Scientific Reproducibility and its Problems
Heterodox Research Society: Alexander Vawter

There exists a critical and long recognized problem in science, around the question of experimental reproducibility. This is a topic of much discussion, especially in more recent discourse as many scientific institutions are under scrutiny, and science is forced to face the reality of its shortcomings while a general sense of stagnation feels more publicly pronounceable. Even the incumbents themselves have come to recognize this problem of experimental reproducibility (a lack thereof), with a fashionable Nature article announcing that 75% of experimental results cannot be replicated by other researchers, and an even more staggering 50% of researchers can't even replicate their own results.

This problem is particularly interesting given its bipartisan recognition by both institutional insiders and outsiders. Nonetheless, allow me to be clear: experimental irreproducibility is the unquestionable result of methodological hubris in the sciences. We have pigeon-holed ourselves into particular methodological approaches, and basic research has been hostile to new ways of conducting business. As such, now as many institutionally funded researchers squeal out of necessity to keep their labs alive, I find this to be a particularly opportune moment to transform the way in which science is done, at both a practical level, and even at a deeper epistemic level.

The Canon View

First let us lay the foundation for public perception of reproducibility through the lens of how the institutions understand it. Inside the scientific establishment, the canonical interpretation of the reproducibility crisis is unsurprisingly narrow: it is treated as a technical quality-control problem, not a philosophical one. Institutions tend to frame irreproducibility as a failure of execution standards rather than a failure of methodological imagination or epistemic assumptions.

This view shows up clearly in flagship coverage like the well-known Nature survey on reproducibility. The article doesn't frame the crisis as a collapse of scientific methodology; instead, it frames the problem as a workflow malfunction: poor reporting practices, inadequate statistical training, pressure to publish, selective reporting, and irreproducible computational pipelines.

According to the Nature survey (2016, N ≈ 1,500 researchers):

70% of scientists reported that they had failed to reproduce another scientist's experiment

50% admitted they couldn't reproduce their own results

60% said the literature “sometimes” or “often” reports questionable statistical significance

Yet more than 70% still expressed confidence in the underlying scientific enterprise. In other words: the establishment is more interested in symptoms rather than causes.

In reality, the reproducibility crisis reveals something far deeper than sloppy documentation or inadequate training. The problem is methodological ossification across multiple dimensions, each reinforcing the others in a system that has become fundamentally hostile to its own stated values. Consider first the ex post facto narrative that has become the standard form of scientific communication. The modern scientific paper takes a messy, nonlinear, intuitive experiment full of blind alleys and retrofits it into a clean narrative that hides the true messiness of science. Science is not the merely perfect, logical task we’ve been taught it to be, or led to believe. It is messy and chaotic.

Understanding this reality, dealing with it, and creating frameworks that allow us to play within this swampy pit are how we actually advance science and communicate it. However, we have reduced our collective narrative of what science is into an algorithm, or a reduction to mere facts. Of course, in that context researchers would be terrified to share the reality of their experiments. To be criticized in the literature would be a tragedy! Methods sections are skeletal by design, with critical steps abstracted away and data pathways obscured, not merely because scientists are lazy documentarians but because the entire culture demands a performance of certainty that bears no resemblance to the actual practice of research.

This connects directly to science’s allergy to new practices. As with many human endeavors that bear fruit in one time period or another, there emerges an attempt by those in or around the caravan to seize control. Science is a good example of a wonderful enterprise often being co-opted for political gain by some of its incumbents. As such, in order for a regime to stay in power, it must have a unifying dogma. In this case for science it has been “the scientific method.” The particular use of the word “the” scientific method is rather damning as the trademarking of a singular method definitionally kills attempts at new scientific methodologies. Open source has rarely been adopted in any meaningful capacity despite its obvious applications. Sharing of results, while understandable in some cases, maintains a rather low standard for granularity. Researchers operate in constant fear of data leaks or intellectual property concerns. The general culture has become utterly allergic to methodological experimentation, and through its hubris ultimately believes that matters can be resolved by a stronger dose of its own poison.

And then there is hyper-specialization. We live in an era whereby scholars narrow their interests until only a mere handful of people in the world can understand them. Fields develop increasingly arcane nomenclature, which reduces interdisciplinary communication and increases isolation. One seemingly inevitable consequence is the creation of what José Ortega y Gasset called “learned ignoramus,” experts in a very particular area who are unable to see beyond. In some sense establishment science actually doesn’t even want others to replicate results. Indeed, it is a powerful thing to be the lone guard of what many in the culture believe to be the utter truth, so results must solely be of their own making. At a very practical level, even incredibly talented researchers cannot come to understand the sheer jargon of the hyper-specializations of some pursuits. The research community through hyper-specialization has learned from the likes of Heisenberg, creating their own uncertainty principle. As one mindlessly digs deeper and deeper

into a problem, the more impossible it is for anyone to question you. Thus, you will dig and, as we see today, many have worn their shovels to death.

And how do institutions respond to all this? There must be more mandatory reporting guidelines, and more bureaucracy. The solution (of course!) must always be to add another layer of procedural requirements and methodological tightening, as if our brilliant researchers simply aren't filling out the forms correctly. Standardization in theory seems sensible, but rigidity is precisely what prevents the sort of methodological flexibility and pluralism that science demands. What's needed is freedom to experiment with how science itself is conducted: freedom to pull from seemingly unrelated disciplines (the arts and the humanities, a terrifying thought, I'm sure), to bring theorists into proximity with engineers, and to experiment with methodologies beyond the approved canon. Instead, the tightening continues lest we stop it.

Ultimately, The Science™ fundamentally misunderstands its own enterprise. One will never be able to replicate results when the knowledge required to interpret them is locked inside hyper-specialized intellectual silos, when the culture demands performances of certainty over honest documentation, or when any deviation from established procedure is treated as either heresy or a threat to their house of cards. The publication of today is nothing more than retroactive performance, and a rather dull one at that.

The insiders feel that the reproducibility crisis is justification for even more methodological rigidity. We must not go down this terrible spiral. Rigidity itself is the problem and we must provide a clear alternative both in practice and in philosophy.

Liberatio in Modum

Even if we attempt to charitably accept the institutional account of irreproducibility, it still exposes something deeper about how science actually functions. The real problem is not merely that scientists fail to document their work but rather modern science has built a culture and set of incentives that actively resist tools and practices that would expose the real texture of laboratory life.

The institutional response to irreproducibility misunderstands both the nature of discovery and the nature of reproducibility. Methodological rigidity makes results harder to reproduce because it eliminates alternative paths to the same destination. When you mandate a single protocol, reproduction becomes an exercise in perfectly replicating steps that may have been poorly documented, poorly understood, or context-dependent in ways the original researchers didn't recognize. In a very practical sense, methodological pluralism enables reproduction through multiple roads as one obtains the ability to adapt methods rather than merely copy them dogmatically. In reality, there are many existing low-hanging fruit ways in which the scientific enterprise can methodologically expand, and it is my belief that such methods simultaneously offer great promise on the question of reproducibility. Consider what methodological flexibility has enabled historically. In 1669, Hennig Brand boiled down over 50 buckets of human urine hoping to find gold, based on alchemical theories about "life force" in bodily fluids. As a result, he accidentally isolated elemental phosphorus, the first modern element discovered. Of course this example addresses the question of novel discovery more so than reproducibility - however I

think it gets at the broader notion that our self-stultification is a more recent phenomenon. Nonetheless, let us examine some more relevant examples:

During the Human Genome Project, the major public sequencing centers agreed to the Bermuda Principles in 1996. There was a simple rule: any newly sequenced DNA had to be placed in a public database within 24 hours. Labs at places like the Sanger Centre, Washington University, and MIT uploaded their raw sequences daily into GenBank. This meant the genome wasn't held privately by any one group; every lab in the consortium could immediately see the others' data. This became especially important when Celera Genomics entered the race in 1998 with the intention of producing a faster, proprietary version of the human genome. Celera planned to control access through subscriptions and licensing. But because the public consortium was releasing sequence data every day, Celera was unable to claim exclusive ownership over large portions of the genome. Anything Celera assembled later was already in the public record, which legally prevented patent claims. The HGP's open approach also allowed the centers to cross-check each other's work. When two centers happened to sequence overlapping regions, they sometimes found misalignments or base-calling mistakes in each other's data, and they corrected these by comparing their uploads. The rapid, open exchange hardened the final assembly by letting people see what others were doing in real time.

Venturing outside the realm of atoms and into the world of bits, there is of course the familiar example of the Linux kernel, which began as a small personal project by Linus Torvalds in 1991, but became stable because every change to the kernel was made in public. Contributors submitted patches to mailing lists where anyone could review them, and the full commit history eventually moved to Git, which Torvalds created for that exact purpose. The development workflow exposed every modification and this openness produced a kind of resilience that no closed system matched. When a kernel bug appeared, thousands of developers around the world could inspect the exact chain of changes that led to it. Vulnerabilities like the 2003 "do_fork" issue or the later Dirty COW flaw were fixed quickly because the entire history of the affected parts of the kernel was visible. Companies that depended on Linux, such as IBM and Red Hat, relied on this transparency because it meant no single team controlled the codebase's reliability. The system worked precisely because the full lineage of every decision was out in the open, which allowed both scrutiny and rapid correction without relying on any single authority. I do not believe there should be open-source approaches merely for the sake of being open-source; in fact the worst open-source projects are those that unnecessarily trudge out open-source materials for the sake of being publicly perceived as open-source. Great science can happen behind closed doors just as it can happen in full public view. Bell Labs, among many other examples, proved this conclusively.

Bell Labs succeeded through cross-disciplinary integration, creating a truly methodologically plural ecosystem fueled by an engine of cash and reflexive economic incentives. Chemists, physicists, metallurgists, engineers, theoreticians and experimentalists were assembled in close physical proximity. This matters for reproducibility because when theorists and experimentalists work in proximity, they develop the cross-disciplinary fluency required to actually understand each other's work. The theorist learns enough about experimental constraints to recognize when a result seems off and the experimentalist learns enough theory to know which parameters actually matter. More fundamentally, there was a cultural openness of an internal flavor that

fueled all of this. There must be a willingness amongst participants in an organization to accept their shortcomings, and be willing to open up the true breadth of their work to others across domains (if you wish to distinguish domains). Presently, institutional incentives, particularly in academia, do precisely the opposite. Domains are hyper-demarcated and specialized. Theorists and engineers occupy different buildings, publish in different journals, and rarely collaborate meaningfully. There exists a profound physical and cultural isolation of scientists and this is a leading contributor to the development of intellectual silos within said institutions. Even if you have a perfectly documented experiment, there are fewer and fewer people with the cross-disciplinary knowledge required to reproduce them, let alone to find alternative approaches when the original method fails.

The aforementioned approaches are merely examples, not prescriptions, of methodological experimentation that have paid dividends elsewhere. However, looking forward, there are countless methodological frontiers science has barely begun to explore. Crowdsourced global experiments that leverage distributed participation. AI augmentation that reveals patterns invisible to human perception alone. Ascending towards the fringe, perhaps there could even be psychedelic-assisted research that accesses cognitive states traditional methodology ignores. There could be Aesthetic modeling that treats scientific problems as design challenges - and this is something theoretical physicists have done knowingly or not many times - indeed Wolfgang Pauli invented the neutrino as a desperate remedy to resolve the beautiful symmetry of energy conservation - despite there being no existing detection mechanism for such a fickle particle.

Nonetheless, the objections are predictable: worries of intellectual property, loss of competitive edge against other labs, or an unwillingness to capitulate one's faith in "the scientific method". But these objections reveal the problem. Science has narrowed its own orchard by refusing to adopt new ways of conducting experiments. The methodological fallout has led to many problems, reproducibility being one of them. Being open to new models of research, or new ways of conducting day-to-day business may demand characteristics like transparency or ground-truth accountability, or they may not. But such practical experiments will be made by those who construct our wonderful new frontier, and those who refuse to step forth will be left behind, as the many tales of history will show.

Epistemic Freedom

Putting earthly matters aside, there is a substrate which undergirds any practical discussion around the sciences, or how they are conducted, perceived, and thus scored. This fundamental substrate is, of course, epistemology, and I believe that the earthly problems which certainly exist, such as the reproducibility crisis, are birthed out of a broken epistemology.

Modern science has committed itself to a rigid vision of method that is singular, standardized, and homogenous, and it assumes that adherence to a prescribed methodological script guarantees validity. When experiments fail to reproduce, the institutional instinct is to double down on this rigidity by implementing more checklists, more standardization, and even tighter protocols. But this response mistakes the symptom for the cause. What establishment science desires is not merely the reproducibility of results but the reproducibility of both results and process. It seeks to monopolize methodology itself, to establish "the scientific method" as the singular algorithm to truth, the machine that can replace God. By attempting to control and monopolize the process

by which one experiments, by insisting on methodological uniformity, science has made results irreproducible. The very rigidity meant to guarantee validity has produced a crisis of validity. When you force all inquiry through a single methodological channel, you eliminate the natural, opportunistic approaches that would actually allow researchers to reproduce findings. Methodological pluralism is not merely a practical convenience but an epistemological reality. When researchers can be opportunists, pursuing naturalistic and intuitive approaches rather than bureaucratic procedures, they develop genuine comprehension of phenomena. And genuine comprehension enables reproduction. The difficulty of reproducing scientific results is that nobody actually understands the process taken to achieve such results beyond some abstract procedural description, a description divorced from the more messy reality. And if one does not understand the real path taken to results, one will never understand the results.

Historically, scientific inquiry has never been, and cannot be, captured by a single method. Both culturally it has been viewed as, and in practice has been, an inherently pluralistic, improvisational, and embodied craft of practitioners, despite its more recent conceptual mutations. It is Einstein who once said that scientists must be "unscrupulous opportunists." Indeed, Kekulé dreamt of benzene rings. Pauli invented the neutrino to preserve symmetry. Fleming noticed mold on a petri dish. These were acts of a sort of opportunism and methodological flexibility that has since been lost, and ultimately this epistemological narrowing we see today is not simply incidental to the reproducibility crisis. It is the cause.

We may one day understand dark matter and dark energy, or perhaps we will learn of ways to travel the cosmos outside the paradigm of chemical rocketry, and such results may be reproducible so that we may all share in the wealth of the discoveries. However, I believe such a future will demand a plural enterprise that is not limited to reductive methods and merely logical steps. Perhaps we will need to draw from mystical areas of inquiry which have often been ridiculed. We will look to the methods of the Sumerians and the Ancient Egyptians, or perhaps some entirely new fusion of domains altogether.

As many continue to critically examine science, we must look seriously in the direction of methodology, as it follows from a broader philosophy of science that is sufficiently realistic in examining the messy, pluralistic nature of inquiry. Many have taken the fairly surface-level path of addressing institutional problems and the necessity of decentralization. I agree with claims addressing the need for the modern lab to be more self-sufficient. However, if you simply port one who follows the prevailing dogma onto a new frontier, they will not simply change their ways. There must be a broader epistemic shift and cultural reframing of science. We can reproduce results more effectively by practically and culturally allowing researchers to pursue new methodological paths and arrive at results through more pure, naturalistic, and ultimately opportunistic means. We must open the floodgates of methodological flexibility.

Alas, in order to reform science *we must first remember it*.