is the last name, first name, middle initial, and additional qualifiers separated by commas, e.g. Smith, Richard, J, Jr.

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES). Self-explanatory.

8. PERFORMING ORGANIZATION REPORT NUMBER. Enter all unique alphanumeric report numbers assigned by the performing organization, e.g. BRL-1234; AFWL-TR-85-4017-Vol-21-PT-2.

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES). Enter the name and address of the organization(s) financially responsible for and monitoring the work.

10. SPONSOR/MONITOR'S ACRONYM(S). Enter, if available, e.g. BRL, ARDEC, NADC.

11. SPONSOR/MONITOR'S REPORT NUMBER(S). Enter report number as assigned by the sponsoring/ monitoring agency, if available, e.g. BRL-TR-829; -215.

12. DISTRIBUTION/AVAILABILITY STATEMENT. Use agency-mandated availability statements to indicate the public availability or distribution limitations of the report. If additional limitations/ restrictions or special markings are indicated, follow agency authorization procedures, e.g. RD/FRD, PROPIN, ITAR, etc. Include copyright information.

13. SUPPLEMENTARY NOTES. Enter information not included elsewhere such as: prepared in cooperation with; translation of; report supersedes; old edition number, etc.

14. ABSTRACT. A brief (approximately 200 words) factual summary of the most significant information.

15. SUBJECT TERMS. Key words or phrases identifying major concepts in the report.

16. SECURITY CLASSIFICATION. Enter security classification in accordance with security classification regulations, e.g. U, C, S, etc. If this form contains classified information, stamp classification level on the top and bottom of this page.

17. LIMITATION OF ABSTRACT. This block must be completed to assign a distribution limitation to the abstract. Enter UU (Unclassified Unlimited) or SAR (Same as Report). An entry in this block is necessary if the abstract is to be limited.

Final Report for "Discovering the Extent of Estimable Prediction (DEEP) in Science and Technology", FA9550-15-1-0162

Abstract
We proposed to establish mathematical and empirical foundations regarding the nature and extent of quantifiable prediction in science and technology (S&T), the central question of science policy, a fundamental challenge for complex systems research, with the potential to dramatically improve the productivity and focus of science. Research based on this program yielded a wave of relevant discoveries published in *Nature*, *Science*, *PNAS*, *Nature* subfield journals, every major sociology outlet, and top venues in research policy, social, computer and information science. Moreover, we have drafted two forthcoming books from Cambridge University Press and Princeton University Press, and many more articles that will be published in the coming year. These works review the state of the art in science and technology prediction, but also probe and exceed those limits by predicting science & technology success and failure, career and team productivity and influence, the disruptiveness and popularity of novel idea and technology combinations, team and community conflict, and a host of indicators that predict future focus and impact. Moreover, this research generated new public data, and the development and calibration of new models that allowed us to push the limits of science and technology prediction in unanticipated ways. Finally, the grant project formed the basis of several other funded projects, including a Minerva award funded by AFOSR, which builds on these foundations to promote a flourishing and productive science of science and innovation that will advance the national and global interest.

Our proposal promised to explore three levels of foundations of science and technology (S&T) regularities that could be identified and potentially predicted: 1) the symbolic foundation tracing the distribution and dynamics of concepts, claims, and described components in the corpus of published S&T—what becomes important and impactful; 2) the behavioral foundation of attention, effort, communication and influence beneath the flow of S&T symbols—how does the behavior of researchers, teams, and institutions shape and reveal the unfolding importance of discoveries and inventions; and 3) the material foundation of natural and fabricated reality that scientists and engineers seek to uncover—how existing S&T reveal hints about their own future. By using and developing tools that spanned network science, machine learning and computational social science, our proposed program yielded new answers to the following questions:

- Can we foresee new scientific discoveries and technological inventions? If so, what are the quantifiable signals of an impending scientific or engineering breakthrough?
- Is the long-term impact of a scientific discovery or technological breakthrough predictable? If so, to what degree and how soon can we confidently predict impact? What are the key drivers behind uncovered predictability?
- What is best predicted in the domain of S&T—what represents the optimal balance of risk and reward in a funded research portfolio? Scientific and technical discoveries? Successful scientists, engineers, or research teams? Fruitful S&T fields or research strategies?

We explored and uncovered preliminary answers to these questions along with many other entailments that push our understanding of prediction in science, technology and related areas, published across three *Nature* articles, two *Science* articles, two *PNAS* and *Nature Physics* articles, four *Nature Human Behaviour* articles, and one article in the following journals: *Nature Communications*, *American Sociological Review*, *American Journal of Sociology*, *Research Policy*, *eLife*, *Social Science Computer Review*, *Scientific Data*, *Journal of the Association for Information Science and Technology*, *AISTATS*, *IEEE Computer Graphics and Information*, *Journal of Infometrics*. Moreover, our explorations have led us to draft two forthcoming books from Cambridge University Press and Princeton University Press, and more than ten additional articles that will be well published, which probe the limits of prediction in science and technology.

We organize our discussion of these investigations and associated discoveries into the following groups:

How failure predicts success

Human achievements are often preceded by repeated attempts that fail, but little is known about the mechanisms that govern the dynamics of failure. In the following pieces, we develop a simple but powerful one-parameter model that mimics how successful future attempts build on past efforts and identifies a phase transition separating the dynamics of failure into regions of progression or stagnation. Above the critical point, agents exploit incremental refinements to systematically advance towards success, whereas below it, they explore disjoint opportunities without a pattern of improvement (see abstracted figure 2 and 3 below).

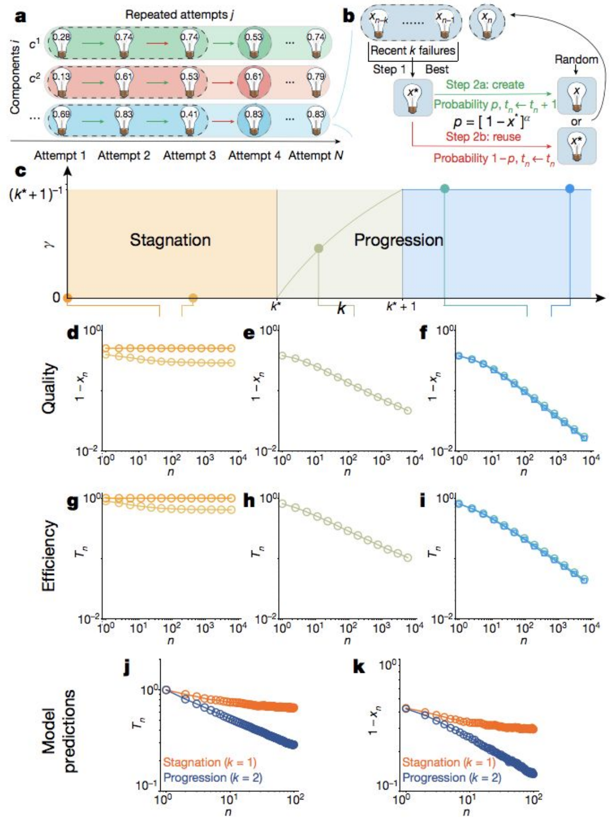


Fig. 2 | The *k* model. a, We treat each attempt as a combination of many independent components (c^i). For attempt *j*, each component *i* is characterized by an evaluation score x^i_j, which falls between 0 and 1. The score for a new version is often unknown until attempted, hence a new version is assigned a score, drawn randomly from the range 0–1. **b**, To formulate a new attempt, one can either create a new version (with probability *p*, green arrow) or reuse an existing version by choosing the best one among past versions x^* (with probability $1-p$, red arrow). $P(x \geq x^*) = 1 - x^*$ captures the potential to improve on prior versions, prompting us to assume $p = (1-x^*)^\alpha$ where $\alpha > 0$ characterizes the propensity of an agent to create new versions given the quality of existing ones. **c**, The analytical solution of the model reveals that the system is separated into three regimes by two critical points k^* and k^*+1. The solid line shows the extended solution space of our analytical results. **d–i**, Simulation results from the model ($\alpha = 0.6$) for quality (**d–f**) and efficiency (**g–i**) trajectories for different *k* parameters, showing distinct dynamical behaviour in different regimes. All results are based on simulations averaged over 10^4 times. **j, k**, A phase transition around k^* predicts the coexistence of two groups that fall in the stagnation and progression regimes, respectively.

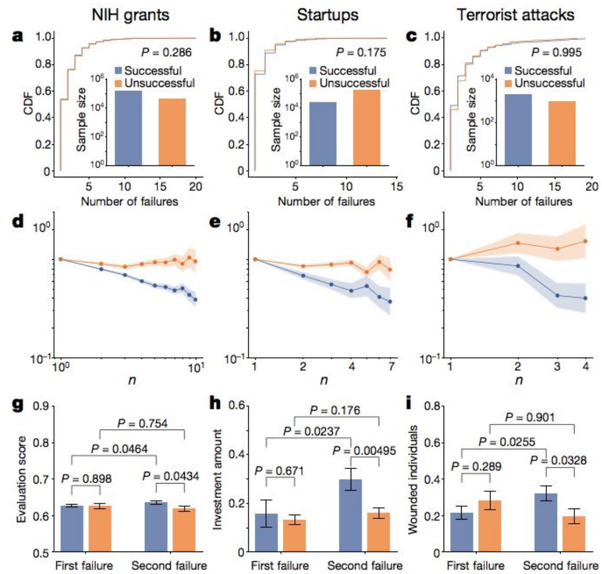


Fig. 3 | Testing model predictions. a–c, Cumulative distribution function (CDF) of the number of consecutive failures before the last attempt for successful and unsuccessful groups. To eliminate the possibility that agents were simply in the process of formulating their next attempt, we focus on cases for which it has been at least five years since their last failure. In each of our three datasets, the two distributions are statistically indistinguishable (Kolmogorov–Smirnov test for samples with at least one failures). For clarity, here we show results for less than 21 failures (Supplementary Information 5.2). Inset, the sample size of successful and unsuccessful groups, showing their size is of a similar order of magnitude. **d–f**, Early temporal signals separate successful and unsuccessful groups. **d**, $n = 43{,}705$ (successful), 15,132 (unsuccessful). **e**, $n = 2{,}455$ (successful), 16,656 (unsuccessful). **f**, $n = 446$ (successful), 321 (unsuccessful). For each group, we measure the average inter-event time between two failures $T_n \equiv t_n/t_1$ as a function of the number of attempts. Dots and shaded areas are mean ± s.e.m. measured from data (Supplementary Information 5.3). All successful groups manifest power-law scaling $T_n \sim n^{-\gamma}$ (Extended Data Fig. 2). The two groups show distinguishable temporal dynamics for $n = 2$. Two-sided Welch's *t*-test; $P = 3.02 \times 10^{-4}$, 7.18×10^{-3}, 9.42×10^{-2} for comparisons of successful and unsuccessful groups in **d**, **e**, **f** respectively. This temporal scaling is absent for unsuccessful groups. **g–i**, Performance at first attempt appears indistinguishable between successful and unsuccessful groups that experienced a large number of consecutive failures before the last attempt (at least 5 for D_1, 3 for D_2 and 2 for D_3, two-sided Welch's *t*-test), but becomes distinguishable at the second attempt (two-sided Welch's *t*-test). Whereas performance improves for the successful group (one-sided Welch's *t*-test), this improvement is absent for the unsuccessful group (one-sided Welch's *t*-test). Data are mean ± s.e.m. **g**, $n = 628$, 145, 571, 123 (from left to right). **h**, $n = 248$, 1,332, 237, 1,312 (from left to right). **i**, $n = 231$, 173, 229, 174 (from left to right).

The model makes several empirically testable predictions, demonstrating that those who eventually succeed and those who do not may initially appear similar, but can be characterized by fundamentally distinct failure dynamics in terms of the efficiency and quality associated with each subsequent attempt. We collected large-scale data tracing repeated attempts by investigators to obtain National Institutes of Health (NIH) grants to fund their research. Together, these findings unveil detectable yet previously unknown early signals that enable us to identify failure dynamics that will lead to ultimate success or failure. Given the ubiquitous nature of failure and the paucity of quantitative approaches to understand it, these results represent an initial step towards the deeper understanding of the complex dynamics underlying failure.

Yin, Yian, Yang Wang, James Evans & Dashun Wang. 2019. "Quantifying dynamics of failure across science, startups, and security." *Nature* 575: 190-194.

We also explored the role of setbacks in a successful scientific career in the context of junior scientists applying for National Institutes of Health R01 grants. By focusing on proposals fell just below and just above the funding threshold, we compare near-miss with narrow-win applicants, and find that an early-career setback has powerful, opposing effects. On the one hand, it significantly increases attrition, predicting more than a 10% chance of disappearing permanently from the NIH system.

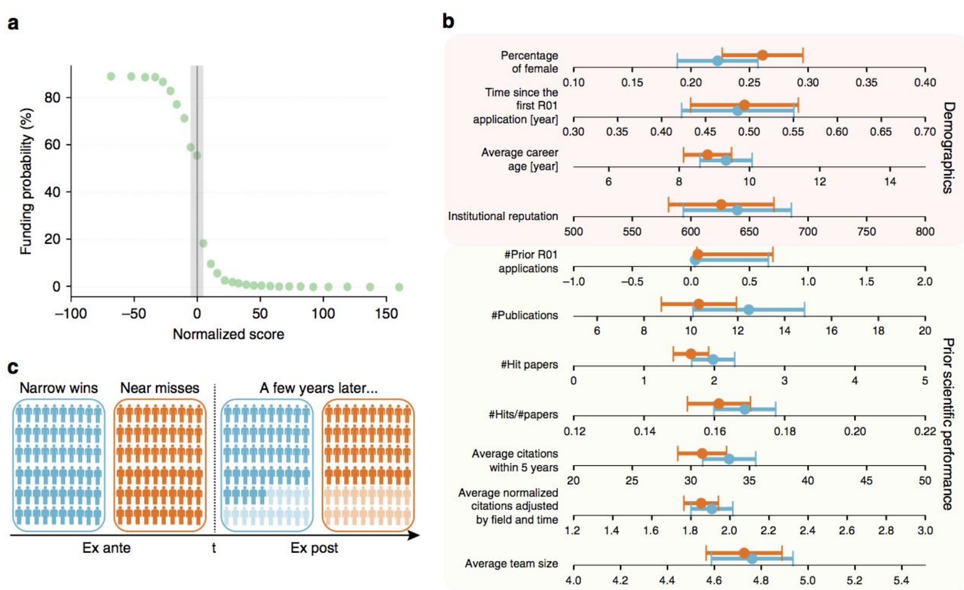


Fig. 1 Pre-treatment comparisons between the narrow-win and near-miss applicants. **a** Relationship between normalized score and award status. Funding probability shows a clear transition around the funding threshold. We focus only on junior PIs whose normalized scores lie within the range (−5, 5), the shaded gray area, which includes 561 narrow-win and 623 near-miss applicants in our sample. **b** Pre-treatment feature comparisons between the near-miss and narrow-win group. We compared 11 different demographic and performance characteristics. The features are defined as follows (from top to bottom): (1) percentage of female applicants; (2) number of years since the first R01 application; (3) number of years since the first publication; (4) institutional reputation, measured by the number of R01 grants awarded to an institution between 1990 to 2005; (5) number of previous R01 applications; (6) number of publications prior to treatment; (7) number of prior papers that landed within the top 5% of citations within the same field and year; (8) probability of publishing a hit paper; (9) average citations papers received within 5 years of publication; (10) citations normalized by field and time;[34] and (11) average team size across prior papers. We see no significant difference between the two groups across any dimension we measured; Error bar represents the 95% confidence interval. **c** An illustrative example of the underlying process. Solid color indicates people who remained active, whereas shaded color denotes the fraction that disappeared from the NIH system. Blue and orange indicate narrow-win and near-miss applicants, respectively

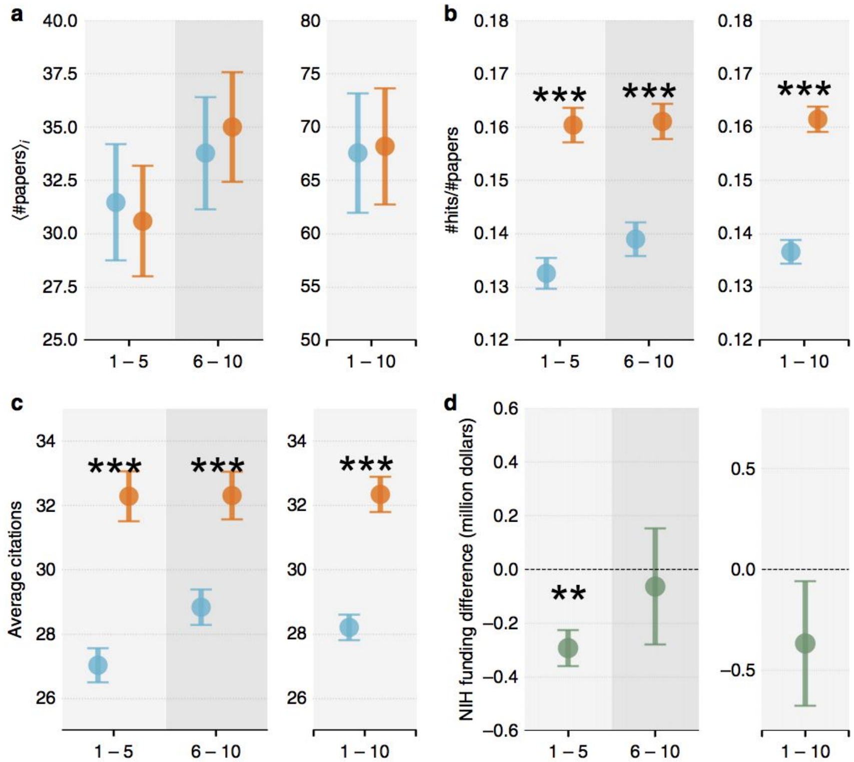


Fig. 2 Comparing future career outcome between near misses (orange) and narrow wins (blue). **a** The average number of publications per person. **b** Near misses outperformed narrow wins in terms of the probability of producing hit papers in the next 1–5 years, 6–10 years, and 1–10 years. Note that there appears a slight performance improvement for the narrow-win group in the second five-year period, but the difference is not statistically significant (χ^2-test p-value > 0.1, odds ratio = 1.05). **c** Average citations within 5 years of publication. The near-miss applicants again outperformed their narrow-win counterparts. To ensure all papers have at least 5 years to collect citations, here we used data from 1990 to 2000 to avoid any boundary effect. **d** Funding difference between the near-miss and narrow-win group from the NIH (near misses minus narrow wins). ***$p < 0.001$, **$p < 0.05$, *$p < 0.1$; Error bars represent the standard error of the mean

Yet, despite an early setback, individuals with near misses systematically outperform those with narrow wins in the longer run. Moreover, this performance advantage seems to go beyond a screening mechanism, suggesting early-career setback appears to cause a performance improvement among those who persevere. Overall, these findings are consistent with the concept that "what doesn't kill me makes me stronger," which may have broad implications for identifying, training and nurturing junior scientists.

Yang Wang, Benjamin F. Jones, and Dashun Wang. 2019. Early-Career Setback and Future Career Impact, *Nature Communications*.

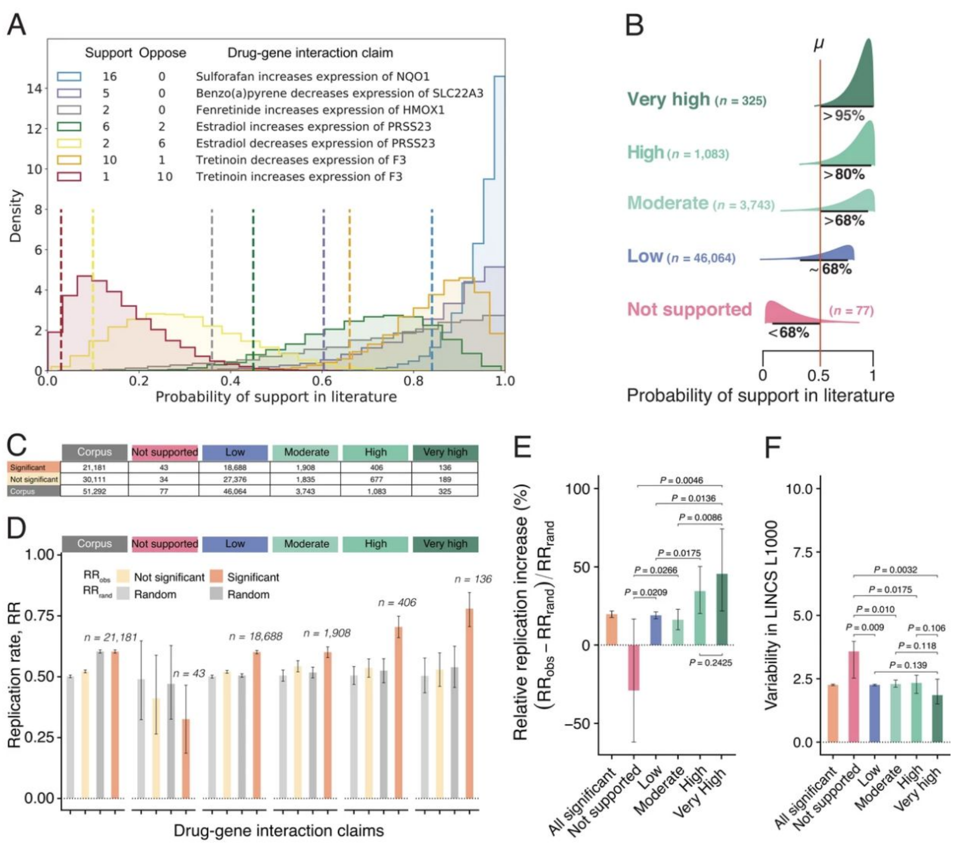
Finally, we explored the nature of failed innovation and its consequences.

He, Zhongyang, Zhen Lei, Yang Wang, and Dashun Wang. "Diamond in the rough: Quantifying failed innovation endeavors." (Pending NIH approval for submission)

The Limits of Prediction

We also explore the limits of robustness and the degree to which the scientific literature predicts its own future knowledge at scale for the first time. We identify a large sample of published drug-gene interaction claims curated in the Comparative Toxicogenomics Database (for example, benzo(a)pyrene decreases expression of SLC22A3) and evaluate these claims by connecting them with high-throughput experiments from the LINCS L1000 program. Our sample included 60,159 supporting findings and 4253 opposing findings about 51,292 drug-gene interaction claims in 3363 scientific articles. We show that claims reported in a single paper replicate 19.0% more frequently than expected, while claims reported in multiple papers replicate 45.5% more frequently than expected. We also analyze the subsample of interactions with two or more published findings, and show that centralized scientific communities, which use similar methods and involve shared authors who contribute to many articles, propagate less replicable claims than decentralized communities, which use more diverse methods and contain more independent teams. Our findings suggest how policies that foster decentralized collaboration will increase the robustness of scientific findings in biomedical research.

Danchev, Valentin, Andrey Rzhetsky & James Evans. 2019. “Centralized communities more likely generate non-replicable results.” *eLife* 8:e43094 DOI: 10.7554/eLife.43094



	Corpus	Not supported	Low	Moderate	High	Very high
Significant	21,181	43	18,688	1,908	406	136
Not significant	30,111	34	27,376	1,835	677	189
Corpus	51,292	77	46,064	3,743	1,083	325

Estimates of claim replication as a function of the probability of support in the literature and generalizability across high-throughput experiments.

(**A**) Posterior distributions of probability of support in the biomedical literature for a sample of seven DGI claims for which there are at least two findings (supporting and/or opposing). Note that the top three claims receive only supporting findings in the literature, whereas the fourth and fifth claims are opposites (so papers that support the fourth claim oppose the fifth claim, and vice versa), and likewise for the sixth and seventh claims. We obtained model estimates for each claim by performing 10,000 Markov chain Monte Carlo (MCMC) sampling iterations (see Materials and methods). For each claim, we summarize the probability of support (dashed vertical line) using the lower bound on the one-sided 95% posterior credible interval: this value ranges from 0.84 for a claim that is supported by 16 findings and opposed by no findings, to 0.03 for a claim that is supported by one finding and opposed by 10 findings. (**B**) DGI claims in the literature can be categorized into one of five classes of support (Very High; High; Moderate; Low; Not Supported) on the basis of distributions like those in panel A; the number of claims included in each class is shown in brackets. (**C**) Number of DGI claims that are significant (second row) and not significant (third row) at the 0.05 level in the LINCS L1000 dataset for the whole corpus (second column) and for each of the five classes of support in the literature (columns 3–7). (**D**) Observed replication rates (RR_{obs}) and expected replication rates (RR_{rand}) for claims that are significant and non-significant in the LINCS L1000 dataset for the whole corpus (left) and for each of the five classes of support in the literature. (**E**) The relative replication increase rate ($RRI = 100 \times \frac{RR_{\text{obs}} - RR_{\text{rand}}}{RR_{\text{rand}}}$) for claims that are significant in the LINCS L1000 dataset (left) and for each of the five classes of support in the literature. (**F**) Variability (coefficient of variation) in the LINCS L1000 dataset across cell lines, durations and dosages for claims that are significant in this dataset (left) and for each of the five classes of support in the literature. Statistical significance and error bars were determined by bootstrapping (see Materials and methods). All error bars represent 95% CI.

We also explored how breakthrough discoveries and inventions involve unexpected combinations of contents including problems, methods, and natural entities, and also diverse contexts such as journals, subfields, and conferences. Drawing on data from tens of millions of research papers, patents, and researchers, we construct models that predict more than 95% of next year's content and context combinations with embeddings constructed from high-dimensional stochastic block models, where the improbability of new combinations itself predicts up to half of the likelihood that they will gain outsized citations and major awards. Most of these breakthroughs occur when problems in one field are unexpectedly solved by researchers from a distant other. These findings demonstrate the critical role of surprise in advance, and enable evaluation of scientific institutions ranging from education and peer review to awards in supporting it.

Shi, Feng and James Evans. "Science and Technology Advance through Surprise"

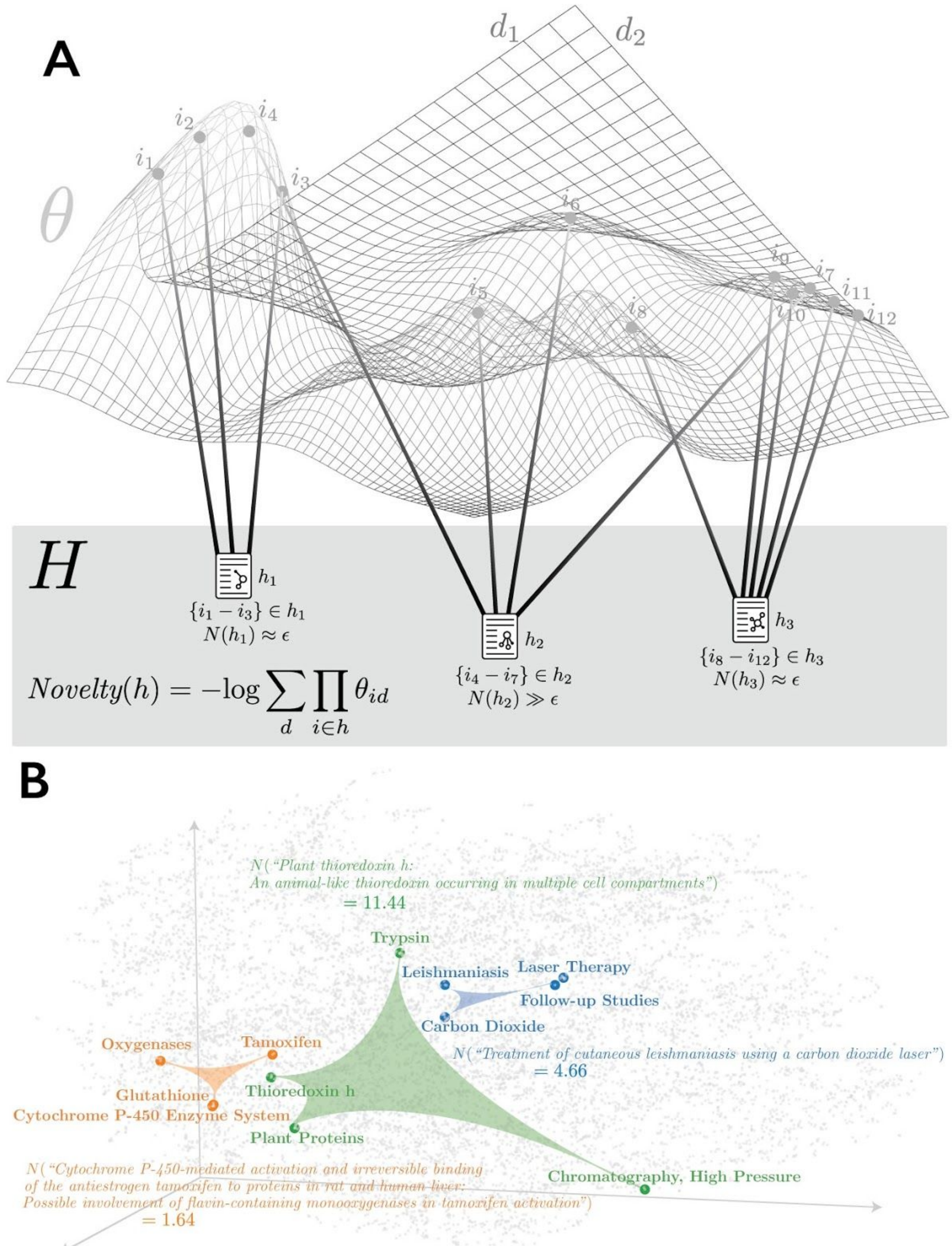


Figure 1. (A) Illustration of the manifold inscribing all topics θ and an evaluation of three articles or patents (hyperedges $h_{1\text{-}3}$) in terms of their novel combinations. Articles/patents h_1 and h_3 represent projects that combine scientific or technical components near one another in θ, making each of high probability and low (ε) novelty—similar to many related papers from the past. By contrast, paper h_2 draws a novel combination of components unlike any paper from the past, making it of low probability and high ($\gg \varepsilon$) novelty. **(B)** Actual three dimensional projection of the manifold best inscribing all MeSH codes from MEDLINE articles

in our analysis. Also included are MeSH terms in the most novel article (blue), the least novel article (orange), and a random article in between (green) among all articles including four MeSH terms.

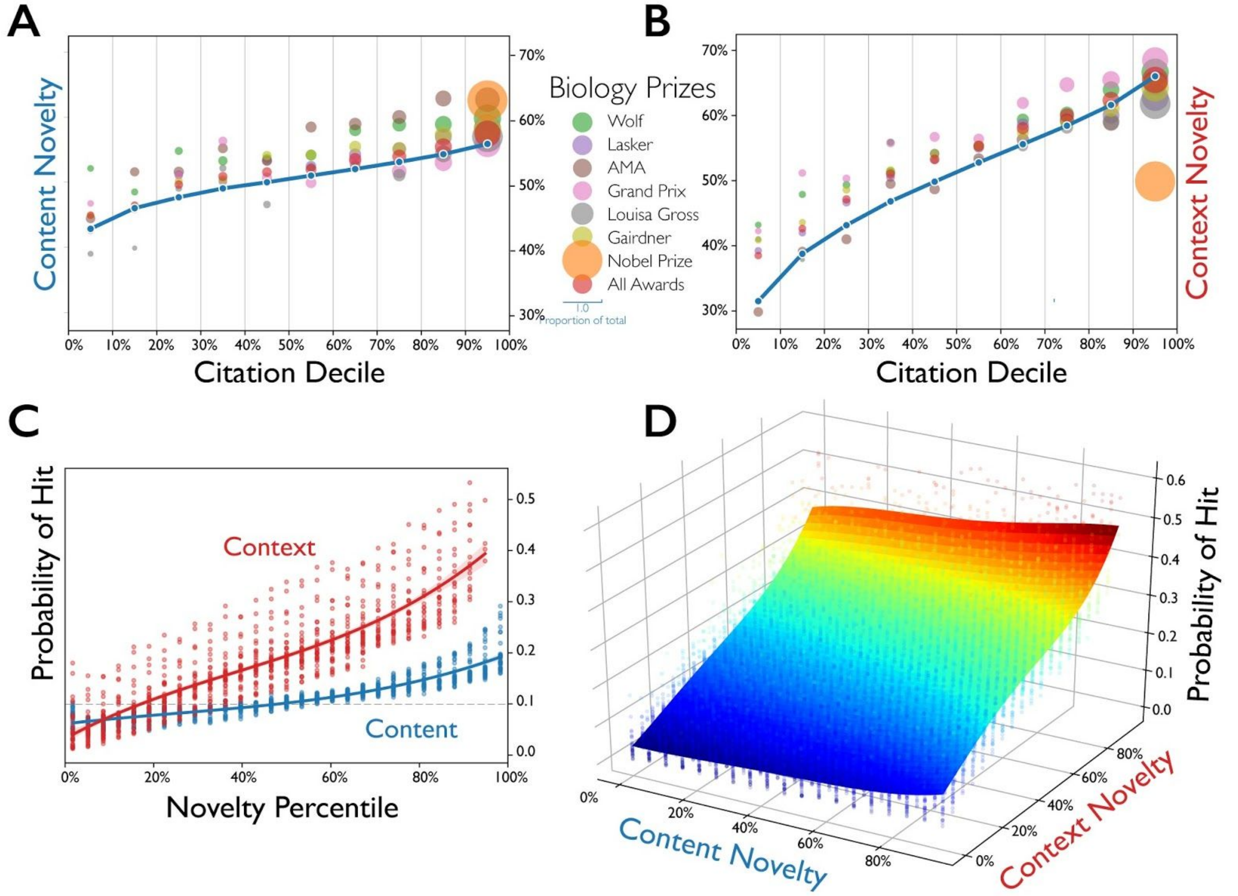
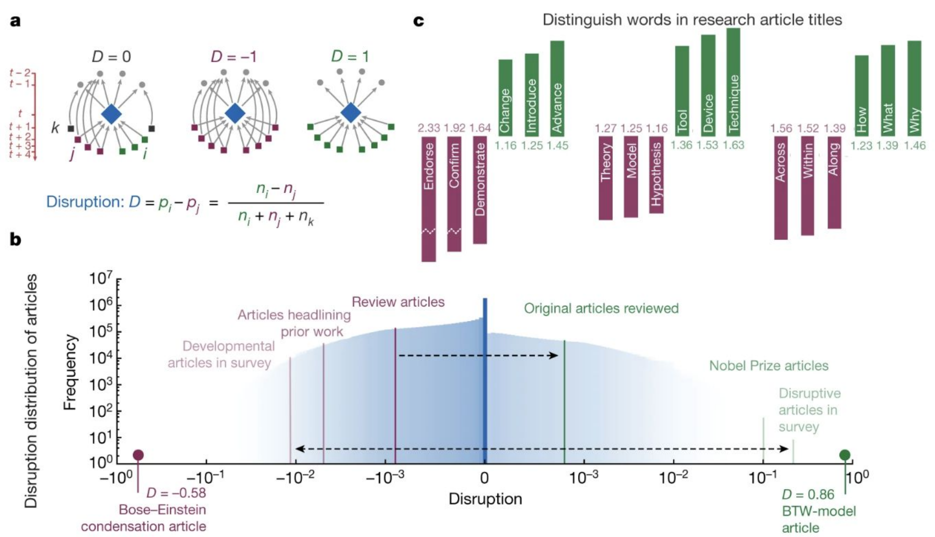


Figure 2: Average content and context novelty for each decile of citations, tracing a monotonic rise; Including average for Nobel prizes in Physiology or Medicine, Chemistry (first row). Probability of being a hit paper as a function of content and context novelty separately (row 3-5, left) and jointly (row 3-5, right). Third row shows results for MEDLINE data, fourth row for APS data, and bottom row for the USPTO data. For each, bivariate distribution of content and context novelty across articles or patents on the left.

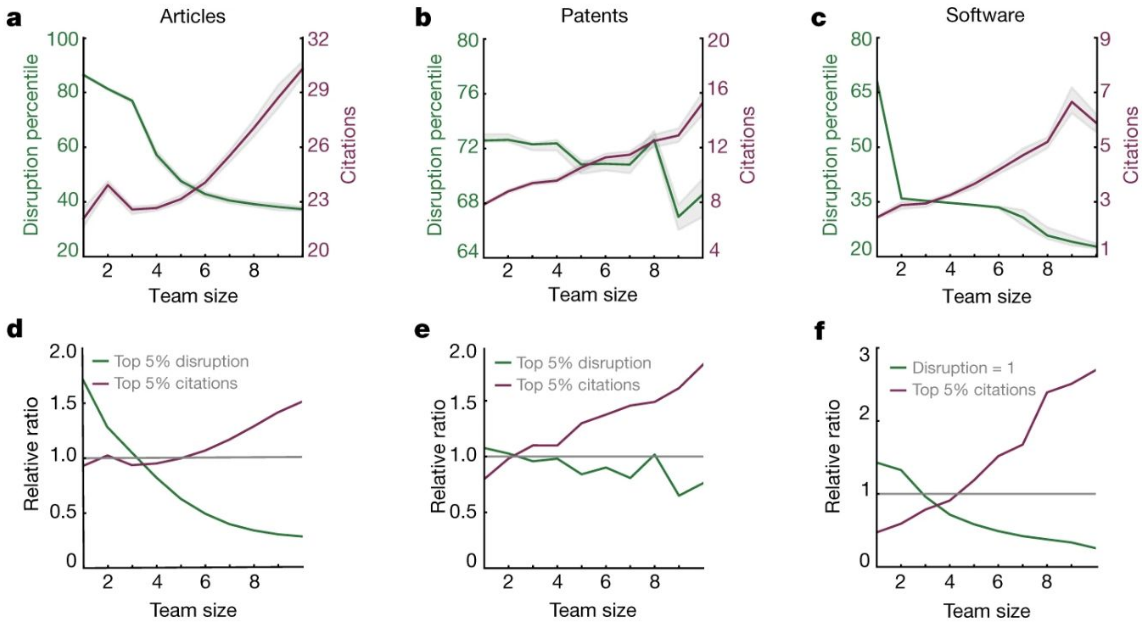
Predictability in Scientific Careers and Teams

One of the most universal trends in science and technology today is the growth of large teams in all areas, as solitary researchers and small teams diminish in prevalence. Increases in team size have been attributed to the specialization of scientific activities, improvements in communication technology, or the complexity of modern problems that require interdisciplinary solutions. This shift in team size raises the question of whether and how the character of the science and technology produced by large teams differs from that of small teams. Here we analyse more than 65 million papers, patents and software products that span the period 1954–2014, and demonstrate that across this period smaller teams have tended to disrupt science and technology with new ideas and opportunities, whereas larger teams have tended to develop existing ones. Work from larger teams builds on more-recent and popular developments, and attention to their work comes immediately. By contrast, contributions by smaller teams search more deeply into the past, are viewed as disruptive to science and technology and succeed further into the future—if at all. Observed differences between small and large teams are magnified for higher-impact work, with small teams known for disruptive work and large teams for developing work. Differences in topic and research design account for a small part of the relationship between team size and disruption; most of the effect occurs at the level of the individual, as people move between smaller and larger teams. These results demonstrate that both small and large teams are essential to a flourishing ecology of science and technology, and suggest that, to achieve this, science policies should aim to support a diversity of team sizes.

Wu, Lingfei, Dashun Wang & James A. Evans. 2019. "Large Teams Develop Science and Technology, Small Teams Disrupt It." *Nature* 566: 378-382. [Cover Article]



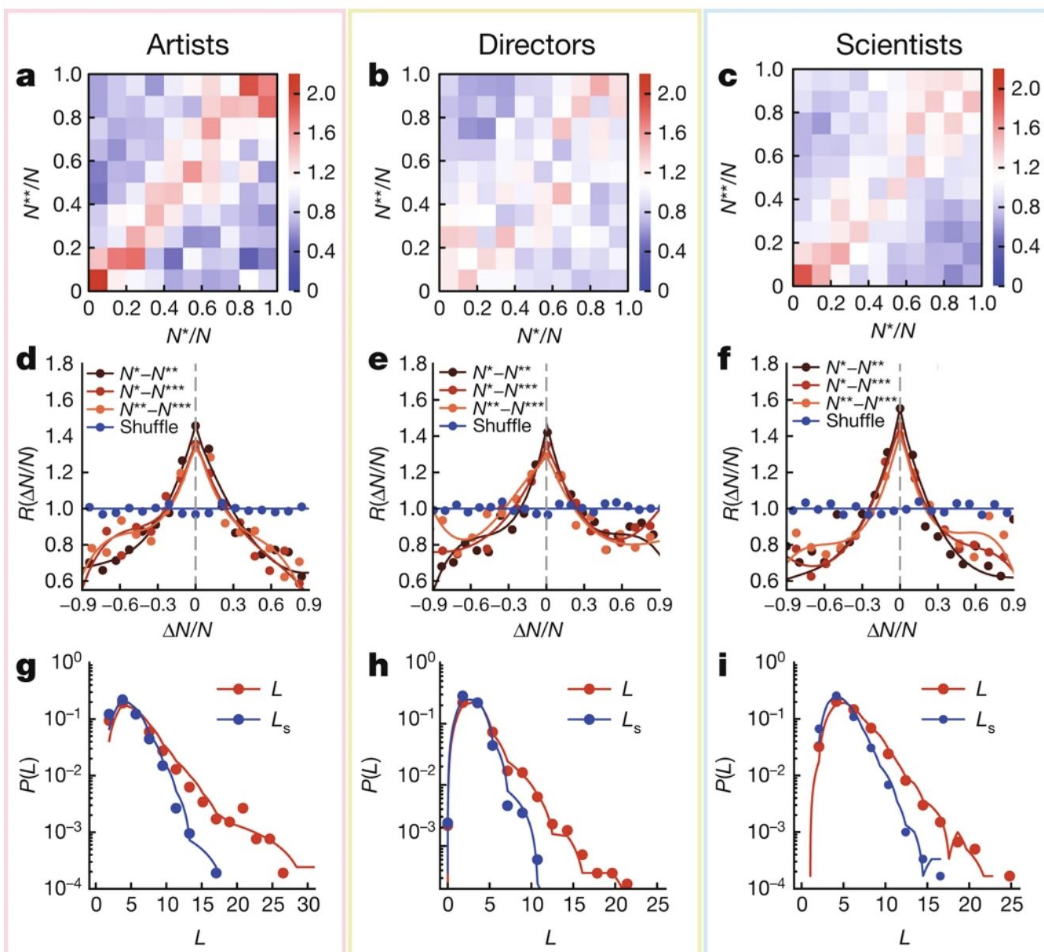
a, Simplified illustration of disruption. Three citation networks comprising focal papers (blue diamonds), references (grey circles) and subsequent work (rectangles). Subsequent work may cite the focal work (*i*, green), both the focal work and its references (*j*, red) or just its references (*k*, black). Disruption, *D*, of the focal paper is defined by the difference between the proportion of type *i* and *j* papers $p_i - p_j$, which equals the difference between the observed number of these papers $n_i - n_j$ divided by the number of all subsequent works $n_i + n_j + n_k$. A paper may be disrupting ($D = 1$), neutral ($D = 0$) or developing ($D = -1$). **b**, The distribution of disruption across 25,988,101 WOS journal articles published between 1900 and 2014. On this distribution, we mark the BTW-model ($D = 0.86$, top 1%) and Bose–Einstein condensation articles ($D = -0.58$, bottom 3%) along with several samples used to validate *D* (Methods, Supplementary Tables 1–3). This includes (1) 104 'disruptive' articles (disruption mean $E(D) = 0.215$, top 2%) and 86 'developing' articles ($E(D) = -0.011$, bottom 13%) nominated by a surveyed panel of 20 scholars across fields; (2) 877 Nobel-prize-winning papers published between 1902 and 2009 ($E(D) = 0.10$, top 2%); (3) 22,672 review articles ($E(D) = -0.0009$, bottom 46%) and 1,338,808 original research articles that they review ($E(D) = 0.0008$, top 23%); and (4) 148,303 articles that headline prominent prior work by mentioning one or more cited authors in the title ($E(D) = -0.0049$, bottom 24%). **c**, We select titles from 24,174,022 articles published between 1954 and 2014 and assign them to one of two groups, disrupting ($D > 0$) or developing ($D < 0$) articles. For the 1,033,879 words observed in both groups, we calculate the ratio of frequency in disrupting versus developing articles, *r*. We visualize differences in the content and writing style between these two groups in terms of verbs, nouns, and adverbs and prepositions (from left to right). To facilitate comparison, we visualize *r* in green if $r > 1$, and $1/r$ in red otherwise.



a–**c**, For research articles (24,174,022 WOS articles published between 1954 and 2014), patents (2,548,038 US patents assigned between 2002 and 2014) and software (26,900 GitHub repositories uploaded between 2011 and 2014), median citations (red curves, indexed by right *y* axis) increase with team size whereas the average disruption percentile (green curves, indexed by left *y* axis) decreases with team size. For all datasets, we present work with one or more citations. Teams of between 1 and 10 authors account for 98% of articles, 99% of patents and 99% of code repositories. Bootstrapped 95% confidence intervals are shown as grey zones. Extended Data Figure 3a shows that observed relationships hold for two orders of magnitude of team size. **d**–**f**, As in **a**–**c** but for extreme cases rather than for average behaviour. Relative ratios compare the observed proportion of teamwork being extremely (top 5%) disruptive or impactful (measured with citations) against a constant baseline (grey line $y = 1$), which indicates a situation in which the most disruptive and impactful work is distributed equally across team sizes. We find that the probability of observing papers, patents and products of highest impact increases with team size (Kolmogorov–Smirnov statistics and probabilities for all team sizes plotted in Extended Data Fig. 2f), whereas the probability of observing the most disruptive work decreases with team size (t-statistics and probabilities for all team sizes plotted in Extended Data Fig. 2c). For example, **d** shows that the percentage of top 5% disruptive papers depends on team size, with 8.6% contributed by single authors and only 1.4% contributed by teams of ten authors. This posts relative ratios of 8.6/5 = 1.72 and 1.4/5 = 0.28, respectively. For software, 69% of the codebases have disruption values that equal 1; we therefore use this maximum value instead of the top 5%.

We also explored hot streaks in science—periods during which an individual's performance is substantially better than his or her typical performance. Little was previously known about whether they apply to individual careers. Here, building on rich literature on the lifecycle of creativity, we collected large-scale career histories of individual scientists, tracing the scientific publications they produced. We find that hit works within a career show a high degree of temporal regularity, with each career being characterized by bursts of high-impact works occurring in sequence. We demonstrate that these observations can be explained by a simple hot-streak model, allowing us to probe quantitatively the hot streak phenomenon governing individual careers. We find this phenomemon to be remarkably universal. The hot streak emerges randomly within an individual's sequence of works, is temporally localized, and is not associated with any detectable change in productivity. We show that, because works produced during hot streaks garner substantially more impact, the uncovered hot streaks fundamentally drive the collective impact of an individual, and ignoring this leads us to systematically overestimate or underestimate the future impact of a career. These results not only deepen our quantitative understanding of patterns that govern individual ingenuity and success, but also may have implications for identifying and nurturing individuals whose work will have lasting impact.

Lu Liu, Yang Wang, Roberta Sinatra, C. Lee Giles, Chaoming Song, and Dashun Wang. 2018. "Hot Streaks in Artistic, Cultural, and Scientific Careers". *Nature*.

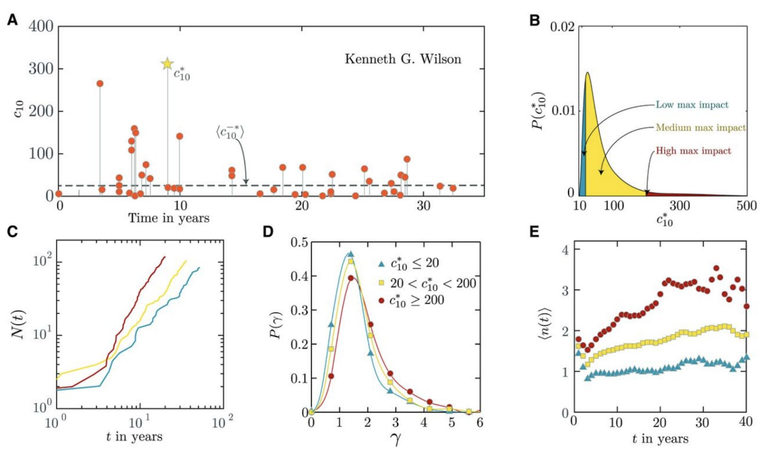


a–**c**, $\phi(N^*, N^{**})$, colour coded, measures the joint probability of the two highest-impact works within a career for artists (**a**), directors (**b**), and scientists (**c**). $\phi > 1$ indicates that two hits are more likely to colocate than would be expected at random. **d**–**f**, $R(\frac{\Delta N}{N})$ measures the temporal distance between highest-impact works relative to the null model's prediction. Real careers show a clear peak around 0 (red dots), which is well captured by the hot-streak model (solid lines). Different shades of red correspond to different pairs of hit works. Blue dots denote the same measurement but on shuffled careers, and blue lines are predictions from shuffled careers generated by our model. **g**–**i**, The distribution of the length of streaks $P(L)$ for real careers and $P(L_s)$ for shuffled careers. The hot-streak model (red lines) and its shuffled version (blue lines) closely reproduce $P(L)$ observed in real (red dots) and shuffled careers (blue dots).

Despite the frequent use of numerous quantitative indicators to gauge the professional impact of a scientist, little was previously known about how scientific impact emerges and evolves in time. In this paper, we quantify the changes in impact and productivity throughout a career in science, finding that impact, as measured by influential publications, is distributed randomly within a scientist's sequence of publications. This random-impact rule allows us to formulate a stochastic model that uncouples the effects of productivity, individual ability, and luck and unveils the existence of universal patterns governing the emergence of scientific success. The model assigns a unique individual parameter Q to each scientist, which is stable during a career, and it accurately predicts the evolution of a scientist's impact, from the h-index to cumulative citations, and independent recognitions, such as prizes.

Roberta Sinatra, Dashun Wang, Pierre Deville, Chaoming Song, and Albert-Laszlo Barabasi. 2016. "Quantifying the evolution of individual scientific impact", *Science*, 354: 6312.

Wang, Dashun. "The Science of Career: When Do You Do Your Best Work. *Scientific American* (With the editor. Accepted by the editorial board).



(**A**) Publication history of Kenneth G. Wilson (Nobel Prize in Physics, 1982). Horizontal axis indicates the number of years after the scientist's first publication, and each vertical line corresponds to a research paper. The height of each line corresponds to c_{10}, that is, the number of citations the paper received after 10 years (sections S1.3 and S1.6). The highest-impact paper of Wilson was published in 1974, 9 years after his first publication, and it is the 17th of his 48 papers; hence, $t^* = 9$, $N^* = 17$, and $N = 48$. (**B**) Distribution of the highest-impact paper $P(c^*_{10})$ across all scientists. We highlight in blue the bottom 20% of the area, corresponding to low maximum impact scientists $(c^*_{10} \le 20)$; red area indicates the high maximum impact scientists (top 5%, $c^*_{10} \ge 200$); yellow corresponds to the remaining 75% medium maximum impact scientists $(20 < c^*_{10} < 200)$. These cutoffs do not change if we exclude review papers from our analysis (see figs. S4 and S36). (**C**) Number of papers $N(t)$ published up to time t for three scientists with low, medium, and high impact but with comparable final number of papers throughout their career. (**D**) Distribution of the productivity exponents γ (*18*). The productivity of high-impact scientists grows faster than does that of low-impact scientists. (**E**) Dynamics of productivity, as captured by the average number of papers $\langle n(t) \rangle$ published each year for high-, average-, and low-impact scientists. $t = 0$ corresponds to the year of a scientist's first publication.

We also sought to understand quantitatively how scientists choose and shift their research focus over time, because it affects the ways in which scientists are trained, science is funded, knowledge is organized and discovered, and excellence is recognized and rewarded. Despite extensive investigation into various factors that influence a scientist's choice of research topics, quantitative assessments of mechanisms that give rise to macroscopic patterns characterizing research-interest evolution of individual scientists remain limited. Here we perform a large-scale analysis of publication records, and we show that changes in research interests follow a reproducible pattern characterized by an exponential distribution. We identify three fundamental features responsible for the observed exponential distribution, which arise from a subtle interplay between exploitation and exploration in research-interest evolution. We developed a random-walk-based model, allowing us to accurately reproduce the empirical observations. This work uncovers and quantitatively analyses macroscopic patterns that govern changes in research interests, thereby showing that there is a high degree of regularity underlying scientific research and individual careers.

Tao Jiaz, Dashun Wang, and Boleslaw K. Szymanskiz. 2017. "Quantifying patterns of research-interest evolution". *Nature Human Behaviour* 1: 0078.

Yian Yin, and Dashun Wang. 2017. "The time dimension of science: Connecting the past to the future". *Journal of Informetrics* 11.2: 608-621.

We also published several other papers on how conflict in crowds could lead to higher performance if the diversity is correlated with the nature of relevant content.

Shi, Feng, Misha Teplitskiy, Eamon Duede, James A. Evans. 2019. "The Wisdom of Polarized Crowds." *Nature Human Behaviour*, Mar 4: 1.

Predictability from Components

We also have pieces published or under review that explore the ways in which the components of science and technological systems predict the future, whether the systems are best characterized by component substitutions, as in technological platforms, or complex combinations, as in science.

Ching Jin, Chaoming Song, Johannes Bjelland, Geoffrey Canright, Dashun Wang, "Emergence of Scaling in Complex Substitutive Systems". *Nature Human Behaviour*, 2019. [Cover Article]

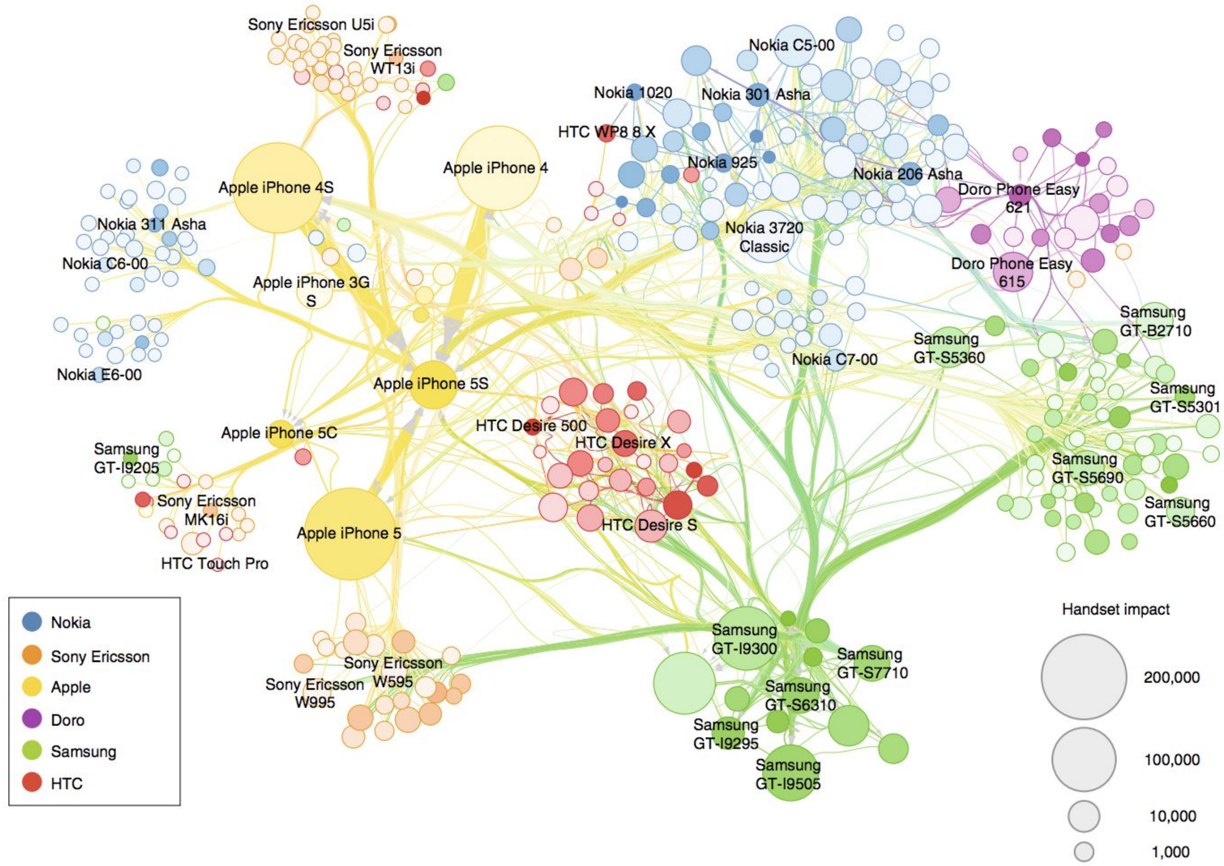


Fig. 2 | Empirical substitution network. We used the backbone extraction method[42] to construct a substitution network, capturing substitution patterns among handsets aggregated within a 6-month period (January–June 2014). Each node corresponds to one type of handset released before 2014 by one of the six major manufacturers. The node size captures its popularity, measured by the number of users of the particular handset at the time. Handsets are coloured based on their manufacturers (node colouring), which fade with the age of the handsets. If users substituted handset *i* with *j*, we add a weighted arrow pointing from *i* to *j*. The link weight captures the total substitution volumes between two handsets within the 6-month period. As the full network is too dense to visualize, here we only show the statistically significant links as identified by the method proposed in ref. [42] for a *P* value of 0.05. We colour the links based on the colour of the substituting handset. The network vividly captures the widespread transitions from feature handsets to smartphones. Indeed, most cross-manufacturer substitution links are either yellow or green, indicating their substitutions by iPhones or Android handsets. Substitution patterns are also highly heterogeneous. A few pairs of handsets have high substitution volumes, for example, between the successive generations of iPhones, but most substitutions are characterized by rather limited volumes. The structural complexity shown is further coupled with a high degree of temporal variability. Indeed, the system turns into a widely different configuration every year, even for the most dominant handsets (Supplementary Fig. 18b–e).

Shi, Feng and James Evans. "Science and Technology Advance through Surprise"

Scientific Success

We published several pieces that explore extreme scientific success in the context of major scientific awards, revealing that Nobel laureates reveal some patterns that are the same as the rest of scientists in some ways, but they work and achieve in contexts that are become quite different.

Jichao Li, Yian Yin, Santo Fortunato, and Dashun Wang, "Nobel laureates are almost the same as us". *Nature Reviews Physics*, 2019.

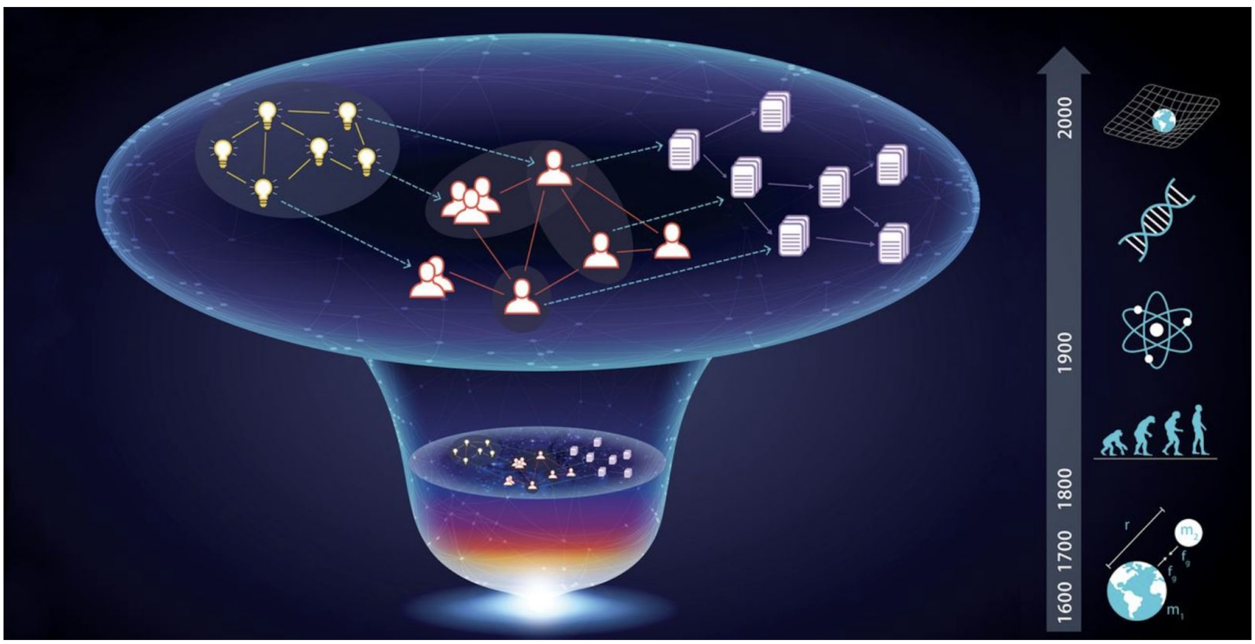
Jichao Li, Yian Yin, Santo Fortunato, and Dashun Wang, "A dataset of publication records for Nobel laureates”. *Scientific Data*, 2019

Li, Jichao, Yian Yin, Santo Fortunato, and Dashun Wang. “Scientific Elite revisited: Patterns of productivity, collaboration, authorship, and impact.” (submitted to *Nature Human Behaviour*)

Surveys of the Science of Science and New Scientific Directions

We published several reviews of the science of science, and others that drew on emerging science of science tools and approaches to review other critical scientific and technological fields, including physics and artificial intelligence.

Fortunato, Santo, Carl T. Bergstrom, Katy Börner, James A. Evans, Dirk Helbing, Staša Milojević, Alexander M. Petersen, Filippo Radicchi, Roberta Sinatra, Brian Uzzi, Alessandro Vespignani, Ludo Waltman, Dashun Wang, Albert-László Barabási. 2018. “Science of science.” *Science* 359(6379): eaao0185, doi: 10.1126/science.aao0185.



Science can be seen as an expanding and evolving network of ideas, scholars, and papers. SciSci searches for universal and domain-specific laws underlying the structure and dynamics of science.

Federico Battiston, Federico Musciotto, Dashun Wang, Albert-Laszlo Barabasi, Michael Szell, Roberta Sinatra. 2019. “Taking census of physics”, *Nature Reviews Physics*, 1, 89–97.

Foster, Jacob G., and James A. Evans. 2019 “Promiscuous Inventions: Modeling Cultural Evolution with Multiple Inheritance.” *Beyond the Meme*. Ed. William C. Wimsatt and Alan Love. Minneapolis: University of Minnesota Press.

Sinatra, Roberta, Pierre Deville, Michael Szell, Dashun Wang, and Albert-Laszlo Barabasi. 2015. “A Century of Physics”, *Nature Physics*, 11.10:791-796. [Cover Article].

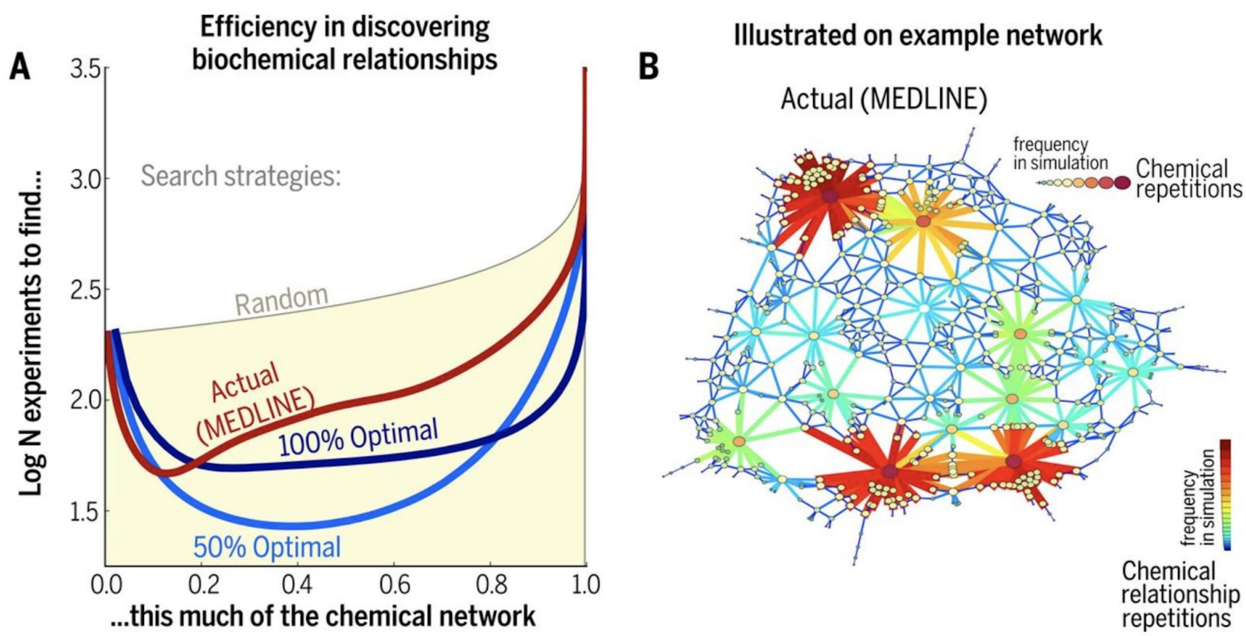
Wang, Dashun and Albert-Laszlo Barabasi, *The Science of Science*. (Cambridge University Press, Forthcoming, 2020).

Foster, Jacob and James Evans, *Knowing*. (Princeton University Press, Forthcoming, 2021)

Improving the Scientific Method through Prediction

We published several papers that proposed improvements to the scientific method through improved prediction. For example, we explored how we could use the scientific literature to better predict what are the most fruitful areas for future experimentation.

Rzhetsky, Andrey, Jacob Foster, Ian Foster and James Evans. 2015. “Choosing Experiments to Accelerate Discovery.” Issue cover: “Engineering of Biology and Medicine”. *Proceedings of the National Academy of Sciences* 112(47):14569–14574, doi:10.1073/pnas.1509757112



Choosing experiments to accelerate collective discovery.
(A) The average efficiency rate for global strategies to discover new, publishable chemical relationships, estimated from all MEDLINE-indexed articles published in 2010. This model does not take into account differences in the difficulty or expense of particular experiments. The efficiency of a global scientific strategy is expressed by the average number of experiments performed (vertical axis) relative to the number of new, published biochemical relationships (horizontal axis), which correspond to new connections in the published network of biochemicals co-occurring in MEDLINE-indexed articles. Compared strategies include randomly choosing pairs of biochemicals, the global (“actual”) strategy inferred from all scientists publishing MEDLINE articles, and optimal strategies for discovering 50 and 100% of the network. Lower values on the vertical axis indicate more efficient strategies,

showing that the actual strategy of science is suboptimal for discovering what has been published. The actual strategy is best for uncovering 13% of the chemical network, and the 50% optimal strategy is most efficient for discovering 50% of it, but neither are as good as the 100% optimal strategy for revealing the whole network. (B) The actual, estimated search process illustrated on a hypothetical network of chemical relationships, averaged from 500 simulated runs of that strategy. The strategy swarms around a few "important," highly connected chemicals, whereas optimal strategies are much more even and less likely to "follow the crowd" in their search across the space of scientific possibilities. [Adapted from (15)]

We also explored how we could use prediction to improve the social sciences, but designing better surveys that would anticipate and focus on only the things that are not already known.

Sengupta, Nandana, Nathan Srebro & James A. Evans. 2019. "Simple Surveys: Response Retrieval Inspired by Recommendation Systems. *Social Science Computer Review*.

Katariya, Sumeet, Lalit Jain, Nandana Sengupta, James A. Evans, Robert Nowak. 2018. "Adaptive Sampling for Coarse Ranking." *AISTATS*.

Sengupta, Nandana, Madeleine Udell, Nathan Srebro and James Evans. "Matrix Factorization for Missing Value Imputation."

Predictive Signals in Text & Citations

We produced a number of investigations that explored how signals within the text and citations of science could be used to trace and predict future influence in science, with striking results. We found that more ambiguous works were more likely to generate debate that integrated new fields, that text could allow us to diagnose the biases embedded in citations, and that we could use these signals to predict not only typical but also atypical scientific publications.

McMahan, Peter & James A. Evans. 2018. "Ambiguity and Engagement". *American Journal of Sociology* 124(3): 860-912.

Gerow, Aaron, Yuening Hu, Jordan Boyd-Graber, David M. Blei, James A. Evans. 2018. "Measuring Discursive Influence across Scholarship." *Proceedings of the National Academy of Sciences*, doi: 10.1073/pnas.1719792115.

Zhongyang He, Zhen Lei, and Dashun Wang. 2018. "Modeling citation dynamics of 'atypical' articles". *Journal of the Association for Information Science and Technology*.

Improving Review through Prediction

A newer stream of research is focused on using prediction to better understand and improve the review process associated with publishing, grant allocation and promotion.

Teplitskiy, Misha, Daniel Acuna, Aida Elamrani-Raoult, Konrad Kording, and James A. Evans. 2018. “The Sociology of Scientific Validity: How Professional Networks Shape Judgement in Peer Review”. *Research Policy* 47(9): 1825-1841.

Wang, Yang, Travis Hoppe, B. Ian Hutchins, George M. Santangelo, James Evans and Dashun Wang. “New Ideas & Approaches Discussed for NIH Funding Only If Scientists’ Old Ideas Succeed”. (Pending NIH approval for submission)

Acuna, Daniel E, Misha Teplitskiy, James Evans & Konrad Kording. “Should journals allow authors to suggest reviewers?”

Improving Prediction in Social Science

Finally, we applied insights from science to improve prediction in the social sciences, which have allowed us to refine cutting edge methods that allow us to “predict the past”—turning text and implicit references into indicators of deep cultural quantities like opinions and associations that social and cultural analysts have never been able to identify before. We believe that these can be extended to gather critical information in time and security constrained settings where it cannot be elicted..

Kozlowski, Austin, Matt Taddy and James Evans. 2019. “The Geometry of Culture: Analyzing the Meanings of Class through Word Embeddings.” American Sociological Review 84(5): 905-949.

Börner, Katy, Olga Scrivner, Mike Gallant, Shutian Ma, Xiaozhong Liu, Keith Chewning, Lingfei Wu, and James A. Evans. 2018. “Skill discrepancies between research, education, and jobs reveal the critical need to supply soft skills for the data economy.” *Proceedings of the National Academy of Sciences*

Shi, Feng, Yongren Shi, Fedor Dokshin, James Evans & Michael Macy. 2017. “Can We Agree on Science? Measuring the Ideological Alignment of Science with Book Co-purchase Data.” *Nature Human Behavior* 1(0079), doi: 10.1038/s41562-017-0079.

Nan Cao, Yu-Ru Lin, Fan Du, and Dashun Wang. 2015. “Episogram: Visual Summarization of Egocentric Social Interactions”. *IEEE Computer Graphics and Applications*.

Wu, Lingfei, Linzhuo Li, and James Evans. forthcoming. “Social Connection and Cultural Collapse: Hyperbolic Embeddings of Networks and Text.” *Poetics*: http://arxiv.org:1807.10216.

Sun, Jiachen, Ling Feng, Jiarong Xie, Xiao Ma, Dashun Wang, and Yanqing Hu. “Revealing the Predictability of Intrinsic Structure in Networks”.

Teplitskiy, Misha, Julianne St. Onge & James Evans. “How Firm is Sociological Knowledge? Reanalysis of GSS findings with alternative models and out-of-sample data, 1972-2012.”

Chu, Johan and James Evans. “Too Many Papers: Slowed Canonical Progress in Large Field of Science.” socarxiv.org:10.17605/OSF.IO/JK63C

The Future

This grant enabled other successful applications, including a Minerva program that will fund science of science community activities as well as research that builds on these foundations in the years to come. We credit this grant with major advances in the science of science that have formed the backbone of a new field of science with vast and direct implications for all fields of science.

Bibliography

Yin, Yian, Yang Wang, James Evans & Dashun Wang. 2019. “Quantifying dynamics of failure across science, startups, and security.” *Nature* 575: 190-194.

Kozlowski, Austin, Matt Taddy and James Evans. 2019. “The Geometry of Culture: Analyzing the Meanings of Class through Word Embeddings.” American Sociological Review 84(5): 905-949.

Danchev, Valentin, Andrey Rzhetsky & James Evans. 2019. “Centralized communities more likely generate non-replicable results.” *eLife* 8:e43094 DOI: 10.7554/eLife.43094

Sengupta, Nandana, Nathan Srebro & James A. Evans. 2019. “Simple Surveys: Response Retrieval Inspired by Recommendation Systems. *Social Science Computer Review*.

Wu, Lingfei, Dashun Wang & James A. Evans. 2019. “Large Teams Develop Science and Technology, Small Teams Disrupt It.” *Nature* 566: 378-382. [Cover Article]

Yang Wang, Benjamin F. Jones, and Dashun Wang. 2019. Early-Career Setback and Future Career Impact, *Nature Communications*.

Ching Jin, Chaoming Song, Johannes Bjelland, Geoffrey Canright, Dashun Wang, “Emergence of Scaling in Complex Substitutive Systems”. *Nature Human Behaviour*, 2019. [Cover Article]

Jichao Li, Yian Yin, Santo Fortunato, and Dashun Wang, “Nobel laureates are almost the same as us”. *Nature Reviews Physics*, 2019.

Jichao Li, Yian Yin, Santo Fortunato, and Dashun Wang, “A dataset of publication records for Nobel laureates”. *Scientific Data*, 2019

Federico Battiston, Federico Musciotto, Dashun Wang, Albert-Laszlo Barabasi, Michael Szell, Roberta Sinatra. 2019. “Taking census of physics”, *Nature Reviews Physics*, 1, 89–97.

Shi, Feng, Misha Teplitskiy, Eamon Duede, James A. Evans. 2019. “The Wisdom of Polarized Crowds.” *Nature Human Behaviour*, Mar 4: 1.

Foster, Jacob G., and James A. Evans. 2019 "Promiscuous Inventions: Modeling Cultural Evolution with Multiple Inheritance." *Beyond the Meme*. Ed. William C. Wimsatt and Alan Love. Minneapolis: University of Minnesota Press.

McMahan, Peter & James A. Evans. 2018. "Ambiguity and Engagement". *American Journal of Sociology* 124(3): 860-912.

Börner, Katy, Olga Scrivner, Mike Gallant, Shutian Ma, Xiaozhong Liu, Keith Chewning, Lingfei Wu, and James A. Evans. 2018. "Skill discrepancies between research, education, and jobs reveal the critical need to supply soft skills for the data economy." *Proceedings of the National Academy of Sciences*

Teplitskiy, Misha, Daniel Acuna, Aida Elamrani-Raoult, Konrad Kording, and James A. Evans. 2018. "The Sociology of Scientific Validity: How Professional Networks Shape Judgement in Peer Review". *Research Policy* 47(9): 1825-1841.

Gerow, Aaron, Yuening Hu, Jordan Boyd-Graber, David M. Blei, James A. Evans. 2018. "Measuring Discursive Influence across Scholarship." *Proceedings of the National Academy of Sciences*, doi: 10.1073/pnas.1719792115.

Katariya, Sumeet, Lalit Jain, Nandana Sengupta, James A. Evans, Robert Nowak. 2018. "Adaptive Sampling for Coarse Ranking." *AISTATS*.

Fortunato, Santo, Carl T. Bergstrom, Katy Börner, James A. Evans, Dirk Helbing, Staša Milojević, Alexander M. Petersen, Filippo Radicchi, Roberta Sinatra, Brian Uzzi, Alessandro Vespignani, Ludo Waltman, Dashun Wang, Albert-László Barabási. 2018. "Science of science." *Science* 359(6379): eaao0185, doi: 10.1126/science.aao0185.

Lu Liu, Yang Wang, Roberta Sinatra, C. Lee Giles, Chaoming Song, and Dashun Wang. 2018. "Hot Streaks in Artistic, Cultural, and Scientific Careers". *Nature*.

Zhongyang He, Zhen Lei, and Dashun Wang. 2018. "Modeling citation dynamics of 'atypical' articles". *Journal of the Association for Information Science and Technology*.

Shi, Feng, Yongren Shi, Fedor Dokshin, James Evans & Michael Macy. 2017. "Can We Agree on Science? Measuring the Ideological Alignment of Science with Book Co-purchase Data." *Nature Human Behavior* 1(0079), doi: 10.1038/s41562-017-0079.

Tao Jiaz, Dashun Wang, and Boleslaw K. Szymanskiz. 2017. "Quantifying patterns of research-interest evolution". *Nature Human Behaviour* 1: 0078.

Yian Yin, and Dashun Wang. 2017. "The time dimension of science: Connecting the past to the future". *Journal of Informetrics* 11.2: 608-621.

Roberta Sinatra, Dashun Wang, Pierre Deville, Chaoming Song, and Albert-Laszlo Barabasi. 2016. "Quantifying the evolution of individual scientific impact", *Science*, 354: 6312.

Rzhetsky, Andrey, Jacob Foster, Ian Foster and James Evans. 2015. "Choosing Experiments to Accelerate Discovery." Issue cover: "Engineering of Biology and Medicine". *Proceedings of the National Academy of Sciences* 112(47):14569–14574, doi:10.1073/pnas.1509757112

Roberta Sinatra, Pierre Deville, Michael Szell, Dashun Wang, and Albert-Laszlo Barabasi. 2015. "A Century of Physics", *Nature Physics*, 11.10:791-796. [Cover Article].

Nan Cao, Yu-Ru Lin, Fan Du, and Dashun Wang. 2015. "Episogram: Visual Summarization of Egocentric Social Interactions". *IEEE Computer Graphics and Applications*.

Forthcoming Manuscripts

Wang, Dashun and Albert-Laszlo Barabasi, *The Science of Science*. (Cambridge University Press, Forthcoming, 2020).

Foster, Jacob and James Evans, *Knowing*. (Princeton University Press, Forthcoming, 2021)

Wu, Lingfei, Linzhuo Li, and James Evans. forthcoming. "Social Connection and Cultural Collapse: Hyperbolic Embeddings of Networks and Text." *Poetics*: http://arxiv.org:1807.10216.

He, Zhongyang, Zhen Lei, Yang Wang, and Dashun Wang. "Diamond in the rough: Quantifying failed innovation endeavors." (Pending NIH approval for submission)

Wang, Dashun. "The Science of Career: When Do You Do Your Best Work. *Scientific American* (With the editor. Accepted by the editorial board).

Li, Jichao, Yian Yin, Santo Fortunato, and Dashun Wangz. "Scientific Elite revisited: Patterns of productivity, collaboration, authorship, and impact." (submitted

to *Nature Human Behaviour*)

Wang, Yang, Travis Hoppe, B. Ian Hutchins, George M. Santangelo, James Evans and Dashun Wang. “New Ideas & Approaches Discussed for NIH Funding Only If Scientists’ Old Ideas Succeed”. (Pending NIH approval for submission)

Sun, Jiachen, Ling Feng, Jiarong Xie, Xiao Ma, Dashun Wang, and Yanqing Hu. “Revealing the Predictability of Intrinsic Structure in Networks”.

Teplitskiy, Misha, Julianne St. Onge & James Evans. “How Firm is Sociological Knowledge? Reanalysis of GSS findings with alternative models and out-of-sample data, 1972-2012.”
Chu, Johan and James Evans. “Too Many Papers: Slowed Canonical Progress in Large Field of Science.” socarxiv.org:10.17605/OSF.IO/JK63C

Labutov, Igor, Hod Lipson & James Evans. “Machine Teaching.”

Acuna, Daniel E, Misha Teplitskiy, James Evans & Konrad Kording. “Should journals allow authors to suggest reviewers?”

Foster, Jacob, Feng Shi and James Evans. “Measuring Novelty by Simulating Discovery.”

Sengupta, Nandana, Madeleine Udell, Nathan Srebro and James Evans. “Matrix Factorization for Missing Value Imputation.”

Shi, Feng and James Evans. “Science and Technology Advance through Surprise”

AFOSR Deliverables Submission Survey

Response ID:12200 Data

1.

Report Type

Final Report

Primary Contact Email
Contact email if there is a problem with the report.

jevans@uchicago.edu

Primary Contact Phone Number
Contact phone number if there is a problem with the report

8722236166

Organization / Institution name

University of Chicago

Grant/Contract Title
The full title of the funded effort.

Discovering the Extent of Estimable Prediction (DEEP) in Science and Technology

Grant/Contract Number
AFOSR assigned control number. It must begin with "FA9550" or "F49620" or "FA2386".

FA9550-15-1-0162

Principal Investigator Name
The full name of the principal investigator on the grant or contract.

James Evans

Program Officer
The AFOSR Program Officer currently assigned to the award

Enrique Parra

Reporting Period Start Date

09/01/2015

Reporting Period End Date

08/31/2019

Abstract

We proposed to establish mathematical and empirical foundations regarding the nature and extent of quantifiable prediction in science and technology (S&T), the central question of science policy, a fundamental challenge for complex systems research, with the potential to dramatically improve the productivity and focus of science. Research based on this program yielded a wave of relevant discoveries published in Nature, Science, PNAS, Nature subfield journals, every major sociology outlet, and top venues in research policy, social, computer and information science. Moreover, we have drafted two forthcoming books from Cambridge University Press and Princeton University Press, and many more articles that will be published in the coming year. These works review the state of the art in science and technology prediction, but also probe and exceed those limits by predicting science & technology success and failure, career and team productivity and influence, the

disruptiveness and popularity of novel idea and technology combinations, team and community conflict, and a host of indicators that predict future focus and impact. Moreover, this research generated new public data, and the development and calibration of new models that allowed us to push the limits of science and technology prediction in unanticipated ways. Finally, the grant project formed the basis of several other funded projects, including a Minerva award funded by AFOSR, which builds on these foundations to promote a flourishing and productive science of science and innovation that will advance the national and global interest.

Distribution Statement

This is block 12 on the SF298 form.

Distribution A - Approved for Public Release

Explanation for Distribution Statement

If this is not approved for public release, please provide a short explanation. E.g., contains proprietary information.

SF298 Form

Please attach your SF298 form. A blank SF298 can be found here. Please do not password protect or secure the PDF The maximum file size for an SF298 is 50MB.

sf298.pdf

Upload the Report Document. File must be a PDF. Please do not password protect or secure the PDF . The maximum file size for the Report Document is 50MB.

Final_AFOSR_Report.pdf

Upload a Report Document, if any. The maximum file size for the Report Document is 50MB.

Archival Publications (published) during reporting period:

Yin, Yian, Yang Wang, James Evans & Dashun Wang. 2019. "Quantifying dynamics of failure across science, startups, and security." Nature 575: 190-194.

Kozlowski, Austin, Matt Taddy and James Evans. 2019. "The Geometry of Culture: Analyzing the Meanings of Class through Word Embeddings." American Sociological Review 84(5): 905-949.

Danchev, Valentin, Andrey Rzhetsky & James Evans. 2019. "Centralized communities more likely generate non-replicable results." eLife 8:e43094 DOI: 10.7554/eLife.43094

Sengupta, Nandana, Nathan Srebro & James A. Evans. 2019. "Simple Surveys: Response Retrieval Inspired by Recommendation Systems. Social Science Computer Review.

Wu, Lingfei, Dashun Wang & James A. Evans. 2019. "Large Teams Develop Science and Technology, Small Teams Disrupt It." Nature 566: 378-382. [Cover Article]

Yang Wang, Benjamin F. Jones, and Dashun Wang. 2019. Early-Career Setback and Future Career Impact, Nature Communications.

Ching Jin, Chaoming Song, Johannes Bjelland, Geoffrey Canright, Dashun Wang, "Emergence of Scaling in Complex Substitutive Systems". Nature Human Behaviour, 2019. [Cover Article]

Jichao Li, Yian Yin, Santo Fortunato, and Dashun Wang, "Nobel laureates are almost the same as us". Nature Reviews Physics, 2019.

Jichao Li, Yian Yin, Santo Fortunato, and Dashun Wang, "A dataset of publication records for Nobel laureates". Scientific Data, 2019

Federico Battiston, Federico Musciotto, Dashun Wang, Albert-Laszlo Barabasi, Michael Szell, Roberta Sinatra. 2019. "Taking census of physics", Nature Reviews Physics,
1, 89–97.

Shi, Feng, Misha Teplitskiy, Eamon Duede, James A. Evans. 2019. "The Wisdom of Polarized Crowds." Nature Human Behaviour, Mar 4: 1.

Foster, Jacob G., and James A. Evans. 2019 "Promiscuous Inventions: Modeling Cultural Evolution with Multiple Inheritance." Beyond the Meme. Ed. William C. Wimsatt and Alan Love. Minneapolis: University of Minnesota Press.

McMahan, Peter & James A. Evans. 2018. "Ambiguity and Engagement". American Journal of Sociology 124(3): 860-912.

Börner, Katy, Olga Scrivner, Mike Gallant, Shutian Ma, Xiaozhong Liu, Keith Chewning, Lingfei Wu, and James A. Evans. 2018. "Skill discrepancies between research, education, and jobs reveal the critical need to supply soft skills for the data economy." Proceedings of the National Academy of Sciences

Teplitskiy, Misha, Daniel Acuna, Aida Elamrani-Raoult, Konrad Kording, and James A. Evans. 2018. "The Sociology of Scientific Validity: How Professional Networks Shape Judgement in Peer Review". Research Policy 47(9): 1825-1841.

Gerow, Aaron, Yuening Hu, Jordan Boyd-Graber, David M. Blei, James A. Evans. 2018. "Measuring Discursive Influence across Scholarship." Proceedings of the National Academy of Sciences, doi: 10.1073/pnas.1719792115.

Katariya, Sumeet, Lalit Jain, Nandana Sengupta, James A. Evans, Robert Nowak. 2018. "Adaptive Sampling for Coarse Ranking." AISTATS.

Fortunato, Santo, Carl T. Bergstrom, Katy Börner, James A. Evans, Dirk Helbing, Staša Milojević, Alexander M. Petersen, Filippo Radicchi, Roberta Sinatra, Brian Uzzi, Alessandro Vespignani, Ludo Waltman, Dashun Wang, Albert-László Barabási. 2018. "Science of science." Science 359(6379): eaao0185, doi: 10.1126/science.aao0185.

Lu Liu, Yang Wang, Roberta Sinatra, C. Lee Giles, Chaoming Song, and Dashun Wang. 2018. "Hot Streaks in Artistic, Cultural, and Scientific Careers". Nature.

Zhongyang He, Zhen Lei, and Dashun Wang. 2018. "Modeling citation dynamics of 'atypical' articles". Journal of the Association for Information Science and Technology.

Shi, Feng, Yongren Shi, Fedor Dokshin, James Evans & Michael Macy. 2017. "Can We Agree on Science? Measuring the Ideological Alignment of Science with Book Co-purchase Data." Nature Human Behavior 1(0079), doi: 10.1038/s41562-017-0079.

Tao Jiaz, Dashun Wang, and Boleslaw K. Szymanskiz. 2017. "Quantifying patterns of research-interest evolution". Nature Human Behaviour 1: 0078.

Yian Yin, and Dashun Wang. 2017. "The time dimension of science: Connecting the past to the future". Journal of Informetrics 11.2: 608-621.

Roberta Sinatra, Dashun Wang, Pierre Deville, Chaoming Song, and Albert-Laszlo Barabasi. 2016. "Quantifying the evolution of individual scientific impact", Science, 354:
6312.

Rzhetsky, Andrey, Jacob Foster, Ian Foster and James Evans. 2015. "Choosing Experiments to Accelerate Discovery." Issue cover: "Engineering of Biology and Medicine". Proceedings of the National Academy of Sciences 112(47):14569–14574, doi:10.1073/pnas.1509757112

Roberta Sinatra, Pierre Deville, Michael Szell, Dashun Wang, and Albert-Laszlo Barabasi. 2015. "A Century of Physics", Nature Physics, 11.10:791-796. [Cover Article].

Nan Cao, Yu-Ru Lin, Fan Du, and Dashun Wang. 2015. "Episogram: Visual Summarization of Egocentric Social Interactions". IEEE Computer Graphics and Applications.

New discoveries, inventions, or patent disclosures:

Do you have any discoveries, inventions, or patent disclosures to report for this period?

Yes

Please describe and include any notable dates

New discoveries are marked by the publications of major papers above.

We pursued a patent for the model in the most recent Nature paper on the model for failures predicting success.

Do you plan to pursue a claim for personal or organizational intellectual property?

Yes

Changes in research objectives (if any):

N/A

Change in AFOSR Program Officer, if any:

N/A

Extensions granted or milestones slipped, if any:

1 year no-cost extension

AFOSR LRIR Number

LRIR Title

Reporting Period

Laboratory Task Manager

Program Officer

Research Objectives

Technical Summary

Funding Summary by Cost Category (by FY, $K)

	Starting FY	FY+1	FY+2
Salary			
Equipment/Facilities			
Supplies			
Total			

Report Document

Report Document - Text Analysis

Report Document - Text Analysis

Appendix Documents

2. Thank You

E-mail user

Dec 10, 2019 12:35:46 Success: Email Sent to: jevans@uchicago.edu