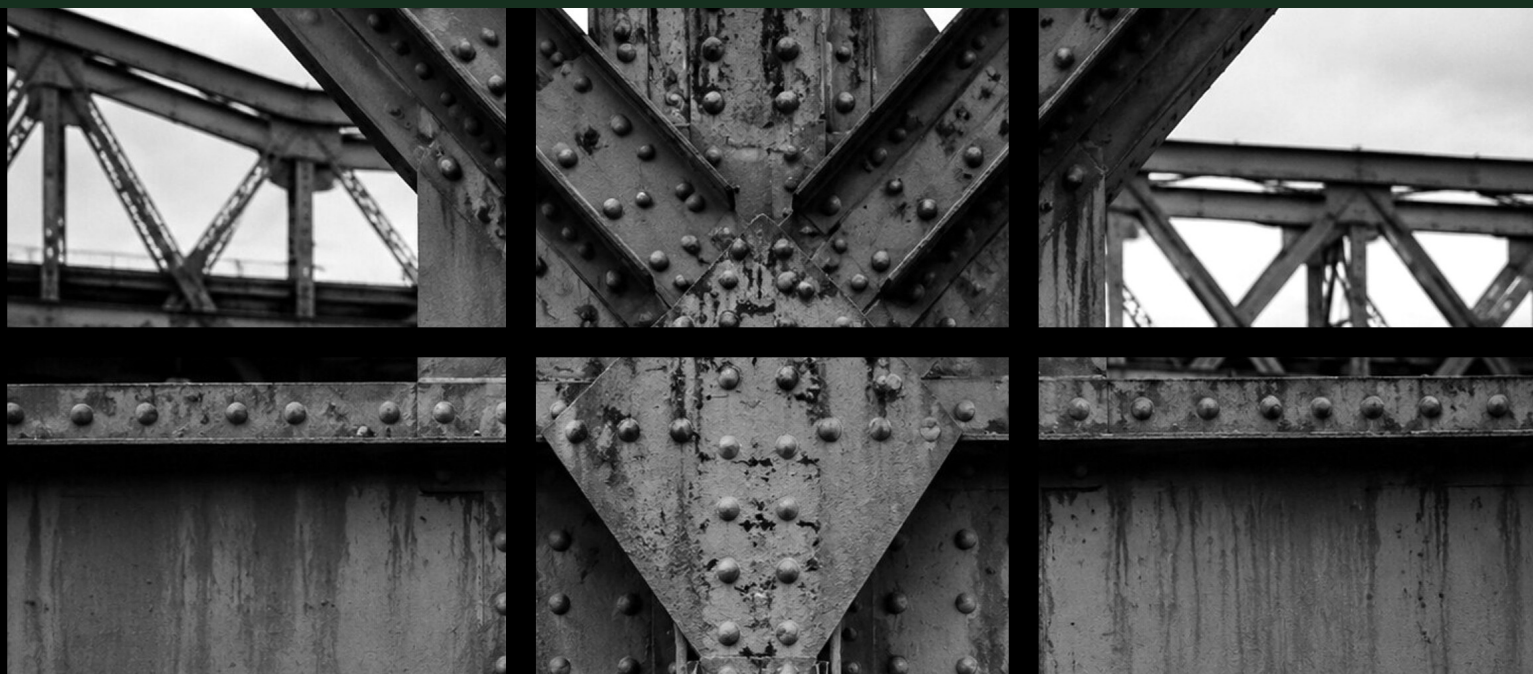


BEYOND THE VISIBLE CUT

The Wrong Waste: Why Lean Thinking Stalls When It Reaches the Programme

How a discipline forged on the factory floor loses its way in the governance of large-scale change



Giovanni Leonardi

April 2010

Contents

1. Executive Summary	3
2. The efficiency that made nothing faster	3
3. Where the word "waste" comes from	4
4. The seven wastes, translated	4
5. But governance is not waste	5
6. Why we keep cutting the wrong thing	6
7. What Lean actually offers a programme	7
8. The distance between intent and reality	8

The Wrong Waste: Why Lean Thinking Stalls When It Reaches the Programme

Programme · No. 151

Originally written April 2010 | Re-edited and re-rendered for the WhereBeyond Edition, 2025

“A board that convenes, receives a report, notes the report, and defers the decision to the next board is not control – it is waste wearing the costume of control.”

1. Executive Summary

Lean thinking has reached the programme office, and it has arrived at the worst possible moment to be understood properly. Nearly two years after the credit markets seized, organisations are under standing instruction to do more with less, and Lean – with its clean promise to eliminate waste – reads to a tired executive like a licence to cut. So programmes are put on a diet. The highlight report shrinks from forty pages to four; the weekly board is halved; three documents are merged into one. The team applauds the new leanness. Six months later the programme is no faster than it was.

This essay argues that the failure is not one of effort but of translation. Lean was forged on the factory floor, where waste is physical and visible – a bank of unsold inventory between two machines, a bin of defective parts. Lifted into a programme, whose real substance is decisions, coordination and the movement of information, the word waste now points at nothing the eye can see. So we attack the waste we can see, the visible artefacts of governance, and leave untouched the waste that is actually slowing us down: the decision waiting three weeks for a board that will only defer it, the analysis sitting half-finished because its sponsor has moved on, the handoff between two functions that neither of them owns.

Examined honestly, a programme's governance is very largely a queuing system, and the deepest thing Lean ever taught manufacturing was not the elimination of paperwork but the management of flow. That is precisely the part we have left on the factory floor. What follows works through what the seven wastes actually look like at programme scale, takes seriously the objection that governance is not waste but control, and ends on the uncomfortable distance between what we say we are doing when we "go Lean" and what we are in fact doing.

2. The efficiency that made nothing faster

Picture a programme that has been told to become efficient. The director is a capable, conscientious person who has read the same articles everyone has read, and who understands that the mood of the organisation has changed: the years of expansive budgets are over, and the watchword now is value. She convenes her leads and declares a war on waste. Within a month the forty-page highlight report is a four-page dashboard. The fortnightly programme board, which used to run to three hours, is capped at ninety minutes. A layer of documentation – the ones nobody could quite remember the purpose of – is quietly retired. By any visible measure the programme is leaner. The reporting

burden is lighter, the meetings are shorter, the sponsors are pleased.

And nothing goes faster. The change requests that took nine weeks to clear three governance boards still take nine weeks. The dependency on the integration team that held up two workstreams last quarter holds them up again this quarter. The business case for the second release, submitted in January, is still moving through its approvals in April. Whatever the programme's problem was, it was not the length of the highlight report — and yet the length of the highlight report is precisely what got fixed.

This is the pattern that recurs wherever Lean arrives in a programme without being understood: we cut what we can see, because what we can see is the only thing we know how to cut. The paradox is that the cutting is not wrong, exactly — a forty-page report probably was over-processed — it is simply beside the point. The real waste was never in the room where the decisions to cut were being made.

3. Where the word "waste" comes from

It is worth remembering how literal the idea of waste was at its origin. On a production line, waste is a physical thing you can walk up to and touch. It is the stack of half-finished sub-assemblies piled between one machine and the next because the second machine cannot keep pace with the first. It is the defective component that must be scrapped or reworked. It is the operator walking twenty feet to fetch a tool that should have been at arm's reach. The great contribution of the Toyota tradition, and of the writers who later carried it into the Western vocabulary, was not to invent efficiency — every manager wants efficiency — but to make waste visible and therefore attackable. Once you can see the pile of inventory growing between two machines, you cannot un-see it, and you are compelled to ask why it is there.

A programme has no floor to walk. Its inventory is invisible. The half-written business case sits in a drafts folder; the decision awaiting sign-off sits in an inbox; the requirement that has been agreed but not yet built sits, in effect, in a warehouse that no one can point to. You cannot see the queue, so you do not manage the queue, and because you do not manage it, it grows. The pile of work-in-progress between two workstreams is every bit as real as the pile between two machines, and every bit as expensive — it is capital tied up, risk accumulating, decisions ageing — but because it throws no shadow on the wall, it escapes attention entirely.

The consequence is that when a programme sets out to remove waste, it removes the only waste it can find: the tangible artefacts. Documents are tangible. Meetings are tangible. So documents and meetings are cut, and everyone feels the satisfying sensation of having done something, while the vast invisible inventory of stalled decisions and partially-done work sits exactly where it always sat.

4. The seven wastes, translated

The discipline of Lean has always been specific about the forms waste takes; the classic account names seven. The exercise that almost no programme performs is to translate each of them faithfully out of the factory and into the world of large-scale change. It is worth doing carefully, because the translation is where the misunderstanding lives.

Manufacturing waste	Its true programme equivalent
Overproduction	Producing governance artefacts, reports and analyses that exist because the process demands them, not because a decision needs them
Waiting	A decision, an approval or a dependency sitting idle in a queue — by far the largest and least visible waste in any programme
Inventory	Partially-done work: business cases drafted but not approved, requirements agreed but not built, code written but not integrated
Over-processing	The forty-page report where four pages would do; the three-board approval for a reversible decision; governance polished far beyond the risk it manages
Transport and handoffs	Work passed between functions — analysis to design, design to build, build to test — each handoff a chance to lose context and add delay
Motion	The effort of finding information that should be to hand: chasing the current version, reassembling context, hunting for who decided what
Defects	Rework caused by ambiguity — features built from a misread requirement, integration that fails because two teams assumed different things

Set out this way, one line should trouble any honest reader: waiting is the largest waste in almost every programme, and it is the one our efficiency drives never touch. We attack over-processing, because the forty-page report is visible and mildly embarrassing. We rarely attack waiting, because waiting has no artefact to hold up and no owner to hold responsible.

Consider a single change request, followed honestly from the moment it is raised to the moment it is implemented. In the programmes I have watched at close quarters, the arithmetic is remarkably consistent. The request spends perhaps two days actually being worked on — assessed, estimated, designed, built. It spends the remaining sixty-one days waiting: waiting to be triaged, waiting for the impact assessment, waiting for the board that meets fortnightly, waiting for the board that meets monthly to ratify the board that meets fortnightly. Two days of work inside sixty-three days of elapsed time is a process efficiency of around three per cent. Now ask which of those two numbers our war on waste was aimed at. We shortened the two days of touch time by writing a slightly terser assessment. We never once looked at the sixty-one days of wait.

The uncomfortable measure of a programme is not how hard its people are working but how long its work is waiting. Touch time is where we look; wait time is where the waste lives. A programme running at three per cent process efficiency does not have a productivity problem — it has a queuing problem, and no amount of working faster on the two per cent will fix the ninety-seven.

5. But governance is not waste

Here the argument must slow down and meet its strongest objection honestly, because there is a version of everything above that is dangerous nonsense, and any experienced

hand will have felt the danger rising.

The objection runs like this. Control exists for a reason. Anyone who has been in the room when an ungoverned programme comes apart — the spend that ran months past the point anyone was watching, the go-live that failed because no one owned the decision to stop it, the audit that found no trail where a trail was legally required — knows that governance is not bureaucratic decoration. It is the immune system of a large endeavour. Stage gates exist because programmes that do not stop to check whether they should continue tend not to stop at all. Boards exist because decisions that are no one's job to make in particular become everyone's disaster in general. To wander through this apparatus calling it waste is the naive enthusiasm of someone who has never had to clean up after its absence. Lean, in the wrong hands, is simply a sophisticated argument for removing the brakes.

This objection is correct, and it is the reason the whole enterprise is delicate. But it is correct about governance, and the argument here is not against governance. It is against a specific confusion — the confusion of governance with its rituals. The distinction is the entire game.

A board that convenes, weighs a genuine question, and makes a decision that changes what happens next is not waste; it is the most valuable hour in the programme. A board that convenes, receives a report, notes the report, and defers the decision to the next board is not control — it is waste wearing the costume of control. The stage gate that genuinely asks "should this continue?" and is empowered to answer "no" is the immune system doing its work. The stage gate that has never once stopped anything, that exists to be passed rather than to decide, is pure over-processing that has learned to call itself assurance.

Lean thinking, properly understood, does not remove the brakes. It asks of every governance step exactly the question the factory asks of every step on the line: does this transform the work — does a decision get made, a risk get retired, a direction get set — or does it merely move the work along while adding delay? The steps that transform are precious and should be protected. The steps that only move things along, that exist because they have always existed, are the waste — and they are waste precisely because they masquerade as control, which is what makes them so much harder to see than a forty-page report.

6. Why we keep cutting the wrong thing

If the real waste is so much larger than the visible waste, why does the pattern persist? Why, programme after programme, organisation after organisation, do we reach so reliably for the document and so rarely for the queue? The forces are structural, not personal, which is why exhortation never fixes them.

- The measurement asymmetry. You can count pages, and you can count meeting hours, effortlessly. Decision latency — the number of days a typical decision spends waiting — is almost never measured, because nobody's system captures the moment a decision becomes possible as distinct from the moment it is made. What gets measured gets managed, and what cannot be seen cannot be measured, so the visible waste is attacked and the invisible waste is not even named.
- The framing of the moment. In a downturn, efficiency is heard as cost, and cost means things that sit on a ledger — headcount, licences, the printing bill, the day rate of the people in the long meeting. Flow does not appear on any ledger. The sixty-one days a change request spends waiting cost the organisation a great deal, but the cost

is diffuse, deferred and unattributed, so it never enters the conversation about savings.

- The ownership vacuum. Every function owns its own work. No one owns the white space between functions — the handoff, the queue, the waiting. The delay lives in the gaps, and gaps have no manager, no budget line and no name on a slide. A waste that belongs to everyone belongs to no one, and so it endures.
- The comfort of visible activity. This is the deepest force and the hardest to unlearn. A fully-booked board, a busy team, a high utilisation figure — these feel like control and progress. But one of Lean's most counter-intuitive lessons, learned the hard way in manufacturing, is that a resource run at a hundred per cent utilisation does not maximise throughput; it guarantees queues. The busier every part of the programme looks, the longer the work waits between them. We mistake the appearance of full effort for the reality of fast delivery, and so we optimise for the very thing that is slowing us down.

“We do not have a programme that is working too slowly. We have a programme whose work is waiting too long, run by people who are busier than they have ever been.”

None of these forces is a failure of intelligence or of will. They are the ordinary physics of large organisations, and they explain why a genuinely capable director, genuinely trying to remove waste, ends up removing the report and keeping the delay. The intent is transformation; the mechanism available to her points at the wrong target.

7. What Lean actually offers a programme

If Lean is not, at programme scale, a licence to cut documents, what is it? Stripped back to what made it powerful on the factory floor, it is a discipline of flow — and every one of its instincts translates, if we are willing to do the translation rather than the cost-cutting.

1. Make the invisible visible. The first move is not to cut anything but to see. Put the decisions currently in flight on a wall where everyone can count them. Show the queue of change requests with the date each one entered the queue, so that ageing becomes undeniable. There is a growing interest, in the more thoughtful corners of the field, in visualising the flow of work in exactly this way — borrowing the factory's habit of making the pile of inventory impossible to ignore. A programme that can see its own queues will not tolerate them for long.
2. Limit work in progress. A programme that starts everything finishes nothing. The instinct under pressure is to open more workstreams, to show that every priority is being addressed; the effect is to spread scarce decision-making and integration capacity so thin that everything waits on everything else. The counter-instinct — deliberately capping the amount of work allowed to be in progress at once — feels like doing less and in fact delivers more. I have seen a programme cut its concurrent workstreams from eleven to four and increase what it actually shipped, because for the first time the things that were started could be finished before the next thing began.

3. Treat the decision as the unit of work. In a factory the unit is the part; in a programme it is the decision. Measure decision latency the way a plant measures cycle time. Ask, of every governance body, how long a typical decision waits in its queue, and whether it is empowered to decide or merely to recommend to something else that decides. A programme that manages its decisions as carefully as a plant manages its parts will find most of its sixty-one days of waiting hiding in the gaps between boards.
4. Pull, do not push. Push systems start work because the plan says it is time; pull systems start work when there is capacity to finish it. A programme that pulls — that begins the next piece only when the previous piece has cleared, rather than launching on a date because a Gantt chart demanded it — accumulates far less partially-done inventory and far less of the rework that comes from building on foundations that later shift.
5. Attack the wait, not the touch. The single reorientation that changes everything is to stop trying to make busy people busier and start asking why the work is standing still. Every hour spent shaving touch time is an hour spent on three per cent of the problem. The prize is in the ninety-seven.

Notice that not one of these moves is a cut. Some of them — visualising flow, instrumenting decision latency — actually add a little activity in order to remove a great deal of delay. That is the tell that distinguishes Lean thinking from a cost programme wearing its clothes: a cost programme always subtracts, whereas flow thinking will happily add a small, cheap, visible step if it dissolves a large, expensive, invisible queue.

8. The distance between intent and reality

There is a gap at the centre of all of this, and it is worth naming plainly, because it is the same gap that opens between almost every transformation's intent and its result. When an organisation says we are going Lean, what it intends is a programme that is faster, more responsive, less clogged — genuinely transformed. What it enacts, more often than not, is thinner documents and shorter meetings on top of exactly the same delays. The two are not the same transformation. The distance between them is precisely the distance between attacking the waste you can see and managing the flow you cannot.

Lean thinking was never, at its origin, a cost programme. It reduced cost, handsomely, but it did so as a consequence of relentlessly improving flow — not by hunting for things to remove, but by making the work move so smoothly that the cost fell out on its own. Somewhere on the journey from the factory floor to the programme office we inverted it. We kept the reduction and discarded the flow. We went looking for things to cut and called the looking Lean, and because documents and meetings are the only waste a programme can see, documents and meetings are what we cut.

The factory floor gave us a genuine gift, and we have half-unwrapped it. The half we took — the promise that waste can be found and removed — is the easy half and, on its own, the misleading half. The half we left behind — that the waste worth removing is mostly invisible, mostly delay, and mostly living in the gaps that no one owns — is the half that would actually make our programmes faster. Until we go back for it, we will keep shortening the report, applauding the new leanness, and wondering, six months on, why nothing goes any quicker.