

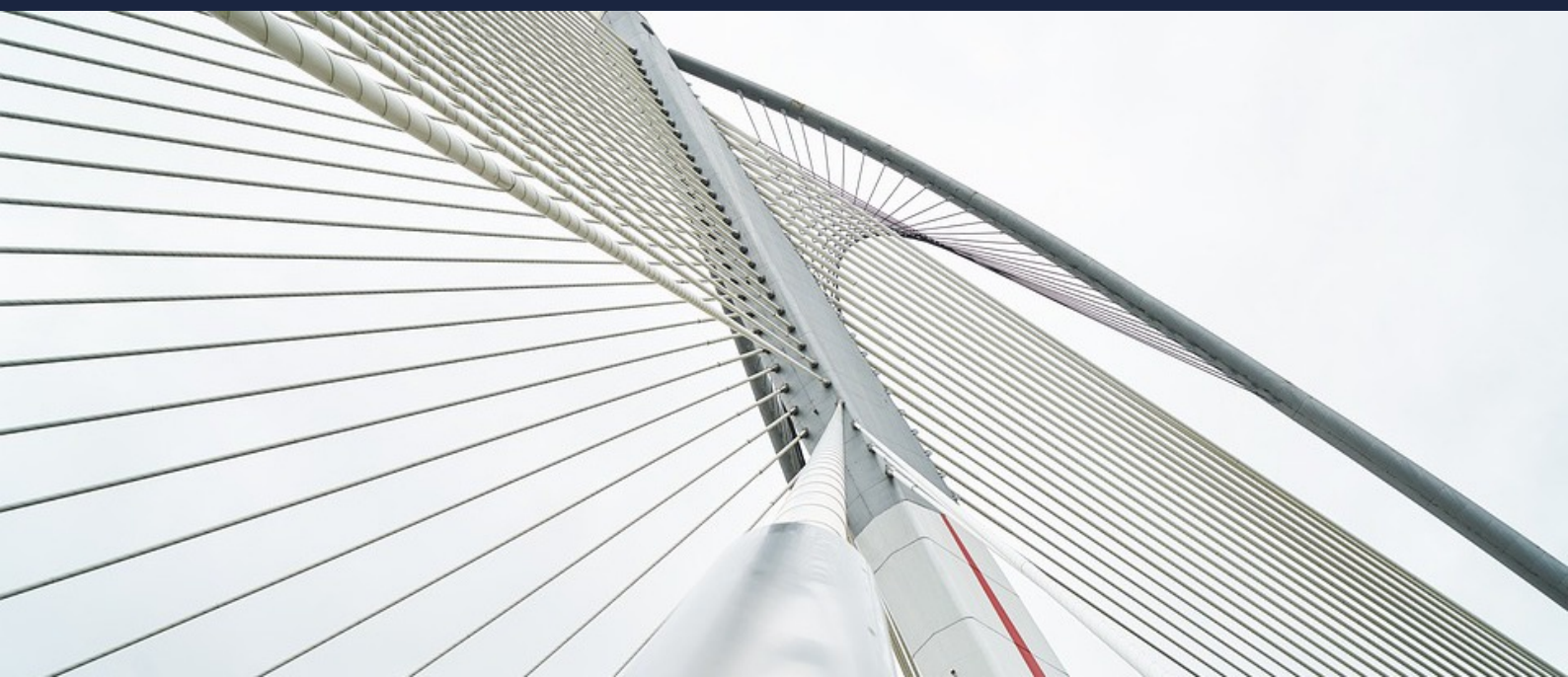
BEYOND HUMAN DOCTRINE

Managing a Team of Humans and Agents — The First Genuinely New Leadership Problem in Thirty Years

On Deskilling, Morale, and the Quiet Redesign of What It Means to Lead

Giovanni Leonardi

November 2025



Contents

1. Executive Summary	3
2. The Allocation Decision Nobody Was Trained For	3
3. A Doctrine Built for Humans	4
4. The Hollowing Problem	4
5. The Review Paradox	5
6. The Morale Problem That Is Not Change Management	6
7. What New Doctrine Might Require	7
8. The Doctrine That Does Not Yet Exist	8

Managing a Team of Humans and Agents — The First Genuinely New Leadership Problem in Thirty Years

Leadership · No. 124
November 2025

“The leadership problem is not how to manage the agents. It is how to lead the humans in the presence of the agents.”

1. Executive Summary

The deployment of autonomous AI agents into working teams has surfaced a leadership challenge that thirty years of management thinking — built entirely for the motivation, development, and fair treatment of humans — has no chapter for. This essay explores what happens when a team is no longer composed entirely of people: the quiet hollowing of developmental roles when agents absorb the work that juniors once learned from; the unsettling inversion of review when the person checking the output is slower than the system that produced it; and the particular strain on morale that comes not from reorganisation or cost-cutting but from watching a craft lose its economic centre of gravity in real time. These are not problems of adoption, integration, or change management. They are problems of leadership — and we are improvising the doctrine as we go.

2. The Allocation Decision Nobody Was Trained For

The moment it became clear that something genuinely new was happening was not dramatic. It arrived during a routine sprint planning session. The team had eight analysts and access to three autonomous agents capable of producing first-draft regulatory impact assessments. The question on the table was simple enough: who takes the next batch of assessments?

The efficient answer was obvious. The agents could produce credible first drafts in a fraction of the time, freeing the analysts for higher-order review and interpretation. But the programme director paused — because four of those eight analysts were in their first two years, and regulatory impact assessment was precisely the work through which they were learning to think like regulators. Hand that work to the agents, and the juniors would move to review roles they were not yet equipped to perform. Keep it with the juniors, and the team would be measurably slower for no reason the client would accept.

This was not a resourcing decision. It was not a technology adoption question. It was a question about what kind of team this would be in two years — and nothing in three decades of leadership training had prepared anyone in the room to answer it.

3. A Doctrine Built for Humans

We have not lacked for leadership thinking. The past thirty years have produced sophisticated frameworks for motivation, engagement, trust, psychological safety, delegation, coaching, and the fair distribution of opportunity. These frameworks share a foundational assumption so obvious it was never stated: that every member of the team is a human being, with needs for growth, recognition, autonomy, and meaning.

The assumption was not wrong. It was simply exhaustive — it covered the whole team, because the whole team was human. Now it does not.

An autonomous agent requires no motivation. It has no career trajectory. It does not experience trust or its absence. It is indifferent to whether its contribution is acknowledged, and it will not leave for a competitor offering more interesting work. It has no stake in fairness, no need for psychological safety, no developmental arc that the leader is responsible for shaping.

This sounds like simplification — one constituency in the team that demands nothing. In practice, it is the opposite. The agent's indifference to everything that leadership doctrine was built to manage does not eliminate those concerns; it concentrates them on the humans who remain. And it does so in ways the existing doctrine never anticipated, because the humans are no longer working alongside other humans. They are working alongside systems that are faster, cheaper, tireless, and — in a growing number of task domains — producing output that is difficult to distinguish from competent professional work.

The leadership problem is not how to manage the agents. It is how to lead the humans in the presence of the agents.

4. The Hollowing Problem

The most consequential decision a leader of a mixed team makes — and the one least discussed — is work allocation. Not the logistics of it, but the developmental implications.

Every profession has work that is simultaneously productive and educational. The junior associate who drafts the contract is producing a deliverable and learning to think in the structures of commercial law. The graduate analyst who builds the financial model is answering a question and developing the instinct for where numbers go wrong. The apprentice engineer who writes the test cases is contributing to quality assurance and learning how the system actually behaves under pressure.

This dual nature of work — productive and developmental — is so embedded in professional life that we rarely name it. It is the mechanism through which expertise is transmitted across generations of practitioners. When an agent takes over the drafting,

the modelling, or the test-writing, the productive function is preserved. The developmental function is not.

The early evidence is specific and observable. In teams where agents handle the bulk of first-draft production, junior team members move to review and refinement roles significantly earlier than their experience would traditionally warrant. The work they review is competent — sometimes more than competent — but they are reviewing it without having internalised the reasoning that produces it. They can spot a formatting error or a logical non sequitur, but they cannot yet feel the structural flaw that an experienced practitioner would catch in the rhythm of a paragraph or the shape of an argument.

Consider one pattern we have watched repeat across several programmes. A junior data analyst, eighteen months into the role, is asked to validate an agent-produced segmentation analysis. She can check whether the methodology statement matches the outputs. She can verify that the numbers add up. What she cannot yet judge is whether the segmentation boundaries are meaningful — because she has never built one from raw data, never fought with ambiguous categories, never discovered through three failed attempts why certain cuts reveal patterns and others obscure them. The review becomes procedural: a checklist, competently executed, rather than an act of professional judgement.

This is not a training problem in the conventional sense. The organisation has not failed to provide courses, mentoring, or structured development. It has inadvertently removed the substrate on which expertise grows: the slow, sometimes inefficient, occasionally wrong work of doing the thing yourself.

The most important thing a leader allocates in a mixed team is not work — it is learning. Every task assigned to an agent is a task a human will not learn from, and no amount of structured review can replace the understanding that comes from production.

5. The Review Paradox

If the hollowing problem is the quiet one, the review paradox is the one that surfaces in meetings. It arrives the first time a senior practitioner is asked to review an agent's output and discovers that the review takes longer than the production did.

The traditional relationship between production and review rested on an implicit hierarchy of speed: the person who produced the work was slower and less experienced; the person who reviewed it was faster and more experienced. Review was efficient precisely because expertise compressed the time required to assess quality. A senior lawyer could read a junior's draft contract and identify the three structural problems in less time than it took the junior to write the first clause.

When the producer is an agent, this relationship inverts. The agent produces a complete draft in minutes. The reviewer — however senior — must still read, assess, and reason about it at human speed. The ratio of production time to review time shifts from something like five-to-one in favour of production to something closer to one-to-three in

favour of review. The bottleneck moves from creation to assessment.

This inversion has implications that go well beyond scheduling. It raises a question that organisations have been remarkably reluctant to confront: what does "review" actually mean when the reviewer cannot match the producer's speed or, increasingly, its breadth of reference?

In practice, three responses are emerging. The first is genuine expert review — slow, thorough, drawing on deep domain knowledge to interrogate not just what the output says but what it omits, misweights, or assumes. This works, but it does not scale; there are not enough experts, and their time is the scarcest resource the organisation has. The second is procedural review — checking against a rubric or a defined set of quality criteria. This scales, but it catches only the errors the rubric anticipated, which are rarely the interesting ones. The third — and this is the one no one wants to name — is performative review: a senior person spending twenty minutes with a document, making a small number of changes to demonstrate engagement, and approving it on a confidence that rests more on the general reliability of the system than on the specific adequacy of this particular output.

The third pattern is not negligence. It is a rational adaptation to an impossible volume of work. But it hollows out the meaning of professional accountability, and it creates a category of organisational risk that existing governance frameworks do not capture: the risk that no one has truly assessed the work, even though every approval box has been ticked.

6. The Morale Problem That Is Not Change Management

Organisations have a well-developed vocabulary for managing the human impact of change. When roles are restructured, when teams are merged, when technology automates a manual process — the playbook exists. Communicate early. Acknowledge the disruption. Offer retraining. Help people see their place in the new structure.

The morale challenge of mixed human-agent teams resists this vocabulary, because what is happening is not restructuring, automation, or process improvement in any form the playbook recognises. It is something more intimate and more difficult to name: the experience of watching your professional craft lose its distinctiveness in real time, while you are still practising it.

A content strategist who spent a decade developing the ability to translate complex propositions into clear, engaging narrative now works alongside an agent that produces serviceable first drafts in seconds. The strategist's skill has not been made redundant — the agent's drafts require shaping, judgement, and occasionally wholesale rethinking. But the gap between what the agent produces unaided and what the strategist adds has narrowed enough to be visible, and it is narrowing still. The strategist knows this. The organisation knows this. Neither has a comfortable way to talk about it.

This is not the anxiety of automation, where a role disappears and a person is displaced. It is the anxiety of compression — watching the economic distance between expert and adequate shrink, week by week, while remaining in the role. The strategist is still needed. The question that hangs in the room, unasked but felt, is for how long — and whether the

answer is measured in years or in model generations.

"The hardest leadership conversations in mixed teams are not about efficiency, adoption, or integration. They are about what it means to build a career in a craft whose boundaries are being redrawn by a system that has no awareness it is doing so."

Leaders who reach for the standard change management toolkit — reskilling programmes, positive reframing, "focus on what only humans can do" — find that it lands differently here. Reskilling implies that the current skill is becoming obsolete, which is precisely the anxiety being managed. Positive reframing — "you are freed to do higher-value work" — is true but insufficient, because it elides the loss. And the instruction to "focus on what only humans can do" invites an obvious and uncomfortable follow-up: what, specifically, is that, and will it still be true in eighteen months?

The honest answer is that we do not know. And the failure of nerve — the unwillingness to say that plainly, to sit with the uncertainty alongside the team rather than papering over it with reassurance — is the leadership failure we are seeing most often in the organisations furthest along this curve.

7. What New Doctrine Might Require

We are improvising. The leaders who are handling mixed teams most thoughtfully are not following a framework; they are making it up, reflecting on what works, and adjusting. Several patterns are emerging from the improvisation, though none is yet settled enough to call a method.

The first is deliberate developmental allocation — explicitly reserving categories of work for humans, not because agents cannot do it, but because humans need to. This means accepting a measurable efficiency cost and being willing to defend it to stakeholders who see only the throughput numbers. In one consulting programme, the leadership team ring-fenced all first-draft client recommendations for human analysts, even though agent-produced drafts were faster and often comparable in quality. The rationale was straightforward: constructing a recommendation from evidence was how analysts learned to think about client problems, and that capability was more valuable to the firm over five years than the time saved on any individual deliverable.

The second is transparent uncertainty — being candid with the team about what the leader does and does not know about the trajectory of these changes. This is harder than it sounds, because leadership culture rewards confidence and penalises doubt. But in mixed teams, false certainty is corrosive. People can see the gap between the reassuring message and the reality of their working week, and the dissonance costs more trust than the truth would.

The third is redefining review as a first-class professional skill — investing in it as seriously as organisations once invested in production capability. If review is where human judgement now concentrates its value, then it should be trained, practised, assessed, and recognised accordingly. This means moving beyond procedural checklists to develop the capacity for what one programme manager described as "adversarial reading" — the ability to interrogate a plausible document for the specific ways in which it might be subtly wrong, underpowered, or misleading in its framing.

The fourth — and this is the one that most distinguishes the leaders who are navigating this well from those who are not — is moral seriousness about the team's composition. The decision about what proportion of the team's capacity comes from humans and what proportion from agents is not a procurement decision or a headcount optimisation. It is a decision about what kind of organisation this is, what it owes to the people who work in it, and what it is willing to sacrifice in short-term efficiency for long-term capability and human flourishing.

The leaders who treat it as a procurement decision end up with efficient, brittle teams whose human members are disengaged and whose institutional knowledge is shallow. The leaders who treat it as a question with moral weight end up with teams that are slower on paper but more resilient, more capable of genuine judgement, and — not incidentally — more able to attract and retain the experienced practitioners whose expertise makes agent output trustworthy in the first place.

8. The Doctrine That Does Not Yet Exist

It would be satisfying to close with a framework — a tidy set of principles that codifies what mixed-team leadership requires. But the honest position is that the doctrine does not yet exist, and pretending otherwise would violate the very transparency this essay argues for.

What we have instead is a set of questions that did not exist three years ago, and that every leader of a mixed team is now encountering whether or not they have found language for them:

1. What work in this team exists for its developmental value, and what happens to the team's future capability when that work is handed to a system that cannot learn from doing it?
2. What does meaningful review look like when the reviewer is slower and narrower than the producer — and how does the organisation distinguish genuine assessment from its performance?
3. How do we talk honestly about the compression of professional distinctiveness without either minimising the loss or catastrophising the trajectory?
4. What is this team's obligation to the long-term capability of its human members, and how does that obligation weigh against the short-term efficiencies available to it?

These are not technical questions, and they will not be resolved by better tooling, smoother integration, or more sophisticated agent capabilities. They are leadership questions — questions about what we owe to the people we lead, what kind of organisations we want to build, and what we are willing to pay in efficiency for the

answer.

We have been remarkably good, over thirty years, at developing doctrine for leading teams of humans. We have been less good at noticing when the foundational assumptions of that doctrine have shifted beneath us. They have shifted now — not with the drama of a disruption, but with the quiet insistence of a planning session where the obvious efficient answer is no longer obviously the right one. The question is whether we will develop the new thinking this moment demands, or whether we will keep applying the old playbook to a team it was never designed for, and wonder why the humans in it are struggling in ways that the playbook cannot name.