Science

HOME > SCIENCE > VOL. 393, NO. 6808 > ADVANCING SCIENCE BY DESIGNING FOR SURPRISE

EXPERT VOICES

Advancing science by designing for surprise

JAMES EVANS, CASEY PETROFF, AND GARY KING

Science • 16 Jul 2026 • Vol 393, Issue 6808 • DOI: 10.1126/science.aej4257

Consider a thought experiment. Two graduate students have written identical dissertations: the same experiments, the same data, the same robust findings. One is written as a standard, cautious text that flatters the subfield whose intuitions drove the design and to whom the findings will feel familiar. The other reframes the study for a different, but equally relevant, audience—one holding different suppositions and therefore open to being surprised into shifting its position. The latter approach is more valuable for good reasons: the scientific community and the world will learn more.

The job of a scientist is to explain and predict the world, and a vital part of this objective is to deliver knowledge in a way that updates the understandings and expectations of the largest audience. However, scientists have long treated “framing” as a secondary art form. Students are advised to do rigorous empirical work, sprinkled with creative brilliance for which there is neither definition nor strict formula. But what if audiences’ expectations and understandings could be measured and used to guide what scientists choose to study, what conclusions they draw from the evidence, and how they frame findings for distribution? What if the framing and choice of audience are adjustable parameters of a new, broader theory of inference? The question grows particularly urgent as literature becomes flooded with machine-produced papers written by no one, for no one.

Papers are typically written as cautious confirmations. They marshal evidence for a pre-existing hypothesis, build incrementally on prior work, and frame results as expected extensions of existing knowledge. This style is designed to reflect new findings back to the writers’ home community. But it also slows scientific progress.

Information theory, pioneered by mathematician Claude Shannon, offers a different perspective: only surprises carry new information. If people in an American town have conservative political preferences, and someone demonstrates that they are more likely to vote Republican, little is learned. A paper that confirms what readers already believe tells them little they did not already know and perhaps nothing that changes their worldview. Yet, the incentives of modern science reward this type of uninformative communication. Peer reviewers are drawn from one’s own field and publication norms favor deference to established paradigms. Scientists cite sources in their immediate field roughly 500% more than those in distant fields. This is performance art designed for reviewers rather than a search for the most informative surprise designed to advance knowledge. Thus, most papers are written to be unsurprising, and consequently, have minimal impact. These consequences are linked. The ultimate value that a paper adds to collective knowledge depends not only on what it shows, but on whose worldview it updates, and by how much.

To understand why surprise is the currency of impact requires understanding what science is. Science is not merely a method. It is an emergent, socially organized system that has produced more knowledge over the last 400 years than at any time in human history, by any means. The architecture deserves credit, and within it, the papers that contribute the most are those that update knowledge the most. This structure is governed by rules of cooperation (sharing methods, data, and goals) and decorum (validating empirical claims solely with empirical evidence, and justifying the importance of these claims based solely on whether other scientists will go along). But its primary driver is competition. Progress requires persuading the scientific community to accept new empirical claims. What is less appreciated is that the farther a scientist can move the predominant position of a community, the more people will acknowledge the finding, and the more importance will be attributed to the work. The predominant view is often supported by considerable evidence, which is why moving it demands extraordinary evidence and why the system offers its largest rewards for successes and largest penalties for failures, the farther a claim sits from that view. A researcher who convinces the scientific community, with robust evidence, that objects can travel faster than the speed of light would become famous, and the researcher who presses such claims without convincing evidence would be ignored or sanctioned.

How can surprise be moved from art to science? The answer may lie in the logic of philosopher Charles Sanders Peirce. He described science as encompassing the processes of deduction (extending a rule to a case to get a result) and induction (observing cases and results to form a rule) to understand the world. But he also proposed that discoveries increasingly arise through abduction—the collision between an expectation and a surprising anomaly, leading to a new proposed rule. Abduction begins with observing something that should not have happened under prevailing expectations, but did. It is not only the logic of discovery, but also the logic of communication. Entire formal disciplines are devoted to deduction (mathematics) and induction (statistics), yet abduction has thus far been left to qualitative intuition or dismissed as showmanship.

This is why Hypothesizing After the Results are Known (HARKing) represents an epistemic vice. HARKing erases surprise. A researcher looks at data, spots a surprising pattern, and then creates a hypothesis to match it—but then presents the findings as expected. According to Peirce, this destroys information. By contrast, abductive framing preserves surprise. It makes the field's prior expectations explicit so that the violation becomes legible. It suggests further data collection to confirm the pattern. HARKing hides the anomaly; abductive framing helps people learn from it.

Framing surprise may sound like rhetorical ornamentation—a manipulative trick to make an anomaly more appealing. A 2017 study of crowd wisdom suggests otherwise. The study focused on questions on which majorities are reliably wrong, such as 'Is Philadelphia the capital of Pennsylvania?' (it is Harrisburg). Majority vote fails, but a different procedure succeeds: Ask respondents both what they believe and what they expect others to believe. The correct answer tends to be the one endorsed more often than the crowd predicted because respondents who know the truth also know most others will get it wrong. Surprise, in other words, can be a diagnostic of truth: a crowd becomes wiser when its members recognize where their private, accurate information would surprise everyone else.

Scientific papers are votes in a vast, distributed crowd-wisdom procedure. A paper that confirms what informed readers already expect adds little to collective knowledge. But a paper that presents evidence unanticipated by a relevant audience provokes surprise on a larger scale. The author's task is to create a prediction gap between what informed readers predict and what the evidence shows: Here is what readers expected; Here is what the evidence shows; Here is how the finding should change the audience's understanding.

Recent large-scale analyses of the scientific corpus demonstrate that surprising combinations of research contents (what the paper is about) and contexts (which fields it draws from) are highly predictive of a paper landing in the top 10% of citations. The most surprising papers are roughly five times more likely to be massive hits. Notably, citation impact disproportionately rewards surprises that cross disciplinary boundaries. The most impactful surprises come from "knowledge expeditions" where scientists from one field travel to address problems in a distant other. These outsiders bring alien data patterns, frameworks, and theories that disrupt expectations.

A skeptic might worry that abductive framing invites a reverse-HARKing problem: Are findings selected because they surprise someone, and audiences selected simply because they are easily surprised? Does this not manufacture novelty by dressing ordinary findings with a new appearance? Validating a surprise requires more evidence to be convincing, and the best such evidence is numerous, diverse and, at best, produces consilience. In the 1840s, philosopher William Whewell argued that when independent inductions from disparate phenomena converge on the same generalization, the conjunction is more robust than any single line of evidence. A surprising finding supported by a single, underpowered study is hype. But a surprising finding supported by independent methods, diverse datasets, distant researchers, and intersecting theoretical mechanisms is the most valuable kind of scientific contribution. Consilience makes the surprise difficult to dismiss as an artifact of clever framing. This implies a calibration norm: The intensity of surprise framing must track the strength and breadth of the evidence. This involves the author identifying the audience whose priors the evidence most violates, confirming the evidence can bear that rhetorical weight, and staking reputation on the claim.

This kind of strategic framing belongs inside the tent of formal inference. Consider statistical inference. Decades ago, missing data was considered a nuisance of preprocessing in which all blank responses were set to the mean or the respondent was deleted and formal statistical inferences were made as if there were no gaps in the data. This missing data problem was later formalized, quantified, and incorporated as a core part of modern inference. The same evolution of formal statistical inference occurred over many years with sample selection, data preprocessing, measurement error, among other methods. The theory of inference keeps expanding the tent.

Framing, audience selection, and strategic shifts in the research question represent a new frontier to be formalized by the disciplines of statistics and data science, as well as the methodological subfields within disciplines. Scientists must move beyond telling students and colleagues to "add brilliance." They need to systematically map how changing the question of interest and intended audience to maximize surprise can solve problems so completely that the problems cease to exist or advance knowledge in important ways. This will require methods for systematically measuring the predominant positions from different fields. A model of a field's perspective is an explicit, quantitative representation of what that community collectively accepts, expects, and would find surprising. Large language models trained on a field's literature already approximate such expectations, and modern statistical tools can calibrate those approximations against surveys of scientists, their citation behavior, and their data, turning audience priors from mystery into measurement. Once audience expectation is brought fully inside the theory of inference—treating what they believe as a measurable input to discovery—scientists can begin to formally establish its properties, optimize methods and research procedures, and adjust research efforts to achieve it.

To formalize this program, questions need to be addressed. Of all possible audiences, how do we identify where a researcher's tentative empirical results would be the most relevant, surprising, and impactful? Once found, how best does one frame a question that would result in maximum impact? And how best to represent uncertainty if the inference includes not merely determining the quantity to be estimated and the question asked, but the framing of the question and the audience addressed?

A paper that changes nothing about how people think is uninformative and uninteresting. Science advances through the systematic, rigorous deployment of the unexpected. It is time for the scientific community to learn how to design—and write—for surprise.