

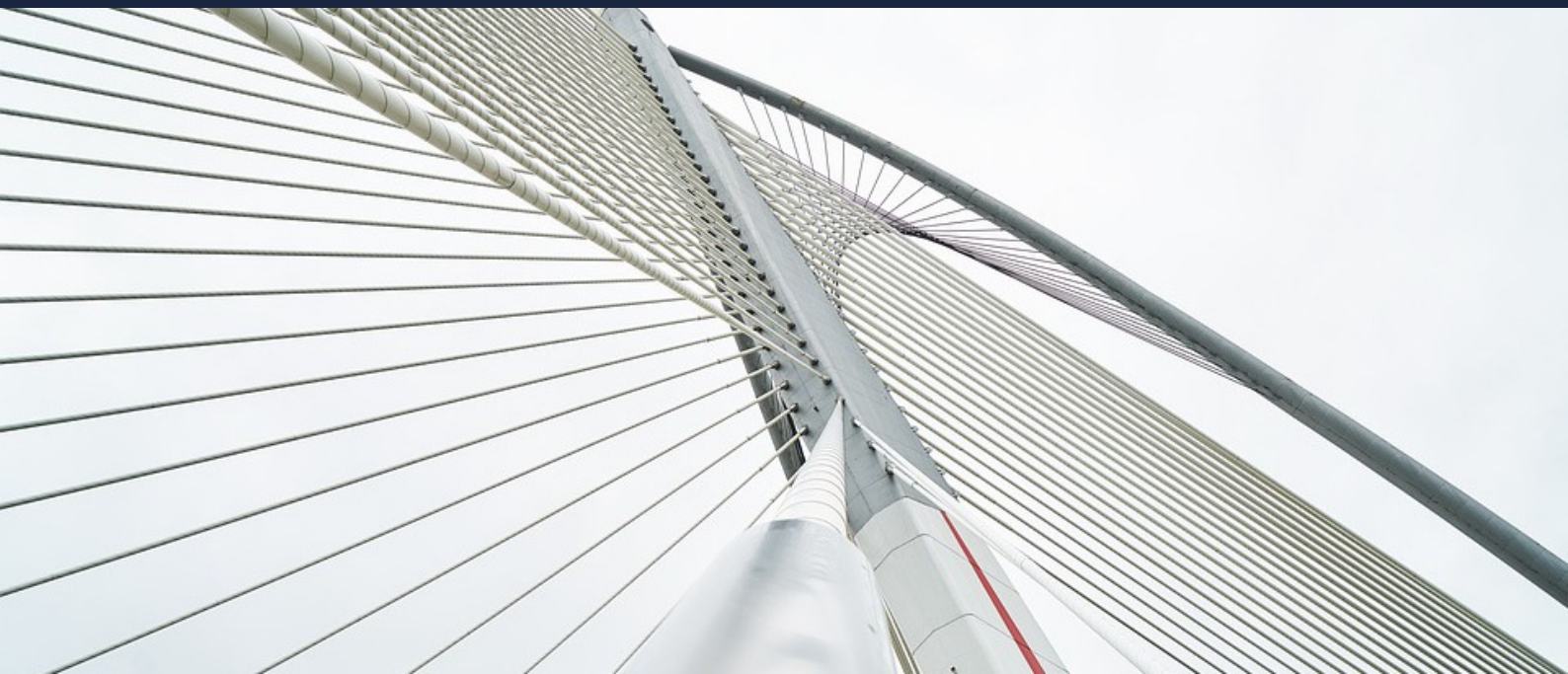
BEYOND CONVICTION

What I May Be Wrong About

Falsifiability Turned Inward on Twenty-Six Years of Writing About Enterprise Delivery

Giovanni Leonardi

September 2026



Contents

1. Executive Summary	3
2. The Gap That Closes the Library	3
3. I May Be Wrong About the Pace	4
4. I May Be Wrong About the Constants	5
5. I May Be Wrong About Formation	6
6. I May Be Wrong About Europe	7
7. The Practitioner's Own Instrument	8
8. Checkable, Not Right	9

What I May Be Wrong About

Leadership · No. 128
September 2026

"Doubt that cannot be tested is indistinguishable from performance."

1. Executive Summary

A library of two hundred and eighty papers built on conviction owes its reader one page of structured doubt. This essay names five specific uncertainties that a fair reader should hold against the body of work: that the pace of agent-era adoption may not follow the curve the operational papers assume; that the five constants treated as organisational physics may bend under machine-mediated context; that designed apprenticeship may not save a craft whose accidental formation is dying; that the European regulated-excellence thesis may be partly hope dressed as analysis; and that the author's own experience systematically over-weights the failures lived and under-weights the ones merely read about. For each, the essay states the evidence that would settle it and the date by which silence starts to mean something. The close is the library's real epistemology: these papers were never meant to be right — they were meant to be checkable.

2. The Gap That Closes the Library

The last paper in this library was supposed to be yesterday's. The reversed-reading essay closed the sequence with a structural conceit — reading the library backwards to see what changed — and it felt like an ending. It was not. A body of work built on conviction, on the repeated claim that experience reveals patterns worth stating, has one conspicuous gap: it has never said where the convictions might be wrong.

That gap is not modesty deferred. It is an obligation ducked. Two hundred and eighty papers of argued position carry an implicit promise — that the author has thought hard enough to know where the argument is weakest. This page delivers on that promise, late.

What follows are five named doubts. They are not hedges, not the pro-forma "limitations" section of an academic paper, and not the false modesty of a practitioner who wants credit for humility without the cost of specificity. They are the structural places where I believe a fair-minded reader should hold the work to account — where the evidence assembled over twenty-six years might break under weight not yet applied to it.

"Self-doubt that stays atmospheric, that never names its object or its resolution, is not doubt at all. It is decoration."

For each doubt, I state what the library claims, why I may be wrong, what evidence would settle the question, and the date by which silence becomes an answer. This is falsifiability turned inward. If the exercise sounds clinical, that is the point. Doubt that cannot be tested is indistinguishable from performance.

3. I May Be Wrong About the Pace

The library's operational doctrine — its papers on agent-era governance, the restructuring of delivery around machine capability, the new operating models — rests on an assumption about adoption pace. Not that autonomous agents will arrive: the technology is here in the autumn of 2026, and debating its existence would be absurd. The assumption is about the rate at which enterprises will reorganise around it — that within the next three to five years, the majority of large organisations will have moved from experimenting with agent capabilities to embedding them in the operational fabric of how work gets done.

The papers that built the agent-era operating doctrine did so from observable signals: the speed of capability improvement, the pattern of enterprise pilot programmes, the shifting language in board-level strategy documents. Those signals were real. But signals of early adoption and signals of structural reorganisation are different phenomena, and confusing the two is a classic practitioner error — one I have warned against in other contexts and may have committed here.

The assumption could break in three ways. Regulation could slow it: the EU AI Act's operational provisions are being interpreted by national authorities now, and an aggressive reading of the high-risk classification system could make agent deployment in regulated industries so compliance-heavy that the economics invert. A serious incident could stall it: a single high-profile failure of autonomous decision-making in a safety-critical context, the kind that reaches front pages and triggers parliamentary inquiries, could reset the political clock by years. Or the economics may simply not arrive: the cost of deploying, monitoring, and governing agent systems at enterprise scale may remain high enough, for long enough, that the returns never clear the hurdle rate for organisations still carrying legacy operating models and legacy debt.

If the pace assumption breaks, the damage to the library runs in both directions. The operational papers — those laying out governance for agent-era delivery, the capability models, the redesigned programme frameworks — will have over-built for a future that came slower than expected. And the sceptical papers — those arguing that the fundamentals do not change even as the tools do — will have under-claimed: if adoption truly stalls, the old methods will not merely persist as a foundation beneath innovation. They will remain the entire building, and the papers that framed them as bedrock will have undersold their continuing dominance.

The evidence that would settle it: enterprise adoption surveys showing agent-system integration in operational – not experimental – processes across at least forty per cent of large organisations in regulated industries. The date by which silence means something: the end of 2028. Sub-twenty per cent operational deployment by then would confirm the pace assumption was premature and the library's operational doctrine ahead of its evidence.

4. I May Be Wrong About the Constants

Across twenty-six years of writing about delivery, governance, and organisational change, five propositions have functioned as this library's load-bearing walls – claims treated not as arguments to be won but as facts of the terrain, reliable enough to build upon without re-proving in each paper.

Stated plainly, they are: that governance is a decision-making system, not a reporting mechanism; that culture absorbs strategy unless the strategy addresses culture explicitly; that the executive sponsor's primary role is protection, not direction; that the practitioner's craft is irreducible – it cannot be fully codified, only developed through situated experience; and that complexity demands simplicity of method, because elaborate process in volatile conditions produces compliance without competence.

These five propositions held for twenty-six years. They held through waterfall and agile, outsourcing waves and insourcing corrections, digital transformation and cloud migration and data-first and platform-first and whatever label came next. They held because they described something about human organisations – how people make decisions under uncertainty, how authority flows through social structures, how skill is transmitted between individuals, how groups respond to complexity – that did not depend on the particular technology or methodology of the moment.

The question I have not adequately faced is whether they hold when the organisation is no longer primarily human-run.

An enterprise in which machines hold most of the operational context – where an agent system has processed every document, every meeting transcript, every metric, and can surface the relevant precedent faster than any human governance board – may bend rules I have treated as physics. If a machine can genuinely synthesise the context that a human governance board currently synthesises, badly and by committee, then perhaps governance should become a reporting mechanism: one where the machine reports its decisions and the human reviews exceptions. If culture is absorbed by the speed of machine-mediated adaptation, the cultural-inertia thesis weakens. If craft knowledge can be captured and replicated at sufficient fidelity by machine systems that learn from practitioner behaviour, the irreducibility claim fails – and with it, the apprenticeship thesis that depends upon it.

I do not believe this has happened yet. But the honest position is that the constants were tested in human-run organisations, and the next decade will test them in something else. I have been confident enough to build an entire analytical framework on their stability. I should be honest enough to name the conditions under which they would fall.

The evidence that would settle it: a sustained pattern — not a single experiment — in which organisations that delegate governance synthesis to machine systems outperform those retaining human-committee governance on decision speed, decision quality, and programme-outcome realisation, without increased failure rates in the domains that matter. One constant falling would not invalidate the others, but it would break the frame that treats them as a set. The date by which silence means something: the end of 2030. The constants accumulated their evidence over twenty-six years; the counter-evidence needs at least four years of operational data in genuinely machine-mediated environments to carry weight. Shorter than that, and we are measuring novelty effects, not structural change.

5. I May Be Wrong About Formation

The library makes a sustained argument — developed across its papers on craft, capability, and the practitioner's development — that the accidental apprenticeship which historically formed delivery practitioners is dying, and that a designed apprenticeship must replace it. Senior practitioners once formed junior ones through proximity: the slow transfer of judgement that happens when someone watches how a programme director handles a failing supplier, reads a room, or decides what not to escalate. That proximity is vanishing. Remote and hybrid work, flatter structures, faster career movement, and the sheer operational pressure that pushes seniors into execution rather than formation all conspire against it. The designed apprenticeship is the proposed answer: deliberate, structured exposure to the situations that develop judgement, because the accidental version can no longer be relied upon.

The doubt is simple and serious: some crafts died when their apprenticeships did.

The history of the skilled trades carries uncomfortable precedents. Certain forms of stonemasonry, of furniture-making, of glasswork — these did not merely change when their apprenticeship systems collapsed. They disappeared as living disciplines and survived only as heritage recreations. The tacit knowledge — the hand-feel, the situational recognition, the things that could be demonstrated but never adequately written down — proved genuinely untransferable through any mechanism other than the one that was lost. The designed replacement, however well-intentioned, was never as good as the accidental original, because the accidental original worked precisely because it was not designed. The learning happened in the gaps, in the unplanned moments, in the things the master did not know they were teaching.

Delivery management may be one of these crafts. The judgement that distinguishes a senior practitioner — the ability to read the political undercurrents of a steering committee, to sense when a risk register is being used as a shield rather than a diagnostic tool, to recognise the precise moment a programme has silently shifted from recovery to managed decline — may be knowledge that cannot survive the transition from accidental to designed transmission. If so, the library's formation papers are not merely optimistic. They describe a path that leads somewhere else entirely: to a different, thinner form of competence that resembles the old one from the outside but lacks its internal weight.

I want to be wrong about this. The designed-apprenticeship argument is among the most carefully constructed in the library, and it carries real hope — that the craft can be

preserved and extended through deliberate action rather than accident. But wanting to be wrong is not evidence of being right, and the historical precedents are too pointed to set aside.

The evidence that would settle it: a cohort of practitioners formed primarily through designed apprenticeship who, when placed in genuinely ambiguous, politically complex delivery situations, demonstrate the same quality of judgement — measured not by process compliance but by outcome and by the assessment of experienced peers — as practitioners formed through the old accidental proximity. The test must involve situations of real difficulty, not routine delivery, because the doubt concerns specifically the upper registers of professional judgement. The date by which silence means something: the end of 2031. Formation is slow work. The first serious designed-apprenticeship cohorts are only now being structured in a handful of organisations, and it will take five years before they face the situations that reveal what they actually absorbed.

6. I May Be Wrong About Europe

Several papers in this library — particularly those on regulation, on governance in complex environments, and on strategic positioning — carry an argument about European competitive advantage that has crystallised, over the years, into what I think of as the regulated-excellence thesis. The claim is that Europe's regulatory environment, characterised by its critics as a barrier to innovation and speed, actually creates the conditions for a superior form of organisational competence. Companies that learn to operate within serious regulatory constraints develop governance muscles, risk discipline, and stakeholder accountability that their less-regulated competitors lack — and these capabilities become genuine competitive advantages as the global operating environment grows more complex, more scrutinised, and more demanding of institutional trust.

The thesis rests on patterns I have observed: organisations that treated compliance as a floor rather than a ceiling, that used regulatory requirements as the skeleton for genuine operational discipline rather than papering over dysfunction with documentation. Some of those observations are real. The best-governed European programmes I have encountered were genuinely better than their equivalents in less-regulated environments — not despite the regulation, but because of what the regulation forced them to build.

But honesty requires the counter-case. The regulated-excellence thesis may be partly hope wearing the clothes of analysis. I am a European practitioner. My career has been shaped by this regulatory environment, and my professional instincts formed within it. The argument that this environment produces something better — not merely different, not merely compliant, but actively superior — may be the kind of claim practitioners make about the conditions they know best, because the alternative is admitting that the constraints they navigated for a career were merely constraints, not catalysts.

The evidence is genuinely mixed. For every organisation I have seen use regulation as a scaffold for excellence, I can point — in composite, not by name — to three that achieved compliance without excellence: that built the documentation without building the capability, that passed the audits without developing the judgement. If the ratio is

three-to-one against, the thesis describes an aspiration, not a pattern.

The evidence that would settle it: a rigorous comparison of organisational outcomes — not compliance rates, but actual delivery performance, innovation quality, and long-term resilience — between heavily and lightly regulated environments, controlling for sector and scale. The comparison must be against the best performers in both environments, because the thesis claims that regulation enables the best to be better, not that it lifts the average. The date by which silence means something: this is the most immediately testable doubt. The EU AI Act's operational provisions are taking effect now — the first serious, cross-sector regulatory framework applied to a technology still in rapid development. If European organisations subject to the Act's full provisions demonstrably outperform on delivery quality and risk management by the end of 2028, the thesis holds. If they merely comply, it was hope.

7. The Practitioner's Own Instrument

The four doubts above are structural: they concern specific claims the library makes about the world. This fifth is different. It concerns the instrument producing the claims — the practitioner's own experience, and the systematic distortions that experience introduces.

Every practitioner over-fits to their own history. The programmes I ran, the failures I watched unfold from the inside, the recoveries I participated in — these are not a representative sample of organisational reality. They are a single career's worth of vivid, emotionally weighted data, and the vividness is the problem. The programme that collapsed because the sponsor refused to acknowledge the schedule was fiction. The transformation that succeeded because one middle manager quietly rebuilt the governance framework while the steering committee argued about branding. The portfolio review that exposed eighteen months of undisclosed dependency risk across three supposedly independent initiatives. These are not examples chosen from a balanced evidence base. They are scars, and scars distort perception.

The library leans toward the failures I lived through and under-weights the failures I merely read about. The failure modes I witnessed from the inside — sponsor absence, governance theatre, cultural inertia dressed as process compliance — recur across paper after paper because they are the patterns my pattern-recognition is tuned to detect. The failure modes I encountered only at second hand — technical architecture collapse, vendor lock-in spiralling into operational dependency, data-migration failures cascading into regulatory exposure — appear less frequently, not because they are less important but because they are less mine.

This is not a confession of incompetence. Every experienced practitioner carries this distortion; it is the shadow cast by the very experience that gives their observations value. The question is whether I have acknowledged it sufficiently — whether the library's shape, its recurring emphases, its persistent themes, and its silences reflect a genuine reading of organisational reality or a portrait of one career's particular education.

I am not sure, and that uncertainty is the most honest thing I can say about twenty-six years of writing.

The evidence that would settle it: practitioners from genuinely different backgrounds — different industries, geographies, career paths — reading the library and identifying specific, named blind spots that correlate with experiences I did not have. Not disagreements of emphasis, which are inevitable and uninteresting, but structural absences: entire failure modes, entire patterns of organisational behaviour, that the library does not see because its author was never in the room when they happened. The date by which silence means something: this doubt carries no expiry. The over-fitting distortion is permanent, because the experience that causes it is permanent. The best I can offer is the invitation itself — and the observation that the reader who finds the blind spot has not found a flaw in the library. They have found the next page it needs.

8. Checkable, Not Right

Two hundred and eighty papers, and the question that hangs over all of them is deceptively simple: why should anyone believe this?

Not authority. Not credentials, not years of experience, not the accumulated weight of having-been-there. Authority is asserted, and asserted authority is the weakest possible foundation for the kind of claims this library makes. Not certainty, either. A library that presents itself as certain — that never names where it might break, that treats its own convictions as settled science — is not a body of knowledge. It is a brochure.

The answer is checkability. These papers were never meant to be right. They were meant to be argued from stated experience toward stated conclusions, with enough visible mechanism that a reader can follow the reasoning, test it against their own observations, and determine for themselves where it holds and where it doesn't.

The claims are specific enough to be wrong. The reasoning is exposed enough to be challenged. The experience is described in enough textured, composite detail — anonymised but operationally concrete — that a practitioner from a different background can recognise whether they have seen the same patterns or different ones.

That is the epistemology, and it is the only one worth defending: not that the author saw clearly, but that the author wrote clearly enough to be checked.

The five doubts above are a first serious attempt at self-checking. They are incomplete — five is not an exhaustive inventory of the ways a twenty-six-year practice might have gone wrong, and a fairer author would probably find more. But they are specific, they carry their falsifiers, and they come with dates. A library that stakes its claim on checkability must eventually submit to the check.

The reader who finds the errors — who demonstrates that the pace was wrong, that the constants bent, that the apprenticeship died, that Europe was hope, that the scars distorted the map — has not disproved the library. They have finished the work the author started.

A conviction library that cannot name its doubts is a brochure. This page is the difference.