The Performance of Depth

Prestige, incentives, and the reproduction of intellectual conformity

Henry Shin · First published 19 October 2026 · Revised 3 November 2026

Refacte · web edition

Appointments, journals, grants, and citations help universities evaluate work they cannot examine afresh in every decision. Those signals also influence what researchers choose to produce.

This essay asks how reliance on those signals can affect independent judgment: the ability to explain why a method is appropriate, identify what would undermine an assessment, and revise a conclusion when its grounds change.

Researchers rely on questions, methods, and standards received from others. Independent judgment includes recognizing when those resources apply and when they need revision. A new theory, a careful replication, an improved instrument, or a counterexample can each require that judgment.

Institutions can favor the absence of such judgment. They can also constrain people who possess it. A researcher who sees the weakness of an established agenda but cannot afford to leave it occupies a different position from someone unable to examine that agenda. The historical problem concerns both the selection of intellectual habits and the conditions under which independent judgment can affect the work.

Following an evaluation through three steps makes the problem more precise. Which signs are used to allocate an opportunity? How do candidates adapt their work to those signs? Once successful candidates become evaluators, what keeps the signs from becoming the standard itself? These questions distinguish an individual misuse of reputation from a process through which a standard reproduces itself.

A century of certification

The history begins well before university rankings and digital citation dashboards.

In March 1903, William James published “The Ph.D. Octopus,” describing a gifted scholar whose appointment was threatened after his prospective employer discovered that he lacked a doctorate. James and his colleagues considered the candidate exceptionally capable. The institution nevertheless insisted on the certificate. His abilities had already persuaded the people assessing him; what remained unsatisfied was the requirement for a recognized sign of those abilities.[1]

James supplies an early witness to a recurring substitution. A credential intended to record accomplishment can become a condition for recognizing accomplishment at all. Intellectual life then becomes partly organized around acquiring the sign.

The subsequent history contains improvements as well as new forms of dependence. Older appointments could rest on informal sponsorship and decisions whose reasons were difficult to inspect. In an archival study of 145 history appointments at sixteen German universities between

1950 and 1985, Julian Hamann found increasing transparency, public advertising, and more extensive discussion of candidates' publications. Formal procedures could expose private favoritism and require more substantive justification.[2]

That improvement identifies the standard against which later developments should be judged. A procedure strengthens intellectual life when it makes relevant reasons easier to examine. It weakens intellectual life when passing the procedure acquires authority beyond the examination actually performed.

Postwar expansion made this distinction increasingly consequential. Research became more specialized, public funding grew, and more decisions had to be made by people unable to examine every relevant claim themselves. Institutions needed ways to allocate money, positions, publication space, and attention under conditions of limited knowledge.

The late 1960s and 1970s brought an important change in how those decisions were justified. Historian Melinda Baldwin traces the elevation of peer review into a public condition of scientific legitimacy. In disputes over US research funding, scientists defended expert review as the appropriate means of reconciling public accountability with scientific autonomy. Refereeing already existed; what strengthened was the claim that passage through the procedure established scientific credibility.[3]

Peer review supplies an important mechanism of criticism. Its defenders also sought to protect scientific autonomy from political intervention. The public defense increased the authority of procedural certification: whether work had passed through an approved process became more powerful evidence about whether it deserved serious consideration.

From the 1980s onward, other changes made professional evaluation more demanding and more pervasive.

Glenn Ellison documented a substantial slowdown in economics publishing between 1970 and 1999, with increasingly extensive revisions accounting for an important part of the change. At the top economics journals he studied, papers that had typically reached acceptance within six to nine months around 1970 faced roughly two-year review-and-revision cycles around 2000. This establishes rising demands in the production of an acceptable article. It does not establish that the ideas became worse: revisions can improve arguments. The historical point is the growing investment required to satisfy the profession's evolving standards.[4]

Academic employment became less protected as well. According to the AAUP's analysis of federal data, contingent appointments rose from approximately 47 percent of US faculty appointments in 1987 to 68 percent in 2022. These figures cover higher education broadly, including teaching-intensive institutions. They document a substantial expansion of employment arrangements with less durable protection, rather than a corresponding change confined to research professors.[5]

The intellectual implication follows from the conditions of dependence. Sustained inquiry becomes harder to protect when continued employment requires frequent renewal, favorable evaluations, and accommodation to immediate institutional needs. Independence has a material cost. A culture's capacity to sustain it depends partly on whether people can afford to exercise it.

During the 2000s and early 2010s, performance assessment also became more deeply embedded in public research funding across European systems. A comparative study of EU member states documents the spread of performance-based institutional allocation. This development differs from competitive project funding, which did not increase uniformly. The consequential change was the growing connection between assessed research performance and the resources institutions received.[6]

Once assessment affects institutional income, its influence travels downward. Universities have incentives to recruit people whose records perform well under the assessment, encourage outputs that receive recognition, and organize internal decisions around the relevant criteria. Individual career strategies and institutional financial strategies become more closely aligned.

Researchers' responses to numerical incentives can be observed directly. A study of 15,037 Italian professors found that adjusted self-citations per paper increased by 9.5 percent after the introduction of citation-based accreditation criteria in 2012. Responses varied considerably, and many individuals did not increase their rates. The rise, including its variation with professors' accreditation incentives, provides concrete evidence that making an indicator consequential for advancement can alter the behavior generating it.[7]

The language of successful research changed too. Among 901,717 successful NIH grant abstracts from 1985 to 2020, the proportion containing at least one adjective coded as promotional rose from 72 percent to 97 percent. Changing linguistic conventions contribute to the result. What the study establishes is a substantial shift in the presentation of funded research.[8]

These observations describe different aspects of an increasingly consequential evaluation environment: more effort devoted to securing approval, stronger connections between assessment and resources, less protected employment across much of the workforce, adaptation to numerical thresholds, and a more promotional language of achievement.

They do not become evidence of intellectual decline merely by being placed together. More demanding research can require longer training, greater specialization, larger teams, and more careful verification. Benjamin Jones's work on the growing burden of knowledge supplies a substantial account of why some of these changes can arise from intellectual accumulation itself.[9]

The criticism concerns an additional development: satisfying evaluative demands acquiring rewards partly independent of examining the work's purposes and grounds. To establish that concern, we need to look at how institutions judge, allocate opportunity, and reproduce their standards.

How recognition reproduces itself

An institution must decide before the eventual value of much research becomes clear. It therefore relies on previous appointments, journal names, grants, citations, recommendations, and recognized authorities. Such signals help organize decisions that would otherwise be impossible at scale.

They also create a second arena of accomplishment: producing evidence that evaluators will recognize as evidence of accomplishment.

Understanding a difficult problem may require a long period without presentable results. Demonstrating productivity requires a visible sequence of results. Discovering that a standard question is badly framed may make a project harder to classify. Producing a competent extension of a fashionable framework can be immediately legible, even when its larger significance is modest.

As advancement becomes more dependent on legibility, researchers acquire reasons to organize their work around it.

John Meyer and Brian Rowan described a general organizational mechanism behind this possibility. Institutions adopt recognized procedures and structures partly because those forms secure legitimacy, resources, and survival. The formal demonstration that an institution is performing its function can become loosely connected to the activities that fulfill that function. Applied to scholarship, the implication is that a university can become highly competent at demonstrating research performance while judging its intellectual substance less directly.[10]

An author's standing can influence assessment of the work itself.

In a preregistered finance-journal experiment, reviewers recruited through anonymous invitations assessed the same manuscript under a Nobel laureate's name, a relatively unknown researcher's name, or no name. The text remained the same; reviewers were less likely to recommend rejection when it carried the prominent name.[11]

Reputation can carry useful information. The danger is allowing a favorable expectation about an author to perform the work of examining the present argument.

Funding can make an early difference cumulative. In a study of Dutch early-career grants, applicants just above the funding threshold accumulated more than twice as much subsequent research funding over eight years as applicants just below it, excluding the initial award. Their initial evaluations had been nearly identical. Early success nevertheless opened a growing institutional gap, through both subsequent evaluation and differences in willingness to enter further competitions.[12]

An institutional decision helps produce the later record through which that decision appears vindicated. Resources can enable genuine achievement. The error arises when differences in later achievement are treated as though they straightforwardly reveal comparable differences in the original applicants' capacities.

This process becomes culturally consequential when successful researchers acquire the authority to evaluate the next generation. They select candidates, review proposals, assess publications, and define what a promising career should look like. Practices associated with their own success become standards for recognizing excellence in others.

Smaldino and McElreath's "The Natural Selection of Bad Science" makes the selection logic explicit. Their model shows how practices producing rewarded outputs can spread even when they reduce reliability. Laboratories differ in their practices; success affects their reproduction through trainees and institutional continuation; successful practices become more common. A later corrected reimplementation preserved the broad result while changing some replication dynamics.[13]

The model concerns scientific rigor under specified incentives. The proposed extension here concerns habits of judgment: practices that secure advancement also help determine who will assess the next generation.

Committees can sincerely want excellent scholars while repeatedly preferring the most defensible signs of excellence. Candidates can sincerely love their subject while learning to supply those signs. People whose habits fit the selection system advance more readily, while others adapt their habits or leave.

The evidence bears more directly on adaptation to incentives than on changes in researchers' capacity for judgment. Adaptation can nevertheless change who acquires influence: those who thrive under particular demands gain greater power to maintain them.

Conformity need not feel like conformity from within. Someone repeatedly rewarded for asking recognizable questions, using accepted methods, and presenting results in an approved idiom has substantial practical evidence that this is what good scholarship consists of. Adaptation can become conviction.

This process does not require a decline in technical skill. Candidates can become better at the work rewarded by an evaluation while receiving less encouragement to question its purpose. When those who succeed then administer the evaluation, increasing technical demands can coexist with a narrowing of the judgments that influence the research agenda.

The cost of an independent agenda

George Akerlof's critique of economics identifies one expression of this danger. He argues that a preference for methodological "hardness" can direct attention away from important questions difficult to approach through favored methods. A discipline can maintain rigorous standards within its chosen projects while allowing those standards to distort its choice of projects.[14]

The loss often takes the form of work never attempted, questions quietly abandoned, or doubts never developed far enough to become visible.

Growth in the literature can reinforce this dependence. More publications create more possible knowledge, but also more material than researchers can examine. Established reference points become useful shortcuts.

Chu and Evans analyzed approximately 90.6 million papers and 1.8 billion citations across 241 subjects between 1960 and 2014. Larger fields displayed more persistent citation canons and greater difficulty for new papers to enter their most-cited ranks. Their findings support an attention bottleneck: as fields expand, established visibility can become more self-reinforcing.[15]

A stable canon can preserve valuable cumulative knowledge. The danger arises when reliance on that canon increasingly organizes attention before new claims receive substantive examination. Work connected to established reference points becomes easier to recognize and cite; departures must overcome both their intellectual difficulty and the cost of making themselves intelligible within an already crowded field.

A culture can become more productive while making its organizing assumptions harder to disturb.

Research risk adds a further pressure. Foster, Rzhetsky, and Evans examined strategies in biomedical chemistry and found that work exploring new relationships was more likely to be ignored, but also more likely to achieve exceptional impact. Their analysis suggests that the additional citation reward was insufficient to offset the added risk of failing to publish. The aggregate distribution of published strategies remained remarkably stable even while the space of possible new relationships expanded.[16]

For the individual researcher, intellectual promise and career viability can pull apart. A direction may be worth pursuing while being a dangerous way to spend the years before the next consequential evaluation.

The timing of recognition matters. Wang, Veugelers, and Stephan found that novel combinations of prior knowledge carried greater uncertainty, greater potential for long-term impact, and delayed recognition. The patterns varied with the degree of novelty and the citation window. Short observation periods could therefore misrepresent exploratory work whose value took longer to become visible.[17]

Funding design can change the work researchers undertake. Azoulay, Graff Zivin, and Manso compared HHMI investigators—supported through arrangements emphasizing flexibility, longer horizons, and tolerance of early failure—with similarly accomplished NIH-funded scientists. Their matched analyses found substantially more high-impact work and greater exploration of new directions among the HHMI investigators. The comparison supports the importance of the funding environment, while leaving room for differences in selection into the programs.[18]

Creativity depends partly on who can survive the interval between beginning uncertain work and establishing its value.

A system may celebrate originality while making that interval financially or professionally intolerable. Researchers then learn to present adventurousness in forms that look sufficiently predictable to support. The rhetoric of exploration can flourish while the conditions necessary for exploration weaken.

There are genuine counterweights. A study of submissions to 49 life-science and physics journals found that more novel manuscripts were more likely to be accepted, especially when combined with conventional foundations. Publication incentives can reward exploration and partly compensate for risks elsewhere in the process.[19]

Publishing a completed result and financing an uncertain attempt remain different decisions. A journal may welcome a surprising discovery once the evidence exists, while a researcher had difficulty securing the resources needed to produce it. Whether later rewards adequately compensate for earlier risks depends on the funding system, the field, and the researcher's position. The journal study rules out a general claim that peer review inherently punishes novelty; it leaves the distribution of risk as a substantive institutional question.

Other findings show why familiar measures of academic success cannot simply be added together as if they identified the same contribution. Li, Tessone, and Zeng found that greater publication productivity was associated with higher average citations but lower average citation disruption, including during particularly productive years within researchers' careers. Wu, Wang, and Evans,

analyzing more than 65 million papers, patents, and software products, found that smaller teams tended toward more disruptive citation patterns, while larger teams tended toward developing existing trajectories.[20,21]

These studies concern research organization and subsequent influence. They distinguish kinds of contribution that an institution may otherwise compress into a general impression of productivity or impact.

The much-discussed historical decline in citation disruption requires similar care. Park, Leahey, and Funk reported a long decline across approximately 45 million papers and 3.9 million patents. A subsequent critique identified database artifacts affecting the measured trend; the authors' reply defended substantial declines under additional analyses. The magnitude and interpretation of the finding remain matters of methodological dispute.[22]

A citation index cannot settle the historical distribution of intellectual depth. The institutional argument rests on independently documented changes in evaluation, employment, presentation, behavior, opportunity, and attention. These describe conditions under which recognized academic achievement can become more advantageous than sustained examination of whether the prevailing agenda deserves its authority.

What becomes dominant

Established leadership also shapes who enters a domain. In a study of the premature deaths of eminent life scientists, Azoulay, Fons-Rosen, and Graff Zivin found that contributions by outsiders to the affected fields increased after a star's death. The new contributions drew on a different scientific corpus and were disproportionately likely to be highly cited. The authors found little support for a simple account of stars directly blocking entry through editorial or funding control; deterrence and the structure of the intellectual community were important parts of the explanation.[23]

Dominance need not mean a numerical majority. People exercise intellectual power when they determine which questions receive attention, which departures appear credible, and which risks a younger researcher can afford to take.

The historical conclusion concerns how the documented changes alter that distribution of power. Less protected employment raises the cost of departure from an established agenda; performance-based allocation makes recognizable outputs more consequential for institutional resources; cumulative advantage helps place successful researchers in positions from which they assess the next generation. These pressures can shape both the work that survives and the standards used to choose its successors.

This can transform capable people as well as favor compatible ones. A researcher begins with independent ambitions, then acquires students, obligations, a laboratory, and a public identity. An unfamiliar direction threatens the machinery supporting both the researcher and others. Responsibilities produced by success narrow the freedom through which the success originally arose.

Where inquiry can continue

The changing conditions inside universities raise a further historical question: where else have the means of inquiry become available? That history is uneven.

During the late twentieth century, many major corporations were retreating from internal published science. Arora, Belenzon, and Patacconi documented this development among American R&D-performing firms during 1980–2006, even while scientific knowledge remained important to their inventions. Corporate research had its own difficulty sustaining long-horizon inquiry.[24]

The later industrial concentration of AI research followed a different trajectory. As deep learning developed during the 2010s, compute, data, engineering infrastructure, and capital became increasingly consequential to the experiments researchers could conduct. The industrial advantage changed the feasible research agenda.

Ahmed, Wahed, and Thompson reported that the share of North American AI PhDs entering industry rose from 21 percent in 2004 to almost 70 percent in 2020. Their account identifies a growing concentration of resources required for important parts of AI research.[25]

A 2026 working paper by Akcigit and colleagues links publication histories to US employer–employee records for approximately 42,000 AI researchers. The nonacademic employment share rose from 48 percent in 2001 to 68 percent in 2019. In their event-study comparison, three years after a move, researchers leaving academia published roughly two-thirds fewer papers, patented substantially more, and earned more than researchers moving between academic jobs. Top industrial earnings accelerated around the deep-learning breakthrough period beginning in 2012; departures toward large incumbents intensified after 2017, particularly among younger researchers. These are observed timings, rather than isolated causal effects of particular inventions.[26]

By 2025, industry produced over 90 percent of the notable AI models tracked by Stanford's AI Index. This establishes a dramatic concentration of frontier development capacity. Recent employment flows followed a different pattern: industry's share of new US and Canadian AI PhD employment remained flat between 2022 and 2024. Concentration of frontier infrastructure and the movement of researchers are related but distinct developments.[27]

The migration of material capability is clearer than the migration of independent judgment. A study of AI researchers moving to industry found that higher measured novelty predicted a lower propensity to leave, and that published novelty fell immediately after a move. Its coverage is limited to researchers who continued publishing, but the finding directly resists the story that the most exploratory minds simply departed and became freer.[28]

Firms can supply extraordinary resources while organizing inquiry around ownership, secrecy, deployment, and commercial strategy. Universities can offer greater agenda freedom while lacking the resources to execute particular projects. Intellectual opportunity changes with the combination.

Alternative institutes make the organizational question explicit. Janelia, opened in 2006, emphasizes internal funding, small individual research groups, shared infrastructure, and leaders who remain active researchers. Arc, announced in 2021, offers renewable eight-year funding to core

laboratories while maintaining substantial university partnerships. Focused Research Organizations fund defined efforts to build scientific tools and infrastructure that conventional academic and commercial arrangements have difficulty supplying.[29,30,31]

These are experiments in changing the relationship between inquiry and institutional survival. They support different time horizons or forms of coordination while retaining demanding evaluation. Arc's founding rationale also identifies an excessive premium on novelty relative to robustness and reproducibility. An institution can distort inquiry through compulsive novelty-seeking as well as through compulsive conservatism.

The common question is how much authority the reward system acquires over what deserves to be done.

The important historical development is a redistribution of the conditions under which inquiry is possible. Some opportunities have concentrated in large firms; others have developed through philanthropic support and alternative organizations; many remain within universities. Their intellectual value depends on the resources, time, criticism, and control over the agenda they provide.

Authority beyond the record

A distinguished appointment, a publication record, and technical expertise supply relevant evidence. Their meaning depends on the work assessed, the conditions under which it was produced, and the judgment now being requested. A record can justify confidence in one capacity while leaving another insufficiently examined.

Independence becomes visible through sustained intellectual conduct: locating the assumption doing the work, distinguishing a result from its interpretation, identifying what would change a conclusion, and correcting the affected claim when an objection succeeds. A brief encounter or an unanswered message cannot settle those capacities. Repeated engagement with the substance can reveal them.

The same discipline applies to criticism of academia. A claim of originality requires an argument. Rejection does not establish depth. An outsider can become as attached to an image of neglected genius as an insider becomes attached to an image of recognized authority.

What gives the institutional problem its force is the capacity of recognition to reproduce the evidence supporting its own judgments. Publications justify appointments; appointments lend credibility to evaluations; evaluations allocate resources; resources make further publications possible. The sequence can sustain excellent work. It can also make an inherited standard increasingly difficult to examine, because each successful passage through it supplies another reason to trust those who administer it.

One way to examine this process is to follow a decision through the cycle: what was assessed, which evidence counted, what resources followed, and whether later success was used to justify the original assessment. That would test the proposed mechanism more directly than a general judgment about academia's depth.

Sources and notes

The historical argument distinguishes dated institutional changes, studies of mechanisms, and the cultural inference drawn from their interaction. Notes identify each source's relevant scope.

1. William James. The Ph.D. Octopus. *Harvard Monthly* (March 1903). A contemporary criticism of credential-based hiring, used here as a historical witness to the substitution of certification for judgments of demonstrated ability. ↩

2. Julian Hamann. The making of professors: Assessment and recognition in academic recruitment. *Social Studies of Science* 49(6): 919–941 (2019). Archival study of 145 history appointments at sixteen German universities, 1950–1985. It traces scholarly, formal, and organizational judgments, alongside increasing public advertisement and documentation of decisions. ↩

3. Melinda Baldwin. Scientific Autonomy, Public Accountability, and the Rise of "Peer Review" in the Cold War United States. *Isis* 109(3): 538–558 (2018). DOI: 10.1086/700070. Historical analysis of journal and funding refereeing, especially US controversies in the late 1960s and 1970s. The change concerns peer review's public authority as a condition of legitimacy; refereeing itself was much older. ↩

4. Glenn Ellison. The Slowdown of the Economics Publishing Process. *Journal of Political Economy* 110(5): 947–993 (2002). Documents lengthening submission-to-acceptance times at leading economics journals during 1970–1999, with more extensive revisions contributing substantially. The study identifies increasing demands on accepted papers; the quality and intellectual value of those demands require a further judgment. ↩

5. American Association of University Professors. Contingent Appointments and the Academic Profession. Revised 2023; statistical information updated 2024. AAUP analysis of federal data reports contingent appointments at approximately 47% of US faculty in 1987 and 68% in 2022. Coverage includes teaching-intensive institutions as well as research universities. ↩

6. Thomas Zacharewicz, Benedetto Lepori, Emanuela Reale, and Koen Jonkers. Performance-based research funding in EU Member States—a comparative assessment. *Science and Public Policy* 46(1): 105–115 (2019). DOI: 10.1093/scipol/scy041. Compares European systems using funding data for 2000–2014 and national policy accounts. Performance-based institutional allocation spread, with substantial national variation; competitive project funding did not increase uniformly. ↩

7. Giovanni Abramo, Ciriaco Andrea D'Angelo, and Leonardo Grilli. The effects of citation-based research evaluation schemes on self-citation behavior. *Journal of Informetrics* 15(4): 101204 (2021). Analyzes 15,037 Italian professors over 2008–2017. Adjusted self-citations per paper increased 9.5% after the 2012 accreditation criteria, with considerable variation among individuals and differences associated with their incentives. ↩

8. Neil Millar, Bojan Batalo, and Brian Budgell. Trends in the Use of Promotional Language (Hype) in Abstracts of Successful National Institutes of Health Grant Applications, 1985-2020. *JAMA Network Open* 5(8): e2228676 (2022). DOI: 10.1001/jamanetworkopen.2022.28676. Analysis of 901,717 funded NIH grant abstracts, 1985–2020. The share containing at least one adjective coded as promotional rose from 72% to 97%; the study concerns public-facing language and acknowledges linguistic change. ↩

9. Benjamin F. Jones. The Burden of Knowledge and the "Death of the Renaissance Man": Is Innovation Getting Harder? *Review of Economic Studies* 76(1): 283–317 (2009). Develops and tests a burden-of-knowledge account of longer learning periods, greater specialization, and teamwork. It supplies an explanation rooted in intellectual accumulation for changes that should not automatically be interpreted as decline. ↩

10. John W. Meyer and Brian Rowan. Institutionalized Organizations: Formal Structure as Myth and Ceremony. *American Journal of Sociology* 83(2): 340–363 (1977). Organizational theory explaining how recognized formal structures secure legitimacy and resources while becoming decoupled from ongoing activity. Its application to the relationship between academic performance and understanding is the essay's interpretation. ↩

11. Jürgen Huber, Sabiou Inoua, Rudolf Kerschbamer, Christian König-Kersting, Stefan Palan, and Vernon L. Smith. Nobel and novice: Author prominence affects peer review. *Proceedings of the National Academy of Sciences* 119(41): e2205779119 (2022). Preregistered field experiment presenting an identical finance manuscript under different author identities. The comparison discussed here uses the conditions with anonymous review invitations. ↩

12. Thijs Bol, Mathijs de Vaan, and Arnout van de Rijt. The Matthew effect in science funding. *Proceedings of the National Academy of Sciences* 115(19): 4887–4890 (2018). Comparison near a Dutch early-career funding threshold. Over eight years, initial winners accumulated more than twice the subsequent research funding of near-winners, excluding the initial award; later application behavior also contributed to the gap. ↩

13. Paul E. Smaldino and Richard McElreath. The natural selection of bad science. *Royal Society Open Science* 3(9): 160384 (2016). Paul E. Smaldino and Richard McElreath. Correction to: 'The natural selection of bad science' (2016) by Paul E. Smaldino and Richard McElreath. *Royal Society Open Science* 10(9): 231026 (2023). See also Florian Kohrt, Paul E. Smaldino, Richard McElreath, and Felix Schönbrodt. Replication of the natural selection of bad science. *Royal Society Open Science* 10: 221306 (2023). A model of selection among research practices under publication incentives. A corrected reimplementation changes some replication dynamics while preserving the broad result. The extension from methodological rigor to habits of judgment is an inference developed in the essay. ↩

14. George A. Akerlof. Sins of Omission and the Practice of Economics. *Journal of Economic Literature* 58(2): 405–418 (2020). A methodological critique of economics: preference for "hard" approaches can divert attention from important questions less amenable to those approaches. Used for its account of research agenda selection. ↩

15. Johan S. G. Chu and James A. Evans. Slowed canonical progress in large fields of science. *Proceedings of the National Academy of Sciences* 118(41): e2021636118 (2021). DOI: 10.1073/pnas.2021636118. Analyzes 90.6 million papers and 1.8 billion citations across 241 subjects, 1960–2014. Larger fields show more persistent citation canons and reduced entry into the most-cited ranks. Canon persistence and substantive intellectual progress are distinct. ↩

16. Jacob G. Foster, Andrey Rzhetsky, and James A. Evans. Tradition and Innovation in Scientists' Research Strategies. *American Sociological Review* 80(5): 875–908 (2015). Studies biomedical chemical-relationship research strategies during 1983–2008. Exploratory strategies carry greater risk and exceptional-reward potential. The published strategy mix remained remarkably stable as the space of possible new relationships expanded. ↩

17. Jian Wang, Reinhilde Veugelers, and Paula Stephan. Bias against novelty in science: A cautionary tale for users of bibliometric indicators. *Research Policy* 46(8): 1416–1436 (2017). Measures novelty through new combinations of cited journals and examines citation outcomes over different windows. The degree of novelty matters: short windows can miss later advantages, without establishing a universal citation trajectory for exploratory work. ↩

18. Pierre Azoulay, Joshua S. Graff Zivin, and Gustavo Manso. Incentives and creativity: evidence from the academic life sciences. *The RAND Journal of Economics* 42(3): 527–554 (2011). DOI: 10.1111/j.1756-2171.2011.00140.x. Compares HHMI investigators with similarly accomplished NIH-funded

scientists using matching/weighting and longitudinal comparisons. Flexible, longer-horizon support is associated with more high-impact work and exploration; entry into the programs was not randomized. ↩

19. Misha Teplitskiy, Hao Peng, Andrea Blasco, and Karim R. Lakhani. Is novel research worth doing? Evidence from peer review at 49 journals. *Proceedings of the National Academy of Sciences* 119(47): e2118046119 (2022). Across 49 life-science and physics journals, greater measured novelty was associated with higher acceptance, particularly alongside conventional foundations. Novelty uses combinations of cited journals; parts of the analytic sample require locating a subsequently published version. ↩

20. Heyang Li, Claudio J. Tessone, and An Zeng. Productive scientists are associated with lower disruption in scientific publishing. *Proceedings of the National Academy of Sciences* 121(21): e2322462121 (2024). DOI: 10.1073/pnas.2322462121. Finds an association between greater publication productivity, higher average citations, and lower average citation disruption, including variation within careers. These measures distinguish patterns of output and influence; they do not by themselves establish causes or intellectual depth. ↩

21. Lingfei Wu, Dashun Wang, and James A. Evans. Large teams develop and small teams disrupt science and technology. *Nature* 566(7744): 378–382 (2019). Analyzes more than 65 million papers, patents, and software products. Smaller and larger teams show different patterns of subsequent citation disruption and development. The measure concerns citation patterns associated with team organization. ↩

22. Michael Park, Erin Leahey, and Russell J. Funk. Papers and patents are becoming less disruptive over time. *Nature* 613(7942): 138–144 (2023). DOI: 10.1038/s41586-022-05543-x. Vincent Holst, Andres Algaba, Floriano Tori, Sylvia Wenmackers, and Vincent Ginis. Dataset artefacts can partially drive the measured decline in disruption. *Nature* 656(8127): E7–E13 (2026). DOI: 10.1038/s41586-026-10787-y. Michael Park, Erin Leahey, and Russell J. Funk. Reply to: Dataset artefacts can partially drive the measured decline in disruption. *Nature* 656(8127): E14–E21 (2026). DOI: 10.1038/s41586-026-10788-x. The original study reports declining citation disruption in approximately 45 million papers and 3.9 million patents. The 2026 exchange disputes how dataset artifacts affect the size and interpretation of that trend. The essay treats the exchange as an unresolved methodological dispute. ↩

23. Pierre Azoulay, Christian Fons-Rosen, and Joshua S. Graff Zivin. Does Science Advance One Funeral at a Time? *American Economic Review* 109(8): 2889–2920 (2019). Studies premature deaths of eminent life scientists and subsequent publication in their research areas. Outsider contributions increased, often drawing on different literature. The evidence does not reduce the mechanism to direct editorial or funding obstruction by the stars. ↩

24. Ashish Arora, Sharon Belenzon, and Andrea Patacconi. The decline of science in corporate R&D. *Strategic Management Journal* 39(1): 3–32 (2018). DOI: 10.1002/smj.2693. Examines US R&D-performing firms over 1980–2006. Their participation in published science declined even while scientific knowledge remained valuable for invention. The measure concerns internal published research, not all corporate research activity. ↩

25. Nur Ahmed, Muntasir Wahed, and Neil C. Thompson. The growing influence of industry in AI research. *Science* 379(6635): 884–886 (2023). DOI: 10.1126/science.ade2420. Policy Forum assembling evidence about industry's growing AI research resources and influence. The doctoral employment figures describe new North American AI PhDs entering industry, rather than the employment share of all AI researchers. ↩

26. Ufuk Akcigit, Craig A. Chikis, Emin Dinlersoz, and Nathan Goldschlag. Attention (And Money) Is All You Need: Why Universities Are Struggling to Keep AI Talent. NBER Working Paper 34964 (2026). DOI: 10.3386/w34964. Working paper linking publication histories to US employer–employee records for approximately 42,000 AI researchers. It documents employment and earnings shifts and compares publication and patenting changes around moves. The post-2017 departure pattern toward large

incumbents is especially pronounced among researchers under forty. Nonacademic employment is broader than employment in technology firms. ↩

27. Stanford Institute for Human-Centered Artificial Intelligence. *The 2026 AI Index Report*, chapters Research and Development and Education (2026). The first chapter reports industry's share of notable models in 2025; the education chapter reports employment of new US and Canadian AI PhDs through 2024. Model-development concentration and graduate employment flows have different populations and time windows. ↩

28. Roman Jurowetzki, Daniel S. Hain, Kevin Wirtz, and Stefano Bianchini. The private sector is hoarding AI researchers: what implications for science? *AI & SOCIETY* 40(5): 4145–4152 (2025). Tracks AI researchers across academia and industry using publication histories. Higher measured novelty is associated with a lower propensity to leave, and published novelty falls initially after a move. The observed output excludes research that firms do not publish. ↩

29. Howard Hughes Medical Institute, Janelia Research Campus. Our Model (undated institutional account; accessed 19 October 2026). See also Janelia Turns 10 (2016). The model page describes internal funding, small individual groups, shared infrastructure, and active bench scientists as group leaders. The anniversary account dates the campus opening to October 2006. These sources describe institutional design and history. ↩

30. Patrick Hsu, Silvana Konermann, and Patrick Collison. Introducing Arc Institute. Arc Institute, 15 December 2021. Founding statement describing renewable eight-year funding, university partnerships, and technology centers. It also identifies an excessive premium on novelty relative to robustness and reproducibility. Used as evidence of an organizational design and its stated rationale. ↩

31. Convergent Research. About FROs (undated institutional account; accessed 19 October 2026). Describes nonprofit teams building public tools, datasets, and scientific infrastructure over defined periods, with specified technical milestones and external review. These organizations address coordination and infrastructure bottlenecks, rather than offering unrestricted research support. ↩