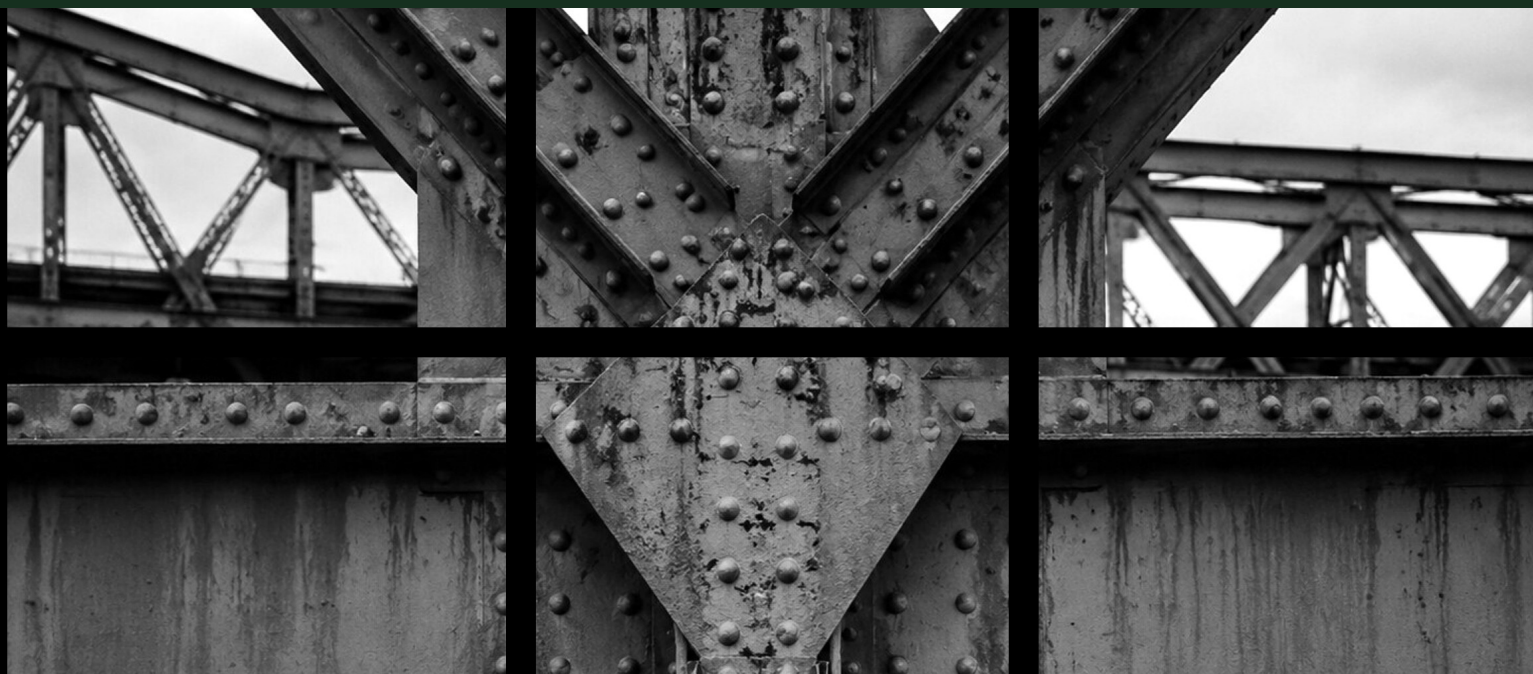


BEYOND THE GREEN RAG

When KPIs Become Targets — Goodhart's Law in Programme Management

A Practitioner's Perspective on Metric Distortion in Complex Delivery Environments



Giovanni Leonardi

July 2007

When KPIs Become Targets – Goodhart's Law in Programme Management

Programme · No. 99

Originally written July 2007 | Re-edited and re-rendered for the WhereBeyond Edition, 2025

“The moment a programme manager realises that the metric has become the objective, it is already too late – the behaviours it was designed to prevent are already embedded in the delivery culture.”

1. The Law That Explains Everything

Charles Goodhart's observation – that any measure which becomes a target ceases to be a good measure – was originally directed at monetary policy. But it has found its most fertile ground in programme management, where the pressure to demonstrate control through metrics creates precisely the conditions under which Goodhart's Law operates with devastating efficiency.

The pattern is so consistent that it barely qualifies as an observation any more. A programme establishes KPIs. The KPIs are tied to governance reporting. The governance reporting is tied to executive attention. Executive attention is tied to consequences – funding decisions, career progression, organisational reputation. And so the KPIs, which were designed to provide an honest signal about programme health, become instead the thing that the programme optimises for, regardless of whether optimising for the metric actually advances the programme's objectives.

This is not a subtle phenomenon. It is visible in almost every large programme I have encountered. And yet it persists, because the alternatives – governing without metrics, or governing with metrics that nobody is held accountable for – feel worse.

2. How It Manifests

The specific forms vary by sector and programme type, but the underlying mechanics are remarkably stable. Consider some of the patterns that recur:

Milestone completion rates. A programme reports that it has hit 85% of its milestones on time. This sounds healthy. But drill into the data and a different picture emerges. Milestones have been redefined – what was originally a single milestone representing a working system has been decomposed into twelve smaller milestones representing documentation deliverables, design approvals, and test plans. The programme has hit twelve milestones, each representing a fraction of the original intent, and missed the one that mattered. The completion rate is arithmetically correct and strategically meaningless.

Defect counts. A programme tracks defects as a quality indicator. The target is to reduce defects over time. And defects do indeed fall – not because quality is improving, but because the classification of what constitutes a defect has been progressively tightened. Issues that would have been logged as defects in month one are now logged as change requests, observations, or known limitations. The metric improves; the quality does not.

Resource utilisation. A programme measures the percentage of its team that is productively deployed. The target is high utilisation. And so people are assigned to tasks whether or not the tasks are ready to be worked on, because having people visibly busy is more important than having people productively engaged. The utilisation metric hits its target while the programme burns through its budget on activity that adds no value.

The milestone completion rate is arithmetically correct and strategically meaningless. The programme has optimised for the metric, not for the outcome the metric was designed to represent.

3. The Structural Incentive

It is tempting to attribute this to dishonesty or incompetence, but that misses the point. The people gaming the metrics are usually acting rationally within the incentive structure they have been given. When a programme manager's credibility — and by extension their career — depends on presenting a healthy set of KPIs to a steering committee, they will naturally find ways to present healthy KPIs. This is not deception in any meaningful sense; it is adaptation.

The deeper problem is that most programme governance structures create a binary incentive: either the metrics are good and the programme continues to receive support, or the metrics are bad and the programme faces intervention. There is no reward for honest reporting of difficulty. A programme manager who reports accurately that the programme is behind schedule, over budget, and struggling with quality will receive scrutiny, challenge, and additional oversight. A programme manager who reports that everything is on track — supported by carefully curated metrics — will receive continued trust and autonomy.

The rational choice is obvious. And it is the wrong choice for the organisation, which needs honest signals far more than it needs comfortable ones.

4. The Compounding Effect

What makes this particularly damaging is that it compounds over time. Early in a programme, the gap between the metric and the reality is small — a milestone redefined here, a defect reclassified there. The programme manager may not even recognise what they are doing as gaming; it feels more like pragmatic interpretation. But each small adjustment makes the next one easier to justify, and the cumulative effect is a progressive divergence between what the metrics say and what is actually happening.

By the time the programme reaches a point of genuine crisis — a missed delivery, a budget overrun that cannot be absorbed, a quality failure that reaches the customer — the metrics have been telling a positive story for so long that the crisis comes as a shock to the governance bodies that were supposed to be monitoring the programme's health. The steering committee asks, reasonably: how did we get here when the dashboard was green last month? The answer is that the dashboard has been green every month, and it has been wrong every month, and the governance structure that relied on it was not governing at all.

5. The Audit Response

The typical organisational response to this discovery is to add more metrics, on the theory that a wider net will catch more problems. This is precisely backwards. More metrics create more targets, which create more opportunities for Goodhart's Law to operate. The programme now has to optimise for twenty metrics instead of five, which consumes more management attention, generates more reporting overhead, and produces an even more elaborate edifice of curated data that obscures rather than illuminates the programme's actual condition.

The audit function, well-intentioned as it is, often reinforces this dynamic. An audit that finds metric manipulation will recommend stronger controls, more frequent reporting, and additional oversight — all of which increase the pressure on the programme manager to present clean metrics, which increases the incentive to game them. The cycle tightens rather than breaks.

“The moment a programme manager realises that the metric has become the objective, it is already too late — the behaviours it was designed to prevent are already embedded in the delivery culture.”

6. What the Better Programmes Do Differently

The programmes that manage to maintain honest measurement — and they do exist, though they are not the majority — share several characteristics that are worth examining.

First, they separate the measurement conversation from the accountability conversation. The programme reports its metrics in one forum; the discussion about what to do about them happens in another. This matters because it removes the immediate threat response. A programme manager who can present a deteriorating metric without immediately facing a challenge to their competence is far more likely to present it honestly.

Second, they reward early warning, not clean dashboards. The programmes that sustain honest reporting are the ones where a programme manager who raises a risk early is treated as someone doing their job well, not as someone who has failed. This requires a fundamental shift in governance culture — from one that punishes bad news to one that punishes late news.

Third, they use leading indicators rather than lagging ones. Most programme KPIs measure what has already happened — milestones hit, budget spent, defects found. By the time these metrics indicate a problem, the problem is already baked in. Leading indicators — team confidence surveys, decision velocity, dependency resolution rates — are messier, harder to game, and far more useful as genuine signals of programme health.

Finally, they accept ambiguity. The programmes that resist Goodhart's Law are the ones where the governance bodies are willing to sit with uncertainty rather than demanding false precision. A programme manager who is asked are we on track? and is allowed to answer broadly yes, with two areas of concern that I want to talk through is operating in a healthier governance environment than one who is required to answer with a single RAG

status.

7. The Deeper Lesson

Goodhart's Law in programme management is not a measurement problem. It is a governance problem, and ultimately a leadership problem. Metrics can only ever be proxies for the things we actually care about — working systems, satisfied users, realised benefits, sustainable capability. When we treat the proxies as the things themselves, we create an elaborate performance of programme management that substitutes for the real thing.

The organisations that understand this do not abandon metrics. They hold them lightly. They treat them as starting points for conversation rather than endpoints of assessment. They invest in the judgement of their programme leaders rather than in the sophistication of their measurement frameworks. And they build governance cultures where the question what is actually happening? is asked with genuine curiosity rather than prosecutorial intent.

This is harder than building a better dashboard. But it is the only approach that Goodhart's Law cannot defeat.