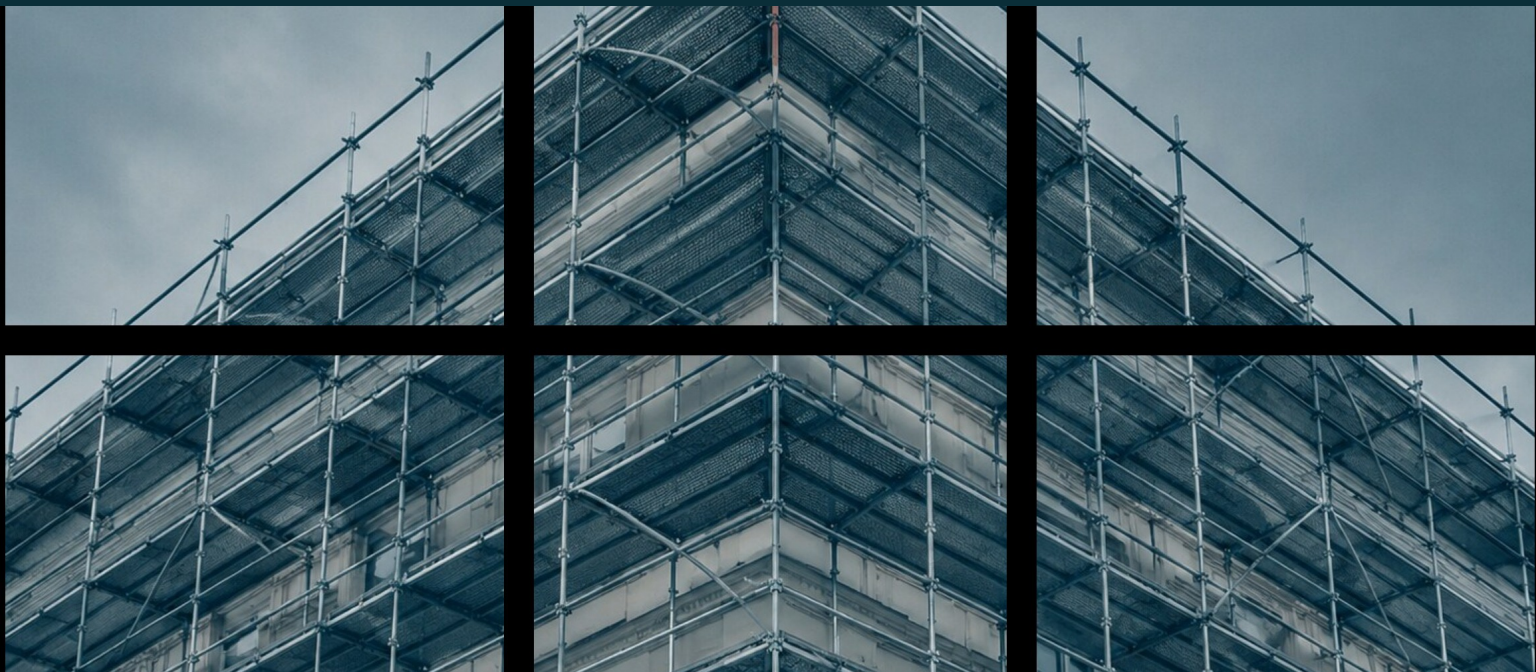


BEYOND THE RATING

# Climbing the Wrong Ladder: Why Maturity Models Measured the Wrong Things

How a decade of process orthodoxy taught organisations to certify their conformance and mistake it for capability



Giovanni Leonardi

October 2006

# Contents

|                                             |   |
|---------------------------------------------|---|
| 1. Executive Summary                        | 3 |
| 2. The Week of the Rating                   | 3 |
| 3. What the Ladder Actually Counts          | 4 |
| 4. Why We Reached for the Number            | 4 |
| 5. Where Conformance and Value Part Company | 5 |
| 6. The Things That Have No Level            | 6 |
| 7. The Counter-Currents                     | 7 |
| 8. What Maturity Might Mean Instead         | 7 |

# Climbing the Wrong Ladder: Why Maturity Models Measured the Wrong Things

Transformation · No. 87

Originally written October 2006 | Re-edited and re-rendered for the WhereBeyond Edition, 2025

---

*“A maturity model can tell you, with real precision, that an organisation follows its processes; it can tell you almost nothing about whether those processes are worth following.”*

---

## 1. Executive Summary

---

For most of the past decade, the maturing of an organisation has been treated as something that can be read off a ladder. Reach Level 3 and you are defined; reach Level 5 and you are optimising. The staged maturity model — the Capability Maturity Model, its integrated successor, and now the newer models reaching beyond software into projects, programmes and whole portfolios — gave the improvement community something it had never had: a single figure that boards could grasp and procurement departments could demand. This essay argues that the figure measured the wrong thing. Maturity models are, at heart, instruments for assessing conformance to defined process. They are precise about whether an organisation does what it says it does, and nearly silent about whether what it does is worth doing. That silence is not a flaw at the edges; it is the centre of the matter, because the qualities that separate a transformation which delivers from one which merely completes — judgement, the honest pursuit of benefit, the willingness to change course when the plan meets the world — have no level and cannot be appraised. What follows traces why we reached for the ladder anyway, what it genuinely gave us, where conformance and value quietly part company, and what a more honest idea of maturity would have to measure instead.

## 2. The Week of the Rating

---

I have watched an organisation celebrate a maturity rating. The appraisal team — external, credentialed, methodical — had spent the better part of a fortnight in the building, working through the evidence: process descriptions, tailoring guidelines, meeting minutes, the artefacts each process area is meant to produce. On the final afternoon they delivered the finding. The organisation was, formally, Level 3. Defined. The word went round the floor before the appraisers had reached the car park. There was an email from the divisional director within the hour and a cake by Friday.

What made the scene stay with me was not the celebration but what sat underneath it. That same quarter, the organisation had shipped a replacement customer system the branch staff quietly refused to use, keeping their old spreadsheets alive on a shared drive. Its flagship integration programme was nine months late and had stopped speaking about benefits at all, retreating to the safer vocabulary of milestones and deliverables. None of this appeared in the rating, and none of it could have. The appraisal had not asked those questions. It had asked, with real rigour, whether the organisation followed its defined processes — and the honest answer was that it did. It followed them faithfully, all the way

to a system nobody wanted.

That gap — between an organisation that reliably does what it says and one that reliably does what matters — is the subject of this essay. It is not an argument that measurement is folly, nor a fashionable sneer at process. It is an argument that we have spent a decade measuring the maturity of the machinery and calling it the maturity of the enterprise.

### 3. What the Ladder Actually Counts

---

The maturity model, in every form the field has produced, rests on a single premise: that the way to improve an organisation's results is to improve the discipline with which it follows its processes. The Capability Maturity Model gave that premise its familiar shape — five stages rising from the initial level, where delivery depends on individual heroics, through repeatable and defined, up to the managed and optimising levels where process is controlled quantitatively and improved continuously. Its integrated successor tidied the picture, absorbed the competing variants that had grown up around it, and in its latest revision offered a choice between the staged representation, which awards one overall level, and the continuous one, which profiles capability area by area. The vocabulary has since spread well beyond its origins in software engineering. There are maturity models now for project management and for service management, and — most ambitiously — a new model from the United Kingdom's public-sector guidance, which reaches up from individual projects into programmes and entire portfolios, promising to tell a department how mature its delivery capability is across the board.

Strip away the differences and the same instrument is at work in all of them. Each model defines a set of practices a competent organisation ought to perform — requirements are managed, configurations controlled, risks logged and reviewed, suppliers governed, decisions recorded. Each then assesses, through documentary evidence and interview, the degree to which the organisation genuinely performs those practices, institutionalises them, and holds to them under pressure. The output is a position on a ladder.

This is a real and useful thing to know. An organisation that cannot reliably control a change to its own requirements is in difficulty, and a model that surfaces the fact is doing honest work. But notice exactly what is being counted. The unit of measurement is conformance to defined practice. A high rating certifies that the organisation does what its processes say, consistently, and can prove it. It does not, and structurally cannot, certify that the processes lead anywhere worth going. A maturity model can tell you, with real precision, that an organisation follows its processes; it can tell you almost nothing about whether those processes are worth following.

### 4. Why We Reached for the Number

---

It is worth being fair to the ladder, because the reasons we climbed it were good ones, and pretending otherwise would falsify the history. Before the maturity movement, getting better was largely a matter of assertion. A delivery director might believe the organisation had improved, but could not show it, could not compare this year with last, and could not compare one division against another. The maturity model changed that. It gave improvement a common vocabulary, so that people in different functions and different countries could mean the same thing by a defined process. It made improvement legible to the people who hold the money, because a movement from Level 1 to Level 3 is a story a finance committee can follow and therefore fund. And in shops where delivery genuinely

depended on which individuals happened to be in the room, the models did something undeniably valuable: they wrote down what the best people did by instinct, reduced the variance between teams, and let competence survive the departure of the heroes. I have watched a chaotic development group steadied by exactly this discipline — its firefighting quietened, its estimates made to mean something for the first time. The models were not a confidence trick. They worked on the thing they were built to work on.

Around that genuine value, though, gathered a set of forces that had less to do with improvement than with the appetite for a number. Boards and executive committees, asked to oversee delivery they did not themselves understand, found in the maturity level a single figure to put on a slide. The climate of the moment magnified the appeal: in the years since the accounting scandals and the compliance legislation that followed them, control had become the watchword, and a model that visibly evidenced control was welcome for that reason almost regardless of what it controlled. And procurement turned the level into currency. It became ordinary for an invitation to tender to name a minimum maturity rating as a gate, and the large offshore delivery houses learned the lesson quickest of all, presenting Level 5 appraisals as proof of superiority in bid after bid. Once a rating is worth money, the incentive shifts quietly from improving to being seen to have improved — and an entire apparatus of appraisal, tailoring and preparation grows up to serve the second thing while speaking the language of the first.

## 5. Where Conformance and Value Part Company

---

Here is the difficulty at the heart of it. The moment a proxy becomes a target, it begins to drift from the thing it was standing in for. Conformance to defined process was only ever a proxy for the outcome we actually cared about — capable, reliable delivery of things worth having. While the level stayed a private diagnostic, proxy and target moved together. Once the level became a prize — funded, procured, celebrated — they came apart, because it is entirely possible to raise the proxy without touching the target, and a great deal cheaper to do so.

An organisation optimising for the rating learns to produce the evidence the appraisal consumes. Processes are documented to precisely the standard the model rewards; artefacts appear at each gate because the gate expects them; the meeting is held and minuted because the practice requires a meeting to have happened. None of this is fraud. It is, in a way, worse than fraud, because it is sincere. The organisation comes to believe that the immaculate defect record and the fully populated risk register are the improvement, when they are only its shadow.

Consider a programme I will keep deliberately composite. By every process measure it was exemplary. It passed each of its staged gateway reviews. Its milestone reporting was the envy of the portfolio: in its final year it recorded ninety-four per cent of planned milestones met on or close to their dates. Its risk log was current, its change control tight, its documentation complete; it would have appraised well against any model you cared to bring to it. And when it closed, the benefits case that had justified three years of funding — the reduction in processing cost, the retirement of the legacy estate, the improvement in customer turnaround — was delivering perhaps a third of what had been promised, with no one able to say when the remaining two-thirds might arrive, because the benefits had not been revisited since the business case was signed off at the outset. The programme had met every process obligation and missed the point. The maturity of its machinery was never in doubt. The maturity of its purpose had never once been assessed.

*“An organisation can be flawless in the doing and mistaken in the what.”*

## 6. The Things That Have No Level

If the models measure conformance well and value poorly, it is worth asking what precisely falls outside their reach – because the omissions are not random. They share a family resemblance. The things a maturity model cannot see are exactly the things that resist being written down as a repeatable practice.

Consider judgement. The decision that matters most in a transformation is often the decision to stop – to abandon a design that is technically on track but heading for a place the organisation no longer wants to be. No process area rewards that decision; several actively penalise it, because stopping shows up as a variance against plan. A mature organisation, in the ladder's sense, is one that executes its plan faithfully. But faithful execution of the wrong plan is not maturity; it is well-organised waste, and the model cannot tell the two apart.

Consider benefits. There is a discipline, still young and still fighting for room, that treats the realisation of benefit as an active responsibility running long after delivery – tracking whether the promised value actually appears, and holding someone accountable when it does not. The better programme guidance has begun to insist on it. But benefits realisation sits uneasily on a maturity ladder, because it cannot be evidenced in an appraisal week. Its proof arrives months or years after the artefacts are all in order, by which time the rating is long since awarded and the team dispersed.

Consider adaptation. Every model prizes stability of process, and stability is a virtue when the environment is stable. When the ground is moving – a shifting market, a technology that changes what is possible, a strategy revised above the programme's head – the virtue inverts, and the organisation most faithful to its defined process is the slowest to respond. There is an argument, growing louder in the development community, that the whole edifice of heavy up-front definition is mismatched to work whose requirements cannot be known in advance, and that responsiveness to change should count for more than adherence to a plan. It is a contested argument, and its more zealous advocates would throw away disciplines worth keeping. But it has located something real that the maturity frame has no way to score.

And consider temperament – the disposition to tell an uncomfortable truth in a status meeting, to surface a failing early rather than manage its appearance until the next gate. This is perhaps the single greatest determinant of whether a transformation recovers from trouble, and it is entirely invisible to appraisal. You cannot institutionalise candour by writing it into a process description.

| What the ladder counts                     | What the ladder cannot see                      |
|--------------------------------------------|-------------------------------------------------|
| Whether processes are defined and followed | Whether they are the right processes            |
| Artefacts produced at each stage           | Whether anyone downstream uses them             |
| Risks logged and reviewed                  | Whether the organisation acts on what it learns |
| Milestones met                             | Benefits actually realised                      |

| What the ladder counts        | What the ladder cannot see                     |
|-------------------------------|------------------------------------------------|
| Consistency and repeatability | Judgement, adaptation, and the courage to stop |

## 7. The Counter-Currents

None of this is going unnoticed. Two currents are pulling against the orthodoxy at once, and it is not yet clear which, if either, will correct it.

The first is the reaction from the development floor — the growing insistence that working outcomes matter more than the documentation of process, and that responding to change should outweigh following a plan. It carries an obvious appeal to anyone who has watched a programme conform its way to failure. But it arrives with its own hazard. In its harder forms it reads as a licence to abandon discipline altogether, and it is easy to imagine an organisation using it as cover to unlearn the genuine gains the maturity movement delivered. The honest position is uncomfortable: the orthodoxy measures the wrong thing, and its loudest critics risk measuring nothing at all.

The second current runs the other way. Maturity thinking is not retreating; it is expanding. The newest models carry the staged-level logic up out of software and into programmes and portfolios — precisely the territory where the gap between conformance and value is widest, because it is at portfolio level that the question are we doing the right things finally overwhelms the question are we doing things right. Whether that expansion is progress or the original error written larger is, at this moment, a genuinely open question. Applied as a private diagnostic — a mirror an organisation holds up to itself — the extension could do real good. Applied as another prize to be won and demanded in tenders, it will reproduce at the level of the whole enterprise the very decoupling it inherited.

The strongest defence of the models must be met head-on, because it is not weak. Its proponents argue that mature organisations do deliver better — that the correlation between high ratings and good outcomes is real and well attested, and that the failures described here are failures of misuse, not of the instrument. There is truth in the first half. Capable organisations tend both to conform and to deliver, so the correlation is unsurprising. But correlation is doing the persuasive work that a causal claim cannot support: it does not follow that raising conformance produces value, and the prize dynamics described above sever even the correlation, because they let an organisation raise its rating precisely without improving its delivery. As for misuse — when an instrument is misused nearly everywhere it is used, at some point the design has to answer for the pattern. A measure that is this easy to game, and this attractive to game, is not merely being held the wrong way round.

## 8. What Maturity Might Mean Instead

I do not think the answer is to burn the ladders. The diagnostic instinct behind them is sound, and the discipline they instilled in genuinely chaotic organisations was real and hard-won. The error was not in measuring; it was in measuring the safe, countable, near thing — conformance — and allowing it to stand in for the difficult, deferred, contestable thing we actually wanted, which is value delivered and judgement well exercised.

A more honest idea of maturity would begin further downstream. It would ask not whether an organisation follows its processes but whether it can tell when a process has stopped serving its purpose and has the nerve to change it. It would treat the realisation of benefit, not the production of artefacts, as the evidence that matters, and it would accept that such evidence arrives too late to fit inside an appraisal week — which is precisely why it is worth waiting for. It would prize the organisation that stops a failing programme over the one that delivers it immaculately. None of this reduces cleanly to a number on a slide, and that is rather the point: the things most worth knowing about an organisation's maturity are the things that were left off the ladder because they would not sit still to be counted.

*The safe thing to measure and the thing worth measuring are rarely the same, and the whole history of the maturity model is the story of choosing the first and hoping it would stand in for the second.*

Where this settles, I cannot say from here. The reaction on the development floor may harden into something disciplined or dissipate into slogans; the new portfolio-level models may become honest mirrors or merely grander prizes. What seems clear, standing in the middle of it, is that an organisation which has learned to pass its own audits has learned something — but not the thing it most needs to know about itself. We became fluent, this past decade, in the grammar of process. We are still, by and large, illiterate in the harder language of whether any of it was worth doing. Until a model can read that, a maturity rating will remain what it has quietly been all along: an accurate answer to a question we should never have mistaken for the important one.