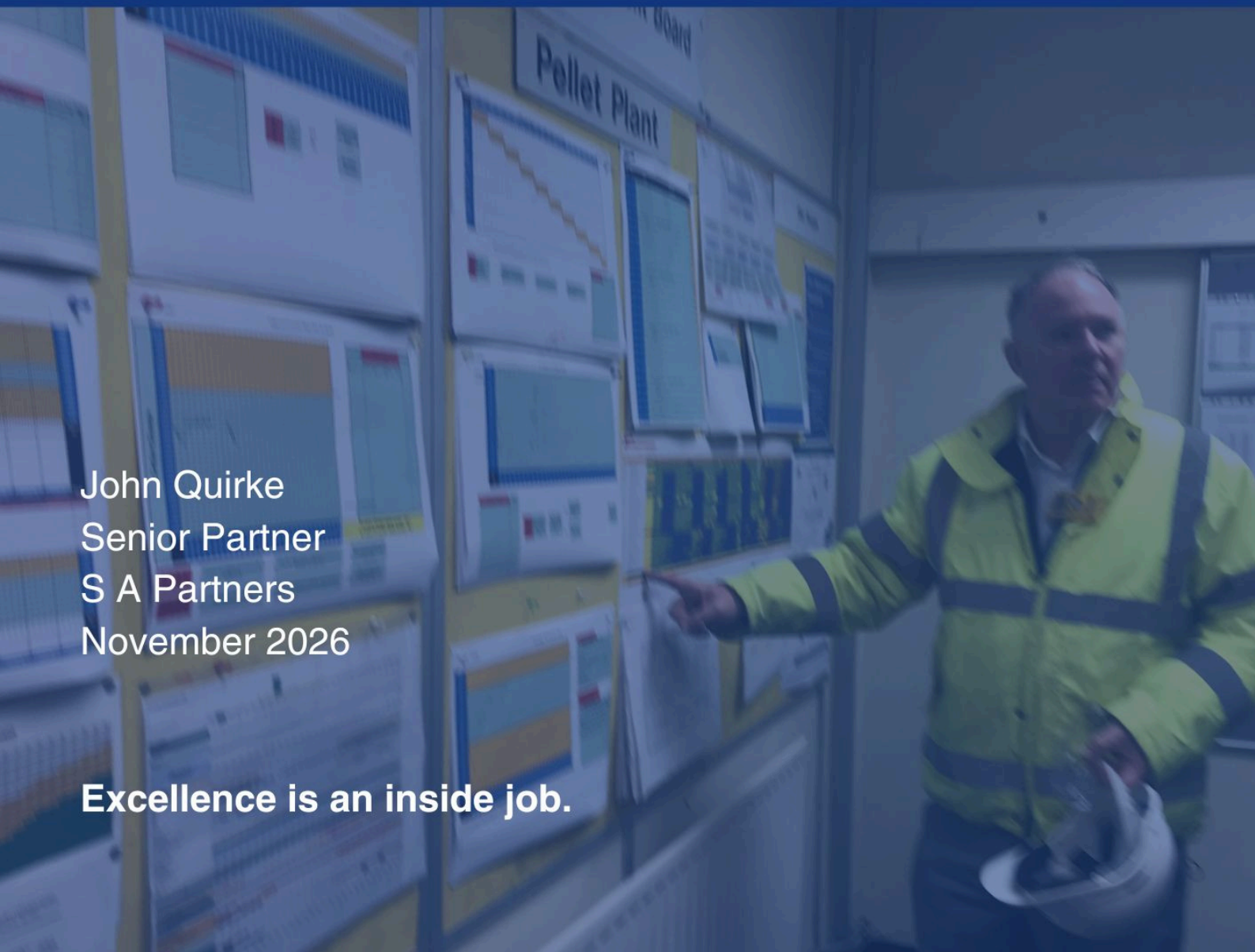
sa partners

WHITE PAPER

MEASURING THE EFFECTIVENESS OF AN IMPROVEMENT SYSTEM



John Quirke
Senior Partner
S A Partners
November 2026

Excellence is an inside job.

EXECUTIVE SUMMARY

Busy is not the same as better

Most organisations measure their improvement system by how busy it is, not by whether it works. Ideas per person, projects per employee, belts certified and claimed savings all count activity. None of them tells leaders whether the organisation is getting better at the things that matter, faster than it did before.

Russell Ackoff warned that the righter we do the wrong thing, the 'wronger' we become. An improvement system tuned to hit participation quotas becomes very efficient at producing participation. It can do this while problems recur, standards drift and the gap to customer need stays open. As the Shingo Model reminds us, purpose and systems drive behaviour, and that behaviour can be ideal or less than ideal.

Four ideas

01 Every system must have an aim, and the aim specifies its own measures. Each clause of a well-written aim is a promise that can be tested.

02 Treat the improvement system as a process in its own right. Measure its flow, quality, capability, alignment and the leader behaviour that sustains it.

03 Anchor on two outcome tests. The repeat occurrence rate of problems already "solved", and the improvement half-life of key business metrics.

04 Make leaders the custodians of wisdom about the work. Improvement culture is a constant quest to understand the work more deeply. Leaders create the conditions for that quest, and the effectiveness review is where they practise it.

Call-out: Used together, these turn the periodic system review from a progress report into a genuine test of whether the system is doing the right things, not merely doing things right.

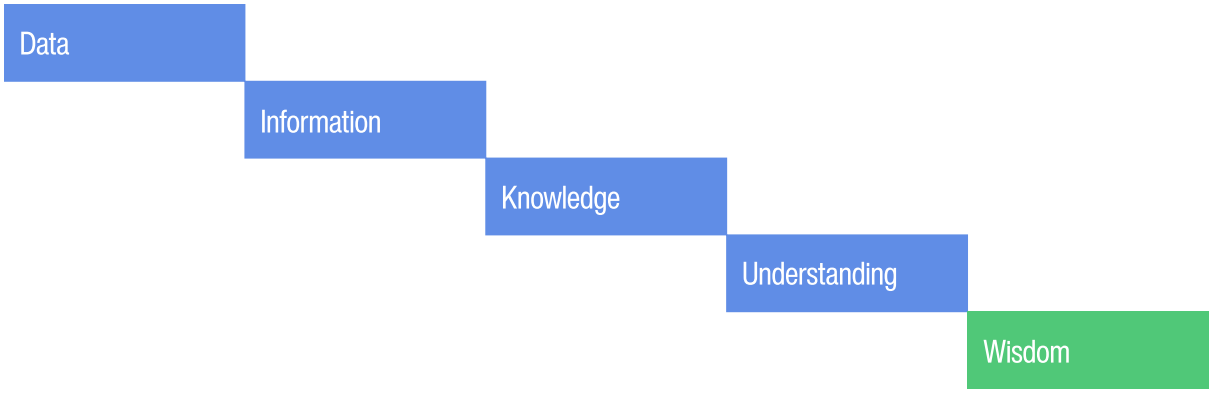
01 INTRODUCTION — The righter we do the wrong thing

> *"The righter we do the wrong thing, the 'wronger' we become." Russell L. Ackoff*

Ackoff drew a sharp line between efficiency and effectiveness. Efficiency is doing things right; effectiveness is doing the right things. His point was not that efficiency is unimportant, but that efficiency applied to the wrong purpose compounds the error. Each gain in competence carries the organisation further in the wrong direction, and makes it harder to turn back.

Improvement systems are particularly exposed to this trap. They are usually launched with a broad, humane aim about engaging people in improving their work. They are then governed through narrow activity counts, because counts are easy to collect and easy to report. Over time the system optimises for the count, and the count quietly replaces the aim.

Ackoff also described a hierarchy running from data, through information, knowledge and understanding, to wisdom. Data and information tell us what is happening. Knowledge tells us how to do things; understanding tells us why. Wisdom concerns whether what we are doing is worth doing at all: it is the level at which effectiveness is judged. It is also the level at which we can plan change effectively, because we understand what is worth changing.



Most improvement system reporting stops at information. It shows how many ideas were raised and how many projects closed. The question this paper addresses sits higher up Ackoff's hierarchy: how would leaders know whether their improvement system is making the organisation wiser about its own work?

02 PURPOSE — Every system must have an aim

The starting point for any assessment of an improvement system is its defined aim. Deming's insistence that a system without an aim is not a system applies with particular force here. The aim is not a slogan for the noticeboard; it is the specification for the system's own measures.

> *"Engage, equip and empower every team member to improve their own work every day, focused on site priorities, with the support of all leaders."*

Read carefully, that typical aim makes six distinct promises. It promises engagement, capability, autonomy, a daily rhythm, alignment to priorities, and leadership support. Each is a claim that can be tested.

In practice, organisations with an aim of this kind often measure only one of the six promises, engagement, and measure it weakly through participation counts. The other five go unexamined. The aim then describes one system while the measures govern another.

A system must also have an inbuilt process of self-reflection on its real and measurable impact. In Deep Excellence the improvement system is one of the signature systems, alongside strategy deployment, key work systems and the leadership system. Each signature system either enables the behaviours the organisation says it wants or quietly contradicts them. The aim, and the measures that flow from it, are where that alignment is either designed in or left to chance.

03 THE PROBLEM — Why activity measures mislead

Participation counts tell you that people are touching the system, not whether the system works. Four failure patterns occur across organisations:

The quota contradicts the aim. A target of two ideas per person per year is incompatible with an aim that says "every day". It sets the expectation at an annual event. Quota-based idea targets reliably produce trivial ideas submitted in the final month of the year.

The unit of improvement is wrong. A target of one improvement project per employee frames improvement as a project. It pulls the system toward the A3 and DMAIC end of the spectrum and away from small daily improvement at the point of work. The aim and the measure end up describing different systems.

Nothing tests alignment, durability or capability. Few improvement systems measure whether improvements address site priorities, whether they stick, or whether people are becoming more capable. Those three are precisely what connect the system to Quality, Delivery and Cost results.

Savings are counted twice. Summing claimed project savings is the most common bridge to financial results, and the least reliable. Claimed savings almost always exceed what appears in the profit and loss account, and the gap erodes leadership trust in the whole system.

Beneath all four sits a well-known dynamic: when a measure becomes a target, it ceases to be a good measure. Once an idea count is attached to an individual or team target, it measures compliance. People learn to feed the count rather than improve the work, which is Ackoff's warning in operational form.

Leadership measures deserve the same scrutiny. Key Behavioural Indicators that place obligations on leaders are a real strength, and a clear response standard (for example, every idea acknowledged within five working days and decided within ten) is good practice. But a standard of one gemba walk per month for senior leaders is a compliance threshold, not a behaviour. It describes the minimum a leader can do without being noticed, not the presence that builds an improvement culture.

Finally, many governance frameworks include a periodic "system effectiveness" review without defining what effectiveness means. Without that definition the review defaults to a progress report on the counts. That gap is the question this paper sets out to close.

04 THE FRAMEWORK — Measure the improvement system as a process

Treat the improvement system as a process in its own right, with its own value stream, and measure it as you would any process. That means flow, quality, capability, alignment and the behaviour that sustains it. Then use the rate of improvement in business metrics, not their level, as the bridge to results.

The framework below derives each measure from a clause of the aim. The example measures are illustrative; each organisation should choose its own from its aim.

Aim element	What to measure	Example measures
Every day	Flow of the improvement process	Median idea-to-implementation lead time; ideas ageing beyond 30 days; ratio of quick-win improvements to A3 projects
Empower	Autonomy at the point of work	% of improvements implemented within the team's own authority without escalation; pulse question "Can I change how I do my work without asking permission?"
Equip	Capability, not training attendance	% of team members who have led a problem to verified root cause unaided; coach-to-learner ratio; time from first training to first independent improvement
Focused on priorities	Alignment	% of active improvements explicitly linked to a site or value stream KPI gap; % of breakthrough objectives with improvement work against them
Improve (does it stick?)	Quality of improvement	% of improvements still in place at 90 days (gemba audit); % that updated standard work, and how long revision takes; repeat occurrence of problems already "solved"
Support of all leaders	Leader behaviour	Response-time KBIs; time leaders spend coaching at the gemba; % of rejected ideas where the reason was explained to the originator

Two measures to anchor on

Repeat occurrence rate. This is the proportion of nonconformances, deviations, CAPAs or downtime events that recur after a problem-solving activity was closed against them. It is the most direct test of whether problem solving works, it is usually already captured in the quality system, and it maps straight onto quality and cost. A system that closes A3s while the same failure modes return is a reporting system, not an improvement system.

Improvement half-life of key metrics. This is the time taken to halve the gap between current performance and the ideal on measures such as scrap, on-time-in-full or cycle time. It avoids the attribution trap of summing claimed savings and asks a simpler question: is the organisation improving faster than before? It is described in detail later in this paper.

A caution on numbers: Keep the set small and focused on what is critical to the business. Six to eight measures are enough, chosen so that at least half are leading indicators a team can act on weekly.

Call-out: Keep idea counts off individual and team targets. System measures exist for learning and belong in the effectiveness review; Quality, Delivery and Cost results remain the accountability layer. Confusing the two is how a learning system becomes a compliance system.



05 LEADERSHIP — Improvement as a constant quest for wisdom about the work

An improvement culture is not a programme of ideas and projects. It is an organisation's constant quest to understand its own work more deeply: what the work is for, how it really flows, where it fails and why. Measures can show whether that quest is alive, but only leaders can create the conditions in which it happens.

Tier 4 Leadership review	↑ Issues escalate up
Tier 3 Site	↑ Issues escalate up
Tier 2 Area/value stream	↑ Issues escalate up
Tier 1 Team/cell	↓ Support and priorities flow down

Knowing what the work is for: Deep Excellence begins with essence: defining, with precision and honesty, what the organisation exists to do and what the work at its best actually looks like. This is the leadership contribution at the top of Ackoff's hierarchy. If leaders cannot say what good looks like, the improvement system has no direction, and every gain in its efficiency risks doing the wrong thing righter. The practical test is alignment. When leaders are clear on essence and priorities, the share of improvements linked to a real performance gap rises. When they are not, improvement drifts toward what is easy to change rather than what matters.

Seeing and hearing the culture: Deep Excellence holds that culture can be seen and heard in the specific, observable detail of daily work. A culture of deep excellence is one where the gap between espoused values and enacted values is small. Improvement systems expose that gap quickly: an organisation that says it values every idea, but leaves ideas unanswered for weeks, has told its people what it really values. This is why leader behaviour sits inside the measurement framework rather than beside it. Response times, reasons explained for rejected ideas, and time spent coaching at the gemba are the visible evidence of whether leaders mean what the aim says.

Matching the conversation to the person: The language leaders use at the point of work shapes whether people feel able to improve it. Deep Excellence stresses matching the style of conversation to the competence and confidence of the person. A new team member needs direction and teaching; an experienced one needs a coaching question and then space. A leader who tells when they should ask removes the learning, and with it the capability the system is meant to build.



05 LEADERSHIP — Improvement as a constant quest for wisdom about the work

Holding up the mirror: Before asking the organisation to change, Deep Excellence asks leaders to examine how their own behaviour contributes to the current state. Applied to the improvement system, the questions are uncomfortable but direct. Which of our own habits make it harder for people to improve their work? Which of our systems reward the count rather than the improvement?

The effectiveness review as a practice of wisdom: The periodic system effectiveness review is where this leadership work becomes concrete. Run as a progress report, it confirms that the counts are on track. Run as a practice of wisdom, it asks whether the system is doing the right things: whether problems stay solved, whether improvement is accelerating, and whether people are more capable than a year ago. In this sense the frontline owns doing things right, and leaders own the question of whether they are the right things. Ackoff's warning is ultimately addressed to leaders, because only they can change the aim, the measures and the conditions that determine what the organisation becomes good at.



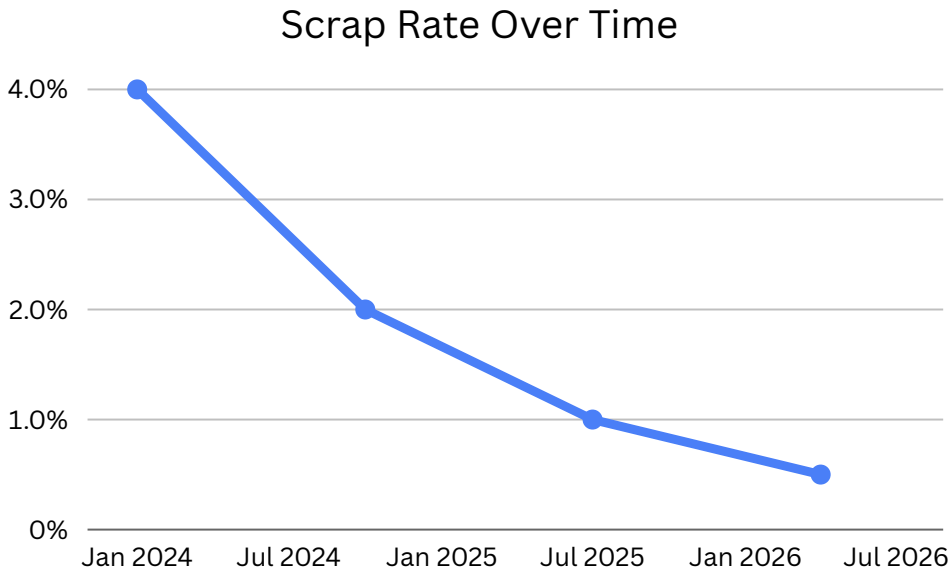
06 THE BRIDGE TO RESULTS — Improvement half-life

The improvement half-life is the most honest link between an improvement system and business results. It replaces the question "how much did our projects save?" with "are we improving faster than we used to, and fastest where the system is strongest?"

Origin: Arthur Schneiderman developed the concept as Vice President of Quality and Productivity Improvement at Analog Devices, and published it in Quality Progress in April 1988 ("Setting Quality Goals"). Analog's chairman, Ray Stata, described its use in Sloan Management Review in 1989, framing the slope of the half-life curve as the rate of organisational learning. The approach underpinned the corporate scorecard Analog built in the late 1980s, which Kaplan and Norton later drew on for the Balanced Scorecard. If the half-life measures the rate at which an organisation learns about its own work, it is a direct measure of the quest for wisdom described in the previous section.

The core observation: Schneiderman noticed that when a defect-type metric is worked on by a sustained, well-functioning PDCA process, it does not fall in a straight line. It falls exponentially: in each fixed period, the remaining gap to the ideal halves. That fixed period is the half-life. If scrap is 4% against an ideal of zero and the half-life is nine months, expect roughly 2% after nine months, 1% after eighteen and 0.5% after twenty-seven. The absolute gains shrink, but the rate of improvement stays constant. That is the signature of continuous improvement working as a system rather than as a series of projects.

scrap 4%, 2%, 1%, 0.5% at months 0, 9, 18, 27



The logic is Pareto applied repeatedly. Teams remove the largest remaining cause, and the next largest is now a smaller share of a smaller total. Each cycle removes a roughly constant proportion of what remains.

How to calculate it: Express the metric as a gap to the ideal: defects, scrap, unplanned downtime, OTIF misses, or lead time minus value-add time. Plot the natural log of that gap against time; if improvement is continuous, the points fall on a roughly straight line. The half-life is ln 2 (about 0.69) divided by the slope of that line. About a year of monthly data gives a credible line, and a spreadsheet trend line is sufficient.

What the shape of the line reveals

A straight declining line: the system is working as designed.	**A line that flattens:** improvement has stalled. The easy causes are gone and the system lacks the capability or cross-functional reach to tackle the next layer.	**A staircase, with flat periods punctuated by drops:** improvement is driven by projects or events, not daily work. This is the likely signature of a "one project per employee per year" target.	**No consistent slope:** the metric is being moved by noise or external factors rather than the improvement system.

Comparison adds further insight. If value streams or shifts with stronger system measures (shorter lead times, higher sustainment, better alignment) are not improving faster than the rest, the theory behind the system design is wrong. That is worth knowing.

What drives the half-life: Schneiderman found the half-life depends mainly on two forms of complexity: technical (how hard the problem is) and organisational (how many functions must cooperate to fix it). A simple problem owned by one team may have a half-life of a few months. A technically difficult problem that crosses functions, suppliers or customers can take well over a year. Organising around value streams rather than functions reduces organisational complexity and so shortens half-lives.

Limitations: The method works best for metrics with a clear ideal and many contributing causes. It is weaker where performance is constrained by capital, equipment physics or demand; reaching those floors requires breakthrough redesign, not incremental PDCA. Low volumes create noise, and changes in product mix or definitions can mimic improvement. It is also only as good as the definition of the ideal. If the "ideal" is this year's negotiated target, the method measures progress to a number, not to the true gap. Used with those caveats, it is one of the few measures that tells a leadership team whether its improvement system is a genuine engine or a reporting activity.

08 CONCLUSION — When problems stay solved

An improvement system is effective when problems stay solved, when the organisation improves faster than it did before, and when its people become steadily more capable of improving their own work. Activity counts can show none of this; at best they show that the machinery is turning.

The measures in this paper are designed to work with the intentional design of the system, shaping behaviour and, over time, culture. When each clause of the aim has a measure behind it, the effectiveness review becomes a genuine test of the system design rather than a progress report. When leaders own that review as a practice of wisdom about the work, they avoid Ackoff's trap: becoming ever more skilled at doing the wrong thing.

REFERENCES

Ackoff, R. L. (1989). From data to wisdom. Journal of Applied Systems Analysis, 16, 3–9.

Deming, W. E. (1994). The New Economics for Industry, Government, Education (2nd ed.). MIT Press.

Quirke, J., Packham, J., Grogan, S. and Cutliff, B. (2023). Deep Excellence. S A Partners

Schneiderman, A. M. (1988). Setting quality goals. Quality Progress, April, 51–57.

Schneiderman, A. M. The half-life/complexity matrix. schneiderman.com

Shingo Institute (2025). The Shingo Model (v15.4). Utah State University.

Stata, R. (1989). Organizational learning: the key to management innovation. Sloan Management Review, 30(3).

DIAGNOSTIC QUESTIONS FOR LEADERS

These questions can open the next system effectiveness review. Each one tests whether the system is doing the right things, not only doing things right.

01 Aim. Can every clause of our improvement system aim be matched to a measure? Which promises in the aim are currently untested?

02 Rhythm. Do our targets describe daily improvement, or an annual event? How long does a typical idea take to go from raised to implemented?

03 Autonomy. What share of improvements can teams make within their own authority? Would our people say they can change their work without asking permission?

04 Capability. How many people have led a problem to verified root cause without help in the last year? Is that number growing?

05 Alignment. What share of active improvement work is linked to a real performance gap? Can every breakthrough objective point to improvement activity against it?

06 Durability. How many improvements are still in place after 90 days? How often do problems we have "solved" come back?

07 Rate. What is the half-life of our three most important performance gaps? Is it shortening, and is it shortest where the system is strongest?

08 Leadership. Are our leader standards set at the level of compliance or of presence? What would our people say we value, judged by how we respond to their ideas?

09 Mirror. Which of our own habits, as leaders, make it harder for people to improve their work?

S A Partners — Consulting and accredited training for business excellence.

S A Partners LLP, Cedar House, Hazell Drive, Newport, NP10 8FY

+44 (0)2920 881014

info@sapartners.com

www.sapartners.com

© 2026 S A Partners