TULSER

No. 15 · © 2026 Tulser

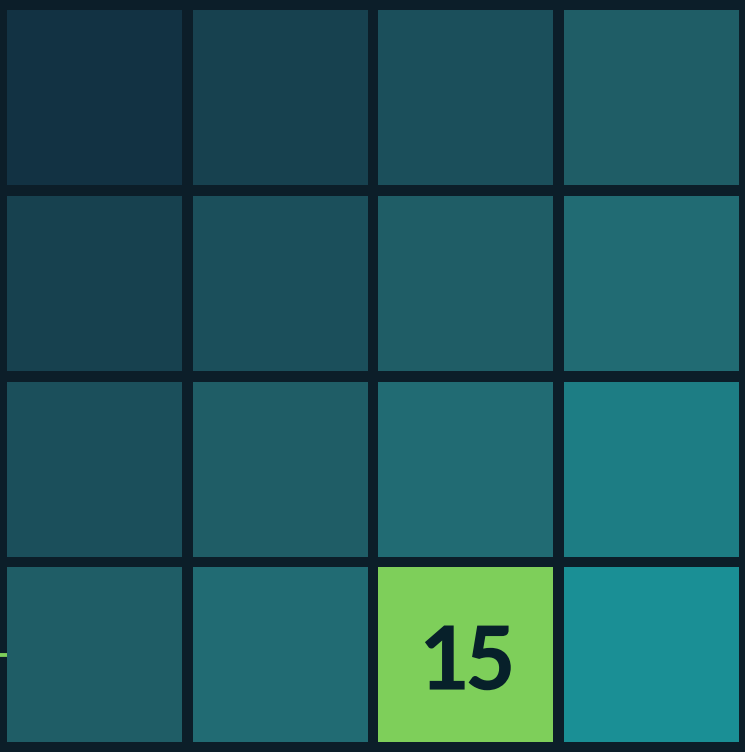
WHITE PAPERS · PERFORMANCE CONSULTING

BOTH DOORS · THE ASSISTANT

A request names a solution before anyone has named the cause.

Performance *augmentation*

AI strengthens what a professional already does. A manager or staff member who owns a business result, and the learning advisor or HR business partner who receives a request for training, can now turn that request into a grounded diagnosis of business, performance and cause, with four reports drafted and ready. What the assistant prepares, what stays the person's to decide and sign, and how it rates its own reliability when the data runs out.



JOS ARETS

with Vivian Heijnen

IN THIS PAPER

The request names a solution.
The diagnosis names the cause.

THE QUESTION

A manager owns a result moving the wrong way, or a learning advisor receives a request for training. Both are about to spend on a solution nobody has traced to a cause. What can an assistant take off their hands, what stays theirs to decide and sign, and how far can anybody trust a diagnosis when the data is thin?

FOR

Managers and staff who own business results, the learning advisors and HR business partners who receive the requests, and the risk, data and legal colleagues asked to sign the capability off.

READ

22 minutes · two worked cases (one at a confidence of 4, one at an 8) and a one-page instrument.

01	The request names a solution	Thirty hotels, no reliable data, and a diagnosis that still moved the decision.
02	Two audiences, one tool	What managers and staff get back, and what L&D and HR get back.
03	What it does, in five moves	From the request to four reports, and how it rates its own reliability.
04	Human in the loop	Everything stands ready; the person decides how, what, and whether to sign.
05	What it delivers	One diagnosis in four reports, the same tool at a 4 and at an 8, and the dividend in both columns.
06	Grounding, governance and three decisions	For the person who signs: the corpus, the structure, the controls.
07	What this does not claim	Four limits, and a one-page test to run before you buy or build.

Training is the small invoice.
The time it takes is the large one.

“Train all thirty hotels” *on data nobody trusts*

A request arrives to roll out service-excellence training across thirty properties before peak season. Guests seem less satisfied, staff turnover is high, and the regional managers say outright that the brand-standard audit scores are gamed and not to be trusted. There is, in their own words, no data anyone would rely on.

	HOW IT GOES NOW	WITH THE TOOL ON THE DESK
The request	“Service-excellence training, all 30 hotels, before peak season.” The date drives everything.	The same request, in a single field. Nobody is told they asked the wrong question.
The data	None trusted. The audit scores are gamed, so the request proceeds on anecdote: some properties are “clearly loved”, others not.	It asks what exists at all, even if flawed: online reviews, OTA scores, complaint volumes. Not reliable data: any data.
The causes	Three functions each blame their own: HR the pay, operations a new reservations system, the VP “inconsistent service culture”.	Those three become candidate causes, each labeled a hypothesis, each with the reason it is only that.
The comparison	·	The loved properties share two visible features: low turnover and long-tenured floor managers. That variation is the strongest signal available.
What goes back	A quote and a plan to train thirty hotels in three months.	Pause the rollout. Establish a baseline and look hard at what the strong properties do: diagnosis rated *4, working hypothesis*.

The same request, read two ways. With no reliable data the tool does not bluff and does not refuse; it names what it can, says how sure it is, and makes the measurement the first task. The figures are illustrative; the analysis shown is a real run of the assistant.

What it did
with nothing solid to stand on

The interesting part is not that the tool found an answer. It is what it did when there was almost nothing to find one with, which is the situation a real analysis is usually in.

It refused to convert anecdote into a finding. The environmental causes (a central-reservations system double-booking at peak, below-market pay at the high-turnover sites, unstable floor management, the franchise structure that resists a mandated rollout) went in as hypotheses, each with its reason for being only that: reported by operations but not yet verified against system data; named by HR but not yet checked against local wage figures. And it put the disagreement between HR, operations and the VP to use rather than resolving it, because a three-way blame is itself a signal about where authority and information sit.

It also named the missing measurement as a cause in its own right. The audit scores are gamed and distrusted, so the group cannot locate the gap, track a change, or make the case for any fix, which is underperformance sustained, not a data inconvenience.

> **AND THIS IS WHERE THE ASSISTANT STOPS**
>
> It set the training request beside the turnover and the reservations data nobody had joined up, and it rated the whole diagnosis **4: working hypothesis**, because the baseline was missing. **It did not decide the pay was the cause.** A person has to spend a day at the two loved properties, sit with the floor manager, watch a handover, and find out whether the strong sites carry a habit the others could copy. What the machine removed was the weeks that used to sit between the request and that visit, which is exactly why the visit so rarely happened.

So the honest first deliverable is not a training plan and not a verdict. It is a measurement the group does not yet have, and a short diagnostic at the best and worst properties before a single euro is committed. That is a smaller, truer answer than the one that was asked for, and it is available in an afternoon rather than a quarter. And because a run costs nothing, the same analysis is worth making on the request that turns out to be justified: a yes with evidence attached is worth as much as a no.

The same instrument, *on either side of the same reflex*

AI does not replace the professional's judgement; it strengthens what they already do: augmentation, in the sense of the series this number grows out of (Arets, 2023). Two audiences meet the same reflex from opposite sides, and the same tool serves both.

IF YOU ARE	WHAT YOU CAN HAND BACK	WHAT YOU STOP HAVING TO BE
A manager or staff member who owns a business result	The cause of your own number, established from data the organization already holds, in an afternoon rather than a quarter, and a remedy chosen against that cause rather than against the loudest theory in the room.	The person who accepts an expensive fix because there is no time to test whether it reaches the cause.
A staff specialist quality, safety, OD, improvement	A gap cut three or more ways, so the argument is about where it sits rather than about who is senior, plus the unit inside your own organization that does not have the problem.	The person whose analysis is weighed by their rank rather than by their evidence.
A learning advisor	The result behind a training request, the condition that is binding, and the remedies that can reach it, which is sometimes, and defensibly, a course.	The person who either says yes to something that cannot work, or says no with nothing to say it with.
An HR business partner	The measure the director is already reporting, and what stands in front of it, so you arrive with the same kind of object everybody else brought.	The one taking the order in a room where everyone else brings figures. The destination piece of this series has the four positions this moves between.

THE CLAIM, STATED EXACTLY

Not that anyone becomes a consultant by being handed a tool. What they get is the method, available when the request lands, with the retrieval and the arithmetic already done. **The judgement stays theirs**, which measure is right, which condition is binding, what the missing data means, and whether to sign. Section 04 draws that line and says why it does not move.

From the request
to the reports, in five moves

The tool guides a conversation, not a form. The hotel request from page one, carried through the five moves, and what the person sees on the screen at each.

THE MOVE	WHAT IS ON THE SCREEN
1 · The request	A question in plain language, in a single field: no form to fill in first. “Is training the fix for our falling service scores?”
2 · The guidance	It asks back, on numbers rather than wishes: is anything measured at all, even imperfectly (reviews, OTA scores, complaints)? Not reliable data. Any.
3 · The data	Documents, dashboards and loose figures go in. Where there are none to trust, establishing a baseline becomes the first deliverable rather than an obstacle: number 06 carries the four moves for exactly this situation.
4 · The gap and its causes	It switches from chat to diagnosis. Each candidate cause is classified by *factor* and *level*, and given a *confidence*: hypothesis, until something verifies it.
5 · The reports	Not a report for a drawer but a presentation, in four versions: Full, Exec, Manager, Slides. The person edits what does not fit and downloads it.

Five moves. The assistant does the retrieval, the classifying and the drafting; every screen above still ends in something a person reads, checks, and can overrule.

How sure it is,
and how it shows a source

Two mechanisms separate this from a system that always sounds certain. The first is that it rates its own diagnosis. The second is how it puts a source in front of you.

A NUMBER FOR HOW SURE IT IS

Every diagnosis carries a rating (in the hotel case, **diagnostic confidence 4, a working hypothesis**) and it names itself for what it is: the most plausible explanation built from what is available, coherent but not yet triangulated against hard data. It even says what would lift the number, and it is not more interviews: it is one focused round of data collection, a satisfaction proxy pulled, two strong floor managers observed, the double-booking rate obtained, any of which raises confidence within a week. The scale runs from 2, a provisional framing, to 10, triangulated; section 05 shows the same instrument reaching an 8 on a case with two years of data. Once the data is entered the full report is there in minutes; when the figures cannot be got, the number says so rather than a confident diagnosis pretending. A high score on thin data is not on offer, and that is the point.

The second is the source. Take a claim an analysis might produce: *orders above the approval threshold wait an average of eleven hours.* One tool ends that sentence with a reference; *Source: Order policy, section 4.* Another ends it with the line it actually read: *"Orders exceeding €25,000 require regional approval; approval is available between 09:00 and 13:00 on working days."* The first tells you where it says it got the sentence; the second lets you see whether the source says it, and here it does not, because the policy sets the threshold and the window and says nothing about eleven hours.

When the data runs out, it does not guess.
It tells you how sure it is.

Everything stands ready.
The person decides how and what

Eight tasks in one analysis. The assistant prepares all of them and presents the result; the line between machine and person is not about current capability and will not move as models improve, because it is drawn on who holds the consequence.

THE TASK	THE ASSISTANT PREPARES	THE PERSON DECIDES
Ordering the analysis	Holds the sequence, refuses to run ahead of it, and says which question is unanswered.	Whether to run it, and when a blank is a finding worth reporting on its own.
Testing a request	Asks the four questions from number 12 and records the blanks.	Has the conversation. The four questions are easy; the room is not.
Working the numbers	Computes the gap, cuts it three ways and recomputes every figure in the reports.	Which measure is the right one, which nothing in the data can settle.
Reading a process	Summarizes an event log, names the queues, reports variants and conformance.	Goes and looks at the work, because the log records what the system was told.
Naming candidate causes	Proposes causes with a factor, a level and a confidence, each with the evidence for it attached.	Establishes which is binding, by asking the performer. This is the part that does not transfer.
Matching remedies	Rules out remedies that cannot reach the established cause, and names alternatives.	Chooses among the survivors, on cost, appetite and who will actually do it.
Choosing the report	Drafts all four versions (Full, Exec, Manager, Slides) ready to present.	How much to show whom, what to cut, and whether to sign it. The signature is the point.
Remembering	Retains what previous analyses established here, so the next one starts higher.	What was learned, and what was an accident of one case.

The assistant does the parts that are laborious and rule-governed; the person does the parts where being wrong has a cost somebody bears.

The reason *for that line*

The line is not drawn out of caution, and it is not a concession to today's technology that will move as models improve. It is drawn on who bears the consequence of being wrong.

One distinction sharpens where it falls. An assistant can be **active**: it answers what it is asked, retrieves, computes, drafts on demand. It can also be **proactive**: noticing that a question is still unanswered, that the measure changed definition in April, that this organization established the same cause eighteen months ago. Proactive is where most of the value sits, and also where the temptation lies, because an assistant that anticipates well starts to look like one that could decide. It cannot, for the reason below.

An analysis that lands on the wrong cause costs an organization a program, a quarter and some of its patience with the function that proposed it. The decision that follows is management's, or the team's, but in either case a person's, which is Deming's point about where responsibility for the system sits, and it does not move because an instrument got better (Deming, 2000). Everything in the right-hand column of the table has that property: the choice of measure, the judgement that a condition is binding, the decision to report a finding a director will not want, the signature. Those are not the hard parts because they are computationally hard. They are hard because being wrong about them has a cost that falls on a person.

> **WHICH IS WHY IT IS BUILT TO BE LESS PERSUASIVE THAN IT COULD BE**
>
> It shows the passage rather than the summary. It states what it is unsure about in the same breath as the claim, and rates the whole diagnosis when the data is thin, rather than in a footer nobody reads. Where two sources disagree it reports the disagreement instead of resolving it silently. And it refuses to write a conclusion the evidence does not carry, which is the behavior most likely to annoy the person using it and the one most worth keeping.

One diagnosis, *four reports for four rooms*

The largest thing it delivers is not a saving; it is a grounded diagnosis. It holds the whole method, will not run past an unanswered question, and states its confidence in what it has produced. That diagnosis comes out in four forms, one for each room it has to enter.

THE REPORT	FOR	WHAT IT LEADS WITH
Full	The analyst, and the record.	The forming diagnosis: the gap, and each cause with its factor, level and confidence.
Exec	The person who signs.	The bottom line, the money, the decision, with the diagnostic-confidence rating in plain sight.
Manager	The person who acts.	What to do next, this week, in numbered steps, and what to watch for.
Slides	The room.	The deck, ready to present. In the hotel case: three causes, one wrong fix.

Same diagnosis, four framings. The person chooses which room to enter, edits what does not fit, and signs: the tool drafts, it does not decide what to show whom.

At a four and at an eight, *the same instrument*

The hotel group scored a 4 because there was almost nothing to stand on. Pointed at a global specialty-chemicals group (17,800 people, 31 sites, two years of records) the same instrument lands at an 8. The distance between those two numbers is what the rating is for.

	THE HOTEL GROUP	THE CHEMICALS GROUP
The request	Service-excellence training, thirty hotels, before peak season.	A €14.6M ($16.6M) two-year culture transformation (workshops for 2,300 managers, coaching, speak-up training) because "people are afraid to challenge their managers."
The data	None trusted. The audit scores are gamed; the gap is anecdote.	Two years of indicators across 31 sites, over 300 structured interviews, a 140-person management survey, a joint review of all 19 serious events, a completed pilot.
The diagnosis	Environmental causes, every one of them labeled a hypothesis.	Structural and process causes, each with its factor and level, and one finding established outright: knowledge is not the problem.
Confidence	4: working hypothesis.	8: substantiated. The report names what keeps it off 10, and the two steps that reach a 9 in four to six weeks.
The first deliverable	A measurement the group does not have.	Redirect the majority of the €14.6M to structural and process interventions; keep a smaller leadership component for after those conditions change.

The number is not a verdict on the tool; it is a verdict on the data. Same method, same discipline, what changed is what the organization could put on the table. In both cases the figures are illustrative and the runs are real.

The scale, as the tool states it, runs in five steps: 2, a provisional framing (the problem framed, little or no data yet; 4, a working hypothesis) the most plausible explanation from what is available; 6, indicative (the direction clear, worth acting on with caution; 8, substantiated) strong evidence across multiple sources, one focused step from 10; and 10, triangulated (hard numbers, qualitative evidence and independent sources confirming each other). An 8 is strong enough to act on. What the two steps toward a 9 buy is precision, not direction.

The eight, in detail
and what it ruled out

A culture program was on the table because people were afraid to challenge their managers. The evidence read differently: the fear was real, and it was downstream of structures nobody had put beside each other.

The consequence structure rewards production over risk prevention: plant-manager pay weights EBITDA at 35 percent and output at 25, with safety at 10; of 26 leaders promoted in three years, 22 had top-half financial results and 9 had top-half safety-climate scores. Corrective-action management has effectively broken down: open actions grew from 4,800 to 13,700, average closure from 41 to 96 days, and 15 of the 19 serious events involved a known abnormality. Near-miss reporting fell by more than a third while employees still identified 92 percent of deliberately planted anomalies. People see the risks. Reporting them stopped paying.

And one finding carries the label **established** rather than hypothesis, triangulated across simulations, incident reviews, assessment records and the interviews: **knowledge is not a material cause.** Training completion runs at 99.3 percent, practical assessments at 94.7, and in simulated abnormal conditions 91 to 97 percent of operators act correctly. The proposed workshops address a knowledge framing of a problem that is not a knowledge problem: the argument of this whole series, in the tool's own labels, in the client's own records.

The comparison cuts both ways. One plant under the same corporate rules keeps justified stop-losses out of shift results and closes high-risk actions in 19 days against a company median of 83: near-misses up 28 percent, serious events down 43. And a fourteen-month pilot of exactly the proposed workshops lifted survey scores nine to fourteen points while shutdowns rose 7 percent: they listen better, said an employee, and the operating constraints have not really changed. The learning measure moved and the business measure did not: number 11's distinction, in the client's own words.

> AND THE DIVIDEND, IN BOTH COLUMNS THIS TIME
>
> The **Performance Dividend Calculator**, number 13 builds it in full, could size only one column for the hotel group, and said so. Here it has both: the €14.6M ($16.6M) that need not go to workshops that cannot reach the causes, and a gap costing **€118M ($134M) a year** whose attainable share is anchored on what the outlier plant already achieves. Those are the two columns of number 13 (the spend that stays, and the result that returns) produced by the tool, before anyone signs.

What the room sees, *from the deck the tool drafted*

The Slides version of the same diagnosis, as it left the tool. Two slides carry the whole argument, and they are worth describing, because the room decides on them.

SLIDE ONE · THE QUESTION

"Fear culture, or structural failure?" The request arrived framed as culture; the title slide turns that framing into a question the evidence can answer, which is the analysis's first act of discipline in front of an audience.

SLIDE TWO · THE ANSWER

Employees know what ought to be reported, and see it happen: the client's own survey percentages sit on the slide. They choose not to report, because reporting stopped paying. A system problem, not a skills problem: the argument of this series, in the deck's own words.

The person presenting edits what does not fit, and signs. That is the division of labor in one scene: the assistant drafted the deck and defended every figure on it; the human decided what the board would see.

It prepares everything.
The person decides what to present, and signs.

Grounded in sources, *or not at all*

From here on, the paper is for the person who signs. One design decision governs everything else, and it is not a technical preference. It is whether the assistant answers out of a model's absorbed training, or out of documents it retrieves and can point at, sentence by sentence.

The first produces fluent, confident, plausible text with no way for anybody to check it. Applied to a method, where the whole value is that a finding can be traced back to a record, that is not a weaker version of the same product. It is a different product, and an actively dangerous one, because a fast wrong diagnosis travels further inside an organization than a slow right one.

So the assistant answers only from a defined corpus: the method as this series sets it out, the reference library behind it, and the organization's own material where it has been made available. It may reason, and it may not invent, and where the corpus has nothing to say, the honest output is that it has nothing to say.

> **WHICH MAKES THE CORPUS THE PRODUCT**
>
> Everything that matters about such a system is upstream of the model: which documents are in, who maintains them, when they were last checked, and what happens to an answer when a policy is superseded. That is unglamorous work and it is where the quality lives. A vendor demonstration shows you the model; the question to ask is about the corpus. And the choice of underlying model is a safety decision rather than a procurement one: retrieval does not repair a weak reasoner, it hands one more material to be confidently wrong about.

A method is a structure, *not a pile of documents*

What sits behind the assistant matters more than the model in front of it, and it is not a folder of white papers. A method is a structure of dependencies, which is why the tool classifies a cause by *factor*, physical resources, structure and process, information, and by *level*, rather than describing it in prose. Those are Gilbert's conditions and Rummler's levels, made operational.

The third question cannot be answered before the second, because it needs the measure. A factor rules certain remedies out and leaves others standing: Gilbert's structure, a set of relations rather than a list (Gilbert, 1978). A cause at one level cannot be reached by an intervention at another, which is Rummler and Brache's point about levels (Rummler & Brache, 2013). Those are relations between objects, and a system that stores only the text has thrown them away, which is why an assistant built on a document search can quote this series accurately and still let a user skip the step that made the quotation mean something.

WHAT THE ASSISTANT IS ASKED	WHAT ANSWERING IT REQUIRES
"What is the desired level in this contract?"	A lookup. One document, one passage, one answer, and most questions in a real analysis are this.
"Given what we have established, what may we conclude?"	A walk over the structure: which questions are answered, which factor that leaves standing, and which remedies that rules out.
"Has anybody here established this before?"	Memory of previous analyses in this organization, held as findings attached to causes rather than as documents in a folder.

Three kinds of question, two kinds of retrieval. Ask any product the second one; the first is easy and every product does it, and the second is where a method either survives the move into software or quietly does not.

Governance, *in three decisions*

Anything of this kind used in professional decisions now sits inside a governance frame, and in Europe inside a regulatory one still being phased in. The paperwork will be substantial. The substance is three decisions.

THE DECISION	WHAT IT MEANS IN PRACTICE	WHAT IT ANSWERS A YEAR LATER
Keep the log	Every answer retains what was retrieved, which version of the corpus and which model produced it, and what the person did next.	"On what basis did we conclude that?": asked after somebody has left, about a decision that affected people.
Show the passage	The source sentence travels with the claim, in the answer. Not a link, not a footnote, not an appendix.	"Did the source actually say this?": checkable in ten seconds rather than in an afternoon.
Name who signs	One person owns the analysis and its consequences. The assistant is an instrument they used, and the report says so.	"Who decided?", and the answer is a name rather than a system.

Three decisions. An organization that makes them is most of the way to any documentation a regulator will ask for; one that does the documentation without them has a folder rather than a control.

THE ORDER TO DO THIS IN

Design for the three decisions first and read the regulation second. The European rules for systems of this kind arrive in stages and the timetable has already been amended once, so check the current dates rather than quoting a presentation, but the direction will not change: traceability, an explanation a person can understand, and accountability that rests with a human being. Done the other way round it becomes a compliance project and leaves the actual control, a person who can check a claim against its passage, unbuilt.

What this *does not claim*

Four limits, and the tool states the first itself: it delivers diagnostic advice, not a guarantee of results.

- **It replaces neither the L&D team nor the consultant.** It makes the diagnosis routine instead of weeks of specialist work. The method is fifty years old; what is new is the access, not the analysis, and a person still has to establish which cause binds and own the decision.
- **It is only as good as the data it is given.** With the figures in place the reports are ready in an afternoon; without them the assistant asks for more, and where they cannot be got it rates its own confidence down rather than guessing. A high score on thin data is not on offer.
- **It is far too early for results claims.** Assistants of this kind have existed a handful of years. There is no synthesized evidence about what they do to the quality of professional analysis, and this paper makes none, what it sets out are design commitments, checkable today, and an account of our own use, which is experience and not evidence.
- **The part that matters most cannot be automated.** Establishing which cause is binding requires somebody to watch the work and ask the person doing it. That does not transfer, and no amount of retrieval reaches it.

SUMMARY

The method is sound, fifty years old, and was rationed because it cost a week and a practitioner. That constraint has lifted: the assistant carries the retrieval, the sequence and the memory, drafts four reports, and rates its own confidence when the data is thin. Every judgement somebody has to answer for stays with a person. The change worth reporting is not the days saved. It is that a grounded diagnosis is now within reach of the people who receive the requests and own the results, which is where it should always have been.

RUN IT YOURSELF

Six lines,
before you buy or build anything

Thirty minutes, on any assistant being proposed to you: ours included. Line three is the one that separates a product from a demonstration.

1 · What does it retrieve from

Which corpus, maintained by whom, updated when? An assistant answering from the model's training data is a different product with a different failure rate.

2 · Show me a wrong answer

Ask the vendor for one. A vendor who has never seen their system be wrong has not measured it, which is itself the answer.

3 · Does a claim carry its passage

Not a link to a document: the sentence the claim rests on, in the answer. Then check three of them by hand against the source.

4 · What does it do on thin data

Ask for a diagnosis it cannot support. A system that produces a confident one anyway, rather than asking for more or rating itself down, is not safe at any accuracy rate.

5 · What is logged, and who can read it

Retrieved sources, model version, what the user did next. If a decision is questioned in a year, what can be reconstructed?

6 · Who signs

Name the person who owns the consequence of a wrong analysis. If the answer is "the system", stop here.

WHERE THIS PAPER SITS

Two problems, *one method, sixteen papers*

	BUSINESS PROBLEM	TRAINING REQUEST
Owned by	The manager or director who owns the result.	L&D and the HR business partner.
What it prevents	Spending on the wrong intervention: a system, a reorganization, a hire.	Spending on a program that cannot reach the cause.
What it returns	The fix that moves the number, chosen on evidence.	The budget and the working hours that were not spent.
Method used	The same three analyses and five questions, which is why one capability serves both, and why the diagnosis no longer has to be bought in case by case.	

One method, two returns, and a standing capability instead of a standing consultancy line.

Two models carry the method. Gilbert's six conditions ask what kind of cause; Rummler and Brache's three levels ask where it sits. Both are sound and this series uses both, and it leads with Gilbert, because his unit is the accomplishment, which is what an organization already counts, and because his six conditions are the categories an organization can actually change.

Where the method *comes from*

Two further pieces close the set and carry no number, because each points at all the numbered papers. Before the guide sits Arets L&D Business Models, the destination piece for the people who receive the requests: four models a learning function can run, and the route from HR-driven to business-driven. Last sits Solutions that cannot work, the field guide, ordered by the sentence that lands on your desk: seventeen forms a performance problem arrives in, each with an address in one of the numbered papers. Three orders in all: the destination you choose, the question you are on, and the form that landed on your desk.

HOW THIS SERIES CAME ABOUT

Human performance technology travels the way a craft does: from hand to hand. Vivian Heijnen and Jos Arets learned the analysis from Dennis Mankin, in San Diego, more than twenty years ago, and have kept in contact with him ever since. Both remain grateful to him: for the rigor of the method as he taught it, and for the leadership he has given performance consulting over a long career.

What followed was twenty years of use. Vivian took the method into organizations in the Netherlands and well beyond them, and what accumulates over that many analyses is largely tacit: the cut nobody thought to make, the question that opens a defensive room, the finding that turns out to be an artifact of the data. Making that knowledge explicit, case by case, is what this series is for. Jos wrote it down as it went, in two books and more than forty articles and blog posts, most of them in Dutch, and one seven-part English series, *Is AI the bicycle of the mind?*, which number 15 grows out of. The books: *Liever (g)een training* with Boudewijn Overduin, *Kostbaar misverstand* with Vivian Heijnen, which this series takes a good deal of its thinking from.

None of that is reprinted here. It has been reworked: brought up to date, tested against the research that has appeared since, and dropped where the evidence no longer carries it. Theory that has never met a client is a position; practice that cannot account for itself is a story. This series is what remains when each is held to the other. Performance Consultant AI carries the same discipline to the desk where the decision is actually made.

About the cases. Every worked case in this series comes from our own practice. Names are changed and details altered so that no organization is recognizable, and the arithmetic reconciles. That is said once, here, rather than under each case: client work is confidential, and a case that named anybody would be a breach rather than a proof.

THE SERIES

Sixteen papers, *six parts*

Number 01 is the map and number 02 is the foundation. From number 05 on, every number is depth on exactly one step of the method. The filled block is the paper you are holding; the two unnumbered pieces close the set.

I · WHAT HPT IS

The discipline, where it came from, and why the reflex it replaces survives.

- 01 The performance improvement process
- 02 Human performance technology
- 03 Environment before capability
- 04 The system and the performer

II · THE BUSINESS ANALYSIS

Question 1.

- 05 Business needs and performance needs

III · THE PERFORMANCE ANALYSIS

Questions 2 and 3.

- 06 Gap analysis
- 07 The exemplary performer

IV · THE CAUSE ANALYSIS

Questions 4 and 5.

- 08 Cause analysis: the three levels
- 09 Cause analysis: the six conditions

V · THE DECISION

From cause to solution, and what counts as proof.

- 10 Intervention selection
- 11 Learning metrics and business metrics

VI · THE FUNCTION

The request, the money, the implementation, the assistant and the value.

- 12 The solution request
- 13 The cost of non-performance
- 14 Implementation and maintenance
- **15 Performance augmentation**
- 16 The Value Flip Calculator

Arets L&D Business Models · the destination piece, after 16
No number: the four positions a function can hold, and the move the sixteen papers make simple.

Solutions that cannot work · the field guide, at the end
No number: it points at all sixteen. The same material, ordered by the sentence that lands on your desk.

Business problem: a result is moving the wrong way · Training request: a solution has already been chosen · Both doors: after the decision

REFERENCES

Where every claim *in this paper comes from*

Arets, J. (2023). *Is AI the bicycle of the mind?* [Seven-part series on AI, learning and performance]. eLearning Industry. https://elearningindustry.com

Gilbert, T. F. (1978). *Human competence: Engineering worthy performance.* McGraw-Hill.

Deming, W. E. (2000). *Out of the crisis* (2nd ed.). The MIT Press. (Original work published 1986)

Rummler, G. A., & Brache, A. P. (2013). *Improving performance: How to manage the white space on the organization chart* (3rd ed.). Jossey-Bass. (Original work published 1990)

This number deliberately cites no research about assistants of this kind. The published work is a few years old, it is not yet synthesized, and the figures in it belong to product versions that no longer exist. Number 02 sets out the evidence rule this series follows and why; the three books named earlier are listed in full there.