

# Metascience Manquè

## Five Years of Never Met a Science

Kevin Munger

Date: Jul 17, 2025

URL: <https://kevinmunger.substack.com/p/metascience-manque>

In a total coincidence, today is the 5th anniversary of this blog \*and\* the 100th post! I obviously haven't been hitting the platform-optimal once-a-week minimum but, as a compromise, I do at least spend way too much time writing these things.

My first post promised that this blog would be about three things

- social science methodology (and how it's changing because of the internet), including my recent focus on meta-science, temporal validity and quantitative description
- political communication theory (and how it's changing because of the internet),
- the practice of culture and politics (and how it's changing because of the internet)

Politics, the internet and metascience — the three themes of the blog — have all changed quite a bit since then!

That first is mostly complaints about Twitter, a habit I still heartily endorse.

what is the *ideal* online platform for social scientists to collaborate, share knowledge and tell jokes? I'm confident it's not Twitter; Twitter is a for-profit corporation with a goal that is not to design the ideal platform for social scientists. Further, Twitter is constantly changing, in terms of the userbase and the platform itself; even if it *were* ideal at one point, it won't stay that way.



pretty sure this is not the ideal platform for academics

Over time, my posts have been less often about metascience and more often about internet/media theory. Not at all ironically, my most popular post provides a theory of why that might be the case: I am a content creator on a platform with highly visible audience metrics, and I have been adapting my content to this audience as operationalized by those metrics.

Indeed, Substack (the platform) continues to roll out new audience tracking metrics, and at least appears to be rolling out new audiences — the fact that the platform is growing lifts all 'stacks, giving each author the impression that they're doing better and better, driving us to produce more and more content.

This can be read cynically, as the platform extracting more and more free labor from its producers in order to sell more ads — and that's true to an extent. But each of us is also writing for our own purposes, and given the network effects of the platform, it is genuinely the case that we all benefit the more people open substack.com on their browser rather than x.com.

But another reason I've been blogging about metascience less is that this work falls in a gray area in terms of incentives. It's not popular enough to grow my audience — but since it's not a pdf that's been transubstantiated by "peer review," I don't get to put it on my CV or include it on the exhaustively detailed five-year report from my department to external reviewers.<sup>1</sup>

Five years ago, I was convinced that the manifest absurdity of this process would cause it to collapse under the weight of its contradictions; that we might be able to adopt institutions better adapted to the media technology at present. Like generations of internet-pilled academic reformers before me, I was wrong. The for-profit pdf reigns supreme, except now they've figured out how to make us pay by the article.

So, it is with a heavy heart that I've been paring away the rhetorical flourishes and self-satisfied "jokes" from my metascience writing to try to get it pdf and DOI-compliant.

One such pdf has been promoted to the status of "working paper": Temporal Generalization, with Drew Dimmery. I think this is the best and most important paper I've ever written.

But the rest of the work has turned out to be slow going. It turns out that writing breezy, semi-coherent blog posts about what is going wrong with science is quite a bit easier than making a well-situated, self-contained argument for what science should be doing instead. As I get older, I am forced to accept that there is some wisdom in the institutions we have inherited. This is infuriating.

So, today I present a blog-tier account of what has been going wrong not with science but with *metascience*. The rest of this post is stitched together from angry asides in what I hope will be a publishable pdf, arguing that the current metascience movement is not living up to its potential — that we are seeing the ascendance and establishment of metascience *manqué*.

Moin Syed provides a useful introduction to the players involved, some of which are sponsored by Elsevier, others of which are not. You can guess where my sympathies lie.<sup>2</sup> The Metascience conference — or any other specific entity — is not a one-to-one match with metascience

manqué, which is instead a tendency to which I am opposed.

## **Metascience Manqué**

The past decade has seen significant attention paid to the replicability of published experimental results. There have been many statistical and institutional reforms proposed. Some have been made possible, made easy, made normative, made rewarding, and made required, in Nosek (2019)'s famous strategy for culture change in science.

Graduate training has shifted; standards of evidence in the seminar room and peer review are increasingly stringent; some journals have adopted registered reports and made pre-registration mandatory.

The birth of metascience, at least in social science, can be traced to the “Replication Crisis” which occasioned the empirical studies of replicability cited above. I argue that the consequent focus on “replication” has prevented the maturation of metascience.

This “metascience manqué”, having raised its head above the standard practice of methodology and the individual study, has too quickly settled into normal science at the most obvious meta level. It is, literally, a *science of science*, with an ontology which takes the individual study as its fundamental unit of analysis.<sup>3</sup> The phenomena of interest to normal scientists disappear in the study of the study.

Metascience methodology has progressively developed ways to study studies – towards the fundamentally metaphysical assertion that SCIENCE == REPLICATION. This isn't metascience, it's metapositivism: taking the exact same self-confident methodology and unreflective ontology up one level of abstraction.

As the movement enters its teenage years, it would be healthier if it could look beyond the conditions of its birth, to transcend the obsessive focus on replication and its lack. Following Flusser's framework, there are three fundamental metascience questions:

Ontological: What is science? (What are scientists doing?)

Methodological: How could science be? (What could scientists be doing?)

Normative: What should science be? (What should scientists be doing?)

One crucial impediment to this maturation is the complicated intersection between digitally-mediated “informal communication” (Latour and Woolgar, 1986) and the formal institutions constituting social scientific disciplines and subfields. Metascience is premised on the intersection of insights scientists gained from their practices and their accumulated results. The problems identified and the solutions proposed are essentially local rather than universal.

However, these local problems and solutions have spilled beyond their provenance. The commanding heights of *Science* and *Nature* have proved strategically powerful in overcoming the resistance of “local elites” opposed to reforms, but in order to occupy these heights, bold overstatements of generality are necessary.

More pernicious is the overconfidence (natural to expertise) which allows for implicit generalization. Each of us knows what Science is – we do Science every day! We necessarily invest our own practices with broader or even exclusive claims to validity. Each discipline and subfield used to have their own methodological conversations, but these conversations are collapsing towards a single online cacophony.

The use of similar terminology and statistical methods – particularly in the extent to which the computer terminal has become the sole instrument with which social scientists interact with both each other and with their data – has made fundamental differences in disciplinary aims and scope easier to mistake. In the language of Collins and Evans (2002), social scientists have assumed that their interactional expertise is in fact contributory expertise. By this I mean that scientists’ ability to hold methodological discussions across fields gives the (mistaken) impression that the fields are facing the same problems with the same solutions.

This interdisciplinary communication has clearly been progressive, and generally even constructive. There’s no need to re-invent the wheel of, say, best practices for implementing logistic regression in every subfield which makes use of this statistical technique. But there are risks inherent in cross-disciplinary communication, especially when communication becomes universal methodological proscription. Absent mutual understanding, cross-disciplinary communication is not transdisciplinary but simply undisciplined.

## **The Scope Conditions of Science Reform**

“Statistics asks many different kinds of questions, but we confuse our students because the methods often look the same.”  
Recht (2025)

“Science” is changing. Field, van Dongen and Tiokhin (2024) provide a recent overview of the timeline and unintended consequences of what they call the “Science Reform Movement.” They introduce the subject as follows: “The scientific community has entered a challenging era, as originally noted by Wagenmakers (2012).”

The discipline of computer science has experienced world-changing successes over the past 13 years; the rapid development and deployment of COVID-19 vaccines represents a huge victory for biologists; materials science continues to drive down the cost of solar panels. The “scientific community,” writ large, seems to be doing fine.

But social psychology has certainly been having a rough time. Wagenmakers (2012) describes the *annus horribilis* that involved the publication of the famous proof of extrasensory

perception (ESP) in the flagship Journal of Personality and Social Psychology and subsequent mockery. This was, indeed, an embarrassing example of a subfield in distress.

But how easily the example slips out of the subfield of social psychology to the entire “scientific community” in the retelling! I’m not blaming the authors; this example is far from unique. (It wouldn’t be a very useful example if it were unique.) Indeed, the authors go on to make clear the progression of the diagnosis of problems and proposed solutions from the subfield to the larger scientific community. This is what distinguishes the current reform movement as metascience (Peterson and Panofsky, 2023), in contrast to philosophy or sociology of science.

However, I am not alone in arguing that metascience has become sufficiently mature that it cannot afford to ignore these other ways of thinking about science. Intuition and empiricism has started us down this road; what we need more of now is theory. Consider Hug (2022)’s summary of the metascientific literature on the practices of peer review. Many different journals and subfields have begun wondering whether the peer review practices they had taken for granted are the right practices; this a promising first step.

But the consequent literature presents a variety of local descriptive or experimental research about peer review. This unprincipled exploration of the parameter space of peer review cannot be synthesized absent theory about what peer review is trying to do.

I argue that this is a general problem for metascience. Intuition and blind empiricism produce communities adopting potentially incoherent reforms, each making improvements on some dimension of scientific practice without theorizing or measuring their impacts on other dimensions. Without a synthetic theory of a destination, these reforms represent a more or less random walk through the high-dimensional parameter space of scientific practice.

Rubin (2023) defines an alternative strand of metascience – critical metascience – in opposition to what I’ve called normal metascience. This is a promising development, and Rubin (2023) provides many of the same critiques of normal metascience that I have outlined: that it tends to overplay replication, that it tends to homogenize science, and that the intuitions of social psychologists are over-represented. Rubin (2023)’s comprehensive review piece cites dozens of normal metascience papers engaging in these practices, which I will not replicate here — though see his excellent blog, [Critical Metascience](#).

## **Meta-Science and The Problem of Intuition**

Something has been introduced to the glorious edifice of science...and it eats away, silently but perceptibly, at the foundations. The scientists do not trust their own competence to fight this type of danger. The danger comes from outside. Therefore, the scientists ask for help, and it is the philosophers they call. But philosophers are scientifically irresponsible people... philosophers have the dangerous tendency to analyze concepts. The something that surreptitiously started to penetrate the edifice of science, is nonetheless, a philosophical problem and not a

properly scientific one, therefore philosophers become indispensable. Unless scientists start to do philosophy in order to save themselves. And that is what they try. The result may be a sign of authenticity, but it is generally, weak philosophy.

Vilèm Flusser

Philosophy of Science has historically had a top-down relationship with scientific practice. To caricature the situation slightly, Philosophy of Science has been in the business of figuring out how science works theoretically, without much interest in how it works in practice. Longino (1990) describes the confusion between descriptive and proscriptive accounts of science: is science what scientists actually do, or what they should be doing instead?

But we are several decades into the “historical turn” in the philosophy of science, which is now intermixed with the history and sociology of science. This is for the better. Metascience, however, emerges from a different intellectual tradition – more precisely, it has emerged from many distinct intellectual traditions.

Metascience is an effective mechanism for improving science because its practitioners have an intimate understanding of the problems of their disciplines and subfields. This knowledge is often tacit, involving implicit practices and institutions which outside observers might fail to appreciate as important. In contrast to the top-down, Jupiterian philosophy of science, metascience is a bottom-up initiative.

But metascience is, again, many bottom-up initiatives. It has to be; the local, tacit knowledge of the situation necessarily varies across context. The metascientific impulse should not result in a Science Reform Movement, singular, but in many local Science Reform Movements.

There is an structural reason that metascience has jumped so quickly from local problems to the commanding heights of reforming Science writ large. In status-driven, institutionally conservative academia, the top people in a given subfield have various interlocking mechanisms for advancing their vision of how science should be conducted.

It would be a rare full professor to publicly denounce the status quo in which they worked for twenty years; radical reform can only be expected to come from younger generations. But given the slanted field of battle, the more powerful and more coordinated (it’s easier to coordinate on the status quo than on any given alternative) establishment should win every time.

In the language of Schattschneider (1960)’s classic political science text, it is strategically expedient for the reformers to “expand the scope of the conflict” if they are unable to succeed at the local level. By drawing in public outrage and ridicule, and ultimately by publishing in “general interest” journals outside of their subfield. The reformers were successfully able to align themselves with the majority scientific opinion. The “local elite” dominates the subfield, but cannot maintain their minority position against all of Science.

But this strategy has come at the cost of overclaiming, at least in how the results of science reform have been interpreted and taken up. Semi-formal metascience communication on the internet in the form of blogs and micro-blogs has turned all of experimental social science into a “trading zone” (Galison, 1997). This has been immensely productive, but blurs disciplinary boundaries previously enforced by distinct journals. Terminological similarity has been at times infelicitous; at the highest level, the set of activities categorized as “experiments” is extremely heterogeneous in terms of ontological complexity and temporal scope.<sup>4</sup>

A simple example is the concept of the “treatment.” The language clearly invokes the medical context, and indeed the larger causal inference framework comes from the biostatistics tradition. The ideal-typical “treatment” is an application of some physical substance to a patient’s body.

Consider Acetylsalicylic Acid (C<sub>9</sub>H<sub>8</sub>O<sub>4</sub>), aspirin. The patient has a headache; they ingest the treatment, a stable and ontologically specific molecule. The mapping between the linguistic and physical world is tight. Many of the “treatments” deployed by psychologists, political scientists and economists differ from this ideal case in significant but unspecified ways. The idea of “the” “treatment,” we might say, is a disease afflicting the epistemology of experimental social science.

My intuitions about the severity of this disease comes, of course, from my own area of research: social media, and particularly the effects of feedback on social media platforms.

For my initial graduate research project, I conducted an experiment utilizing Twitter bots to address users engaging in racist harassment (Munger, 2017). Following the publication of my findings, several fellow graduate students expressed interest in adapting this methodology for their own research questions. I provided them with my code and insights from my experience.

However, they encountered an unexpected obstacle - their automated accounts were being suspended by Twitter’s detection systems almost immediately. Despite following my protocol precisely, their bots were consistently banned. The irony became clear: the only way to successfully implement the treatment would have been to conduct it during the earlier period when Twitter’s bot detection policies were less stringent. This inability to re-implement the treatment is clearly a distinct phenomenon from the traditional replication crisis, where researchers fail to reproduce previously documented effects. Rather, this case demonstrates how fundamental changes in the platform’s infrastructure made the original experimental design technically impossible.

This challenge highlights a broader issue in contemporary internet research: the rapidly evolving nature of privately controlled digital platforms poses significant obstacles to scientific replication. The technological landscape changes so rapidly that methodologies valid at one point may become entirely unfeasible in relatively short periods (Munger, 2019).

These kind of problems abound in the study of social media. Consider the large, 32-author, \$25 million collaboration between academics and research scientists at Meta during the 2020 US Presidential Election. The team conducted unprecedented descriptive research and a number of

experiments involving the manipulation of the Facebook and Instagram recommendation algorithms, resulting in multiple publications in *Science* as well as other top outlets. This is, as I have argued Munger (2024), the pinnacle of a prominent strand of quantitative social science of social media – and yet the experiments obviously cannot be replicated.

One response to the realization that “replication” cannot mean “perform the same procedure” is to rethink the role of replication in the scientific process – or at least, how the role of replication might vary across different scientific disciplines. “Perform the same procedure” is almost always a coherent concept in (some domains of) natural science and almost never a coherent concept in (some domains of) social science as it currently exists.

## **Against the pdf**

Metascience manqué is in keeping with the “bureaucratic” strand of science reform (Devezer and Penders, 2023; Penders, 2022). The problem this strand has identified, broadly, is that science is unruly and chaotic. From this chaos emerged questionable research practices that allowed for bad science to make it past peer review. The bureaucratic solution is to reduce the chaos by rationalizing and making transparent scientific practice. This means a proliferation of forms and checks both before and after peer review – forms and checks which can be easily verified by a third party.

This desire for purity and control, this faith is “works alone,” I have previously characterized as the “Second Reformation.” This bureaucratic, open-science impulse is a problem, and tends towards authoritarianism. Rather than the methodology of individual studies, where a team of scientists decides what methods to use and is forced to defend this choice during peer review, metascientific methodology takes the behavior of thousands of other scientists and identifies what they’ve been doing wrong.

I think that the bureaucratic impulse is downstream of the more basic issue with metascience manqué outlined in the introduction: the reification of “the study” as the fundamental unit of knowledge generation. Rather than taking the form of science for granted and forcibly contorting the messy reality towards our platonic ideal p-curve, a more mature metascience would think about how to better align the form of science with the technology we now have access to.

The most important constraint, the narrow pipe through which scientific knowledge is forced to squeeze, is the peer-reviewed pdf. Borrowing Duflo’s plumber metaphor, I believe that a crucial task for metascience is to design and install better scientific pipes.

For encoding information about the details of experimental procedure, videos are radically more effective than textual description. pdfs also flatten the information content that goes into them. In my own realm of media effects, the number of experimental subjects described in a pdf ranges from a few hundred to hundreds of millions.

Most importantly, for the procedure for exploring the parameter space of experiments, the requirement that each iteration be written up and the results appended to lengthy theory

sections replete with citations is absurd. Many practitioners realize this; they only bother publishing entire collections of what could, in principle, be published as individual studies.

More generally, as metascience supplants methodology and we begin to rigorously evaluate collections of studies rather than perfecting individual ones, we need to consider how best to format the knowledge generated by each of those studies. The single-team mega-studies are more efficient, but they will never represent the final word on anything. Social science is still very far from CERN, and given my normative commitments, I hope it remains that way. Social science needs to retain its collaborative nature, but for this to happen at the higher level of precision made possible by contemporary experimentation, the medium of scientific communication needs to allow for more interoperability.

This means making the parameter space of experimentation much more explicit. By developing and reporting “controlled ontology” of the relevant causal factors, and cutting out much of the superfluous text, scientific output in this domain can better inform future experimentation and make machine knowledge synthesis more efficient.

This is obviously not a trivial undertaking. So many institutions have agreed upon the pdf as the unit of scientific output, producing credibility for the actions described in the pdf and the scientists who produced it. The citations from the pdf and to the pdf both play important roles in the scientific conversation as it currently takes place – but in my experience, the vast majority of these citations (in both directions) are scientifically unnecessary.

The selection and development of a successor institution will be intensely political and messy; the science that results will be far less unified than the science we currently have. Or rather, than the science we appear to have. Just as the statistical techniques and the verbal descriptions paper over fundamental differences in goals and possibilities, so too does the insistence of the pdf as the only valid medium of scientific communication.

According to the anarchist theory of prefiguration, the way towards creating the desired society is to begin with your own organization, to both figure out issues in the implementation and serve as a demonstration that it can in fact work. In the context of metascience, I take this to mean that no pdf can overthrow the reign of pdfs; there will not be a peer-reviewed paper demonstrating that we should switch to a new system.

Instead, metascientific reforms must be institutional reforms. The Journal of Quantitative Description: Digital Media is my largest such contribution to date, but this kind of work is extremely difficult for non-tenured academics to pursue — the mainstream institutions of scientific “credit” are not designed to evaluate this kind of contribution.

An effective metascience must operate outside of the realm of the for-profit peer-reviewed pdf. The anarchist is not opposed to tactical opportunism, and should operate within existing institutions when expedient. But the effort to rein in the exciting, free-wheeling metascience movements we have observed over the past 15 years — to discipline them — will only ever produce a metascience manqué.



## References

- Camerer, Colin F, Anna Dreber, Eskil Forsell, Teck-Hua Ho, Jürgen Huber, Magnus Johannesson, Michael Kirchler, Johan Almenberg, Adam Altmejd, Taizan Chan et al. 2016. "Evaluating replicability of laboratory experiments in economics." *Science* 351(6280):1433–1436.
- Camerer, Colin F, Anna Dreber, Felix Holzmeister, Teck-Hua Ho, Jürgen Huber, Magnus Johannesson, Michael Kirchler, Gideon Nave, Brian A Nosek, Thomas Pfeiffer et al. 2018. "Evaluating the replicability of social science experiments in *Nature and Science* between 2010 and 2015." *Nature human behaviour* 2(9):637–644.
- Collaboration, Open Science. 2015. "Estimating the reproducibility of psychological science." *Science* 349(6251):aac4716.
- Collins, Harry M and Robert Evans. 2002. "The third wave of science studies: Studies of expertise and experience." *Social studies of science* 32(2):235–296.
- Devezer, Berna and Bart Penders. 2023. "Scientific reform, citation politics and the bureaucracy of oblivion." *Quantitative Science Studies* 4(4):857–859.
- Field, Sarahanne, Noah van Dongen and Leo Tiokhin. 2024. "Reflections on the Unintended Consequences of the Science Reform Movement." *Journal of Trial & Error* 4(1).
- Fonseca, Fred. 2022. "Data objects for knowing." *AI & SOCIETY* 37(1):195–204.
- Galison, Peter. 1997. *Image and logic: A material culture of microphysics*. University of Chicago Press.
- Hug, Sven E. 2022. "Towards theorizing peer review." *Quantitative Science Studies* 3(3):815–831.
- Latour, Bruno and Steve Woolgar. 1986. "Laboratory Life: The Construction of Scientific Facts.".
- Leonelli, Sabina. 2018. *Rethinking reproducibility as a criterion for research quality*. In *Including a symposium on Mary Morgan: curiosity, imagination, and surprise*. Emerald Publishing Limited pp. 129–146.
- Longino, Helen E. 1990. *Science as Social Knowledge: Values and Objectivity in Scientific Inquiry*. Princeton University Press.
- Munger, Kevin. 2017. "Tweetment effects on the tweeted: Experimentally reducing racist harassment." *Political Behavior* 39(3):629–649.
- Munger, Kevin. 2019. "The limited value of non-replicable field experiments in contexts with low temporal validity." *Social Media+ Society* 5(3):2056305119859294.
- Munger, Kevin. 2024. "What Did We Learn About Political Communication from the Meta2020 Partnership?" *Political Communication* pp. 1–7.
- Nosek, Brian. 2019. "Strategy for culture change." *Center for open science* 11.
- Penders, Bart. 2022. "Process and bureaucracy: Scientific reform as civilisation." *Bulletin of Science, Technology & Society* 42(4):107–116.
- Peterson, David and Aaron Panofsky. 2023. "Metascience as a scientific social movement." *Minerva* pp. 1–28.
- Recht, Benjamin. 2025. "A Bureaucratic Theory of Statistics." *Observational Studies* 11(1):77–84.
- Rubin, Mark. 2023. "Questionable Metascience Practices." *Journal of Trial Error* .
- Schattschneider, EE. 1960. "The Semisovereign People (New York: Holt, Rinehart and Winston, 1960)." *SchattschneiderThe Semi-Sovereign People1960* .

Simons, Daniel J. 2014. "The value of direct replication." *Perspectives on psychological science* 9(1):76–80.

Wagenmakers, Eric-Jan. 2012. "A year of horrors." *De Psychonoom* 27:12–13.

---

1. I'll write something later about my experiences in European academia — but I can confirm that thus far the bureaucratic oversight is just as asinine and annoying as I had been led to expect.

2. I should say that I, too, was rejected from 2025's Metascience conference, despite have presented at the 2023 version and being the only political scientist (of whom I am aware) to have produced [a peer-reviewed pdf explicitly about metascience](#).

3. I am inspired here by Fonseca (2022)'s discussion of Data Science as literally a "science of data," something which takes the data as ontologically real rather than representative of or summarizing reality.

4. Leonelli (2018) makes the problem of linguistic cross-disciplinary confusion explicit in the case of reproducibility: "Remarkably, despite the variable interpretations of the term that are immediately visible through such basic considerations, the framing of reproducibility used in top science journals today remains narrowly linked to a particular understanding of what good research should look like. Most typically, reproducibility is associated with experimental research methods that yield numerical outcomes."