Accelerating Performance
Consulting

ARTICLE

Human Systems

Agency in an Agentic World

Colin Price and Sharon Toye

October 2026

Part of our series on leadership and organisational performance

On 4 July 2012, in an auditorium at CERN outside Geneva, a physicist put a graph on the screen and the room rose to its feet. Peter Higgs, who had predicted the particle nearly half a century earlier, was seen wiping his eyes. Thousands of people from scores of countries had built a 27-kilometre machine to catch something that exists for less than a billionth of a trillionth of a second. They caught it.

Now think of the pothole at the end of your road. It has been reported. It may have been inspected, ringed in spray paint, patched and reported again. It is probably still there.

We can discover the Higgs boson. We struggle to fix the road outside our house.

The gap between those two achievements cannot be explained by intelligence, money or goodwill alone. What matters is our ability to organise them into a system that acts with purpose. That ability must become the centre of leadership. **The human system is the leader's principal product.**

By human systems we mean the arrangements through which people organise collective effort: the purposes, relationships, rules, incentives and decisions that turn individual contributions into a shared result. Businesses, hospitals, governments, churches, sports teams and the alliances between them are all human systems. Their quality sets the ceiling on what we can achieve together.

This paper begins an attempt to forge a new science of organisation for the age of AI. We need to understand how human systems can combine growing machine capability with judgement, legitimate authority and a meaningful place for people. The stakes are both practical and profound: whether greater intelligence helps us accomplish worthwhile things together, and whether we remain able to direct and correct what it does.

We are privileged to spend our working lives inside other people's organisations. We recently spent a week at CERN and came away inspired. Colin has worked with the English National Health Service (NHS) for many years, and it inspires and frustrates us in equal measure. Its commitment to care is extraordinary; it changes lives every hour of every day. Yet in England, patients who are well enough to go home too often stay in hospital beds because the pieces needed to get them there cannot be brought together in time. Outside the hospital, the pothole deepens while responsibility passes from one institution to the next.

We usually describe leadership through the person at the top: their vision, judgement, authority and charisma. These matter. But even an exceptional leader cannot personally supply what is missing from hundreds of decisions made by people they will never meet. And with machines supplying analysis on demand, the claim to be the cleverest person in the room has never been weaker.

The centre of gravity of leadership shifts from intellect towards wisdom. Intelligence is becoming more widely available through specialists, partners and machines. The scarce capacity is turning it into coordinated, worthwhile action. The leader must arrange work so that people and machines can exercise judgement and deliver dependable results together.

For a chief executive, that design has to earn its place in the value-creation plan: revenue up, cost down, cash in, risk out. Under private-equity ownership, the investment period gives those choices a particular urgency. The business must deliver better results soon enough to matter, while building capability that a future owner can rely on. An exhausted organisation with a flattering spreadsheet is a poor inheritance.

A robust system keeps producing worthwhile results as circumstances change, notices when it is failing and can recover. Pay and authority help, but cannot by themselves create shared understanding, reconcile conflicting obligations or make an institution learn.

This changes the question a leader asks. “What should I decide?” still matters. Beside it sits a harder question: “What must be true here for people to make good decisions and act on them when I am not in the room?” A diary full of rescue missions may be a warning about the system we have built.

Agency is the capacity to influence what happens, and the title of this paper gives it a double meaning. Machines are acquiring agency: AI agents can now plan, decide and act with growing independence. People need agency too, and it does not arrive by default. **Agency in an agentic world means increasing people’s ability to shape outcomes as machines acquire more ability to act.** An organisation can become technically more powerful while its people become less able to question, learn or make anything happen.

To the person doing the work, agency comes down to three questions:

Understanding: **can I see how my work helps**? People need to know how their contribution affects the result for somebody else.

Means: **can I get what I need to act**? That includes authority, resources, information and dependable commitments from others.

Voice: **can I raise a problem and get a response**? A concern needs a route to someone able and obliged to act, without punishing the person who raised it.

We will return to these three questions throughout the paper. Where one goes unanswered, capable people can become spectators of outcomes they could have changed. The frustration is often greatest among those who care most: they can see what needs doing and still cannot get it done.

Figure 1 turns these questions into a practical way to examine the system. The point is to find what prevents useful action, then change it.

The three questions of agency

Element	Ask the person doing the work	Possible indicators
Understanding	Can I see how my work helps? Connect the daily task to the result for a customer, patient or citizen.	People can trace their contribution to the end result; priorities are clear at the front line.
Means	Can I get what I need to act? Join authority to information, resources and dependable commitments.	Gaps between responsibility and decision rights; commitments kept; time to resolve exceptions.
Voice	Can I raise a problem and get a response? A serious concern reaches someone able and obliged to act, without punishing the person who raises it.	Time from concern to visible response; near-misses reported before harm occurs.

Figure 1. Three questions of agency and possible indicators.

For companies and their owners, the test is commercial. Does the system help win and serve customers profitably, turn earnings into cash and prevent avoidable loss? The human questions

matter because they shape those results. A team that cannot get a decision, secure a handover or challenge a bad assumption cannot deliver the plan reliably.

> **The human system is the leader's principal product.**

Our capacity to succeed is extraordinary

Start with what goes right, because it goes very right indeed.

Fabiola Gianotti, CERN's Director-General from 2016 to 2025, describes a community of more than 17,500 people of over 110 nationalities in her 2026 account for IMD. Most work for universities and laboratories elsewhere. No common payroll holds them together. Shared scientific ambition, confidence in colleagues and recognition of individual contributions carry much of the weight that hierarchy carries elsewhere.

Gianotti's 2026 account for IMD is revealing about what that demands of a leader. "If the youngest student has the right idea, we follow it," she says, yet someone still has to decide. As Director-General she had to "resist the temptation to go into every technical detail" and lead the institution as a whole. On projects that last decades, she argues, celebrating intermediate milestones helps people see why today's work matters.

This is no leaderless utopia. CERN's Council approves programmes and budgets and appoints the Director-General. ATLAS elects its own leadership; a new team began a two-year term in March 2025. Scientific authority, executive responsibility and public accountability all operate at once. The achievement lies in making them compatible.

Mark Robinson's historical study of CERN adds the politics. Contributions of equipment and expertise tied international ambition to benefits back home. Collective authorship sat alongside recognition of individual work. It shows how common purpose works through arrangements for interests, credit and legitimate authority. A stirring speech would have had rather a lot to do on its own.

CERN makes the three elements visible. People understand the shared scientific endeavour; facilities and commitments give them the means to contribute; and an idea can be heard regardless of its author's seniority. Its leaders sustain those connections, making possible contributions they could never have specified in advance.

NASA offers another example. Thousands of people from 14 countries built the James Webb Space Telescope, folded it like origami inside a rocket and unfolded it a million miles from Earth. In January 2026, NASA reported AI-planned drives by the Perseverance rover on Mars, with engineers checking every command against a virtual replica before transmission. Machine autonomy operated inside an accountable human mission.

Smallpox was declared eradicated in 1980 and the cattle plague rinderpest in 2011. Vaccines mattered; so did surveillance, public cooperation, reliable delivery and people prepared to finish the work long after its glamour had faded. Polio, still unfinished, shows how hard the last mile can be.

Steven Pinker argues, in *The Better Angels of Our Nature* and *Enlightenment Now*, that reason, science and concern for human welfare have produced measurable progress, including long-run declines in several forms of violence. Whatever one makes of the details, the argument places an

obligation on leaders: find out what has improved, and protect the arrangements that made it possible.

Progress is not guaranteed. Uppsala's conflict researchers recorded 65 state-based armed conflicts in 2025, the most in a series that begins in 1946, and around 244,600 deaths from organised violence. Those figures do not settle the long-run debate about violence. They do show how quickly things can get worse. Progress is possible, uneven and reversible. Institutions have to earn it again in every generation.

So our optimism comes with a condition. Extraordinary cooperation is possible. It lasts only where institutions keep renewing the commitments, skills and trust on which it rests. We should study how they do that, and then bring the same seriousness to the ordinary obligations we keep failing to meet.

Ordinary failure has a structure

Most people do not arrive at work hoping to make somebody else's day more difficult. Yet conscientious people can each do their job well while the shared result fails. Responsibility has been divided more precisely than the means of fulfilling it have been connected. Everyone has agency over a task. Nobody has enough agency over the outcome.

Picture a discharge from an English hospital. An 84-year-old woman is declared medically fit to go home on a Tuesday morning. Her medication is not ready until Thursday. Transport is booked for Thursday afternoon, but the care package at home cannot start until the following Monday.

The nurse wants her home safely. The pharmacist wants her medicines right. The care team needs enough people to help her. She would simply like to go home, put the kettle on and sleep in her own bed.

She spends six more nights in a bed she no longer needs, growing weaker, while someone in A&E waits for it. Nobody has the authority and means to secure a complete, safe transition. If community care is unavailable, that requires a decision about capacity and funding. An instruction to "collaborate better" will not create it.

Keep this woman in mind. She is not waiting for a more inspiring mission statement. She is waiting for medicines, transport and care to be ready together. **Understanding** means seeing the whole journey home. **Means** means being able to secure the commitments it requires. **Voice** means being able to flag a missing piece while someone can still do something about it.

Consider a manufacturer. Sales wins an attractive order and earns its bonus. To meet the promised date, operations interrupts a production run and buys emergency freight. A specification dispute delays the invoice. The customer has bought a result that the business has not arranged to deliver profitably. Revenue appears; margin leaks and cash waits.

The salesperson wanted to win. The production manager wanted to keep a promise. Finance wanted an invoice the customer would pay. Each had a reasonable concern. The company needed them to make a dependable commitment together, before the promise left the building. We will return to this order alongside the woman waiting to go home.

For the CEO, the design questions are immediate. Who can commit the business to a delivery date? Who checks the economics of an exception? Who can resolve a dispute that is holding up cash? The hospital and the manufacturer expose the same gap: people understand their own tasks, but lack the means and voice to secure the whole result.

Here is the uncomfortable truth about the pothole. It is not an accident. It is an output.

In January 2025, the UK Parliament's Public Accounts Committee examined local roads in England and found short-term, fragmented funding, unclear accountability and weak knowledge of what spending achieved. Twelve separate funding pots had been used over the preceding decade. Annual allocations pushed councils towards reactive patching rather than better-value prevention. The hole in the road is the physical shape of the way public decisions are organised.

Flagship science and routine services do face different problems. A bounded mission with concentrated funding is easier to protect than maintenance spread across thousands of places and competing needs. We are not claiming that one operating model fits both. We are asking why reliable everyday achievement receives so little of the institutional seriousness we lavish on the exceptional.

The National Audit Office reported that 14.2 per cent of hospital stays in England involved delayed discharge in 2025–26.

The consequences can be far graver. The 2024 final report of the inquiry into the 2017 Grenfell Tower fire in London concluded that all 72 deaths were avoidable. It traced failures across government, industry and the organisations responsible for the building. Residents had raised safety fears for years. Their knowledge had no dependable route to action. Voice existed; its power to change the outcome did not.

Colin's long experience of the NHS has made us wary of treating reorganisation as improvement. The 2012 Lansley reforms brought the language of competition and "creative destruction" to a service whose success depends on interdependence. In our judgement, too much leadership attention, organisational energy and public money went first into implementing those arrangements and then into unpicking them. We certainly got the destruction. The creation is rather harder to find. The purpose of better care endured while the means of achieving it were repeatedly torn up.

That is not the explanation for every poor NHS outcome. COVID, rising demand, management and execution all mattered. But a noble purpose does not make an operating model an innocent bystander. If authority, incentives and responsibilities pull in different directions, dedicated people can work harder and harder while the collective result gets worse. **Good intentions need a system that makes them effective.**

Business has its own version. Sheffield's Audit Reform Lab found no going-concern warning in the final audit reports of three-quarters of the 250 large UK listed companies it studied that collapsed between 2010 and 2022. It raises a useful question: what, exactly, does a system of assurance assure?

Complexity changes the leadership task

AI expands what companies might accomplish while geopolitical fragmentation makes supply, investment and cooperation harder. Government debt and competing public demands affect customers and markets. Climate change shifts costs across businesses, borders and generations. For a chief executive, these pressures arrive as choices about customers, capacity, capital and risk.

The pressures reinforce one another. Deferring maintenance can flatter today's earnings and make tomorrow's disruption worse. A supply shock changes the economics of a promised delivery. Technology can widen options or deepen dependence on a few providers. The system responds to every intervention, sometimes with an enthusiasm its designer did not intend.

Some conflicts are not coordination failures. The suffering in Gaza and the war involving Iran remind us that actors may deliberately pursue incompatible ends. Power and coercion matter. Human systems can be effective at achieving objectives that others experience as catastrophe. The paper's argument concerns worthwhile results, not efficiency at any price.

Climate presents a related difficulty: UNEP's 2025 assessment shows a continuing gap between policy trajectories and agreed ambitions. Science alone cannot settle who pays or how commitments are enforced. The IMF's April 2026 assessment similarly describes debt and spending pressures narrowing governments' choices. Knowing what is needed does not provide the means to deliver it.

As analytical capability becomes more widely available, the demands on leadership shift from intellect towards wisdom. Leaders still need to understand the work and interrogate the evidence. Increasingly, however, their contribution lies in judging which problems deserve attention, which trade-offs are acceptable, whose interests must be heard and when a confident answer deserves another question. Greater capability makes these choices more consequential.

> **The saddest aspect of life right now is that science gathers knowledge faster than society gathers wisdom.**
>
> Isaac Asimov, 1988

Asimov named the gap between what we can know and what we have the wisdom to do. Closing that gap is exactly the purpose of this paper. We want to begin forging a new science of organisation: a better account of how people and machines can turn knowledge into worthwhile collective action, and how we can learn which arrangements help them do it.

The era of treating hierarchies, management by objectives, complex incentive structures, committees and subcommittees as a sufficient answer to organisation is over. These arrangements may still perform specific jobs. They have to earn their place in a system capable of adapting, learning and exercising judgement. A committee does not acquire wisdom by having a subcommittee.

We will have to feel our way towards that system through explanation, experiment and correction. We already have valuable knowledge about behaviour, incentives, power and cooperation. The task is to bring it together around the work, test new combinations and preserve what proves useful. AI makes that inquiry urgent by expanding both what an organisation can achieve and the damage a poor design can do.

Wisdom begins partly with recognising our own limits. A leader cannot personally supply all the judgement an organisation needs. The task is to build a human system in which different experiences inform decisions, assumptions can be challenged, consequences become visible and people learn from what happens. **The leader's wisdom is expressed partly in the quality of judgement the system makes possible.** We have to develop that design with the people who do and receive the work. They know where the official account parts company with Tuesday afternoon.

Andy Jassy, Amazon's chief executive, offers a useful example in his 2024 shareholder letter, published in 2025. These annual letters explain a company's performance, priorities and prospects

to its owners. Jassy describes how a principle about being right had sometimes encouraged people to defend their own ideas. His test is whether leaders seek perspectives that might disprove their beliefs and change their minds when new evidence warrants it. Our reading is simple: wisdom includes making it safe for someone else to improve your answer. Seniority should not confer immunity from a better idea.

We do not need to wait for artificial general intelligence to see the stakes. Tools are entering institutions that already struggle to turn intention into achievement. Poorly organised, AI may leave its promise unfulfilled: faster analysis waiting for the same slow decision. Poorly governed, it may run ahead of our capacity to understand, challenge or stop what it does. A defective objective can become a great deal more consequential when pursued at machine speed. Every grant of machine agency is a decision about human agency: who sets the goal, who can see what is happening and who can stop it.

These pressures do more than add items to a leader's desk. They make consequences harder to isolate. A staffing decision changes service capacity, which changes queues, which changes demand somewhere else. A new target changes behaviour, which changes the information reaching management. A new technology changes who holds expertise and who depends on whom. The system responds to the intervention.

That is why organisations are not machines. A machine's parts do not care what the designer intended. Nobody asks a conveyor belt whether the new strategy has damaged its career prospects. People interpret rules, spot opportunities, protect their identities and adjust to one another. Incentives get gamed. Authority hides error. An efficient process turns brittle when demand varies. Designing a human system means attending to these interactions, especially the ones that make a well-intentioned intervention backfire.

The leadership task is to create enough common purpose, constraint and coordination for distributed intelligence to become collective action, while leaving people room to respond to what nobody foresaw. Understanding, means and voice have to survive the arrival of the clever new colleague, including the one that never sleeps.

Incentives do not settle what people will do

In 1902, the French colonial authorities in Hanoi had a rat problem. They offered a bounty: one cent for every rat tail handed in. The tails poured in. Before long, so did reports of tailless rats running through the city, alive and still breeding. The catchers had worked out that a living rat was worth more than a dead one. Health inspectors later found rat farms on the city's outskirts. The historian Michael Vann, who uncovered the records in the French colonial archives, has made the episode a classic. The scheme delivered exactly what it paid for.

People respond to pay and promotion, but also to status, security, recognition and the avoidance of blame. In our illustrative manufacturer, the salesperson may fear losing a customer to a rival; the production manager may fear being labelled unhelpful. Their choices can make sense locally and still damage the business. Understanding that logic gives a CEO somewhere useful to intervene; it does not excuse the loss.

Most of us can spot a badly designed incentive. It is a little harder when it is paying our bonus.

A measure is rarely the whole achievement. Reward a department for cutting cost and it may push a larger cost onto the customer. Reward cases closed and the hard cases become somebody else's problem. People can satisfy the contract while defeating its purpose.

Bengt Holmström and Paul Milgrom's work on incentives across multiple tasks explains why. Strong rewards for measured output can divert attention from unmeasured quality. The design of the job and the balance of rewards have to be considered together.

Wells Fargo's sales culture urged staff to sell each customer eight products: "eight is great". In 2020 the bank agreed to pay $3 billion to resolve criminal and civil investigations. The US Department of Justice described millions of accounts and products opened without consent or under false pretences between 2002 and 2016, under pressure from unrealistic sales goals. The scale demands scrutiny of the system as well as the individuals. People's capacity to act had been directed at the wrong result.

Volkswagen's emissions scandal, exposed in 2015, took that distortion into engineering. Software recognised emissions tests and altered the car's behaviour; on-road nitrogen-oxide emissions could reach 40 times the US limit. The US case covered around 590,000 vehicles. This was deliberate deception: ingenuity organised around defeating a test. A human system can coordinate effort brilliantly while betraying the outcome it claims to serve.

Meaning and professional standards can help people see beyond the metric. A teacher committed to a child's development, or an engineer committed to safety, has a reason to protect what the target overlooks. But this is not a simple story in which money is bad and purpose is good. Samuel Bowles and Sandra Polania-Reyes, reviewing fifty experiments, describe incentives that sometimes weaken social motives and sometimes strengthen them. Much depends on what people understand the arrangement to mean.

So the answer is not a more elaborate bonus scheme that finally specifies every desirable behaviour. Some of the most valuable contributions cannot be contracted for: telling the boss an uncomfortable fact, helping another team, protecting quality when nobody is watching. Rewards, professional standards and the meaning people find in their work must support one another. Where they conflict, the system teaches people which one really counts.

Return to the order. If sales is rewarded for booked revenue while operations absorbs the cost of promises it did not approve, the argument at the weekly meeting is partly designed in. Match the reward to contribution after avoidable delivery costs, give someone authority over exceptions and make the economics visible early. The target, decision right and information have to work together. A bonus should not require the finance director to develop a sense of humour.

This gives sociology a practical role. George Herbert Mead argued that the self develops through social relationships: we learn who we are partly through the responses of others. Émile Durkheim showed how specialisation could create solidarity through mutual dependence. Applied to work, both suggest that people's willingness to contribute depends on whether their distinct contribution is recognised and has a place in a shared endeavour.

Institutions produce identities as well as outputs. A specialist reduced to feeding answers to a machine may experience the same work very differently from one whose judgement is sought and developed. A newcomer who can see a future in colleagues' careers has something no purpose statement can supply. "Come and try this; I'll stay with you" can be a small act of leadership with a long afterlife. It gives someone a real task, the confidence to attempt it and help when the first attempt goes wrong.

Expertise can lose distinction when a machine reproduces the analysis. Remote colleagues may know one another mainly through scheduled calls. Repeated restructuring can make a career feel less like progress than a series of fresh starts. People need a place to contribute and a future they can recognise, especially when they are new to work.

The data contain warning signs, though not one simple story. Gallup's measure of US employee engagement fell from 36 per cent in 2020 to 31 per cent in 2025. Its global figures are more striking. Engagement among managers fell from 31 per cent in 2022 to 22 per cent in 2025, while among other employees it barely moved, from 20 to 19 per cent. The people we expect to energise everyone else appear to be the ones running out of energy. Yet in Gallup's selected best-practice organisations, 79 per cent of managers were engaged. That comparison is neither representative nor proof of cause, but it does suggest that decline is not destiny. Surveys of pride and meaning, covering different populations, point in different directions. Declaring a seismic collapse would outrun the evidence. Ignoring the pressure on managers would be a mistake.

Nor does ordering everyone back to the office settle the matter. A randomised trial at Trip.com found that the hybrid arrangement it studied improved retention and satisfaction without damaging measured performance. The real question is whether working arrangements sustain relationships, contribution and development. Sharing a postcode can help. People still need a reason, and a way, to work together.

The evidence does not show that people have stopped caring about work. It does justify asking how the design of work supports belonging, development and contribution: the conditions in which agency grows. These are also potential sources of retention, cooperation and capability, whose economic value should be tested rather than assumed.

Meaning supplies energy but no moral guarantee

People see the face of Jesus in burnt toast. Our appetite for meaning is formidable, strong enough to find a pattern almost anywhere. It offers no guarantee that the meaning we find is sound.

As a young consultant, Colin spent a day with the top team of a tobacco company. They seemed sincerely to believe in their purpose: "enabling adult choice". They were selling cancer sticks. Yet they had built a meaning that at least some people could embrace. It showed him how much our search for meaning can accommodate.

This matters because purpose has become an almost automatic prescription for leaders. A strong shared story can sustain difficult work. It can also excuse harm, silence dissent and make outsiders' suffering seem irrelevant. A group's conviction is evidence of commitment, not proof of moral worth.

Meaning becomes tangible when people see whom their work helps. Adam Grant and his colleagues ran a deceptively simple experiment in a university fundraising call centre, where callers spend their evenings asking alumni for money and are turned down again and again. One group of callers spent five minutes with a student whose scholarship their calls had helped to fund. A month later, those callers were spending more than twice as long on the phone, and the money they raised had risen from an average of $186 a week to $503.

Nothing about their pay, script or targets had changed.

They had met the person on the other end of their work.

The practical task is to make contribution visible. A product reaching its customer on time, a problem resolved without another call and a patient getting home safely all give people a reason to feel proud. In the manufacturer, production and finance colleagues should be able to see how their work protects the customer relationship, margin and cash. Recognition matters most when people also have the means to do the job well.

Patagonia carries purpose into ownership: its Purpose Trust holds the voting stock, while the Holdfast Collective holds the non-voting stock and receives distributions after reinvestment. In its 2025 financial year, the company reports repairing 174,799 products and supporting 27,000 hours of paid employee volunteering. It also acknowledges that roughly 85 per cent of its products lack an end-of-life solution. The limit belongs beside the achievement. Purpose needs that honesty.

A strong purpose therefore needs challenge built around it. Who benefits? Who bears the cost? Can outsiders contest the institution's account of itself? Ethics asks whether the ends and means are defensible. Integrity makes truthful reporting, accountability and consistency real under pressure. A leader who creates passionate belief without these protections may have designed a highly effective system for doing harm.

The UK's Infected Blood Inquiry exposes the cost of institutional self-protection. More than 30,000 people were infected with HIV or hepatitis C through blood and blood products, mainly in the 1970s and 1980s; the inquiry attributes more than 3,000 deaths to infected blood. Its 2024 report found disregard for patient safety and autonomy, defensiveness and a lack of candour. An institution devoted to care had protected itself at the expense of those it existed to serve.

Consulting offers no exemption. In December 2024, McKinsey agreed to pay $650 million to resolve US criminal and civil investigations into its work for Purdue Pharma on OxyContin. The US Department of Justice described a deliberate strategy to increase revenue by intensifying marketing to high-prescribing doctors, including those writing unsafe and medically unnecessary prescriptions, despite McKinsey's awareness of widespread abuse. The work helped keep opioid profits flowing at the expense of public health.

For leaders, the question is what counts as a successful engagement. Commercial results and sophisticated analysis cannot settle whether the work should have been undertaken at all.

Good intentions do not make a design innocent, and commercial success does not make every objective legitimate. Sometimes integrity means refusing revenue or accepting a cost. A science of human systems must preserve that moral judgement, rather than treating every efficiently achieved result as progress.

> **A shared purpose can mobilise a system. It cannot be the system's only judge.**

Authority must enable action and admit correction

In December 2020, Handforth Parish Council, a small local government body in England, held a planning meeting on Zoom. Before discussing local planning, members became embroiled in the more fundamental question of whether they were entitled to be having the meeting at all.

"You have no authority here, Jackie Weaver," the chairman told the woman brought in to help proceedings. Weaver moved him into the virtual waiting room. The chairman had the title. Weaver had the button. Appeals to "read the standing orders" did little to restore harmony.

There was even a dispute about the chairman identifying himself on screen as the council's clerk. Weaver observed that anyone could call themselves what they liked: "Please refer to me as Britney Spears from now on." Local government had taken an unexpected turn towards show business.

A later investigation found that Weaver lacked the authority to mute and remove councillors, while acknowledging the difficult circumstances. Even the apparent resolution required another round of governance.

It is easy to laugh. Most leadership teams would prefer to do so before anyone examines their own meetings. Handforth had rules, offices, an agenda and functioning technology. Getting people to work together remained a separate undertaking.

The question behind the farce is serious: does authority help people act together, and can it hear a challenge? Where lives depend on the answer, a failure can be catastrophic.

Late in the afternoon of 27 March 1977, on a fog-bound runway in Tenerife, a KLM Boeing 747 began its take-off roll without clearance. The flight engineer, uneasy, asked whether the Pan Am jumbo might still be on the runway ahead: “Is he not clear, that Pan American?” The captain, one of the airline’s most senior pilots and its chief flying instructor, answered emphatically that it was. The engineer did not press. Seconds later the two aircraft collided. Five hundred and eighty-three people died, in the deadliest accident in aviation history.

The crucial doubt was in the cockpit. It could not climb the authority gradient. Aviation’s response was not merely to look for better captains. It redesigned the human system. Crew resource management, now standard across the industry, trains every crew member to voice concerns and every captain to hear them. Challenge became part of the job.

The NHS Staff Survey in England shows why this matters far beyond aviation. Its 2025 national results cover 751,088 respondents across 206 trusts. Nearly nine in ten, 87.8 per cent, feel their role makes a difference to patients. Only 54.2 per cent feel able to make improvements happen. The responses reveal a gap between the wish to contribute and the means to act. Figure 2 sets out that gap.

Commitment does not ensure the means to act

NHS staff survey item, 2025	Positive responses	What the question concerns
My role makes a difference to patients	87.8%	Significance of contribution
Able to suggest improvements	69.9%	Voice
Able to make improvements happen	54.2%	Practical influence
Involved in deciding changes	48.6%	Participation in decisions

Figure 2. NHS Staff Survey 2025, England trust results. Separate survey items; wording shortened and percentages rounded.

Our two illustrations make the gap concrete. The nurse sees the missing care package; the production manager sees a delivery promise that will destroy margin. Both may care deeply and raise the problem repeatedly. Another appeal for commitment asks them to supply what the organisation has failed to arrange. Useful authority and a timely answer are more valuable than another reminder to be accountable.

Freedom is not the same as agency. People can be left alone without usable authority or help. The two can look identical on an organisation chart and feel completely different in practice. That is why a new science of organisation cannot simply prescribe flatter structures, modular teams or the latest agile label. Each is a design to test against the work. Interdependence does not vanish when a management layer does; informal power may simply become harder to see and challenge. Agency needs means and voice, as well as room to act.

Robert Michels described how even democratic organisations can concentrate power in a small leadership group. In *Political Parties* (1911), his “iron law of oligarchy” traced the advantage created by specialist knowledge, control of information and coordination. Leaders can become harder to challenge or replace, until protecting their position displaces the institution’s purpose. Some football fans may look at Gianni Infantino, president of the world governing body FIFA, and wonder whether the most fiercely defended position is the one in the president’s office.

We treat this as a tendency to confront, not a destiny. Open information, independent oversight, credible succession and the ability to organise disagreement keep authority contestable. CERN’s elected experiment leadership offers one arrangement to study. In an agentic organisation, scrutiny must also reach whoever controls the models, data and permissions. A flat chart can conceal very concentrated power.

Elinor Ostrom offers a different corrective. In her studies of shared resources, from fisheries to irrigation systems, communities sustained cooperation through institutions in which they helped shape the rules, monitored behaviour and resolved disputes. Complex systems could combine several levels of governance. Participation did not remove the need for rules. It made them workable.

The central tension in system design is now visible. Collective action needs enough authority to make commitments dependable. Those affected need enough agency to challenge and improve what authority does. The task is to discover which decisions need a clear hierarchy, which belong near the work and which require cooperation across institutions. The familiar pyramid no longer provides the answer in advance. Neither does turning it upside down.

Wisdom includes recognising when a trade-off is real and when it reflects the way we have organised the work. Clear limits, visible decisions and dependable escalation can give people more freedom while making the result more controllable. Before deciding how much authority to surrender, ask what would make local judgement dependable. Some tensions need a choice; others need a better design.

NASA provides a painful counterpoint to its scientific achievements. On the eve of the Challenger launch in January 1986, engineers at contractor Morton Thiokol warned that the booster seals might fail in the forecast cold. During an internal discussion, the general manager asked the head of engineering to “take off his engineering hat and put on his management hat”, according to testimony recorded by the presidential commission. The recommendation against launch was reversed. Seven astronauts died when Challenger broke apart on 28 January. After Columbia’s destruction in 2003, the investigating board concluded that organisational culture had mattered as much as the foam that struck the wing. Technical brilliance did not ensure that a warning could change a decision.

Consider a familiar hypothetical moment. A manager knows a project is slipping. The boss has promised the date to the board. The manager hopes another week will produce a solution, so the report stays green. Nobody has decided to deceive the organisation for ever. They have decided to survive Thursday. A leader needs to understand that temptation and make an early warning safer, more useful and more consequential. “Thank you for telling me while we can still help” is a useful place to start.

Business is no different. Timo Vuori and Quy Huy's study of Nokia, published in *Administrative Science Quarterly*, traced how the company lost a smartphone battle it had once been leading. Senior managers, anxious about competitors, pressed hard and reacted aggressively to bad news. Middle managers, afraid of those reactions, kept unwelcome facts about the company's technology to themselves. The information Nokia needed to change course existed. It did not reach the people able to act on it.

Colin worked side by side with Stephen Elop, who became Nokia's chief executive in September 2010. Five months later, Elop wrote to every employee. "We too, are standing on a 'burning platform'," he told them, and "we poured gasoline on our own burning platform." He named the organisational causes without euphemism: "We're not collaborating internally," and "I believe we have lacked accountability and leadership to align and direct the company through these disruptive times." The memo did at a stroke what the system had failed to do over years: it put the unwelcome truth in front of everyone at once. It also shows the limit of a single act of candour. By the time the truth was spoken from the top, the room for manoeuvre had narrowed sharply, and in 2013 Nokia agreed to sell its handset business to Microsoft. A leader can announce the truth. Only a well-designed system delivers it early enough to act on.

The Post Office Horizon scandal shows the danger when an institution treats a computer's account as harder to challenge than a person's experience. Around 1,000 subpostmasters were prosecuted between 1999 and 2015 on the strength of faulty system data. The inquiry's July 2025 first volume found that people inside the Post Office and Fujitsu knew, or should have known, Horizon could produce errors; at least 13 people may have taken their own lives. Technology made an erroneous account harder for an individual to contest.

So system design is partly about the movement of truth. Can unwelcome information reach someone able to act? Does raising a problem damage a career? Can an affected person get an explanation and challenge a decision? Does evidence change the judgement, or is it selected to defend a judgement already made? Voice is the element of agency that institutions find easiest to destroy, and most expensive to lose.

AI introduces another way for doubt to disappear. Yuval Atsmon asks whether colleagues using similar AI systems may mistake shared assumptions for independent agreement. The meeting can contain six people and only one line of thought. Leaders should ask what evidence each person brought, which assumptions they tested and what would change their minds. A persuasive alternative generated by the same tool can help exploration, but does not by itself provide independent scrutiny.

> **A warning is useful only if someone can act on it.**

Political leadership adds a further requirement, because citizens can legitimately disagree about the ends themselves. Hannah Arendt distinguished power, people's capacity to act together, from violence, the capacity to compel. The first depends on willing cooperation; the second destroys it. Jan-Werner Müller's account of populism identifies the danger of claiming an exclusive right to represent "the people" and denying opponents' legitimacy. Public grievance deserves a hearing. Treating all dissent as disloyalty disables correction.

Leaders should make an argued judgement about what is right, explain the trade-offs and inspire constructive action. They must also accept constraint, scrutiny and replacement. Better expertise

does not entitle a government to choose everyone's values. More capable technology does not exempt a company from responsibility. A human system must preserve the means of questioning power as well as exercising it.

The leader becomes a designer of collective agency

The leader's distinctive contribution is to make incentives, meaning and authority work together as circumstances change. People need to understand the shared result, have the means to contribute and be able to correct the course. A leader who can take a holiday without the place seizing up has built something worth coming back to.

A CEO starts with an outcome worth achieving and follows it across the business. In the manufacturer, that means the customer's commitment, the economics of delivery and the arrival of cash. Work with the people responsible to align decisions, resources and exceptions, then test the effect on the customer and the accounts. The value-creation plan becomes a set of operating commitments that can actually be kept.

The surgical checklist made this practical. In 2007–08, a team working with the World Health Organization introduced a 19-item checklist in eight hospitals. Team members introduced themselves and voiced anticipated problems before the first incision. Across nearly 7,700 patients, deaths fell from 1.5 to 0.8 per cent and major complications from 11 to 7 per cent. This was a before-and-after study, not a randomised trial; later results have varied. The lesson is to test whether a checklist changes how people work together. Producing the form is the easy part.

The design reaches beyond reporting lines. It may mean changing a contract between organisations, a decision right inside a company, an incentive that makes cooperation costly, or a habit of silence that protects a failing investment. Often it means doing less: removing an approval, a competing priority or a target that causes everyone to optimise the wrong thing. People should feel the improvement in the work they can finish.

The most revealing test is what happens when the leader is absent and conditions change. Do people have the shared understanding and practical means to respond? Can they recognise a conflict, get help and correct an error? Does the system develop their judgement, or train them to wait? Heroics can rescue an episode. Good design makes good delivery less dependent on rescue. Constant rescue is tiring for the rescuer and a poor education for everybody else.

The leader is part of that design. Admitting uncertainty, accepting challenge and respecting decisions genuinely delegated to others show people how to behave. So do reversing a decision without explanation and punishing the bearer of bad news. People read the chief executive more closely than they read the values. The example has to travel through managers and influential colleagues: everyday group habits can support it or quietly undo it. Leaders must practise the behaviour they ask for, especially when the pressure rises.

The ten commitments below turn the argument into practices to test. They are a starting point for the science we are proposing, with each commitment tied to an action and a check on what happened. Keep the three questions in view: can people understand, act and be heard? A new practice earns its place by improving the whole result.

A manifesto for leaders who design systems

1. Design around the result that matters.

Choose a result the customer, patient or citizen would recognise. Trace the decisions and handovers it depends on, and give someone the authority to resolve conflicts across them. Amazon's practice of writing a customer press release and FAQ before building a product makes that starting point concrete.

Check in practice. Follow one recent order or service journey from beginning to end. Where did a successful departmental task leave the customer waiting? Connect the owner's responsibility to the resources and commitments needed to finish the job.

2. Connect authority to the means of delivery.

Choose a decision people are expected to own. Put the information, resources and reciprocal commitments behind it. Haier's micro-enterprises illustrate one attempt to bring decision rights closer to customers; the appropriate arrangement will differ by setting.

Check in practice. Ask the person responsible what they cannot secure without an escalation. Resolve a real gap, then see whether they can complete the work. Giving somebody responsibility without the means to fulfil it is an unusually elaborate way of arranging disappointment.

3. Make incentives support the whole outcome.

Examine a target that looks successful. Find the costs, risks and quality it leaves outside the calculation. Microsoft's move away from forced ranking in 2013 illustrates a change in what individual performance systems encourage; Wells Fargo shows what distorted sales pressure can cost.

Check in practice. Trace one rewarded result into the next team and the customer's experience. Has the work improved, or has the problem moved? Align the measure with usable authority, professional standards and legitimate limits.

4. Give contribution a visible human meaning.

Help people meet, hear from or otherwise understand those their work benefits. Medtronic's patient-and-employee gatherings, a tradition begun in 1960, show one way to make that connection real.

Recognise intermediate progress and share credit fairly. Remember the colleague who prevented the crisis as well as the one who performed magnificently in it.

Check in practice. Ask people to describe a person their work helped and an obstacle they could not remove. Act on the obstacle as well as celebrating the contribution. Keep the institution's purpose open to challenge by those who experience its harms.

5. Give truth a route to action.

Make it clear where a serious concern goes, who must respond and how the person raising it will be protected. Toyota's andon system lets workers signal a problem, summon help and, where necessary, stop the production line before a defect travels further. It connects a warning to the means of putting things right.

Check in practice. Take the last serious front-line concern. Trace who heard it, who acted and what changed. Tell the person who raised it what happened; they have a stake in the answer. A channel that collects worries without resolving them can become an unusually well-documented form of neglect.

6. Build for variation and recovery.

Rehearse a plausible disruption and establish who can change the plan. Keep justified reserve capacity and alternative routes. After the July 2024 CrowdStrike outage, Delta cancelled around 7,000 flights over five days; its chief executive put the cost at around half a billion dollars.

Check in practice. Remove one critical dependency in a tabletop exercise. Can the team sustain a safe service and recover? Test authority and information alongside the technical fallback, and judge the resilience gained against its cost.

7. Specify the bargain between people and agents.

Decide what belongs with employees, in the customer's product or with an external provider. Set the agent's purpose, limits, human review and rights to correct or stop it. Klarna claimed its assistant did the work of 700 customer-service agents in 2024; in 2025 its chief executive acknowledged quality problems and the company began hiring people again.

Check in practice. Follow one automated service from the customer's request to a resolved outcome. Include supervision, correction and failure costs. Make sure a person has both the duty and the means to intervene.

8. Preserve the means of learning.

Identify the experiences through which newcomers develop judgement, and preserve supervised practice and feedback as work changes. As Atsmon argues, expertise is also needed to recognise a plausible answer that is wrong. Automating the first attempt must leave people a way to learn how to judge the result. Shopify's plans for about 1,000 intern places in 2026, up from about 100, illustrate a deliberate investment alongside AI.

Check in practice. Ask a newcomer and their supervisor how they would test an AI-generated answer and where they would seek help. Name the supervised practice, feedback and progression that will develop that judgement. Expertise does not replenish itself out of gratitude.

9. Make ethics and power part of the design.

Set limits on what the institution may pursue and how. Give affected people an explanation, a route to challenge and redress. The congressional investigation of the Boeing 737 MAX crashes, which killed 346 people, exposed production pressure and compromised oversight.

Check in practice. Choose a decision where commercial pressure could conflict with a legitimate obligation. Who can challenge it independently? What work would you refuse? Purpose and rewards need these limits; declaring good intentions does not supply them.

10. Make improvement outlive the leader.

Turn learning into shared practice, renew skills and assets, and develop people who can challenge and replace you. Apple University, founded in 2008, illustrates an attempt to make an organisation's judgement teachable beyond an individual founder.

Check in practice. Step back from one recurring decision. Can the team make it, deliver the result and surface a mistake without your rescue? Resist the temptation to improve the result by taking the decision back. Preserve the capability that made the improvement possible, and check again when conditions change.

These commitments work through complementarity: each becomes more valuable when the others support it. Paul Milgrom and John Roberts explained how organisational practices can raise one another's returns. Erik Brynjolfsson and colleagues show why technology needs complementary organisational investment, which can depress measured productivity before benefits emerge. A good idea can fail because the arrangements it needs are missing.

The manufacturer makes the connection plain. Reward sales for revenue without matching delivery authority and the business may buy growth with margin. Give operations freedom without customer information and it may optimise the wrong work. Automate invoices without resolving specification disputes and the argument merely arrives faster. Figure 3 shows why the commitments belong together. The task is to build a workable combination around a result, not launch ten programmes and hope they get on.

> **In an agentic world, leaders must design for human agency as deliberately as they design for machine capability.**

Why the commitments must work together

Adopted alone	Risk to the whole result	What else the design needs
Incentives for results	People are paid for outcomes they cannot influence, inviting gaming or resignation.	Usable authority and ethical limits. Wells Fargo illustrates distorted sales pressure.
Local freedom	Problems can remain hidden until they are too large to fix.	A dependable route for unwelcome truth. Nokia illustrates the cost of blocked information.
Automation	Initial savings can conceal damage to quality, judgement and future expertise.	Human correction, learning and recovery. Klarna illustrates why quality must stay visible.
Shared meaning	Conviction can excuse harm and silence dissent.	Ethical limits, scrutiny and voice. The tobacco boardroom and Purdue Pharma expose the danger.
Lean efficiency	The expected case works; a shock overwhelms the remaining capacity.	Justified reserves and tested recovery. Delta's disruption illustrates the risk.

Figure 3. Complementarity: how the commitments support one another.

The bargain between people and agents

Imagine a customer whose refund is wrongly refused by an AI agent. The answer is immediate, polished and incorrect. The employee can see the mistake but cannot override it. The apology is beautifully phrased. The money is still missing. For the customer, an instant answer has become a long afternoon.

The machine has acted; the people have lost the means to put things right. The design should let the customer understand the decision, give an employee the means to correct it and make the complaint a useful signal. Six questions turn that requirement into something a leader can inspect.

What are we asking it to achieve? Set the objective and the constraints it must not cross. In the refund illustration, a low payout rate cannot stand in for a fair decision. Check who can alter those rules and whether changes are visible. This is goal-setting.

Can someone see what happened? An employee needs the relevant policy, information and record of the agent's action, including uncertainty where it can be assessed. A fluent explanation is not itself proof. This is observability.

Can a person put it right? Give the employee or affected customer a quick, affordable route to correction, escalation or human review. Specify who can override or stop the agent. This is contestation.

Will people still learn the work? Preserve supervised practice, feedback and exposure to difficult cases. Tomorrow's experts need more than a login and a collection of answers. Give new colleagues someone to ask when an answer looks plausible and feels wrong. This is learning.

Who must resolve the failure? Name the person accountable for the combined service and give them resources and authority to act. The model supplier cannot carry all the institution's obligations. This is accountability.

Did we solve the customer's problem at a defensible cost? Count repeat contacts, corrections, supervision and residual risk alongside the initial saving. A fast refusal followed by three calls has not completed the job. This is the whole-service test.

Klarna illustrates the risk of celebrating local cost savings; Shopify illustrates investment in future capability. The design has to account for both today's service and tomorrow's expertise.

The six tests protect the same three elements of agency. Visible goals and decisions support understanding. Usable authority, accountability and learning provide means. Correction and challenge give people voice. The technology changes; the obligation to make the whole service work remains.

Larry Fink, co-founder, chairman and chief executive of the investment manager BlackRock, raises a further question in his 2026 annual letter to investors: who will share in the wealth AI creates? Fink warns that the gains could concentrate among companies and investors positioned to capture them, and argues for broader ownership.

For us, this makes the bargain with people a practical design question. How will better performance improve pay, skills, career prospects or the quality of work? People invited to help build a more productive business deserve a credible account of their place in it. "You will become more efficient at making yourself unnecessary" is unlikely to fill the suggestion box.

People may welcome a more capable organisation while fearing the loss of expertise, status, relationships or security. Calling this resistance to change can miss what they are trying to protect. Give people time to understand the change, practise unfamiliar work and help shape their future.

Make the upside tangible, preserve what still matters and handle unavoidable losses honestly and fairly. Someone who has spent twenty years becoming useful deserves more than a cheerful invitation to start again.

> **Every grant of machine agency is also a decision about human agency.**

What this changes in practice

At the 2023 Qatar Grand Prix, McLaren set a Formula One world record by changing all four wheels on Lando Norris's car in 1.80 seconds. Some twenty people had defined roles, rehearsed handovers and signals for exceptions. The performance depended on preparation and coordination. Few organisations give their own handovers that care, though some do allow considerably more time for them.

The manufacturer needs that same care before accepting the order. Sales and operations must agree a feasible promise, the commercial owner must see its margin after exceptional costs, and finance must know what evidence the customer needs to pay. A named person must be able to resolve the exception. The desired result is a satisfied customer, a worthwhile margin and collected cash.

For the woman waiting to leave hospital, the equivalent is a shared plan across ward, pharmacy, transport and care provider. Those helping her need dependable commitments, an early warning when a piece is missing and access to a decision about capacity or funding. She and those supporting her should know what happens next, without having to become experts in how the service is organised. The evidence of success is a safe journey home.

For the local road, the design extends across funding, asset information, maintenance planning and accountability for later failure. The parliamentary findings suggest why a short-term repair target crowds out prevention. A better arrangement would give planned maintenance a credible funding horizon and test whether it reduces lifetime cost. That proposition needs evaluation. Announcing a longer plan does not prove that it works.

One idea becomes a working system

A governing idea becomes useful when it changes everyday decisions about money and power. A fine sentence cannot outvote the incentive scheme. CERN connects discovery to facilities, scientific challenge and shared credit. Patagonia connects environmental purpose to ownership and repair. The same test applies closer to home: how far does the declared purpose travel into what people can actually do?

Octopus Energy's customer-service teams provide one company example. Teams of around ten energy specialists look after defined groups of customers, with autonomy to resolve problems. Its Kraken platform gives specialists information and tools intended to let them complete the task in one conversation. The design joins understanding of the customer to practical means of action. The relevant test is whether the customer gets a dependable result, including when the routine fails.

Aravind Eye Care, a network of eye hospitals headquartered in Madurai, Tamil Nadu, in southern India, designs around restored sight rather than the operation alone. Outreach teams help address awareness, transport and cost, bringing people with cataracts to hospital for free surgery. In the year to March 2026, Aravind reported 6.7 million outpatient visits and 475,425 cataract operations, with 1.29 million outpatient visits in its free hospital sections. Visits are not unique patients. The lesson is to organise around the obstacles between a person and the result they need.

These examples make the design visible: funding, work, authority and capability are arranged around a recognisable outcome.

For employers, the hypotheses are commercial as well as human. Does credible progression reduce unwanted departures and the cost of replacing expertise? Do stronger relationships reduce failed handovers? Does visible contribution improve the decisions people make without supervision? Retention and engagement scores are useful evidence about experience. Their effects on service, profit and capability still have to be established.

Design the journey as a whole

Airports bring the argument down to the baggage carousel, and for us this is personal. Sharon and Colin fly almost every week, so airport improvement is dear to our hearts. The passenger wants one dependable journey. Delivering it involves an airport operator, airlines, ground handlers, security teams and border agencies, each with its own responsibilities. The passenger has bought an outcome that no single organisation can provide.

SITA puts a scale on the problem: an estimated 33.4 million bags were mishandled worldwide in 2024, or 6.3 for every 1,000 passengers. Transfers accounted for 41 per cent of the mishandling. The failures cluster at the joins between operations. A passenger waiting in yesterday's clothes is unlikely to be consoled by a successful departmental scorecard.

At Changi, named the world's best airport in Skytrax's 2026 awards, more than 13 kilometres of underground baggage links connect Terminals 1, 2 and 3. Identification and automated scanning connect each bag to its route; an early-baggage store holds up to 2,400 bags until flights open. Movement, information and timing are designed together.

Changi's 2024/25 report also describes joint work by the airport, civil aviation authority, airlines, ground handler SATS and engineering colleagues. Shared data and AI support turnaround planning and flag baggage problems; airside incentives include passenger experience.

The connection to the manufacturer is direct. A warning must reach someone who can respond; a handover must connect two commitments. The passenger wants their bag, the business customer wants the promised delivery and the patient wants to get home. None should need a working knowledge of the organisation chart to find out what happens next.

The leadership lesson is to let the whole journey govern the parts. That means adequate capacity, maintained equipment, trained people and credible recovery when something goes wrong. Those things cost money. So does failure, through tracing, redelivery, compensation, disruption and lost confidence. The purpose of system design is to make the desired result dependable at a defensible total cost.

> **The passenger experiences one journey. The leader must design the system that delivers it.**

Three objections, and our answers

“This is too abstract, and too slow.” Start with a result people recognise. The checklist and pit crew show different ways of making coordination concrete. A first experiment can change one arrangement and test what follows. We can begin learning before somebody has commissioned a logo for the programme.

“We already have purpose, values, agile and OKRs.” Many organisations do. Objectives and key results can clarify priorities, but they cannot by themselves secure the commitments between teams. Purpose without means and voice can become decoration or a shield: the tobacco boardroom believed in its purpose, and the institutions criticised by the Infected Blood Inquiry were devoted to care. The new science has to examine what these practices actually make possible, including what happens when they conflict.

“System design sounds like more bureaucracy.” Good design often removes approvals, competing targets and unnecessary handovers. Another process map is not the prize. People should find it easier to do useful work, get help and correct an error. The paperwork should be able to explain why it is still here.

The design must deliver the investment case

For a CEO and a private-equity owner, the route from system design to value must be explicit. Better coordination can bring a valuable product to market sooner. Dependable delivery protects repeat business and reduces rework, refunds and emergency freight. Resolving disputed invoices releases cash. Removing a bottleneck can increase output before more capital is committed. Each benefit needs an operating change, an accountable owner and a financial consequence. Extra output earns no extra profit if it cannot be sold at an adequate margin.

Put our illustrative manufacturer on a simple set of accounts: £100 million of revenue and £10 million of operating profit. Suppose it removes £2 million of recurring avoidable cost after paying the continuing cost of the improvement. Once transition costs have been absorbed, and provided revenue and service quality hold, operating profit rises to £12 million: a 20 per cent increase.

The hard work is identifying the waste in failed promises, rework and exceptions, then ensuring that the cost actually leaves the accounts. Releasing £2 million of inventory would improve cash separately; it would not create another £2 million of profit. Where EBITDA is the owner’s measure, reconcile the relevant savings to it explicitly. Do not count the same improvement twice because it has acquired a second label.

For the CEO and board, four questions make the investment case concrete: **which operating behaviour changes; which line in earnings or cash moves; what it costs to achieve; and when the benefit becomes dependable.** Compare the evidence with the plan, including the downside case. A slide showing the destination is not evidence that the business has a route there.

The hospital example clarifies a distinction that matters commercially too. Freeing a bed creates capacity; it does not automatically reduce expenditure. Similarly, releasing an hour on a production line creates an option. It becomes value through profitable demand served, expenditure avoided or capital no longer needed. State which benefit is expected and what must happen to realise it.

Under private-equity ownership, pace matters, but the investment period is not a licence to consume the business's future. Deferred maintenance, lost apprenticeships and exhausted people can flatter current results. A future owner inherits those costs as surely as the current owner enjoys the apparent saving. Test whether improvement remains credible when the leadership team changes and the next disruption arrives.

Automation sharpens the point. A tool that saves an hour has created an option. Leaders must turn it into profitable work, better service or lower expenditure, often by changing workflows, staffing plans or external purchases. Count integration, supervision and correction. Keeping the old approvals and reports alongside the new tool can make the business more expensive with impressive speed. Jassy makes a related point in his 2025 letter, published in 2026: AI calls for rethinking the customer experience and how work is organised. That is a design task as well as a technology choice.

Judge reserves and recovery arrangements against the losses they can absorb. Unused capacity is not automatically waste, nor automatically prudent. The question is whether its cost is justified by the risks to customer service, earnings and cash. Make that trade-off visible rather than leaving it to the next emergency.

The same principle reaches into new industries. McKinsey's May 2026 account of biomethane describes a chain joining agricultural waste, plant operations, buyers, gas networks and regulators. Better equipment alone cannot resolve unreliable supply, connection delays or incompatible contracts. Someone has to organise the commitments between them.

The chief executive and finance director should therefore connect each design change to quality, recurring expenditure, profit and cash. Distinguish cash released from profit earned, spending reduced from growth in spending avoided, and durable gains from costs deferred. In *Speed Wins*, we argued for shortening time to value. Here the obligation is to make that value real, repeatable and visible in the accounts.

Beginning a new science of organisation

This is the purpose of the paper: to begin forging a new science of organisation, capable of helping human systems adapt as intelligence becomes more abundant and more able to act. We have to connect what sociology knows about identity, political theory about power, economics about incentives, psychology about judgement and organisation research about work. Computational modelling and network analysis can help us examine how those forces interact. The disciplines offer different views of the same person trying to make something happen. It would help if they compared notes.

We have foundations to build on. Researchers were already modelling networks of people, knowledge and tasks in 2002. Others have shown why predicting a result and explaining what causes it are different jobs. The 2024 "Organizing for Good" paper adds a practical point: organisations with different strengths may each solve part of a problem that none can solve alone. Think of the hospital, care provider and transport service. They need dependable commitments between them; they do not all need to become the same kind of organisation.

Network analysis can reveal how the work actually travels. Who seeks advice from whom? Where does a warning stop? Which handover joins the manufacturer's promise to its capacity to deliver? The relationships cross departments and often extend to suppliers, customers and public bodies. The person keeping the whole thing moving may occupy a modest box on the chart and have a very full inbox. Mapping those connections can help locate a fragile dependency or a missing one.

Computational models let us rehearse alternatives. In an agent-based simulation, simplified people or teams follow explicit rules, exchange information and adapt. For our manufacturer, we could compare delivery decisions made centrally, locally or jointly, then introduce a supply shock. The result depends on the assumptions and needs checking against real behaviour. If everyone in the model shares information promptly and follows every rule, we may have simulated the organisation we wish we had.

Behavioural research keeps the people recognisable. Observation, interviews and experiments can help distinguish a missing skill from a rational fear of blame, or a lack of capacity from a refusal to cooperate. The nurse waiting for a care package and the manager trying to survive Thursday have reasons for what they do. Understanding those reasons gives us something more useful to change than their attitude.

Foresight and practical trials connect possibility to evidence. Rehearse a demand surge, a supplier failure or an AI agent pursuing the wrong objective. Then test a promising arrangement on a bounded piece of real work, with clear responsibilities, safeguards and a way to stop. Follow what happens to quality, cost, learning and people's ability to act. Scenarios expose assumptions; trials show which ones survive contact with the working day.

Calling it a science should make us harder to satisfy, including with our own ideas. We need explanations that can be tested, rival accounts that can be compared and the willingness to abandon a favourite prescription. Human beings adapt to interventions. A design that helps under one set of conditions may fail when demand, expertise, resources or legitimacy change. An elegant diagram cannot settle the argument, however much we enjoyed drawing it.

Our central proposition can be tested. Where delivery depends on several contributors, changing the relationships between their decisions, resources and commitments should improve the whole result more reliably than asking each contributor to improve alone. It could fail because the redesign did not change behaviour, because the binding constraint lay elsewhere, or because coordination cost more than it contributed.

Credible investigation would compare otherwise similar settings, follow changes over time and measure the resources consumed as well as the results produced. It would record unintended harm, costs pushed elsewhere and whether the gains survived the leader's absence. A before-and-after success is inconclusive if demand or funding also changed. We should test when combinations of the ten commitments outperform isolated changes, and whether improvements in understanding, means and voice help explain better delivery.

The human questions deserve the same discipline. When does meaning sustain cooperation, and when does it shield harm from challenge? When does local discretion improve decisions, and what support makes it work? As agents take on tasks, do people become better able to judge, learn and act, or merely more dependent on answers they cannot question? A new science must be curious about those experiences as well as the accounts. People are participants in the design, with knowledge and interests of their own.

Start on Monday

Choose one recurring failure that matters to a customer and the value-creation plan. A late order, a disputed invoice or a complaint that returns three times is a better starting point than an ambition to “transform collaboration”. Start small enough to see the work, but follow it far enough to see its commercial consequence.

Bring the people at the handover together. Include those who make the customer promise, those who deliver it and those who collect the cash. Walk through one recent case and, where possible, hear the customer’s experience. What did each person know, need and try to say? Begin with what made sense to them at the time. Give them room to describe the work before asking them to defend it. Establish the baseline for delay, quality, cost and cash.

Change one arrangement, with its necessary supports. It might be a decision right, a shared commitment or a route for escalation. Name an owner with usable authority. Agree the resources, expected effect, review date and conditions for stopping or revising the trial. A small change may still require sales, operations and finance to keep a promise together. Agree what will stop, shrink or move so people can do the work and learn from it. “Make time for this” is not a resource allocation.

Check the whole result with finance. Did the customer receive what was promised? Did margin improve and cash arrive sooner? What did the change cost, and did the burden move elsewhere? Ask the team whether they understood the result better, had the means to act and received an answer when they raised a problem.

In the manufacturer, an early joint check on a delivery promise might expose a margin-destroying exception before the order is accepted. If the constraint is missing capacity, escalate the investment or commercial decision; do not rename it a communication problem. Keep what works, record what fails and share what others can use. This is how the larger science begins to accumulate: careful explanations, practical trials and results that survive scrutiny. A useful first result is worth more than a splendid launch.

Ten questions for the board

Boards and executive teams can test whether their human system can deliver the investment case with ten questions. Several answers of “no”, or “we do not know”, identify work the CEO and board need to do together. Candour is more useful here than ten reassuring answers agreed in advance.

1. For the three outcomes that matter most to growth, earnings and cash, who owns the whole result and can resolve conflicts across functions?
2. Where do decision rights, resources and commitments fail to meet at the handovers that determine customer delivery?
3. Which targets reward booked revenue, local savings or activity while pushing cost, risk or delayed cash elsewhere?
4. Can a serious concern raised at the front line reach someone with the power and the obligation to act, within a defined time, and is the response visible?
5. When we give work to AI agents, have we decided their goals, constraints, visibility of actions, human accountability and routes for challenge? Whose independent evidence could overturn a plausible answer?
6. Are we preserving the experiences through which our next generation of experts will acquire judgement, including the ability to challenge persuasive AI answers?

7. Do we test our operating model against variation and recovery, or only against the expected case?
8. When purpose collides with commercial or political pressure, what is our actual decision rule, and does everyone know it?
9. Which gains in this year's plan will still be there under a new leader or owner, and which depend on deferred investment or people working beyond a sustainable limit?
10. If the leadership team disappeared tomorrow, would the system still deliver its critical results, and still surface unwelcome truth?

The tenth question is especially revealing for an owner. It tests whether the business has built a dependable capability, or remains reliant on the presence of particular people. Both can produce a good quarter. Only one gives the next leadership team a sound starting point.

Agency in an agentic world

The promise of AI depends on the human systems through which it acts. Those systems must turn capability into useful results and keep its exercise accountable. Otherwise we risk wasting the promise in familiar organisational obstruction, or allowing powerful tools to amplify mistakes and harms beyond our capacity to correct them. Better models do not relieve leaders of choosing defensible ends, making power legitimate or giving people a meaningful place in the work.

Work is one important source of meaning, alongside family, friendship, faith, care and community. Its contribution lies in belonging, developing a craft, exercising judgement and making something better for another person. There is quiet pleasure in being useful to someone who knows your name. Leaders cannot manufacture those experiences by declaration. They can design work in which they are possible, and avoid destroying them in pursuit of an incomplete measure of efficiency.

The shift from intellect towards wisdom becomes practical here. People and machines together hold more intelligence than any individual can command. **The leader brings the judgement to choose worthwhile ends and build a system that can pursue them, learn and accept correction.** For a commercial organisation, that means serving customers, earning a worthwhile return, turning it into cash and adapting as conditions change. People need understanding, means and voice. The CEO remains accountable for the result, its ethics and the capability left behind.

Asimov's gap between knowledge and wisdom will not close itself. We have to begin forging a new science of organisation, learning our way towards human systems worthy of the intelligence now available to them. Go back to that auditorium in Geneva, and then to the pothole at the end of your road. We possess extraordinary capability and still struggle to organise ordinary delivery. That gap is the work.

The human system is the leader's principal product. Its quality is revealed by what people can reliably achieve together, especially when no one is watching.

The achievement may be a discovery that changes physics, or a safe journey home from hospital. For the woman waiting on the ward, it is the familiar sound of her own front door. In each case, people have made something possible for one another. That is work worth designing well.

Sources and further reading

Heather A. Haveman, Joseph T. Mahoney and Elizabeth Mannix (2021). The Evolving Science of Organization: Theory Matters. Academy of Management Review, 46(4), 660–666.

Steven Pinker (2011; 2018). *The Better Angels of Our Nature* and *Enlightenment Now*. See his discussion of violence and account of progress.

Uppsala Conflict Data Program (9 June 2026). Record Number of Conflicts Between States. Findings from its 2025 conflict data.

Massimo Giudici, Alessandro Agosta, Fabrizio Bacchini and Beatrice Enrico (1 May 2026). Scaling Biomethane: The Next Phase of Growth in Renewable Gas. McKinsey.

Fabiola Gianotti (2026). Leadership at the Edge of Discovery, I by IMD. See also CERN's Our Governance and Steering ATLAS Forward with New Management (11 March 2025).

Mark Robinson (2019; online 2018). The CERN Community: A Mechanism for Effective Global Collaboration? Global Policy, 10(1), 41–51.

McKinsey & Company (10 September 2024). What Is Leadership?

NHS Staff Survey (March 2026). 2025 National Results Briefing, pp. 9, 13 and 22. England; fieldwork September–November 2025.

Accelerating Performance Consulting (2025). CEO Challenges, pp. 6, 8 and 11.

US Department of Justice. Volkswagen plea agreement and civil resolutions (11 January 2017); McKinsey resolution concerning Purdue Pharma (13 December 2024).

Infected Blood Inquiry (20 May 2024). Report, Volume 1: Overview.

Robert Michels (1911). *Political Parties*. Selected excerpts, University of Minnesota Duluth, from the 1966 edition via Grusky and Miller (1970).

Colin Price and Sharon Toye. *Speed Wins*. Accelerating Performance Consulting.

George Herbert Mead (1934). *Mind, Self, and Society*, section 26: The Realization of the Self in the Social Situation.

Nicholas Bloom, Ruobing Han and James Liang (2024). Hybrid Working from Home Improves Retention without Damaging Performance. Nature, 630, 920–925. A randomised trial at Trip.com.

UK National Audit Office (9 September 2026). Managing the Flow of Patients through Hospital from A&E. Report summary.

HM Chief Inspector of Prisons (7 July 2026). Annual Report 2025–26. Official summary, covering England and Wales.

Audit Reform Lab, University of Sheffield (2024). Reward for Failure and research announcement (23 May). Study of 250 large UK listed-company collapses, 2010–22.

International Monetary Fund (15 April 2026). War Shock Requires Disciplined Fiscal Reaction, accompanying the April Fiscal Monitor.

United Nations Environment Programme (2025). Emissions Gap Report 2025. Publisher summary.

United Nations (October 2026). OCHA, Humanitarian Situation Report, 2 October; UN, In Islamabad, Guterres Calls for End to US-Iran Conflict (6 October).

Jan-Werner Müller (2016). Interview on populism, Princeton University.

Émile Durkheim (1893). *The Division of Labour in Society*. Selected excerpt on solidarity and specialisation.

Elinor Ostrom (2009). Beyond Markets and States: Polycentric Governance of Complex Economic Systems. Nobel lecture, especially pp. 421–423.

Hannah Arendt (1969). Reflections on Violence. The New York Review of Books.

CERN and NASA. CERN, Higgs Boson; NASA, Webb Team; JPL, Perseverance's First AI-Planned Drives (30 January 2026).

WHO and FAO. WHO, Smallpox and Poliomyelitis; FAO, Rinderpest Eradication.

Aravind Eye Care System. Activity Report 2025–2026, pp. 22–23.

Gallup (2026). US Employee Engagement Declines from 2020 Peak (27 January); State of the Global Workplace 2026. Survey data collected in 2025.

UK Cabinet Office (2026). Civil Service People Survey 2025: Benchmark Results, Table 3, item B47.

Culture Amp and Perceptyx. Culture Amp, Mid-Year 2025 Benchmark Update; Perceptyx, 2026 Benchmark Report. Both draw on customer survey data; Perceptyx covers 2023–25.

UK House of Commons Public Accounts Committee (17 January 2025). Condition and Maintenance of Local Roads in England.

Bengt Holmström and Paul Milgrom (1991). Multitask Principal-Agent Analyses: Incentive Contracts, Asset Ownership, and Job Design. Journal of Law, Economics, & Organization, 7, 24–52.

Samuel Bowles and Sandra Polania-Reyes (2012). Economic Incentives and Social Preferences: Substitutes or Complements? Journal of Economic Literature, 50(2), 368–425. Publisher abstract.

Post Office Horizon IT Inquiry (8 July 2025). Final Report, Volume 1: Official Summary.

Toyota Motor Corporation. Toyota Production System. Corporate account of its operating principles.

Aravind Eye Care System. Community Outreach and Our Story: The Aravind Model.

Patagonia (2022). Ownership Structure and Yvon Chouinard's Founding Account.

Skytrax (18 March 2026). World Airport Awards and Baggage Delivery Ranking. Based on passenger surveys, August 2025–February 2026.

Changi Airport Group. Baggage Handling Systems (1 December 2023); Annual Report 2024/25, pp. 31–32.

SITA (12 June 2025). Baggage IT Insights 2025. Estimates of baggage mishandling in 2024, based on WorldTracer data and IATA passenger traffic.

Patagonia. 2025 Progress Report.

Grenfell Tower Inquiry (4 September 2024). Phase 2 Report, overview and executive summary.

Michael G. Vann (2003). Of Rats, Rice, and Race: The Great Hanoi Rat Massacre, an Episode in French Colonial History. French Colonial History, 4, 191–203.

US Department of Justice (21 February 2020). Wells Fargo Agrees to Pay $3 Billion to Resolve Criminal and Civil Investigations into Sales Practices.

Tenerife runway collision (27 March 1977). SKYbrary accident summary, drawing on the official Spanish investigation.

Presidential Commission on the Space Shuttle Challenger Accident (1986). Report, Volume 1, Chapter 5. **Columbia Accident Investigation Board (26 August 2003).** Final-report announcement.

Timo O. Vuori and Quy N. Huy (2016). Distributed Attention and Shared Emotions in the Innovation Process: How Nokia Lost the Smartphone Battle. Administrative Science Quarterly, 61(1), 9–51. Aalto University summary.

Alex B. Haynes and colleagues (2009). A Surgical Safety Checklist to Reduce Morbidity and Mortality in a Global Population. New England Journal of Medicine, 360(5), 491–499. A before-and-after study in eight hospitals.

Guinness World Records. Fastest Formula One Pit Stop: McLaren, 1.80 seconds, Qatar Grand Prix, 8 October 2023.

Stephen Elop (8 February 2011). The "Burning Platform" Memo, published by Engadget.

Adam M. Grant and colleagues (2007). Impact and the Art of Motivation Maintenance: The Effects of Contact with Beneficiaries on Persistence Behavior. Organizational Behavior and Human Decision Processes, 103(1), 53–67. Grant's account of the study.

Paul Milgrom and John Roberts (1990). The Economics of Modern Manufacturing: Technology, Strategy, and Organization. American Economic Review, 80(3), 511–528.

Erik Brynjolfsson, Daniel Rock and Chad Syverson (2021). The Productivity J-Curve: How Intangibles Complement General Purpose Technologies. American Economic Journal: Macroeconomics, 13(1), 333–372.

Octopus Energy. Interview with Jon Paull on Customer Service, BCG (2020). Market-position data: Ofgem (2024); Kraken account numbers: company reporting (2025).

Colin Bryar and Bill Carr (2021). *Working Backwards: Insights, Stories, and Secrets from Inside Amazon*. St. Martin's Press.

Haier. Company accounts and business-school cases on Zhang Ruimin's Rendanheyi model of self-managing micro-enterprises.

Microsoft (November 2013). Announcement Ending Forced Employee Ranking, Computerworld.

Medtronic. Employee Holiday Program, a tradition begun in 1960.

Delta Air Lines (July–August 2024). Chief Executive's Estimate of Disruption Costs, NPR via TPR (31 July); August securities filing on cancellations.

Klarna (2024–25). Reporting on Its AI-Assistant Claims and Return to Human Customer Service.

Shopify (2025–26). Tobi Lütke's April 2025 memo on AI and headcount; Farhan Thawar on Expanding the Internship Programme, CoderPad.

US House Committee on Transportation and Infrastructure (September 2020). Final Committee Report: The Design, Development and Certification of the Boeing 737 MAX. CNBC report.

Apple University. Published accounts of the programme founded by Steve Jobs in 2008 to teach Apple's history, culture and decision-making.

US Environmental Protection Agency (18 September 2015). Notice of Violation to Volkswagen AG, Audi AG and Volkswagen Group of America.

Isaac Asimov and Jason A. Shulman, editors (1988). *Isaac Asimov's Book of Science and Nature Quotations*, chapter 72, p. 281. Asimov's epigraph on science and wisdom. Attribution checked through Quote Investigator.

Y**uval Atsmon (14 October 2026).** When Knowledge Outruns Wisdom. *Compound Learning*. Commentary on expertise, independent judgement and the risk of apparent consensus.

Handforth Parish Council meeting (10 December 2020). Account and recorded exchanges reported by ITV News, 5 February 2021.

Bevan Brittan / Cheshire East Council (27 May 2021). Handforth Parish Council: Investigation Report, Multiple Complaints against Councillor Burkhill, sections 7.21–7.23. Findings on Weaver's authority and the circumstances of her actions.

Kathleen M. Carley (2002). Computational Organization Science: A New Frontier. *PNAS*, 99 (suppl. 3), 7257–7262. Foundational account of computational agents and social and knowledge networks.

Oliver Baumann, Gerald F. Davis, Sven Kunisch, Jiao Luo and Brian Wu (published January 2024). Organizing for Good—Using Organization Design to Take on Grand Challenges. *Journal of Organization Design*, 12, 165–176. Conceptual introduction to collective action and design across organisational boundaries.

Fink, L. (2026). Annual Chairman's Letter to Investors. BlackRock. The risk that AI gains concentrate among companies and investors, and the case for broader ownership. See also BlackRock's biography of Fink.

Andy Jassy (2025; 2026). Amazon's 2024 and 2025 Letters to Shareholders. Leadership judgement and redesigning work for AI.

Colin Price and Sharon Toye (2017). *Accelerating Performance*. Wiley. Chapters 11 and 16: behaviour change, role modelling, space and dissolving paradox.

Emily Lawson and Colin Price (2003). The Psychology Of Change Management. McKinsey Quarterly, 2003 special edition.

FIFA (2026). FIFA Council. **FairSquare (2024).** Substitute: The Case for the External Reform of FIFA, pp. 79–83.

About the authors

Colin Price

Organisation performance

Colin Price advises boards, chief executives and investors on organisational performance through Accelerating Performance Consulting. He is Chair of itim Group plc and was previously Chair of Co-Company. A former McKinsey Senior Partner, he led its worldwide Organisation Practice for seven years. His executive roles have also included CEO of Heidrick Consulting and COO of Quintet Private Bank.

His work connects strategy, organisational health and leadership behaviour to business results. He is the author or co-author of seven books, including *Accelerating Performance*, with Sharon Toye, and *Beyond Performance*, with Scott Keller.

colin@acceleratingperformanceconsulting.com

Sharon Toye

Leader and team performance

Sharon Toye is a founding partner of Accelerating Performance Consulting and advises chief executives and senior teams on leadership, team effectiveness and lasting behaviour change. She combines commercial experience with psychological insight as a Master Certified Coach with the International Coaching Federation and a psychotherapist registered with the UKCP and BACP. She keeps a low-cost therapy practice for people who could not otherwise afford it.

Previously, she was a founding partner of Co-Company and led the global Team Acceleration practice at Heidrick & Struggles. She holds an MBA from London Business School and master's degrees in organisation psychology and integrative psychotherapy. She is co-author, with Colin Price, of *Accelerating Performance*.

sharon@acceleratingperformanceconsulting.com

Human Systems is part of our series on leadership and organisational performance. All papers are free to read at acceleratingperformanceconsulting.com/read

© Accelerating Performance Consulting Limited. Registered in England and Wales, company number 12224907.