

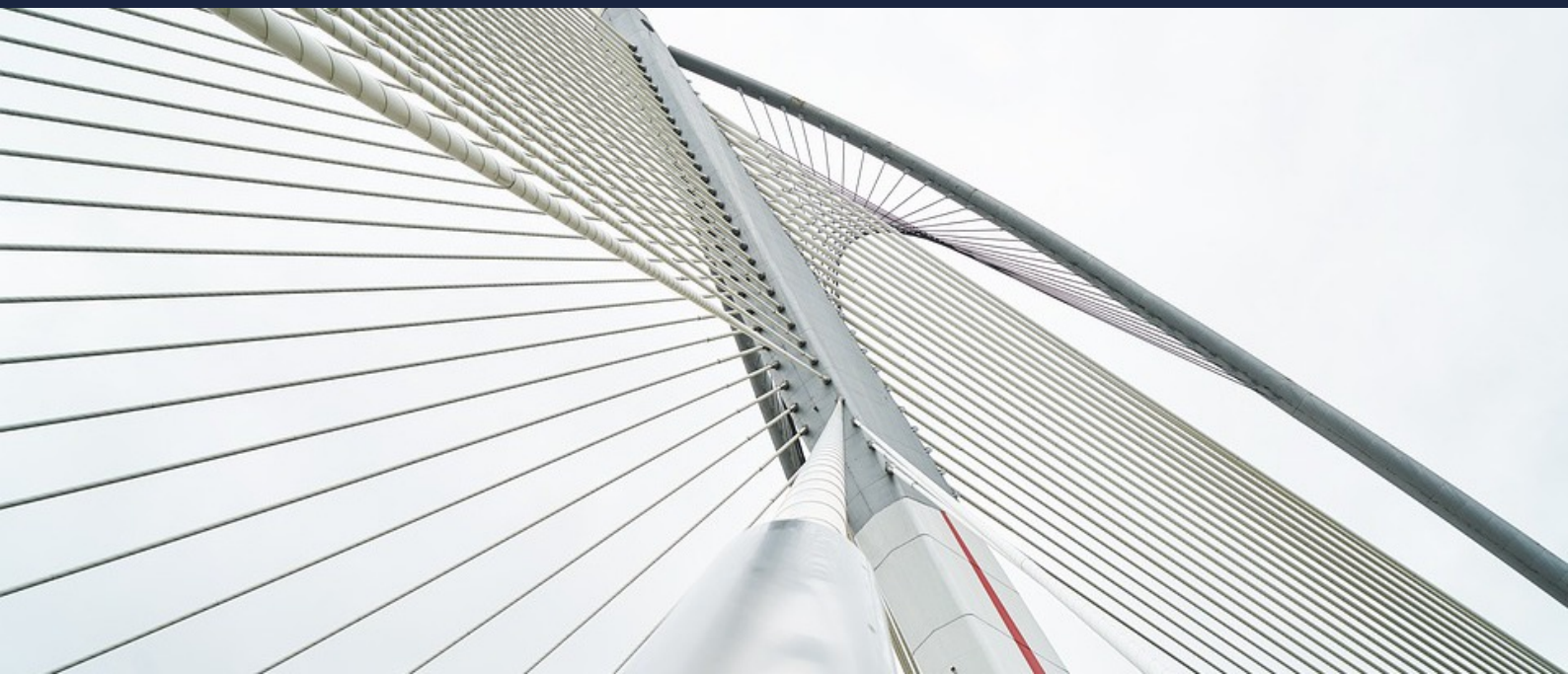
BEYOND THE PYRAMID

The Consultancy Question — What to Buy, What to Keep, and Who Holds the Pen

Engagement Design, Advisory Economics, and the Discipline Both Sides Have Avoided

Giovanni Leonardi

February 2025



The Consultancy Question — What to Buy, What to Keep, and Who Holds the Pen

Leadership · No. 101

Originally written February 2025 | Re-edited and re-rendered for the WhereBeyond Edition, 2025

"Intuition without rigour is opinion, and opinion does not survive a board paper."

1. The Morning After the Kick-Off

There is a particular species of project meeting — held within a fortnight of the engagement letter being signed — where the energy is high and the danger highest. The partner has presented the approach. The workstream leads have been introduced. The governance slide shows a steering committee, a working group, and a set of milestone dates that correspond more closely to the contracting rhythm than to any operational reality. Everyone agrees it is an ambitious programme.

What has not happened — and may never happen, unless someone forces it — is the work that determines whether the engagement will produce anything the organisation can use. No one has named the business owners who will receive each workstream's output and be accountable for acting on it. No one has agreed the format in which recommendations will be delivered — decision-ready papers or the consultant's preferred hundred-page deck. No one has established who holds the pen on the operating model: whether the future state being designed will be the client's own intellectual property, built with advisory input, or a framework the firm will carry to the next client once the logos have been changed.

This is where engagements are won or lost — not in the analysis, not in the strategy, but in the mobilisation. And it is a failure that repeats with remarkable consistency, across sectors and across decades, because it serves almost everyone's short-term interests for it to go unaddressed.

2. Three Chairs, One Table

The consultancy question is rarely examined honestly because honest examination requires holding three uncomfortable perspectives simultaneously. The buyer wants leverage — access to capability the organisation does not have, deployed quickly, with a defined end date. The consultant wants continuity — the engagement extended, the scope widened, the relationship deepened into something that survives the current programme. The internal programme leader, caught between these two, wants the work

done well enough that the board sees progress, without the political exposure of admitting that the organisation could not have done it alone.

Each of these positions is rational. None of them is dishonest. But the interaction between them produces a predictable pathology: the engagement grows to fill the space that internal accountability leaves empty. Where no business owner has been named for a workstream, the consulting team absorbs ownership by default. Where no output format has been agreed, the firm delivers in whatever format maximises its own efficiency — which is rarely the format that enables the client to act. Where scope boundaries lack teeth, the advisory mandate creeps from recommendation into implementation, and the client's own capability atrophies in exactly the area where it was supposed to be built.

I have watched this dynamic from all three chairs. As the executive who commissioned a partner-firm validation of a troubled integration programme — and discovered that the validation team, within six weeks, had quietly expanded from validating the approach to redesigning the workstreams, because no one had defined where validation ended and direction began. As the programme director managing an advisory team whose analysis was exceptional but arrived in a format that required three weeks of internal translation before it could reach the investment committee. And, earlier in my career, as the consultant brought in to provide strategic options for a technology platform — who recognised, halfway through the engagement, that the real problem was not the platform but the governance around it, and that the scope I was stretching towards was the scope the client needed but had not contracted for.

We have all sat at that table. The question is not whether the pattern exists — it is so familiar it barely registers — but whether there is a craft to preventing it, and whether the economics of that craft are about to shift in ways that neither side has fully reckoned with.

3. What the Leverage Model Actually Sells

The traditional consulting model is, at its core, a leverage play. A partnership structure sells senior judgement — pattern recognition accumulated across dozens of engagements, the ability to see what a problem resembles before the client has finished describing it — but delivers it through a pyramid of increasingly junior staff who do the analytical and presentational work that makes the senior insight legible. The partner sees the shape of the answer in the first meeting. The team spends twelve weeks building the evidence base that proves the partner was right.

This is not a criticism. The evidence base matters. Intuition without rigour is opinion, and opinion does not survive a board paper. But it is important to be precise about what each layer of the pyramid contributes, because the economics of the lower layers are changing faster than most firms publicly acknowledge.

The base of the pyramid — the research, the benchmarking, the synthesis of publicly available information, the drafting of slides and position papers — has historically consumed the majority of engagement hours and generated the majority of fee income. It is also the layer most directly exposed to what is happening with generative AI. A task that once required an associate and two analysts working for a week — assembling a market

landscape, synthesising regulatory positions across jurisdictions, drafting a structured options paper — can now be completed, to a credible first draft, in an afternoon by a competent internal team with access to the right tools.

The implications run in two directions simultaneously, and both are uncomfortable.

For the consulting firms, the leverage model that has sustained their economics since the nineteen-eighties is losing its foundation. Clients are beginning to notice — not uniformly, not yet across every sector, but in enough boardrooms that the question is being asked — that they are paying partner-firm day rates for work their own people could now produce, given the right prompts and the right analytical framework. The firms that adapt will unbundle their offering: selling the senior judgement and the pattern recognition explicitly, at a premium, while acknowledging that the production layer has been repriced by technology. The firms that do not adapt will find themselves defending a fee structure that their clients' own tools make increasingly difficult to justify.

For the buyers, the repricing creates an obligation they may not want. If the production work can be done internally, then the buyer must actually build the internal capability to do it — not as an aspiration on a capability roadmap, but as a functioning team with the tools, the access, and the organisational permission to do work that was previously outsourced. This is harder than it sounds. Outsourcing analytical production to a consulting firm is not merely a capability gap; it is often a political choice. The external firm provides air cover, credibility with the board, and a convenient target if the recommendations prove wrong. Bringing that work in-house means accepting accountability for the analysis, not just for the decision that follows it.

4. The Validation Trap

There is a particular category of engagement that illustrates the tension with special clarity: the partner-firm validation. An organisation undertaking a major transformation — a post-merger integration, a technology migration, a regulatory remediation programme — brings in a second firm to validate the approach of the first. The logic is sound: independent assurance, a fresh perspective, a check on assumptions that may have calcified inside the programme team.

In practice, the validation engagement is one of the most difficult to scope and one of the easiest to let drift. The validation firm arrives with its own frameworks, its own pattern library, its own view of what good looks like. Within weeks, the engagement begins to pull towards redesign rather than validation — not because the firm is acting in bad faith, but because genuine validation inevitably surfaces issues that the validator believes it could fix, and because the client, having heard the diagnosis, naturally asks the diagnostician to prescribe. The boundary between "we have assessed your approach and found three material risks" and "we recommend restructuring the programme as follows" is intellectually clear but operationally gossamer.

The consultancy question is not whether to buy advisory. It is whether the organisation has the institutional discipline to define what it is buying — and to hold that definition when the engagement's own momentum works against it.

The organisations that manage validation engagements well do something specific: they define the validation mandate in terms of questions, not capabilities. Not "validate our integration approach" but "answer these seven questions about our integration approach, in this format, by this date, for this decision-maker." The constraint is the discipline. It forces the validation team to work within boundaries that have operational meaning, and it gives the programme director a contractual basis for resisting scope drift when the validator's enthusiasm outpaces the mandate.

5. The Craft of Engagement Design

The consultancy question, properly framed, is not whether to buy advisory — it is how to design the engagement, and there is a craft to this that most organisations have never developed because they have never needed to. When the leverage model was intact and the production work genuinely required external capacity, the engagement could be loosely designed and still produce useful output. The firm would mobilise, the team would work, the slides would arrive, and the client would decide what to do with them. The quality of the engagement design mattered less when the firm's analytical machine was genuinely irreplaceable.

That margin of error is disappearing. As the production layer commoditises, the value of a consulting engagement concentrates in the elements that cannot be commoditised: the framing of the problem, the design of the workstreams, the governance of the engagement itself. And these are precisely the elements that most organisations leave to the firm to define — which is to say, they outsource the design of the very thing they are buying.

Good engagement design has specific, learnable characteristics:

- It begins before the firm is selected, with the client defining not just the problem but the shape of the answer — what decisions the engagement must enable, in what format, by what date, and for which decision-maker.
- It names an internal owner for every external workstream — not a liaison, not a coordinator, but someone accountable for receiving the output, challenging it, and converting it into action.
- It sets scope boundaries with enforceable change control — a mechanism for advisory scope at least as rigorous as the one applied to technology scope, because advisory scope creep is more expensive and harder to reverse.
- It establishes from the outset who holds the pen on the deliverables that will outlive the engagement — the operating model, the target architecture, the governance framework — because these are the assets the organisation will live with long after the firm has moved on.

None of this is revolutionary. Programme managers who have run enough engagements know it instinctively. But it is rarely codified, almost never taught, and in most organisations it lives in the heads of a handful of experienced operators who move on, taking the institutional memory with them.

6. What Cannot Be Bought from a Machine

The temptation, in the current moment, is to frame the consultancy question as a simple substitution: AI replaces the base of the pyramid, so buy less consulting. This misreads both the problem and the opportunity.

What senior advisory genuinely provides — when it is bought well — is not analysis but pattern recognition across a breadth of situations that no single organisation can accumulate internally. The partner who has seen the same operating-model failure in six different industries. The director who knows which governance structure collapses under political pressure and which one holds. The principal who can walk into a stalled programme and identify in half a day whether the problem is technical, political, or structural. This is knowledge that cannot be generated from data, because it was never written down. It exists as accumulated judgement, and it is exactly what organisations need more of as their own analytical capability improves.

The paradox is that better internal capability should lead to better consulting engagements, not fewer. An organisation that can do its own research, build its own evidence base, and draft its own options papers is an organisation that can use a consultant for what a consultant is actually worth: the challenge, the pattern, the external perspective that prevents institutional blindness from hardening into strategic error. The engagement becomes shorter, more focused, and more expensive per day — because what is being bought is genuinely scarce.

But this requires both sides to change. The consultant must be willing to sell a smaller, sharper engagement at the top of the pyramid, rather than padding it with production work the client no longer needs. The client must be willing to pay a higher day rate for genuine expertise, rather than expecting the firm to subsidise its senior people through volume at the base. And both must accept that the engagement, properly designed, will end sooner — which means the institutional relationship must be built on something other than continuous presence.

7. The Question Behind the Question

The consultancy question is, in the end, an organisational maturity question. Organisations that buy consulting badly — that outsource accountability along with analysis, that let the firm define the engagement rather than defining it themselves, that use advisory as a substitute for internal capability rather than a complement to it — do so not because they have chosen bad firms, but because they have not developed the institutional muscle to be good clients.

This matters more now than it did even two years ago, because the economics are shifting and the old arrangement — where the firm's production capacity masked the client's design deficiency — is becoming unsustainable. The firms that understand this are already moving: restructuring their delivery models, investing in the tools that enable their consultants to do more with smaller teams, experimenting with pricing structures that separate the judgement layer from the production layer. The clients that understand it are doing the same: building internal analytical capability, professionalising their engagement-design function, and learning to buy advisory the way they buy any other specialist service — with clear specifications, defined deliverables, and the expectation of a defined end.

“We are not at the end of consulting. We are at the end of a particular model of consulting — one built on leverage, on volume, and on the assumption that the client could not do the production work itself.”

What comes next will be leaner, more expensive per unit, and more valuable per engagement. But only for the organisations — on both sides of the table — that learn the craft the old model allowed them to avoid: the craft of knowing what to buy, what to keep, and above all, who holds the pen.