Reigniting Science

January 22, 2019

Jay Bhattacharya
Stanford University

Mikko Packalen
University of Waterloo

Abstract: New ideas no longer fuel economic growth the way they once did. A popular explanation for stagnation is that good ideas are harder to find, rendering slowdown inevitable. We present a simple model of the lifecycle of scientific ideas that points to changes in scientist incentives as the cause of scientific stagnation. Over the last five decades, citations have become the dominant way to evaluate scientific contributions and scientists. This emphasis on citations in the measurement of scientific productivity shifted scientist rewards and behavior on the margin toward incremental science and away from riskier projects that are more likely to fail, but which are the fuel for future breakthroughs. As attention given to new ideas decreased, science stagnated. We also identify a potential reform that would reverse this trend: broadening how scientific productivity is measured and rewarded. In this reform, bibliometric services such as Google Scholar start measuring also which contributions explore newer ideas, and university administrators and funding agencies start utilizing these new metrics in research evaluation. We demonstrate empirically that measures of novelty are correlated with but distinct from measures of scientific impact. The reform would reignite science by inducing scientists to pursue riskier research paths once again.

Keywords: stagnation; exploration; scientific play; new ideas; scientific impact; novelty.

JEL Codes: O30, O33, O40, H40, I18.

1. Introduction

Economic progress in the advanced world has stagnated in recent decades. As recently as the 1980s, GDP growth rates of four percent or higher were the standard for success. Today, a country that experiences a three percent growth rate in GDP has achieved great success. Real growth rates of one or two percent are the norm when a country is not experiencing a recession. Even many previously fast-developing poorer countries are now experiencing slowing GDP growth rates and diminished expectations for the future. While economists are still debating the causes of this global slowdown in GDP and productivity growth, most now agree that part of the reason must include a slowdown in the rate of discovery of practical new scientific ideas and the development of new technology that those ideas support.[1]

It is fair to ask whether science is actually stagnating. Are we not confronted with evidence to the contrary daily? For example, in biomedicine, we are (once again) seemingly on the cusp of revolutionary technologies that will ultimately (we hope) enable doctors to cure cancer and extend life well into our second century. Moreover, the past two decades have seen astrophysicists train telescopes on black holes and planets orbiting other stars. The recent past has offered us also some very practical advances. For example, the development of techniques to miniaturize semiconductors has provided the engine underlying Moore's law, which predicts that the number of transistors in an integrated circuit will double every two years. These developments have led to computational devices with almost unimaginable power in our pockets.

Systematic evidence, however, points to a decline in scientific productivity. While there are many times more scientists today than in the past, today's advances do not compare favorably to past breakthroughs. In areas where scientific progress is still as robust as in the past, today's discoveries take many times more research effort than did past discoveries. The cost of developing new drugs, for example, now doubles every nine

[1] On global economic and productivity growth slowdown, see Hall (2016), Brynjolfsson et al. (2019), International Monetary Fund (2019), and Gordon and Sayed (2019). On technological stagnation, see Mandel (2009), Thiel (2011), Cowen (2011), Vijg (2011), Gordon (2000, 2012, 2016) and Bloom et al. (2019).

years. This empirical regularity has been referred to as Eroom's law, representing the idea that the forces that govern scientific progress today are in many ways a mirror image of the forces that were at play at the time while Gordon Moore conceived his law. In other scientific areas, such as physics, the decline in scientific productivity appears to be absolute; many physicists themselves view current breakthroughs as less significant than past ones, despite the larger investment in these scientific areas.[2] Even economics, a quantitative social science, has not been immune to concerns about stagnation.[3]

A common explanation for this scientific stagnation is that finding good new ideas has simply gotten harder. Having picked all the low-hanging fruit, the secrets that are still there for scientists to unlock are fewer and harder to find.[4] While this explanation is convenient as it absolves scientists, institutions, and analysts from any responsibility for the slowdown, it is likely wrong. One problem is that this explanation is not unique to our era. This view was common before the scientific revolution, as well as during the late nineteenth century. After the successful extension of classical physical ideas to electricity, magnetism, light, and heat, many physicists argued that physicists had solved physics. On the verge of two great revolutions in physical thinking (relativity and quantum mechanics), Nobel Prize-winning physicist Albert Michelson—one of the duo of physicists who earlier debunked the theory of ether—said in an 1894 speech dedicating the Ryerson Laboratory at the University of Chicago that:

---

[2] On scientific stagnation, see Le Fanu (1999, 2010), Scannell et al. (2012), Bloom et al. (2019), Collison and Nielsen (2018), and Cowen and Southwood (2019).

[3] Caballero (2010) laments the state of macroeconomics and implores researchers to address the pretense of knowledge syndrome by shifting to broad exploration mode. Romer (2015, 2016) argues that macroeconomics has gone backwards, as a culture of deference to authority has led to careless formalism and abandonment of empirical evidence as the coordination device. Heckman and Moktan (2019) worry that the obsession with publishing in high-impact journals incentivizes careerism over creativity, as innovative papers and exploratory odd-ball ideas are unlikely to survive these journals' refereeing practices. Akerlof (2019) and Ruhm (2019) argue that economics excessively favors work where precise results can be produced over work that examines more important questions but yields imprecise results, biasing the profession against new ideas as they tend to be initially poorly understood and formulated. Frey (2003) argues that in economics new ideas are rejected for lack of rigor as they are by necessity less well formulated than well-established ideas, and that this has inundated economics with boring and irrelevant papers.

[4] On ideas becoming harder to find, see Jones (2010), Jones and Weinberg (2010), Cowen (2011), Arbesman (2011) and Bloom et al. (2019).

> "The more important fundamental laws and facts of physical science have all been discovered, and these are now so firmly established that the possibility of their ever being supplanted in consequence of new discoveries is exceedingly remote."[5]

In stark contrast with this dark but once again popular view that we are entering an era of vanishing secrets and diminished expectations for the future, our explanation for the slowdown in research productivity has nothing to do with the paucity of good ideas. On the contrary, we believe that an abundance of ideas remains to be uncovered, ideas that—if given proper attention by the community of scientists—will produce unparalleled breakthroughs.

In this paper, we argue that the root of the slowdown in scientific progress lies in the more quotidian fact that scientific work on truly novel ideas that have the potential to develop into groundbreaking advances is no longer rewarded in the same way it once was. Because scientists respond to incentives just as everyone else does, the reduction in the reward for novel, exploratory, work has reduced the effort devoted to it in favor of pursuing more incremental science which seeks to advance established ideas. Furthermore, as ideas are not born as breakthroughs—they need the attention and revision of a community of scientists to be developed into transformative discoveries—this decrease in scientists' willingness to engage in an exploration of new ideas has meant that fewer new ideas have developed into breakthrough ideas. The underlying change in scientist incentives in turn has been driven by the shift toward evaluating each scientist based on how popular their published work is in the scientific community, with popularity measured

[5] Michelson (1903, pp. 23-24). On the lack of belief in discovery before the scientific revolution, see Wootton (2015). More recent reincarnations of this view include Glass (1971), Horgan (1996), Cowen (2011), Gordon (2016) and Bloom et al. (2019). Silverstein (1999) lists further examples from different fields. Betz (2017) reviews other prominent stagnation theories including increased government regulation and risk-aversion. Another problem with the vanishing secrets theory of stagnation is its self-fulfilling nature. Should scientists increasingly give up on the prospect of uncovering important secrets, it would become even harder for those few who still pursue broad exploration to get funding as their work would then be perceived to be even less feasible than today. See Thiel (2014) and Diamond (2019) for related analyses. The vanishing secrets theory of stagnation also runs against the idea that the arrival of new ideas leads to a combinatorial explosion in the number of possible combinations of ideas and an accompanying increase in available combinations lead to breakthroughs. On combinatorial explosion, see Romer (2019). Stagnation is consistent with combinatorial explosion if scientists at the same time become less willing to try out combinations that involve new ideas.

by the number of times their work is cited by other scientists. We call this shift the 'citation revolution'. This citation revolution has offered a useful way to identify and reward breakthrough science, contributions that have had a large influence on the scientific community.[6] However, it has also dramatically tilted incentives in favor of incremental me-too science (as work in crowded areas tends to gather many citations) over exploration and scientific play, which tends to gather fewer citations but which lays the necessary groundwork for subsequent breakthroughs. The fixation with citations—which popular bibliometric services such as Google Scholar perpetuate today—and the associated shift in scientist incentives and behavior have thus ultimately led to fewer breakthroughs.[7]

The role that exploration and scientific play have in developing new ideas gradually from their infancy to breakthroughs forms a central part of our argument. Accordingly, we present a simple model of the lifecycle of a scientific idea to capture this aspect of scientific production. The model links scientific effort on an idea and the scientific impact of the idea in the three phases of the lifecycle of an idea: exploration, breakthrough, and incremental advance. Since ideas develop slowly in their infancy, early explorative work on an idea typically—and in our model—has little scientific impact. Nevertheless, such work is crucial because it lays the necessary groundwork for the later work on the idea, work that yields the breakthrough.

The model is inspired by prominent examples of scientific breakthroughs in biomedicine. For instance, consider CRISPR—a recent breakthrough in biomedicine and one with the potential for substantial practical patient care benefits. The scientists now predicted to win Nobel Prizes for CRISPR began their first work on the idea after 20 years of exploration and scientific play by other scientists with the ideas that underlie it. This initial work gradually advanced the understanding by biomedical scientists of CRISPR's existence, properties, purpose, and potential uses.

[6] For perspectives on the citation revolution, see e.g. Bensman (2007) and Small (2018).
[7] That scientists too respond to incentives is well demonstrated by Azoulay et al. (2019) who find that a funding mechanism that tolerates early failures yields more novel work and a 97% increase in breakthroughs papers.

Given this crucial aspect of scientific production—that early exploration is indispensable but typically has little impact on the wider scientific community—an excessive reliance on citations in the evaluation of scientists effectively punishes the exploration of new ideas. The incentives created by the focus on citations leads researchers to pursue more established research paths, with stagnant science as the by-product. Our simple model highlights this tension between the exploration of new ideas and citation-based research evaluation. Our analysis also clarifies who exactly stands on the shoulders of giants in science (those who make breakthrough discoveries) and who are the giants (those contemporaries and predecessors who engaged in the early exploration of new ideas). The latter group often has little recognition in the wider scientific community but facilitates the subsequent breakthroughs.

While the "vanishing secrets" theory of stagnation implies that continued stagnation is inevitable, our incentive-based theory of stagnation points to a potential reform to reignite science. Without proper incentives to explore and build on new ideas, we should not expect scientists to try to unlock their mysteries to the same extent that they once did. The reform would change scientist incentives by broadening how scientific productivity is measured and evaluated. Specifically, to encourage more exploration and scientific play, the reform would move us away from measuring and rewarding scientific impact only and toward measuring and rewarding both scientific impact and scientific novelty.

The balance of the paper proceeds as follows. In the first half of the essay, we delineate the link between stagnation and citation-driven science. We first discuss the dominant position that citation counts have come to hold in the evaluation of scientific contributions and scientists. Next, we present a simple model of the lifecycle of a scientific idea that distinguishes the special role that exploration and scientific play have in the development of scientific ideas from their infancy to breakthroughs. The model conveys why we need to reward both exploration and impact. We then revisit the citation revolution in light of our model, focusing on how the citation revolution shifted scientist incentives and behavior away from exploration and toward incremental science.

In the second half of this essay, we examine a potential reform that would reverse the decline in exploration. We first discuss how bibliometric services such as Google Scholar perpetuate the fixation with citations today, and present the rationale to reform these services to measure also other aspects of scientific productivity, such as the propensity to explore new ideas. We then present a constructive way to measure such scientific novelty in practice and show that measures of novelty and scientific influence are empirically distinct. The final section discusses three possible paths forward.

2. Today's One-Dimensional Evaluation of Scientific Contributions

Scientists are measured by concrete metrics that are intended to capture the breadth of their contributions to the scientific enterprise. The rule for research scientists, at least those working in university settings, used to be "publish or perish." This idea, first articulated in the 1940s, implicitly assumes that a productive scientist publishes many papers while an unproductive scientist publishes few.[8] A variant of this rule involves counting the number of papers published in the most prestigious scientific journals.

While useful, in recent decades, the importance of this simple volume metric has faded as another metric focused on measuring how popular a scientist's published papers are in the scientific community has risen in importance. The popularity of a given paper, in turn, is measured by the number of times other scientific papers cite that paper. Scientific journals are now also ranked largely based on their "impact factor," which is a function of the number of citations that papers published by the journal have attracted. An impactful scientist has come to mean a scientist who publishes popular, highly cited papers.

Some modern metrics such as the "h-index," developed by physicist Jorge Hirsh, combine volume and popularity into a single score that is a function of both. The h-index is as important to scientists today as a batting average or slugging percentage is to a professional baseball player, and in a directly analogous way. Success as a scientist, including tenure, promotion, and pay, often depend crucially on this number.

[8] On the origin of publish and perish, see Garfield (1996).

A key criticism that raw citation counts, the h-index, and other measures constructed from citation data have faced in recent years is that they reduce scientific contributions to a number. Even an editorial in Science—the most highly cited scientific journal—concurs.[9] The author laments the prevailing 'impact factor mania,' and the increasing tendency of scientists to work in well-populated research areas and on me-too science. Our concern with citations is a bit different, as we believe that advocates of citation-based metrics are right in arguing that highly cited discoveries are worth celebrating. The issue with citations is not that they reduce scientific contributions to a number.

Rather, our main concern with citations is that they reduce scientific progress to only one number, a number that captures just one important dimension of scientific productivity. While breakthrough advances are worth celebrating, such contributions are not the only discoveries worth celebrating. For scientific progress depends on a steady flow of exploration and experiments with new ideas. When first hatched, new ideas are risky for scientists to work on in the sense that, at the outset, it is hard to distinguish between the ideas that are likely to be good if properly developed and the ideas that will never amount to much. Moreover, transformative discoveries—breakthroughs—depend on the base of knowledge created by this sort of scientific play with risky new ideas.

Because citation-based metrics cannot capture distinctions between scientific exploration of new ideas and incremental work on mature, well-established, ideas, they distort research evaluation and shift scientist incentives away from such exploration and toward incremental science. The resulting decline in exploration and scientific play implies that fewer ideas are developed into breakthroughs, rendering science less vibrant.

[9] See Alberts (2013).

3. The Importance of Scientific Exploration and Play

Crucially, not all eureka moments in science were recognized as such at first. Instead, the ideas only became transformative ideas after other scientists developed them further. Examples from every scientific discipline abound. One particularly instructive example comes from biochemistry.

In 1967, biologists playing around near a Yellowstone National Park hot spring isolated the bacteria that they found thriving there. What drew their interest was the fact that bacteria could survive at such high temperatures. All living organisms require DNA to reproduce, and the key enzyme that enables the copying of DNA is called DNA polymerase. One reason why most bacteria die at high temperatures, such as those found in hot springs, is that excessive heat deactivates DNA polymerase. Unlike other organisms, however, exposure to high heat does not harm the hot springs bacteria. The researchers noticed that one particular bacterium survived in higher temperatures than previously thought possible, and named this bacterium Thermus aquaticus ("hot water"). This curious new fact was published in the Journal of Bacteriology and generated a few citations but did not attract the attention of the wider community of biologists. A decade later, in 1976, other scientists isolated the DNA polymerase—a key enzyme that all living cells use to copy DNA—of the Thermus aquaticus bacterium. This finding, too, was published in the Journal of Bacteriology and had a similarly limited scientific impact.[10]

In the 1980s, biochemist Kary Mullis dreamed up the idea of applying a cell's capacity to copy DNA over and over again on a particular strand of DNA.[11] First, there would be two copies, then four, then eight, and so on. After enough replications, there might be millions of copies of the desired DNA segment. The technology, called polymerase chain reaction ("PCR"), later won Mullis the Nobel prize and revolutionized biomedicine. Without PCR, the human genome project would have been impossible. The earliest versions of this technology required the repetitive heating and cooling of the DNA sample and the

[10] See Brock and Freeze (1969) and Chien et al. (1976).
[11] See Mullis (1994).

continual reintroduction of new DNA polymerase into the solution. This unfortunate fact rendered the technology impractical for widespread use. The key breakthrough happened a few years later when scientists (including Mullis himself) playing with Mullis' idea, learned about the earlier work on the hot springs bacterium and its unique heat resistant DNA polymerase. With this tweak—and no need to repetitively reintroduce DNA polymerase at every replication—PCR was automated so that a machine could run the procedure cheaply. Today PCR is used as an essential tool in every field where analysis of DNA plays a role. But this advance would never have happened without support for the biologists to play around near the Yellowstone geysers and hot springs.

The history of CRISPR, a more recent breakthrough in biomedicine, offers a second illustration of the importance of scientific play and how ideas gradually develop into breakthroughs. In 1987, scientists first noticed *clustered repeated regularly interspaced short repeats* (CRISPR) in genetic material. During the 20 years that followed, work by dozens of scientists demonstrated the presence of CRISPR in different organisms and explored its properties, though in the initial years after its discovery, the biological purpose of CRISPR remained mysterious. By 2007, though, scientists had discovered CRISPR's evolutionary purpose as an adaptive immune system. Cells use these DNA segments as a library of foreign DNA to help immune cells fight off future assaults.

Despite these breakthroughs, CRISPR had yet to influence the wider scientific community. The early work did not receive many citations, and even as late as 2012, leading journals rejected contributions that are now considered key advances. In 2011, it was not clear that CRISPR was the best technology for genome editing; scientists who chose to explore it considered the choice risky. In 2007, the scientists now predicted to win Nobel Prizes for their CRISPR-enabled discoveries had not yet begun their work on it.

Nevertheless, the early work on CRISPR—the two decades of scientific exploration and play by many scientists in relative obscurity—had gradually advanced the understanding of CRISPR's existence, properties, purpose, and potential uses within biomedicine. This knowledge served as the basis for a flurry of scientific activity in the late 2000s and early 2010s that culminated in CRISPR being successfully used for mammalian genome editing in 2012, with considerable scientific influence and potential practical benefits.[12]

There are several key lessons to draw from the stories of PCR and CRISPR. First, it is exceedingly difficult to anticipate which particular advances in science are likely to prove useful for future breakthroughs; it takes a long time for the scientific community to come to understand the properties of each new idea. There is thus a great long run benefit to supporting the exploration of ideas that, at the time, do not have an immediate and apparent application.[13]

The second lesson is that breakthroughs often cannot take place without prior high-risk work. When Isaac Newton—a revolutionary scientist if there ever was one—said that he was "standing on the shoulders of giants," this was not merely false humility. His work built on the foundation of the scientific play of his contemporaries and predecessors, even if nearly everyone today (barring historians of science) would face insuperable difficulties in naming them all or their contributions to physics. It is no accident that Gottfried Leibniz discovered calculus at nearly the same time as Newton—they both benefited from the same scientific community of contemporaries and predecessors; they were both standing on the shoulders of the same giants.[14]

---

[12] For the history of CRISPR, see Yoshizumi et al. (2018), Mojica and Montoliu (2016), Doudna and Sternberg (2017) and Lander (2016) (the last reference is controversial not for its emphasis on the key role that exploration and scientific play had in terms of facilitating the widely celebrated later breakthroughs but for its emphasis on senior scientists over graduate students and postdocs and for its emphasis on some senior scientists over others as well as for its lack of acknowledgement of a potential conflict of interest relating to a patent dispute).

[13] The discovery of the DNA double helix provides another illustration of how raw ideas are when they are first born. The initial response to the discovery was muted in part due to lack of evidentiary support on several dimensions—even Watson and Crick acknowledged the speculative nature of their initial discovery. Subsequent work by many scientists who were willing to explore the new hypothesis in the subsequent years soon filled those gaps. See Olby (2003). The events leading up to the discovery of the double helix by Watson and Crick provide a further illustration of the importance of exploration by a broader scientific community in fueling breakthrough discoveries. See Watson (1968)

[14] The discovery by eventual Nobel winners Barry Marshall and Robin Warren of the role of Helicobacter pylori in pectic ulcers provides another illustration of the value of exploration for later discoveries. Marshall and Warren benefited from 90 years of prior exploration on bacteria that can survive in stomach acid. By the time Warren turned to the topic, the scientific community had largely forgotten the initial exploratory work on the topic. Warren at first faced considerable resistance to his hypothesis, but further experiments by them—including Marshall drinking a slurry of the bacteria and causing himself an ulcer—and by others finally convinced the larger scientific community that their idea was right, and led to a flurry of research to use the finding in clinical practice. The abovementioned work on hot springs bacteria played a role too—it

The third lesson derives from thinking about all the scientific play that does not ultimately prove useful. Most scientific work with new ideas falls into this category, as most new ideas fail in the sense that they do not develop into breakthroughs. While early exploratory work is crucial for subsequent breakthroughs, it is rare for scientists to reward such work with many contemporaneous (or even delayed) citations. Despite the difficulty of distinguishing ultimately useful scientific play from less useful play at the time when this exploration takes place, such work is valuable nonetheless and worthy of support even given its high rate of failure.[15] Scientists who spend their lives on such play need rewards commensurate with the value of their work, which is considerable, even if most of their publications do not generate much attention in the form of citations from other scientists.

In the next section we develop a simple model that captures these key characteristics of how ideas develop and highlights how citation-based evaluation weakens the incentive to explore new ideas. However, before proceeding, it bears mentioning that the distortion of incentives influences not only what scientists do but also who becomes a scientist in the first place.[16] Specifically, the tilt in incentives drives away from science those who would prefer to pursue risky exploration in their work. The eccentric scientist type—the introverts who prefer to work apart from the crowds—has a weakened incentive to become a scientist. Gradually in science, these individuals will be replaced with people who work well in teams but conform more easily to prevailing scientific norms, researchers who tend to eschew exploration in new areas of investigation and instead work in more crowded research areas. To be sure, both types of scientists are valuable. We need scientists who further develop maturing research areas, and we need scientists who explore the unknown in search of establishing new areas of investigation. Nevertheless, this need to make science more appealing to people who prefer to pursue explorative science in favor of relatively well-known research paths, is an important reason to tilt the reward structures in scientific production back toward rewarding novelty and exploration.

---

helped change the previous consensus that bacteria do not survive harsh environments such as in stomach acid or hot springs. See Marshall (2005, 2016) and Warren (2005).

[15] Working scientists often view failure to be common and even as a precursor to success, see Livio (2013), Firestein (2015), Popovian (2016) and Zaringhalam (2016).

[16] On how scientist evaluation mechanisms influence who becomes a scientist, see Osterloh and Frey (2013).

4. The Lifecycle of a Scientific Idea

We now present a simple model of scientific production that describes the lifecycle of an idea. The framework abstracts away many important aspects of the scientific endeavor but still highlights how ideas develop and conveys how citation-based research evaluation has tilted scientists' rewards and behavior away from true innovation. The model also conveys a rationale for measuring and rewarding work that explores new ideas.

At the heart of this model is a division of scientific activity on each scientific idea into three phases: exploration, breakthrough, and incremental advance. Figure 1 illustrates these three phases. As work on an idea in the latter phases builds on earlier work on the idea, we call this framework the "shoulders of giants" model.

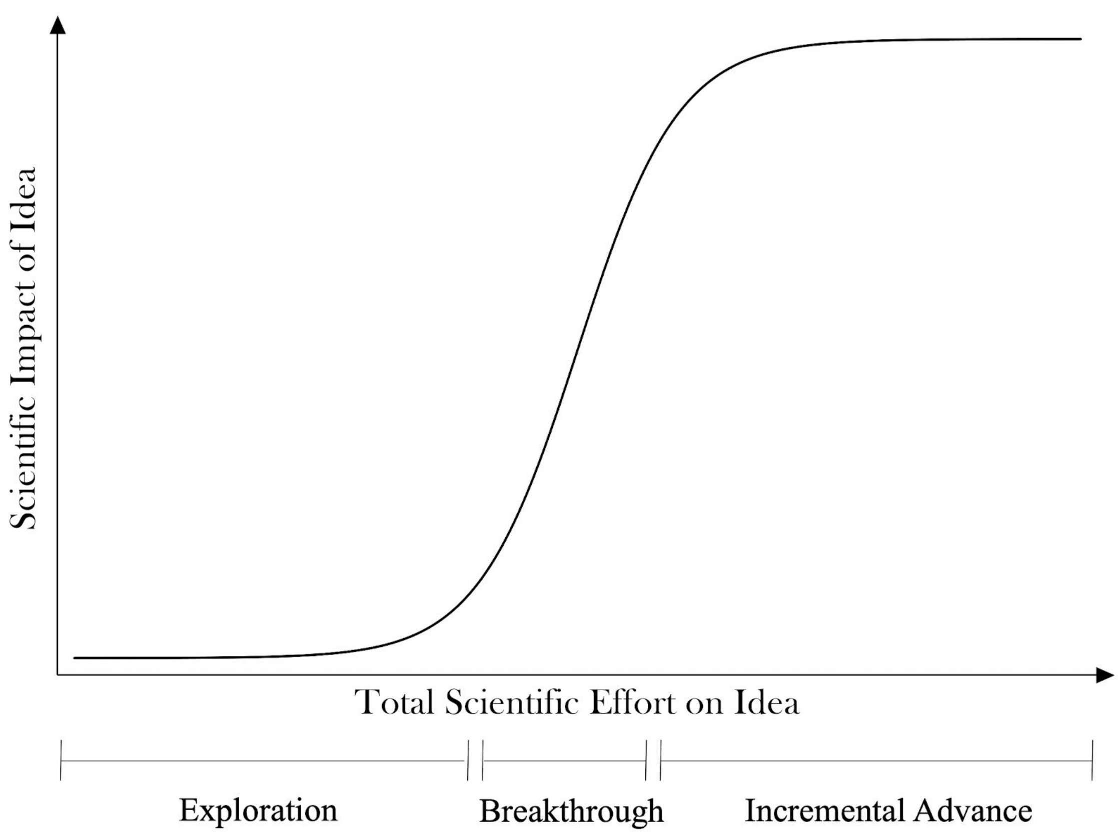

Figure 1. The relationship between scientific effort on an idea and the scientific impact of the idea. The horizontal axis captures the three phases of our model: exploration, breakthrough, and incremental advance. The vertical axis captures the influence of the idea on the scientific community. Work done during the

exploration phase has little observed impact but lays the necessary groundwork for the later breakthroughs.

In this model of scientific production, new ideas occur to scientists all the time during the conduct of their work. These new ideas all share the feature that when they first arrive, they are in an embryonic form that generates no scientific impact whatsoever. Good but yet untested and unexplored ideas are, by themselves, cheap in science and convince no one of anything. Most new scientific ideas die there, with even the scientist who had the original idea giving up on it very shortly after it arrives.

On the other hand, one or maybe a few scientists might see the promise in some of the new ideas and deem them worthy of further exploration. The idea may even entice a broader group of scientists to engage. This early work on the idea initiates the exploration phase. The importance of such attention by a community of scholars in the development of each new idea into a mature advance is neither new nor controversial. However, in citation-based empirical analyses of scientific productivity, this aspect is often forgotten; the analyses focus instead on identifying great scientists without regard for the scientific community that was integral to their success.[17]

During the exploration phase, scientists try out the new idea in different settings and different ways, looking to see what does not work and what looks promising, looking for applications and combinations with other ideas where the new idea might prove useful. Over time and with considerable effort, the cohort of scientists working on the new idea builds up a body of knowledge about it and its usefulness. We label work during the exploration phase "edge science," since researchers work with relatively unknown and untested ideas that are near the edge of the scientific frontier.

Often little of the body of knowledge uncovered during the exploration phase will end up in the published literature since much of it consists of negative results about

[17] Kuhn (1962, 1977) emphasized that ideas are poorly understood when they are first born and need the attention and debate of a community of scientists to be developed into useful advances. For formal theoretical analyses of scientific communities, see Bramoulle and Saint-Paul (2010), Besancenot and Vranceanu (2015) and Akerlof and Michaillat (2019). Foster (1986) used an S-curve to describe the link between effort and performance in technological innovation. The raw nature of new ideas in technological innovation was emphasized by Marshall (1920) and Usher (1929).

contexts in which the idea does not work. When a scientist submits a paper for publication describing a positive result or files a grant application to garner financial support for further exploration, the typical reaction of peer reviewers not involved in the exploration of the idea is skepticism and rejection. When a paper about the new idea is accepted, it will typically attract only a small audience of scientists who cite it.

These properties of scientific work on new ideas are modeled in Figure 1 by the long flat portion of the exploration phase. This work generates little scientific impact (in the sense of cited publications) as a return for the considerable effort that takes place during this early work on an idea. Nevertheless, this work is invaluable as the knowledge gained from such scientific play facilitates the breakthrough phase.

Most scientific ideas never make it to the breakthrough phase, but those that do make scientific careers. The exploration phase has finally produced a promising application of the idea or a fruitful line of experiments that make clear to the scientific community that the new idea is worth its attention. A broader set of researchers start to use the idea in their own work, and someone (not necessarily from the cohort involved during the exploration phase) publishes a paper about the idea in a high-profile scientific journal. A flood of new scientists are drawn to the now better understood and newly popular idea, and a series of high-profile papers are published, and grants awarded to members of this newly enlarged cohort of experts on the idea. The impact of the idea on science from this work is enormous and is modeled in Figure 1 by a sharp rise in the curve during this breakthrough phase. If the idea is important enough or fruitful enough, it may eventually generate a prestigious prize for a scientist publishing about the idea during this phase.

Research conducted during the breakthrough phase is not synonymous with revolutionary research in the sense implied by Thomas Kuhn in his Structure of Scientific Revolutions. Revolutionary ideas fundamentally upend the assumptions on which a scientific discipline operates, relegating much previous research (even breakthrough research) to the status of historical interest only. Hence, for example, the theory of Ptolemaic epicycles plays almost no role in modern astrophysics since the Copernican scientific revolution supplanted it. Though Ptolemy's system was a breakthrough idea in its time, it was supplanted by later developments. Only a select few breakthrough ideas are truly revolutionary in the Kuhnian sense; these are only a small fraction of the ideas that arrive at the breakthrough phase.

Breakthrough science in the context of the shoulders of giants model is also not synonymous with the ideas being capital-T true. As many scientists have evaluated the idea by the time it hits the breakthrough phase, such ideas are more likely to be True than ideas that die in the exploration phase. But even large groups of smart and dedicated people can be wrong about many things. This fact is one reason why scientific influence in our model is not exactly equal to the practical benefit of these ideas, even though in some scientific areas such as biomedicine, the two can be linked quite closely.

Finally, in the incremental advance phase, the idea has matured. There is still considerable work left for scientists to do to work out details that were set aside during the heady days of the breakthrough phase. These details are still scientifically important and will generate further scientific impact, but the ratio of scientific advance to total scientific effort on the idea has decreased considerably. As a general rule, uncertainty about what a scientist is likely to find from putting further effort into this idea has also decreased considerably from the exploration days. Consequently, funding agencies may be eager to award grants to scientists working on the subject. Peer reviewers—a conservative lot if there ever was one—abet this tendency since grant applicants can credibly reassure them the proposed work is likely to produce visible, if marginal, successes.

When the idea has reached near the summit of its potential, further scientific work on it produces little scientific impact. We model this fact in Figure 1 with a flat curve during much of the incremental advance phase. The idea has enriched the work of many scientists, but most gradually recognize that it is time to move on to the next one. The idea has been absorbed into the thinking of the scientific community so thoroughly that few researchers who use the idea bother to cite the scientists who conducted the earlier work. For example, a vanishingly small fraction of published biochemical papers today that involve DNA cite the seminal paper by James Watson and Francis Crick on its molecular structure. Hence, while citations are a useful tool for measuring scientific influence in the short run, the only hope for measuring the long-term impact of an idea is through an analysis of the scientific vocabulary.

Up to now, we have focused on the development and lifecycle of a very successful idea. However, most new ideas will never—even if properly developed and explored—turn into significant breakthroughs. Figure 2 illustrates this variability in the quality of new ideas; it shows the differing potential of four different ideas, imaginatively titled A, B, C, & D. Idea A, if pursued by the scientific community, would result in a ground-shattering breakthrough. Idea B, if pursued, would also produce a nice result, though perhaps not of fundamental significance. By contrast, ideas C & D, no matter how much effort scientists pour into them, would not result in much of anything.

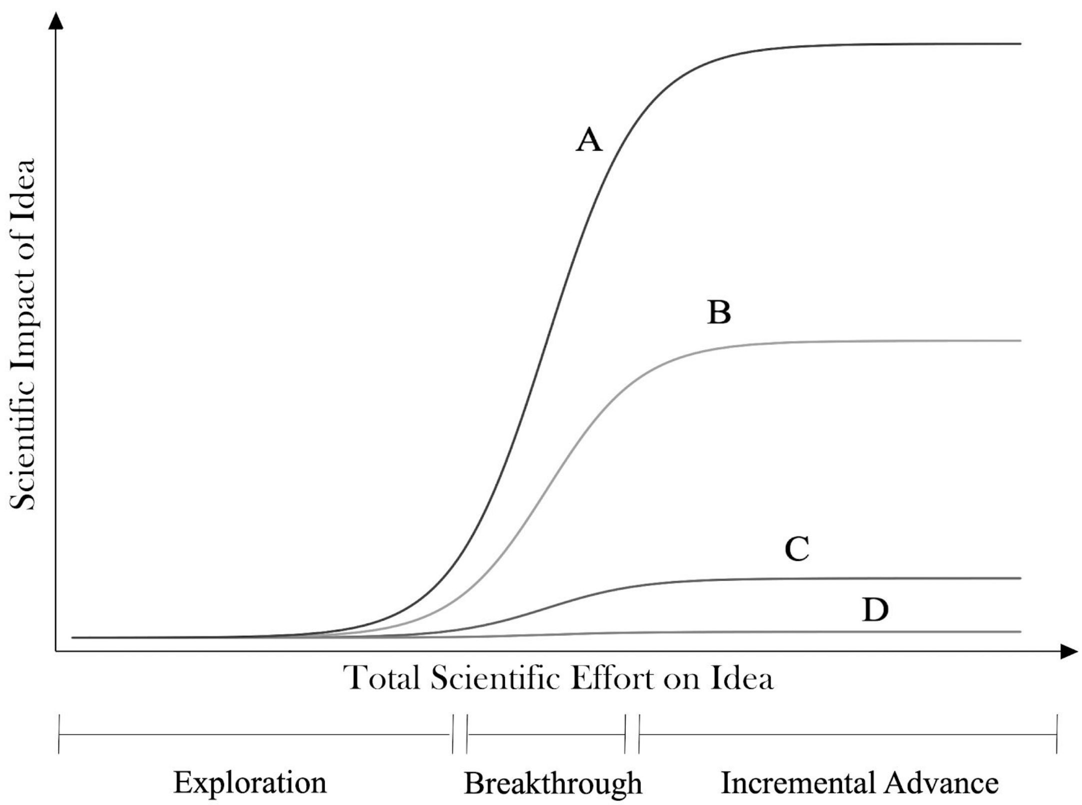

Figure 2. The relationship between scientific effort and scientific impact on four ideas with different potential. Ideas A and B are highly successful if properly developed, whereas ideas C and D do not amount to much, even if properly developed. Ideas that lead to failure in this sense are much more common than successful ideas.

Nevertheless, the early exploration of all the ideas A, B, C, and D, is invaluable. During the early days of scientific investigation of an idea, it is often impossible for scientists to distinguish good ideas from the bad—that is, to distinguish ideas A and B from ideas C and D. It takes some work by scientists to uncover whether an idea is likely to be ultimately fruitful or barren.

Furthermore, even the scientists who work on idea A in its exploratory phase do not generate much immediate scientific impact. Rather, they are the giants upon whose shoulders the later breakthrough phase scientists stand since their work enables the later ground-breaking discoveries. The designation of "giant" extends to all the scientists working in the exploration phase, whether they work on ideas A, B, C, or D; they all advance science whether their results turn out positive or negative by guiding the thinking and actions of future scientists. To further highlight this point, Figure 3 shows the model together with the separate designation of scientific work conducted by giants and work conducted by scientists standing on the shoulders of giants.

Figures 2 and 3 together imply a rationale for measuring and rewarding exploration phase work by scientists, even if the idea does not develop into a meaningful advance. First, rewarding exploration promotes those scientific giants who set the groundwork for future breakthroughs. Second, and just as importantly, such a policy would reward fruitful scientific failures. Of course, this is valuable because, at the time of exploration, there is often considerable uncertainty about whether a particular idea is a good one. It is only by trying out the idea that we gain knowledge about which new ideas are the best new ideas. Without rewards for work on ideas that fail, such an exploration of new ideas is too risky for most scientists to undertake.

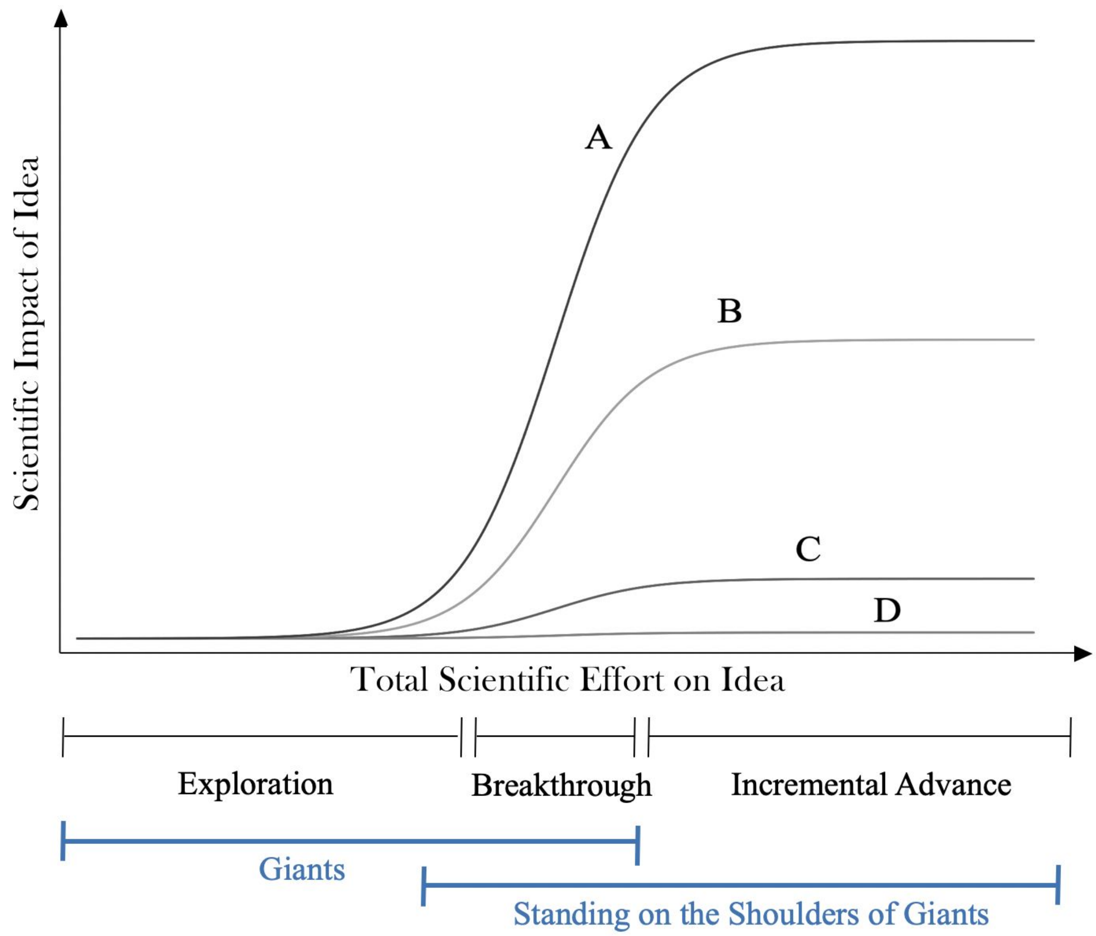

Figure 3. The delineation of scientific effort as giants and as standing on the shoulders of giants. Giants are not only those who work on the idea during the breakthrough phase but also those who investigated and developed the idea during the exploration phase. Furthermore, the status of giant does not depend on whether the idea is ultimately successful.

It is natural to ask which scientist played in retrospect the most important role in the development of a mature and successful scientific idea. There are several plausible candidates, including the scientist who first conceived the idea, the cohort of scientists who played with the idea during its exploration phase, and the scientists who published the most prominent papers during the breakthrough phase. In our view, there is no need to decide who was most important since they were all important.

Crucially, though, in today's science, those who work on the idea during the breakthrough phase will reap the lion's share of the rewards since they are the ones publishing the papers that generate the most citations. Yet their work would have been impossible without the edge science done by scientists during the exploration phase, work that is unlikely to be measured as high impact since it generally does not garner a substantial number of citations.

Furthermore, even work pursued during the incremental advance phase often receives far more citations than comparable work done on the idea during the exploration phase. As a consequence, scientists who work on an important idea after it has matured to the incremental advance phase are usually considerably better rewarded than scientists who worked on the idea during its exploration phase (not to mention scientists who explored ideas that failed). As the idea matures, the relative certainty of making discoveries—however incremental—attracts more and more scientists to the area. As a result, while the scientific impact of work on mature ideas is not necessarily any greater than the scientific impact of work during the exploration phase, there are often many times more scientists working on the idea during the mature phase than during the exploration phase. Because so many are working on the idea, the work will tend to receive many citations, which will attract even more scientists. This cycle leads to an amplified divergence between citations and scientific impact. Thus, the incentive to continue working on the idea is considerable long after the additional value of contributions on it has greatly diminished.

The model thus helps bring to focus a key disconnect: while rewarding exploratory phase science is important for scientific progress, the current practice in science evaluation favors breakthrough and incremental science over exploration. Furthermore, as citation measurement is a valuable tool for rewarding scientists who work during the breakthrough phase, the main issue with citation-based evaluation is that citation counts do not distinguish between early, exploratory, work on an idea and later, incremental, work on an idea. The reliance on citation counts tends to encourage incremental science at the expense of the exploration of new ideas.

The model also points to a potential reform that would increase the incentive to pursue exploration: broadening how scientific productivity is measured and rewarded. The first step of this reform is to build a metric that identifies which scientific contributions represent edge science. The second step of this reform is to use that metric to reward scientists who direct their time and effort to the trying out of new ideas. Before discussing how edge science contributions can be identified in practice (section 7), we first discuss in more detail the rise of citations and its impact on science.

5. As Citations Became Prominent, Science Became More Conservative

Evaluating scientific contributions and scientists was not the original purpose of citation indices. Instead, they were first built to help scientists process the rapidly growing scientific literature. But soon after their introduction, scientific administrators and funders noticed that these indices could also be used to identify influential scientific contributions, influential journals, and influential scientists. A natural consequence of this development was that citations quickly gained a prominent role in determining rewards in science.

Ironically, Eugene Garfield—the scientist primarily responsible for developing the idea of using citations as a measure of scientific productivity in the 1950s—had considerable reservations about their use to evaluate scientific productivity. His obituary, published in *Nature*, reported that:

> "Garfield came to see the impact factor as a mixed blessing, 'like nuclear energy.' Although he felt that citation indexing and the impact factor could be remedies for the limitations of peer review, he was uncomfortable with their misuse as performance indicators."[18]

The citation revolution started to gather steam in the 1970s. Scientists, university administrators, and funding agencies increasingly focused their attention on citations. In

[18] For the obituary, see Wouters (2017). Garfield (1955, 1972) was tenacious with developing and disseminating citation metrics but was not the first to introduce the idea of using citations to rank papers and journals, as the idea had been previously explored by Gross and Gross (1927) and Fussler (1949).

1983, Eugene Garfield noted how commonplace citation analysis had become in research assessment, and emphasized that he had not advocated using citation analyses in evaluating individual scientists. At that point, he felt obliged to give guidance to administrators on how to use citations responsibly in tenure and promotion decisions. Observing the popularity of citation metrics in research evaluation, later commentators—usually focusing on the misuse of journal impact factors rather than citations in general—have often raised the issues in starker terms, even characterizing the situation as "tyranny", "mania" and "obsession".[19]

As citations gained prominence in research evaluation, this tilted incentives in favor of incremental science at the expense of more innovative science.[20] Increased competition for resources has further exacerbated these distortions.[21] It is thus not surprising that science has become more conservative in the decades that span the citation revolution: quantitative evidence from biomedicine shows that scientists are now less likely to try out new ideas in their work. For instance, University of Chicago biologist Andrey Rzhetsky and his colleagues, writing in the Proceedings of the National Academy of Sciences report that "the typical research strategy used to explore chemical relationships in biomedicine... generates conservative research choices focused on building up knowledge around important molecules. These choices [have] become more conservative over time."[22] Another paper in the American Sociological Review by the same team (led this time by UCLA sociologist Jacob Foster) reports even more bluntly that:

> "High-risk innovation strategies are rare and reflect a growing focus on established knowledge. An innovative publication is more likely to achieve

[19] See e.g. Seglen (1997), Lawrence (2003, 2007), Calquhoun (2003), Alberts (2013), Scheckman and Patterson (2013) and Berenbaum (2019).
[20] The recent decade has seen many attempts at alleviating these distortions (see footnote 26).
[21] On increased competition in science, see e.g. Edwards and Roy (2017) who argue that in the last 50 years incentives in science have changed dramatically for the worse as a result of increased competition and emphasis on quantitative metrics. In contrast with our analysis, they do not explicitly tie these developments to novelty of research and the stagnation in scientific progress. Furthermore, while they call for a de-emphasis of quantitative metrics in research evaluation, we call for a more balanced set of metrics to be used in research evaluation.
[22] Rzhetzky et al. (2015).

> high impact than a conservative one, but the additional reward does not compensate for the risk of failing to publish."[23]

In recent decades the largest scientific funding agency, the National Institutes of Health (NIH), has also become less likely to support novel work despite its best efforts to the contrary.[24] The NIH's failure to foster more innovative science in biomedicine is not surprising. Its own funding structures appear to stifle scientific creativity both in terms of novelty and impact.[25] Furthermore, beyond NIH's control, impact factors still dominate the assessment of scientific contributions in hiring, promotion, and salary decisions. And while the NIH strives to evaluate the innovativeness of research proposals separately, it too currently lacks access to metrics to evaluate the innovativeness of researchers' existing contributions and is thus forced to rely largely on impact factors and citation counts in any quantitative evaluation of scientists or grant programs.

Given the changes in incentives toward more conservative science, it is not surprising that research productivity in science has also decreased (as discussed in the introduction). This decline is most apparent in areas such as biomedicine, where direct measures of the practical benefits of science are available. It is also not surprising that science no longer fuels technological innovations at the same robust rate as during earlier eras, as evidenced by the ongoing slowdowns in productivity growth and economic progress.

While others have attributed this stagnation to science getting harder due to the gradual vanishing of valuable secrets and the associated diminishing of research opportunities, we see this change as an unintended consequence of the citation revolution. That science no longer produces breakthroughs comparable to those in the (now quite distant) past is an expected consequence of the shifting of reward structures and research priorities in science in favoring incremental work.

---

[23] Foster et al. (2015).

[24] On increased conservatism in NIH funded research, see Packalen and Bhattacharya (2018).

[25] Azoulay et al. (2011) show that recipients of grants from the Howard Hughes Medical Institute that tolerate early failure better than NIH grants have higher productivity in terms of novelty and impact.

Figure 4 conceptually illustrates the shift in the distribution of scientists' research priorities away from exploration phase science. Before the citation obsession (left panel), a healthy proportion of scientists was willing to engage in true exploration and scientific play with new ideas. Today (right panel), after the shift to the single-minded celebration of high-impact science, scientists are less willing to engage in true exploration. This shift in the distribution of scientists' effort away from exploration and toward incremental science has been costly; as attention to novel ideas has decreased, science has stagnated.

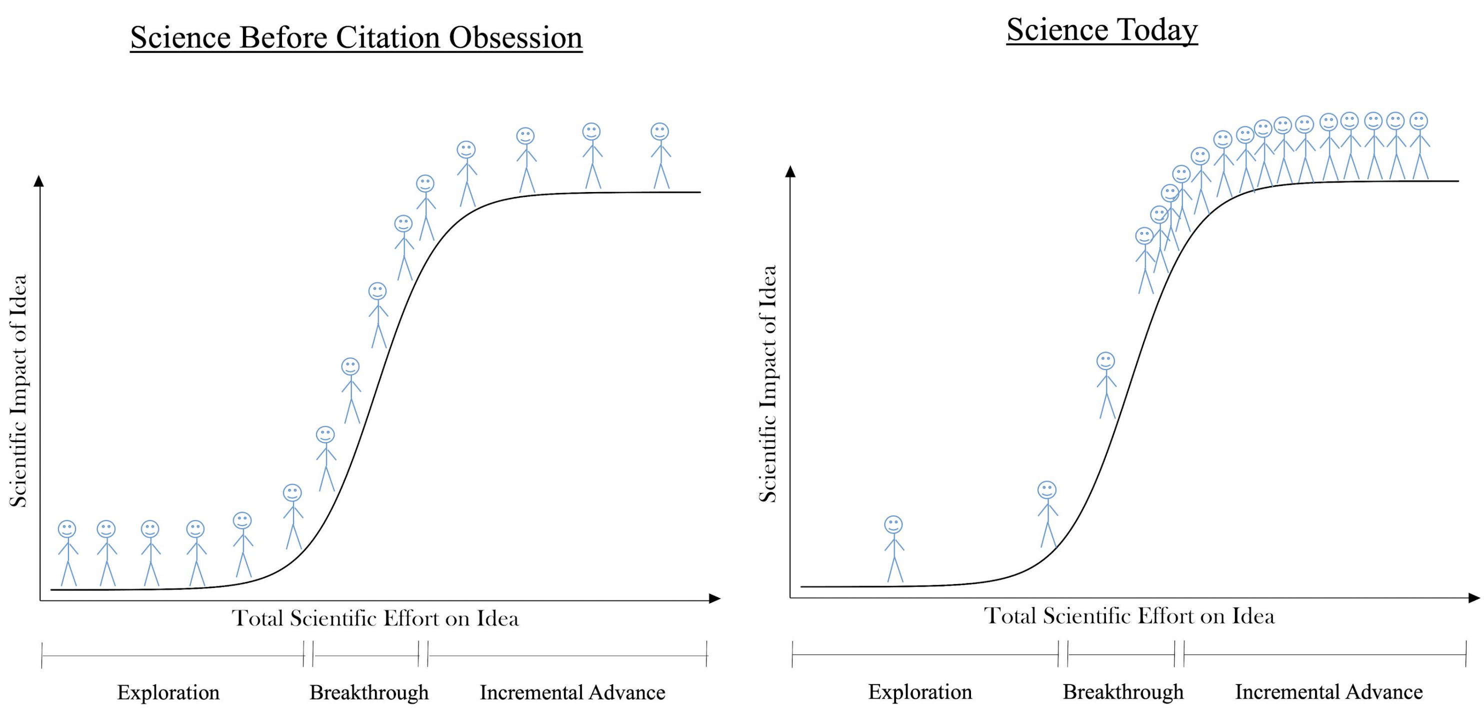

Figure 4. The shift in scientists' effort away from exploration and toward incremental science. Before citations came to dominate research evaluation, more scientists were willing to engage in risky exploration that laid the groundwork for later breakthroughs (left panel). The citation revolution led to a decline in exploration in favor of incremental science. The decline in exploration eventually also decreased opportunities for breakthrough science (right panel).

The potential reform that we identify would change scientist incentives more directly than past reform attempts such as NIH's decision to include 'innovation' as one of many review criteria. Specifically, the reform would measure and provide administrators access to metrics that capture the novelty of contributions in addition to the citation-based impact metrics that administrators already use. A complementary reform by science funders would be to train reviewers on grant review panels to understand that high-impact science is not the only type of science worth funding. Though scientists understand the importance of exploration phase research, the fixation with citations has led much of this generation of scientists to view citations as the only measure of scientific value. This fact, in turn, has doomed the funding prospects of countless innovative grant proposals that do not appear evidently feasible to reviewers. Access to quantifiable metrics on the novelty of each publication and grant proposal, and training reviewers on their purpose, would make it easier for innovative proposals to get funded.

6. Google Scholar and Other Bibliometric Services Perpetuate the Citation Obsession

Today, the fixation with scientific impact, as measured by citations, is on full display on all popular bibliometric services such as Google Scholar and Web of Science. While these services provide a valuable function in that they help researchers identify related works, at the same time, they reduce the value of each scientific contribution to only the citations it has received, and they reduce the worth of each scientist to an index that captures how many citations the scientist's works have generated to date. In Google Scholar, for example, the default view for each scientist shows how many citations the scientist's papers have received (summarized by the total citation count and the h-index) and lists the scientist's papers in a descending order based on the number of citations to them. The alternative view panel, which can be seen by the press of a button, lists the scientist's most recent research papers first. By contrast, information on what kind of science each paper represents—whether the research is exploratory or whether it advances well-established ideas—is not readily available. Nor does Google Scholar show a measure of a scientist's proclivity to engage in the exploration of new ideas.

By presenting information only on scientific influence, the popular bibliometric services Google Scholar, Web of Science, Scopus, and PubMed, as well as their new competitors such as Dimensions and Semantic Scholar, distort scientists' incentives toward incremental science. While their stated purpose is to accelerate scientific discovery, they end up hindering scientific progress. Rather than helping scientists "stand on the shoulders of giants" (which is Google Scholar's motto) or promoting science conducted by exploratory phase giants, these services in their current form end up diminishing the exploration upon which groundbreaking discoveries build. As a result, scientists explore fewer new ideas, which in turn causes breakthroughs to become rarer.

Because citation-focused approaches distort scientist incentives in this way, simply counting papers rather than citations would likely lead to a more innovative scientific ecosystem. However, simply counting papers creates its own distortions in incentives. By rewarding volume alone and not quality, the average quality of scientific work would likely diminish. In any case, given the widespread and easy availability of citation-based metrics and their correlation with scientific impact and contribution quality, it is unlikely that the scientific community would accept a return to a volume-based measure alone.

As we have emphasized repeatedly, measuring scientific influence is useful. However, focus on influence alone ultimately leads to stagnant science. The alternative approach that we identify is that scientific contributions and scientists are evaluated on multiple dimensions, including volume, impact, and novelty. This approach would be more balanced than the present situation, because the kind of science a scientist pursues is as important as how influential or prolific the scientist is. Curiously, as it currently stands, even baseball players are evaluated on many more dimensions than scientists. Yet this comparison is useful for just as sabermetrics serves a useful purpose in baseball, so can a broad enough suite of measures of scientific production help guide scientists, university administrators, and funding agencies.

Our analysis of the state of affairs in science is thus different from the criticism of those who scorn any metric that measures scientific production as 'bean-counting.'[26] The

---

[26] Recent proposed attempts to alleviate the distortions in scientist incentives include the Declaration of Research Assessment (DORA), the Leiden Manifesto, and the Relative Citation Ratio index (Cagan 2013; Hicks

potential reform that we identify does not involve ending the counting of beans in this sense. Rather, implementing this reform leads to counting not only beans (i.e. measuring and rewarding scientific impact) but also fruits and vegetables (i.e. measuring and rewarding also the novelty of scientific work). The goal of this healthy scientific diet is to avoid stagnant science.

The potential reform that we identify seeks to reverse the decline in novelty and exploration in science. The first step in the reform concerns bibliometric services, especially the market leaders Google Scholar and Web of Science. In the reform, these services start including also measures that capture what kind of science each scientific paper pursues. In particular, these services start measuring whether a contribution represents work that tries out new ideas or work that builds on mature, well-established ideas. In the reform, these services also change their scientist-level assessments to include measures that capture to what extent each scientist engages in exploration versus incremental science.

7. Measuring Exploration in Science

While others have suggested that economic and scientific stagnation is an inevitable consequence of diminished scientific opportunities, our message is a constructive one. We

---

et al. 2015; Hutchins et al. 2016). Our analysis of what is at the center of these distortions and the potential reform that we identify stand in marked contrast with this prior work. For we see the key issue in science to be the fixation with citations in general rather than the use of journal-level as opposed to article-level impact measures or excessive reliance on metrics over subjective evaluation. Accordingly, the potential reform that we identify focuses on incorporating also measures of novelty in research evaluation—alongside measures of scientific impact—whereas this prior work advocates for moving from journal- to article-level measures of scientific impact and for diminishing the role of quantitative metrics in research evaluation. As we discuss in the concluding section, attempts to move away from the quantification of scientific contributions are likely to be unfruitful. We also suspect that a shift even further away from using journal-level impact factors and toward emphasizing article-level measures of scientific impact in research evaluation—beyond the shift already facilitated by Google Scholar and other services that have given everyone easy access to article-level citation counts—would have mostly adverse effects in terms of novelty of science unless we simultaneously start measuring also article-level novelty. For unless novelty is also measured, such a shift to increase the emphasis on article-level impact measures in research evaluation would further diminish the incentive to pursue novel work because researchers who are successful in publishing novel but less cited work in high-impact journals would then receive less credit for their work than they currently do.

believe that a decline in scientists' incentives to pursue the exploration of the unknown drives stagnation, and thus rewarding scientific novelty in addition to influence can end the stagnation. Because scientists, like all people, respond to incentives, increasing rewards for novel work will induce scientists once again to pursue exploratory phase research directions more often. Of course, to reward scientific novelty, we must first measure it.

Measuring any aspect of scientific activity is not a simple task. For example, it is well-known that citations are an imperfect and noisy way to measure scientific influence, and yet, citations are a useful measure of scientific influence. Accurately measuring the scientific novelty of a contribution is similarly difficult. Recent advances, though, have rendered the measurement of scientific novelty no more difficult or flawed than the measurement of scientific influence.

Our preferred approach to measuring scientific novelty is based on a textual analysis of research publications. The underlying premise of this approach is three-fold. First, each scientific contribution builds on and advances many ideas. By using an idea and further elaborating on it, scientific work advances our understanding of the idea. Second, the text of a scientific contribution reveals many of the important ideas that the work builds upon and advances. Third, new ideas in science often manifest themselves as new words and word sequences.

Combining these premises, we arrive at the following idea: by indexing the words and word sequences that appear in each scientific paper, we can construct first a list of the ideas that each scientific contribution builds upon and advances. The vintage of each idea can then be determined based on how long ago the idea first appeared in the literature. Having determined the ideas that appear in each paper and the vintage of each idea, we can then identify which research papers try out and advance relatively new ideas. Contributions that build on relatively recent ideas represent novel science, whereas contributions that only build on well-established ideas represent more conventional science. In our prior published research work, we have implemented this idea to measure scientific novelty.[27]

[27] See, for instance, Packalen and Bhattacharya (2019).

Thus, a simple approach already exists for identifying ideas used by scientists and the vintage of those ideas, an approach that allows us to distinguish which papers represent edge science and which do not. Paper-level measures of novelty can then be used to calculate, for example, scientist-level and journal-level tendencies to pursue and promote novel science. Calculating such "edge factors" that capture the average tendency to use novel ideas is analogous to calculating "impact factors" that capture the average tendency to publish influential research papers and which today is commonly used to evaluate scientists, scientific journals, and research institutions.

While the impact factor and edge factors measures are related, they capture two distinct aspects of science. Conceptually, this point is illustrated in Figure 5, which shows the distinct types of science that impact factors and edge factors measure. Impact factors are useful for rewarding work done during the breakthrough phase and during the early part of the incremental advance phase. Impact factors, however, cannot be used to encourage more exploration in science. Edge factors, by contrast, are useful for rewarding work done during the exploration phase. It is thus necessary to have both measures at hand to provide scientists incentives that properly balance the rewards that accrue to those who engage in an exploration of new ideas in their work and the rewards that accrue to those who pursue work that builds on more mature ideas.

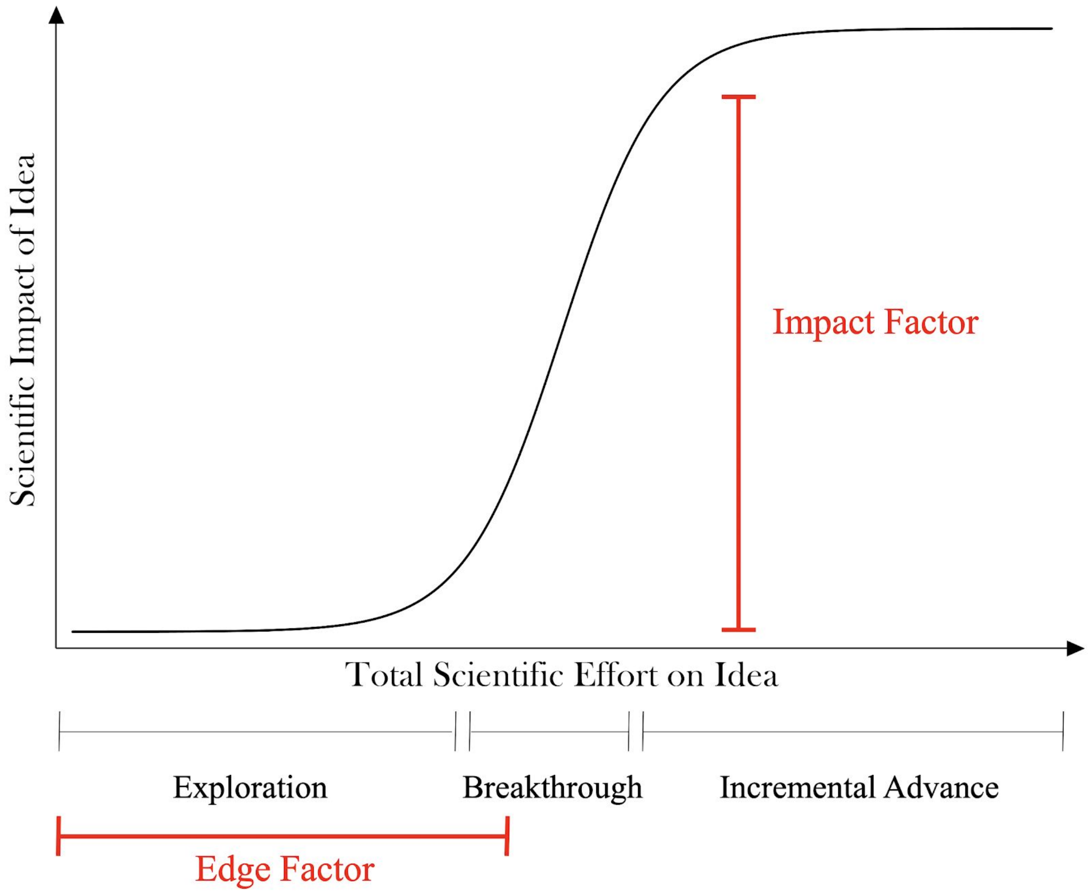

Figure 5. The impact factor and the edge factor capture distinct aspects of science. The impact factor (citation counts) identifies contributions, journals, and scientists that had a large influence on the scientific community. The edge factor identifies contributions, journals, and scientists that explore and develop new ideas, contributions that tend to receive fewer citations but are novel and upon which breakthroughs are built. By measuring and also rewarding scientific novelty as opposed to only scientific influence, we can encourage more scientists to again engage in risky exploration, resulting in healthier science.

Figure 6 illustrates a further fundamental difference between impact and edge factors. While impact factors are mainly useful for rewarding high-impact work on successful ideas, edge factors are useful for rewarding early exploratory work both on successful ideas (ideas A and B) and on ideas that ultimately fail (ideas C and D).

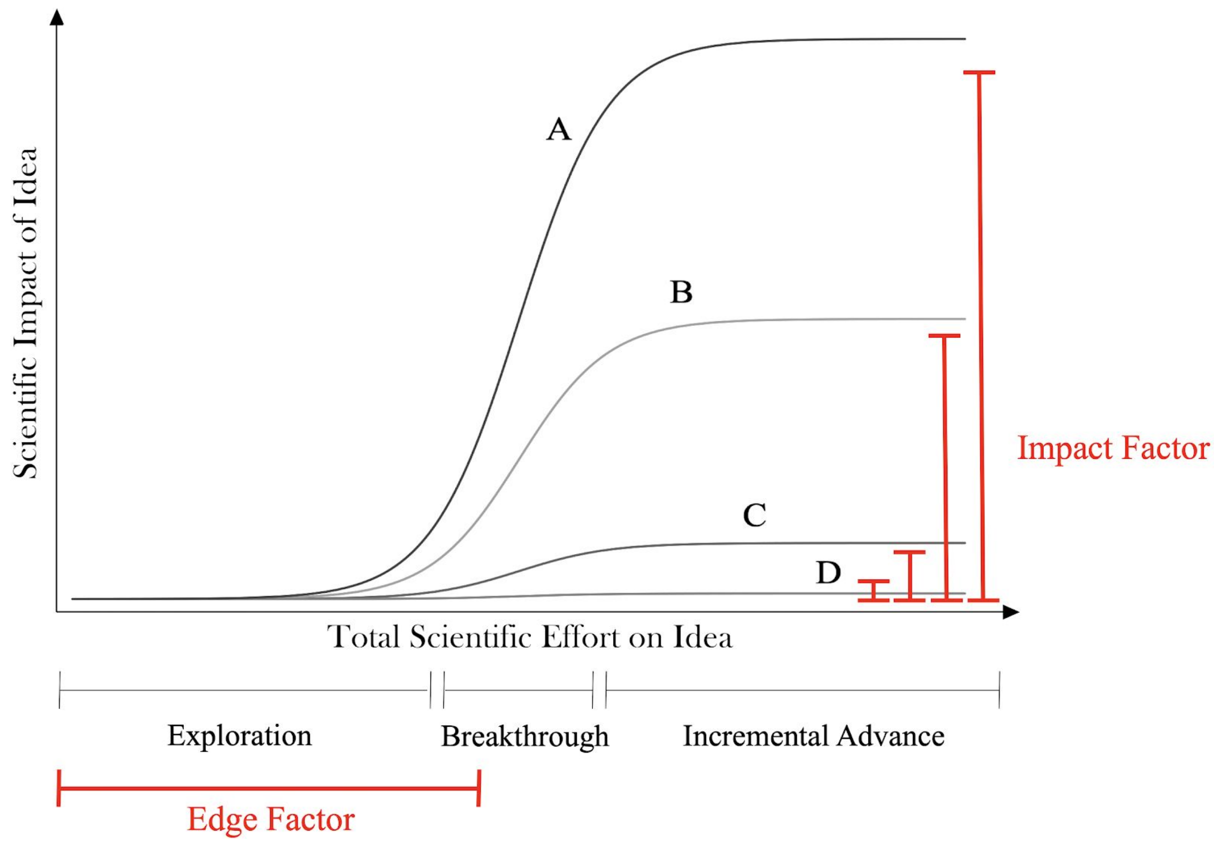

Figure 6. The impact factor and the edge factor for four ideas. The impact factor (citation counts at journal-, scientist-, or article-level) rewards only high-impact work on ideas that succeed. The edge factor rewards early exploratory work on successful ideas (A and B) and on ideas that ultimately fail (C and D). Increasing the tolerance for failure is one key aim of measuring and rewarding scientific novelty separately from scientific impact.

This ability to increase scientists' tolerance for failure is one key objective of measuring and rewarding explorative work on new ideas. Of course, no scientist will purposefully choose to explore an idea that she deems more likely to fail than other available new ideas. The human capital accumulated during explorative work on an idea is more useful for the scientist if the idea is successful. A scientist who embarks on exploratory work thus has an incentive to select the new idea she believes is the most likely to work.

Empirically, measures of novelty are correlated with but distinct from measures of scientific influence. We next illustrate this with data on biomedical research papers. In this analysis, we rely on an analysis of the textual content of every published biomedical research paper indexed in the comprehensive PubMed database. Using the method we describe above, we identify every idea input employed in each paper, and then we determine the vintage of each idea. Using this information, we determine the novelty of the idea inputs for every published peer-reviewed paper in PubMed. We then rank papers by the novelty of their idea inputs. We designate papers that contain newer ideas at the time of publication (within the top 20% of idea vintage recency) as relying on particularly novel ideas and construct an indicator variable of novelty on this paper. We also observe the number of citations each paper received since publication through 2015, which is a measure of the scientific influence of each paper. We then rank papers by their received citation counts and determine the citation percentile of each paper. Finally, we calculate for each citation percentile group what share of papers are novel and what share of papers build on more mature ideas.

Figure 7 depicts the resulting empirical link between citations and novelty. In this figure, the horizontal axis captures the citation percentile, where we group papers to twenty groups that range from the bottom 5% least cited to the top 5% most cited. The vertical axis in turn captures the novelty status of each research paper. The figure illustrates that while novelty rank (edge science) and the citation count rank (scientific influence) are positively correlated, the two measures are distinct. Research evaluation that focuses on scientific influence will mainly reward papers in region 1 as these papers receive the bulk of citations. By contrast, research evaluation that measures and also rewards novelty will reward also papers in region 2, which includes both novel, high-impact papers and novel, low-impact papers.

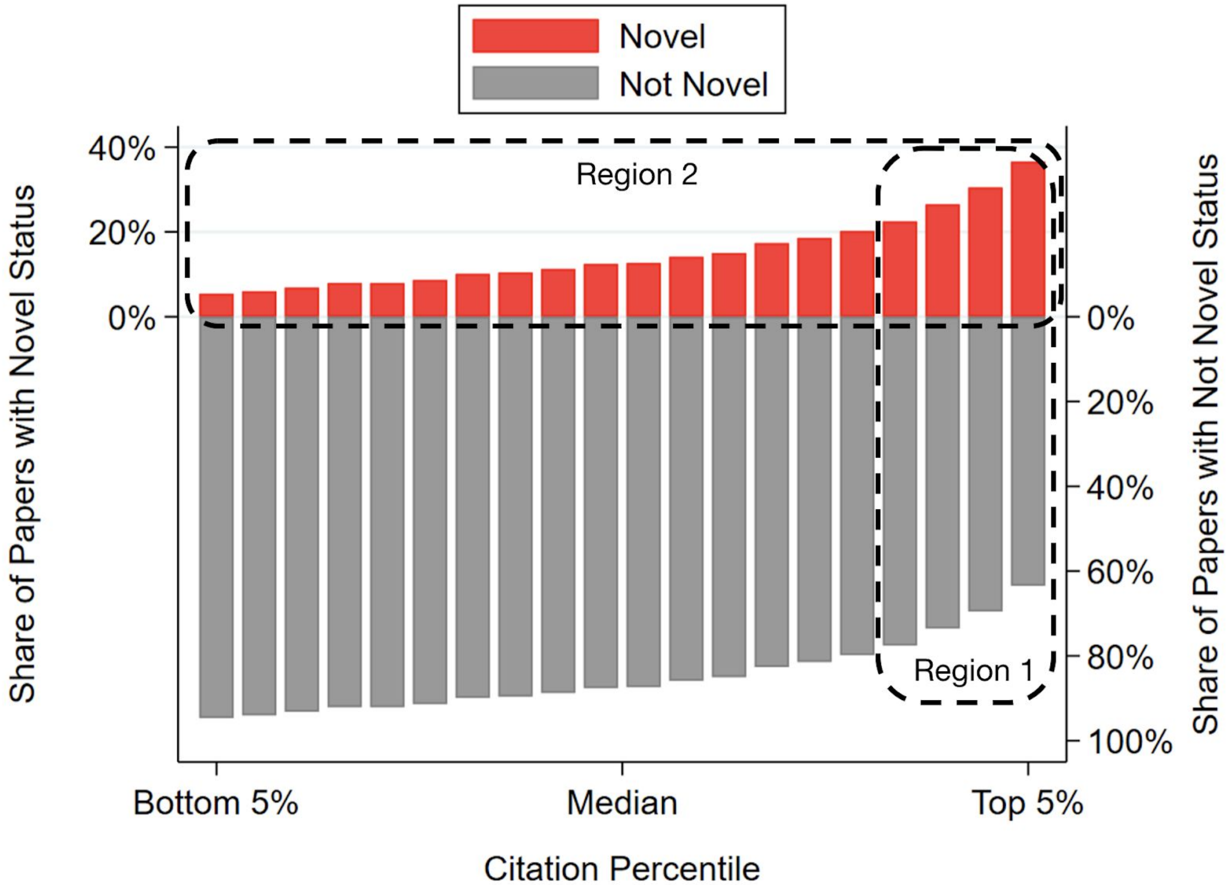

Figure 7. Empirical link between a measure of novelty and a measure of scientific influence. The horizontal axis captures the citation percentile of a research paper. The vertical axis captures whether a research paper is in the top 20% in terms of novelty of idea inputs. We calculate novelty status and citation percentile for 191,354 biomedical papers published in 2001 (PubMed/Web of Science data).

Novelty and scientific influence are empirically distinct for two reasons. First, science that builds on mature ideas does sometimes lead to high-impact contributions, even breakthroughs. This fact underlines the point that incremental science is not useless. The danger with incremental science is not that it serves no purpose; rather, the danger with incremental science is that there is too much of it, that not enough attention is given to new ideas when incentives are tilted too much in favor of incremental science. Second, novel science is not synonymous with breakthrough science. On the contrary, it often leads to failure. This failure comes in many forms, including a lack of publishable work. Even when work on a novel idea results in a publication, other scientists often do not find the idea worthy of pursuit. Alternatively, the idea may not yet be mature enough to have much scientific influence.

Research evaluation that measures and rewards edge science thus increases the incentive to pursue work that results in both potential breakthrough successes and unproductive failures. While rewarding edge science will increase the tolerance for these types of failures in science, it does not mean fetishizing failure. Such incentives will not induce scientists to try out new ideas that they know will certainly fail. As long as research evaluation also rewards scientific impact, the incentive to pursue early work on ultimately successful ideas will still be considerably higher than the incentive to pursue early work on ideas that fail merely for the sake of trying something new. Relative to the status quo, scientists will have a greater incentive to work on those new ideas that they perceive to have the best chance of ultimately leading to transformative discoveries.

As with citation-based measures of scientific influence, the text-based measures of novelty too can have unintended consequences. For example, scientists and journals may be tempted merely to mention new ideas rather than actually incorporate them into their work. For most individuals and journals, the potential reputational costs from writing and publishing frivolously should prevent this sort of gaming. Moreover, future researchers will develop algorithms to detect such behavior, as will new, more robust versions of the edge science measure. These developments will mirror the proliferation of various citation-based indices that have improved on the initial impact measures introduced years ago.[28]

Synonyms can also pose a potential threat to the measurement scheme we propose since this approach treats every unique word or phrase in the text of scientific publications as representing a unique idea. The simple approach we outline above then will work fine as long as two distinct words or phrases never mean the same thing, but we know from common English that that is not the case, and the same holds for scientific and technical vocabularies. In biomedicine, for example, synonyms abound as can be seen from medical

[28] Recent work aimed at improving citation-based measures of scientific influence include Valenzuela et al. (2015), Catalini et al. (2015) and Gerow et al. (2018). That citations are yet a very flawed measure of knowledge flows extends to patent citations, as demonstrated by Arora et al. (2018).

vocabularies, and they certainly have the potential to cause problems for using text to identify edge science.

Fortunately, however, there is a simple solution to this problem, namely the medical thesaurus itself. In particular, the National Library of Medicine in the U.S. compiles a particularly comprehensive medical dictionary that maps millions of biomedical terms together based on whether they refer to the same concept. This metathesaurus, called the Unified Medical Language System ("UMLS"), is collated by professional medical librarians and other experts who scour and combine other medical vocabularies to identify new concepts and place them in their proper place. In our past work, for example, we have taken advantage of this incredible project to inoculate our edge science measure against the problems caused by synonyms. A further advantage is that, since they link each concept to a concept category, the UMLS thesaurus allows us to categorize each idea in each paper based on the new ideas they represent. The UMLS, for instance, enables us to identify papers that explore and develop a recently discovered gene and papers that use a recently introduced research tool.

Thus, one approach for handling synonyms is developing and utilizing large-scale controlled vocabularies. Another option for handling synonyms involves applying recently developed machine learning methods. These methods embed each word and phrase into a vector space, in which "distance" between words and phrases can be measured based upon how often two words or phrases appear near each other in the published literature. To the extent that synonyms are interchangeable, their vector space positions will be nearly identical. The distance of any two words or phrases in the vector space can thus be used to determine which terms are nearly synonymous with one another.

Furthermore, there are fewer synonyms for biomedical terms than one might expect. Indeed, as the purpose of language is communication, it is inefficient to have multiple words and phrases with the same meaning. This fact limits the number of synonyms in natural languages, especially after a word has acquired a well-established, widely used, meaning. Words that represent a scientific idea are thus more likely to change when the underlying idea is still relatively novel. For example, the acronym CRISPR and the corresponding term were invented in 2002 to unify the terminology used until then. This terminological unification happened fifteen years after the first papers on CRISPR had been published and after it had already become an emerging area of investigation, but before its purpose and potential uses had been discovered, and well before the scientists who are now predicted to win Nobel Prizes for CRISPR had begun their work on the idea.

While above we introduced to the reader one particular way of measuring which papers pursue more explorative research directions, this specific approach is not the main thrust of our argument. Rather, our main point is that there is a rationale for measuring also exploration in science and that there already exist ways of doing so.[29] Of course, we are confident that future analysts will develop better ways of measuring edge science. Indeed, should Google Scholar and other bibliometric services start including measures of exploration on them, there will be a strong incentive to further improve upon these metrics just as there is now a strong incentive to continuously improve citation-based measures of scientific influence. We already actively use citation metrics, though they are still imperfect measures of scientific influence. Similarly, the reform that we identify does not entail waiting until we have perfect measures of scientific exploration before they be incorporated to Google Scholar and other popular bibliometric services and as inputs to scientific resource allocation decision.

8. Conclusion

An influential essay written by engineer and science administrator Vannevar Bush in 1945 put forward a grand vision for post-war science. A central focus of this vision was the need to facilitate exploration in science:

[29] Other approaches to measuring novelty include Jones et al. (2008), Azoulay et al. (2011), Kelly et al. (2018), Iaria et al. (2018), Lee et al. (2015), Wang et al. (2016), Jones and Weinberg (2010), Youn et al. (2015), Uzzi et al. (2013) and Boudreau et al. (2016). For the Doc2Vec algorithm, see Le and Mikolov (2014). Giorcelli et al. (2018) utilize a related Word2Vec algorithm to study the flow of new ideas from science to culture. Our own work on measuring new ideas and novelty include Bhattacharya and Packalen (2011), Packalen and Bhattacharya (2015ab, 2017, 2018, 2019) and Packalen (2019).

> "Scientific progress on a broad front results from the free play of free intellects, working on subjects of their own choice, in the manner dictated by their curiosity for exploration of the unknown."[30]

The impetus for this emphasis was that academic freedom had in many ways been lost during the war, as scientific endeavors were redirected to support the needs of the military. This emphasis on scientific play stands in stark contrast with a common refrain about modern science according to which scientists today have gone too far in terms of following their own whims in search of the unknown at the expense of work that places more emphasis on potential practical societal benefits.[31]

Our analysis in this essay, however, indicates that the issue with modern science is more likely the opposite of this common refrain. Scientists now engage in true scientific exploration much less than once envisioned by Vannevar Bush. The rise of citation metrics to prominence in research evaluation shifted scientists' incentives away from the exploration of new ideas and toward incremental science. Because scientists are like everyone else in that they respond to incentives, this shift induced scientists to increasingly pursue research that closely follows well-established research paths and generates only incremental advances, at the expense of the kind of exploration emphasized by Vannevar Bush. The decreased attention to exploring and nurturing new ideas has meant that fewer new ideas have had the opportunity to be developed into breakthroughs, resulting in more stagnant science.[32]

Broadly, we see three potential paths forward for science. Along the first path, science continues on the current citation-focused path that heavily favors incremental work over true exploration. Choosing this route means accepting diminished hopes for the future of scientific and technological progress.

[30] Bush (1945a).

[31] On science as an ivory tower in this sense, see Sarewitz (2016).

[32] To the extent that it was the citation revolution that led to the decline of free play in science and thus to the betrayal of Vannevar Bush's grand vision, this development is somewhat ironic since another Vannevar Bush essay—written that same year—inspired the birth of citation indices. For this essay, see Bush (1945b). For how the essay influenced the development of citation metrics, see Brynko (2007) and Tutterow and Evans (2016).

Along the second path, we stop measuring scientific impact. To us, this path seems infeasible; it stretches credulity that in today's era of relentless quantification, scientists could successfully insist that anything but their own activities can be usefully measured. Moreover, when used properly, quantified measures of research productivity—including citations—can serve a useful purpose in allocating limited research dollars. The issue with using citations and with measuring scientific impact, in general, is not that impact is unimportant. Rather, the issue is that scientific impact has gradually become the only dimension on which scientific research and scientists are evaluated.

The third possible path forward is the reform that we identify and which seeks to put science back on the path envisioned by Vannevar Bush. This reform broadens how scientific productivity is measured and evaluated. Specifically, in addition to scientific impact, we start measuring and evaluating research also what kind of science it represents—whether the work is novel in that it tries out new ideas or more conventional in that it seeks to advance well-established ideas.

Implementing the reform would consist of two steps. In the first step, bibliometric services like Google Scholar end their exclusive focus on citations and scientific influence when summarizing the contributions of scientists, journals, and academic institutions. This change can be implemented relatively quickly. The services would calculate and report measures that capture the novelty of scientific papers. And as we explained in the previous section, measuring novelty is no more difficult or flawed than is measuring scientific impact.

In the second step of the reform, university administrators and funding agencies start shifting reward structures back toward a more balanced model in which rewards depend on both the number citations and on the extent to which the work represents true scientific exploration of the unknown. Soon thereafter, we expect that journal editors will start offering more space for novel contributions as they too will no longer compete only on the scientific influence of the papers that they publish.

Together, these changes would enable scientists once again to pursue the exploration of the unknown without fear of being penalized for it. This will happen in at least three ways. First, the changes in incentives will encourage some current scientists to change their research focus toward more exploration. Second, these changes will encourage some adventurous people, who now choose to stay out of academia for fear that true exploration is no longer appreciated in science, to become scientists in the first place.

Third, by encouraging and rewarding exploration, and by increasing the tolerance for failure, these changes will make it considerably easier to establish new scientific communities that explore and develop new areas of investigation. The birth of such communities, and the debate and scientific play that they facilitate, is a key component of a fruitful scientific enterprise. However, because new areas of investigation are usually small in terms of the number of scientists working in them, the current practice of research evaluation that focuses on scientific impact alone severely punishes scientists working in communities focused on new ideas.

The potential reform that we identify is a large-scale institutional response to the current scientific stagnation.[33] Measuring and evaluating science and scientists in a more balanced way than is currently done would jumpstart a new era of true exploration in science and this exploration would serve as a springboard for new scientific breakthroughs. The reform thus runs counter to the common belief that we live in an era of vanishing secrets and that stagnation is thus inevitable.

The potential reform is also a very low-cost solution; it does not involve allocating more funds for science, and is developed as a response to the fact that science has not scaled well—increased investments in science have not been accompanied by acceleration of progress. By changing what aspects of scientific production are measured and how scientists are evaluated, the reform would only change where scientists direct their attention and where limited research dollars ultimately flow.

The innovation-fueled science that the potential reform would spur might have the follow-on consequence of reversing the longstanding slowdown in productivity growth, which many believe has been driven by the scientific slowdown. However, we do not want

[33] Collison and Nielsen (2018) call for a large-scale institutional response to stagnation but lament the lack of available remedies. Also Romer (2019) calls for fresh meta-ideas for improving science. Cowen and Collison (2019) refer to institutional initiatives as progress engineering and propose a new field of progress studies.

to overpromise; we, of course, do not know what kind of breakthroughs we will get from this different, novelty-driven science. It is possible that the new scientific knowledge that comes from the reform will not come with practical benefits that would also reignite economic growth. Nevertheless, we know for certain that the kind of citation-driven science we have been getting in recent decades has failed to facilitate consistently high economic growth rates in advanced economies, in spite of increased investments in science. By tempering the unintended consequences of the citation revolution, a novelty reformation in science evaluation would thus not only reignite science but might have more tangible benefits as well.

References

Akerlof, G., 2019, "Sins of Omission and the Practice of Economics," *Journal of Economic Literature*, forthcoming.

Akerlof, G. A. and P. Michaillat, 2018, Persistence of False Paradigms in Low-Power Sciences," *Proceedings of the National Academy of Sciences*," 115(52), 13228-13233.

Alberts, B., 2013, "Impact Factor Distortions," *Science,* 340, 6134.

Arbesman, S., 2011, "Quantifying the Ease of Scientometric Discovery," *Scientometrics*, 86(2), 245-250.

Arora, A., S. Belenzon and H. Lee. 2018. "Reversed Citations and the Localization of Knowledge Spillovers," *Journal of Economic Geography,* 18(3), 495-521.

Azoulay, P., Graff Zivin, J. S. and G. Manso, 2011, "Incentives and creativity: evidence from the academic life sciences," *RAND Journal of Economics,* 42(3), 527-554.

Bensman, S. J., 2007, "Garfield and the impact factor," *Annual Review of Information Science and Technology,* 41(1), 93-155.

Berenbaum, M. R., 2019, "Impact factors impacts early-career scientist careers," *Proceedings of the National Academy of Sciences*, 116(34), 16659-16662.

Besancenot, D., and R. Vranceanu, 2015, "Fear of Novelty: A Model of Strategic Discovery with Strategic Uncertainty," *Economic Inquiry,* 53(2), 1132-1139.

Betz, U. A. K., 2018, "Is the Force Awakening?" *Technological Forecasting and Social Change*, 128, 296-303.

Bhattacharya, J. and M. Packalen, 2011, "Benefits and Opportunities as Determinants of the Direction of Scientific Research," *Journal of Health Economics*, 30(4), 603-615.

Bloom, N., Jones, J., Van Reenen, J. and M. Webb, 2019, "Are Ideas Getting Harder to Find?" Manuscript.

Boudreau, K. J., Guinan, E. C., Lakhari, K. R. and C. Riedl, 2016, “Looking Across and Looking Beyond the Knowledge Frontier: Intellectual Distance, Novelty, and Resource Allocation in Science,” *Management Science,* 62, 2765-2783.

Bramoulle, Y. and G. Saint-Paul, 2010, “Research Cycles,” *Journal of Economic Theory,* 145(5), 1890-1920.

Brock, T. D. and H. Freeze, 1969, “*Thermus aquaticus* gen. n. and sp. n., a Nonsporulating Extreme Thermophile,” *Journal of Bacteriology*, 98(1), 289-297.

Brynjolfsson, E., Rock, D. and C. Syverson, 2019, “Artificial Intelligence and the Modern Productivity Paradox: A Clash of Expectations and Statistics,” in Agarwal, A. K., Gans, J. and A. Goldfarb (ed.) *The Economics of Artificial Intelligence: The Agend*a. Chicago University Press.

Brynko, B., 2007, “An Interview with Eugene Garfield—A Lifetime of Achievement and Still Going Strong,” *Information Today,* 24(1), 21.

Bush, V., 1945a, “Science the Endless Frontier: A Report to the President,” United States Government Printing Office, Washington, D.C.

Bush, V., 1945b, “As We May Think,” *The Atlantic Monthly,* 176(1), 101-108.

Caballero, R., 2010, “Macroeconomics after the Crisis: Time to Deal with the Pretense of Knowledge Syndrome,” *Journal of Economic Perspectives*, 24(4), 85-102.

Cagan, R., 2013, “The San Francisco Declaration on Research Assessment,” *Disease Models & Mechanisms*, 6(4), 869-870.

Calquhoun, D., 2003, “Challenging the Tyranny of Impact Factors,” *Nature*, 423, 479.

Catalini, C., Lacetera, N. and A. Oettl, 2015, The Incidence and Role of Negative Citations in Science,” *Proceedings of the National Academy of Sciences*, 112(45), 13823-13826.

Chien, A., Edgar, D. B., and J. M. Trela, 1976, “Deoxyribonucleic Acid Polymerase from the Extreme Thermophile Thermus aquaticus,” *Journal of Bacteriology,* 127(3), 1550-1557.

Collison, P. and M. Nielsen, 2018, “Science is Getting Less Bang For Its Buck,” *The Atlantic*, November 16.

Collison, P. and T. Cowen, 2019, “We Need a New Science of Progress,” The Atlantic, July 30.

Cowen, T., 2011, *The Great Stagnation: How America Ate All The Low-Hanging Fruit of Modern History, Got Sick, and Will (Eventually) Feel Better.* Dutton.

Cowen, T. and B. Southwood, 2019, “Is the Rate of Scientific Progress Slowing Down?” Manuscript.

Diamond, A. M. Jr., 2019, *Openness to Creative Destruction: Sustaining Innovative Dynamism*. Oxford University Press.

Doudna, J. A. and S. H. Sternberg, 2017, *A Crack in Creation: Gene Editing and the Unthinkable Power to Control Evolution.* Houghton Mifflin Harcourt Publishing Group.

Edwards, M.A. and S. Roy, 2017, “Academic Research in the 21st Century,” Maintaining Scientific Integrity in a Climate of Perverse Incentives and Hypercompetition,” *Environmental Engineering Science*, 34(1), 51-61.

Firestein, S., 2015, *Failure: Why Science is So Successful*. Oxford University Press.

Foster, J. G. and Rzhetsky, A. and J. A. Evans, 2015, “Tradition and Innovation in Scientists’ Research Strategies,” *American Sociological Review,* 80, 875-908.

Foster, R., 1986, *Innovation: The Attacker's Advantage*. Summit Books.
Frey, B. S., 2003, "Publishing as Prostitution?" *Public Choice*, 116, 205-223.
Fussler, H. H., 1949, "Characteristics of the Research Literature Used by Chemists and Physicists in the United States," *Library Quarterly: Information, Community and Policy*, 19(1), 19-35.
Garfield, E., 1955, "Citation Indexes for Science: A New Dimension in Documentation through Association of Ideas," *Science*, 122, 108-111.
Garfield, E., 1972, "Citation Analyses as a Tool in Journal Evaluation," *Science,* 178, 471-478.
Garfield, E., 1983a, "How to Use Citation Analysis for Faculty Evaluations, and When Is It Relevant? Part 1," Current Contents, 44, 5-14.
Garfield, E., 1983b, "How to Use Citation Analysis for Faculty Evaluations, and When Is It Relevant? Part 2," Current Contents, 45, 5-14.
Garfield, E., 1996, "What Is The Primordial Reference For The Phrase 'Publish Or Perish'?" *The Scientist*, 10(12), 11.
Gerow, A., Hu, Y., Boyd-Graber, J., Blei, D. M., and J. A. Evans, 2018, "Measuring Discursive Influence Across Scholarship," *Proceedings of the National Academy of Sciences*, 115(1), 3308-3313.
Giorcelli, M., Lacetera, N. and A. Marinoni, 2018, "Does Scientific Progress Affect Culture? A Digital Text Analysis," NBER Working Paper No. 25429.
Glass, B., 1971, "Science: Endless Horizons or Golden Age?" *Science*, 171(3966), 23-29.
Gordon, R. J., 2000, "Does the 'New Economy' Measure Up to the Great Inventions of the Past?" *Journal of Economic Perspectives*, 14(4), 49-74
Gordon, R. J., 2012, "Is U.S. Economic Growth Over? Faltering Innovation Confronts the Six Headwinds," National Bureau of Economic Research Working Paper No. 18315.
Gordon, R. J., 2016, *The Rise and Fall of American Growth: The U.S. Standard of Living Since the Civil War.* Princeton University Press.
Gordon, R. J. and H. Sayed, 2019, "The Industry Anatomy of the Transatlantic Productivity Growth Slowdown," National Bureau of Economic Research Working Paper No. 25704.
Gross P. L. K. and E. M. Gross, 1927, "College Libraries and Chemical Education", *Science*, 66(1713), 385-9.
Hall, R. E., 2016, "The Anatomy of Stagnation in a Modern Economy," *Economica*, 84(333), 1-15.
Heckman, J. J. and S. Moktan, 2019, "Publishing and Promotion in Economics: The Tyranny of the Top Five," *Journal of Economic Literature*, forthcoming.
Hicks, D., Wouters, P., Waltman, L., de Rijcke, S. and I. Rafols, 2015, "Bibliometrics: The Leiden Manifesto for Research Metrics," *Nature*, 520(7548), 429-431.
Horgan, J., 1996, *The End of Science: Facing the Limits of Knowledge in the Twilight of the Scientific Age*. Basic Books.
Hutchins, B. I, Yuan, X., Anderson, J. M. and G. M. Santangelo, 2016, "Relative Citation Ratio (RCR): A New Metric That Uses Citation Rates to Measure Influence at the Article Level," *PLoS Biology*, 14(9), e1002541.

Iaria, A., Schwarz, C. and F. Waldinger, 2018, "Frontier Knowledge and Scientific Production: Evidence from the Collapse of International Science," *Quarterly Journal of Economics*, 133(2), 927-991.

International Monetary Fund, 2019, *World Economic Outlook: Growth Slowdown, Precarious Recovery*. Washington, DC. April

Jones, B. F., 2010, "Age and Great Invention," *Review of Economics and Statistics*, 92, 1-14.

Jones, B. F., Wuchty, S. and B. Uzzi, 2008, "Multi-University Research Teams: Shifting Impact, Geography, and Stratification in Science," *Science,* 322, 1259-1262.

Jones, B. F. and B. A. Weinberg, 2010, "Age Dynamics in Scientific Creativity," *Proceedings of the National Academy of Sciences*, 108(47), 18910-18914.

Kelly, B. T., Papanikolaou, D., Seru, A. and M. Taddy, 2018, "Measuring Technological Innovation Over the Long Run," NBER Working Paper No. 25266.

Kuhn, T. S., 1962, *The Structure of Scientific Revolutions*. Chicago University Press, Chicago.

Kuhn, T. S., 1977, Objectivity, Value Judgment and Theory Choice; in Thomas S. Kuhn, ed., *The Essential Tension*, University of Chicago Press, Chicago, 320-339.

Lander, E. S., 2016, "The Heroes of CRISPR," *Cell*, 164, 18-28.

Lawrence, P. A., 2003, "The Politics of Publication," *Nature*, 422, 259-261.

Lawrence, P. A., 2007, "The Mismeasurement of Science," *Current Biology*, 17(15), R583-R585.

Le Fanu, J., 1999, *The Rise and Fall of Modern Medicine*. Little Brown and Company, London.

Le Fanu, J., 2010, "Science's Dead End," *Prospect*, August 10.

Le, Q. and T. Mikolov, 2014, "Distributed Representations of Sentences and Documents," *Proceedings of the 31st International Conference on Machine Learning*.

Lee, Y.-N., Walsh, J. P. and J. Wang, 2015, "Creativity in Scientific Teams: Unpacking Novelty and Impact," *Research Policy,* 44, 684-697.

Livio, M., 2013, *Brilliant Blunders: From Darwin to Einstein—Colossal Mistakes by Great Scientists that Changed Our Understanding of Life and Universe.* Simon & Schuster.

Mandel, M., 2009, "The Failed Promise of Innovation in the U.S.," *Business Week.*

Marshall, A., 1920, *Principles of Economics.* Macmillan, London.

Marshall, B. J., 2005, "Helicobacter Connections," Nobel Lecture.

Marshall, B., 2016, "A Brief History of the Discovery of Helicobacter Pylori.," in Suzuki, H., Marshall, B. and R. Warren (eds.) *Helicobacter Pylori.* Springer.

Michelson, A., 1903, *Light Waves and Their Uses*. The University of Chicago Press.

Mojica, F. J. M. and L. Montoliu, 2016, "On the Origin of CRISPR-Cas Technology: From Prokaryotes to Mammals," T*rends in Microbiology*, 24(10), 811-818.

Mullis K, 1994, "The Polymerase Chain Reaction (Nobel Lecture)" Angewandte Chemie International Edition 33(12): 1209-13.

Olby, R., 2003, "A Quiet Debut for the Double Helix," *Nature*, 421, 402-405.

Osterloh, M. and B. S. Frey, 2015, "Ranking Games," *Evaluation Review*, 32, 102-129.

Packalen, M., 2018, "Edge Factors: Scientific Frontier Positions of Nations," *Scientometrics*, forthcoming.

Packalen, M. and J. Bhattacharya, 2015a, "New Ideas in Invention," NBER Working Paper No. 20922.

Packalen, M. and J. Bhattacharya 2015b, "Cities and Ideas," National Bureau of Economic Research working paper No. 20921.

Packalen, M. and J. Bhattacharya, 2017, "Neophilia Ranking of Scientific Journals," *Scientometrics* 110: 43-64.

Packalen, M. and J. Bhattacharya, 2018, "Does the NIH Fund Edge Science?" NBER Working Paper No. 24860.

Packalen, M. and J. Bhattacharya, 2019, "Age and the Trying Out of New Ideas," *Journal of Human Capital*, forthcoming.

Popovian, R., 2016, "Dedicated Scientists Driven to Discover Cures" *Morning Consult*. December 5.

Romer, P., 2015, "Mathiness in the Theory of Growth," *American Economic Review*, 105(5), 89-93.

Romer, P., 2016, "The Trouble with Macroeconomics," *American Economist,* forthcoming.

Romer, P., 2019, "The Deep Structure of Economic Growth," blog post, (https://paulromer.net/deep_structure_growth; accessed May 8, 2019).

Ruhm, C. J., 2019, "Shackling the Identification Police?" S*outhern Economic Journal*, 85(4), 1016-1026.

Rzhetzky, A., Foster, J. G., Foster, I. T., and J. A. Evans, 2015, "Choosing experiments to accelerate collective discovery," *Proceedings of the National Academy of Sciences* 112: 14569-14574.

Sarewitz, D., 2016, "Saving Science," *The New Atlantis: A Journal of Technology and Society*, Spring/Summer, 5-40.

Scannell, J. W., Blanckley, A., Boldon, H. and B. Warrington 2012, "Diagnosing the decline in pharmaceutical R&D efficiency," *Nature Reviews Drug Discovery*, 11, 191-200.

Scheckman, R. and M. Patterson, 2013, "Science Policy: Reforming Research Assessment," *eLife*, 2, e00855.

Seglen, P. O., 1997, Why the Impact Factor Should Not Be Used for Evaluating Research," *BMJ*, 314, 497.

Silverstein, A. M., 1999, ""The End Is Near!": The Phenomenon of the Declaration of Closure in a Discipline," *History of Science*, 37, 407-425.

Small, H., 2018, "Citation Indexing Revisited: Garfield's Early Vision and Its Implications for the Future" *Frontiers in Research Metrics and Analytics*, 3, DOI 10.3389/frma.2018.00008.

Thiel, P., 2011, "The End of The Future," *National Review*. October 3.

Thiel, P., 2014, *Zero to One: Notes on Startups, or How to Build the Future / Peter Thiel with Blake Masters*. Crown Business.

Tutterow, C. and J. A. Evans, 2016, "Reconciling the Small Effect of Rankings on University Performance with the Transformational Cost of Conformity," in (ed.) *The University Under Pressure,* Emerald Group Publishing Limited, 265-301.

Usher, A. P., 1929, *A History of Mechanical Inventions.* Mcgraw-Hill, New York.

Uzzi, B., Mukherjee, S., Stringer, M. and B F. Jones, 2013, "Atypical Combinations and Scientific Impact," *Science,* 342, 268-472.

Youn, H., Strumsky, D., Bettencourt, L. M. A. and J. Lobo, 2015, "Invention as a Combinatorial Process: Evidence from U.S. Patents," *Journal of the Royal Society Interface*, 12, 20150272.

Yoshizumi, I., Krupovic, M. and P. Forterre, 2018, "History of CRISPR-Cas from Encounter with a Mysterious Repeated Sequence to Genome Editing Technology," *Journal of Bacteriology*, 200, e00580-17.

Valenzuela, M., Ha., V. and O. Etzioni, 2015, "Identifying Meaningful Citations," AAAI Workshop: Scholarly Big Data.

Vijg, J., 2011, *The American Technological Challenge: Stagnation and Decline in the 21st Century*. Algora Publishing.

Wang, J., Veugelers, R. and P. Stephan, 2016, "Bias Against Novelty in Science: A Cautionary Tale for Users of Bibliometric Indicators," National Bureau of Economic Research Working Paper No. 22180.

Warren, J. R., 2005, "Helicobacter—The Ease and Difficulty of a New Discovery," Nobel Lecture.

Watson, J. D., 1968, *The Double Helix*. Atheneum Press.

Wootton, D., 2015, *The Invention of Science: A New History of the Scientific Revolution*. Harper.

Wouters, P., 2017, "Eugene Garfield (1925-2017)," *Nature*, 543, 492.

Zaringhalam, M., 2016, "Failure in Science Is Frequent and Inevitable—and We Should Talk More about It," Scientific American blog post, (https://blogs.scientificamerican.com/guest-blog/failure-in-science-is-frequent-and-inevitable-and-we-should-talk-more-about-it; accessed May 8, 2019).