A TECHNOLOGICAL AND INTELLECTUAL HISTORY

History of Diagnostics, Observability, and Debugging in Management

Managerial evidence and the examination of correction

Seven parts and 35 chapters

Including a chapter on Critique of Management Reason:
Management Bits and Tips from Can't

Concept by Dmitry Vostokov, DumpAnalysis.org

Developed with AI assistance (GPT-6 Astra Max and Fable 5.1 Extra) and checked against the sources listed in this article.

Version 5 | November 2026

Historical synthesis with 74 references,
chronology, glossary, and evidence taxonomy

Abstract

Management leaves artifacts. A ledger records a transaction. A production chart represents work that should have been completed. A meeting minute preserves a decision but may omit the disagreement preceding it. A dashboard displays a result whose definition has changed since the previous reporting period. In this history, we examine how such artifacts became available for analysis, how managers and other participants investigated discrepancies, and how they attempted to correct the conditions that produced them.

We follow selected developments from administrative records and accounting through scientific management, human relations, statistical quality control, operations research, organizational diagnosis, cybernetics, and organizational learning. Later chapters consider performance measurement, accident investigations, process mining, software operations, evidence-based management, and algorithmic management. These developments overlap. The introduction of a computer did not remove the need to compare reports, and the introduction of a new management theory did not remove the consequences of an old incentive.

The main distinction is between an organizational condition, the evidence available about that condition, and the explanation constructed from that evidence. A failure can occur at any of these levels. We may correct a report while leaving the work unchanged. We may improve a local measure while making the overall service worse. We may also obtain a better explanation without yet having the means or authority to implement a remedy. Diagnostics, observability, and debugging help us keep these activities separate and examine their relationships.

Seven parts and 35 chapters are followed by a selected chronology, glossary, evidence taxonomy, and references. The final part connects the history with our pattern-oriented diagnostics, includes a worked organizational investigation, and discusses my book Critique of Management Reason: Management Bits and Tips from Can’t. The discussion also establishes the relationship between that later title and the earlier Management Bits project cited by Jerzy Kociatkiewicz and Monika Kostera in their 2012 paper on Sherlock Holmes and managerial rationality.

How to read this history

Let us examine management through five questions: how does an organization notice a problem, preserve its evidence, construct explanations, intervene, and check the consequences? These questions connect distant practices without making them historically identical.

We use diagnostics for investigating and interpreting a condition, observability for the arrangements that make relevant evidence available, and debugging for correcting a specified defect and checking the result. Diagnosis is established organizational vocabulary. Observability and debugging also serve here as comparative terms; they do not imply that an organization is a program or that every internal state can be reconstructed.

Managers are not the only diagnosticians. Workers, customers, auditors, researchers, trade unions, and affected communities can identify conditions missing from a management report. The selected history

includes several institutional and geographical settings, without claiming a complete world history or a single line of descent.

Dates normally identify original publications or events. References distinguish translations, reprints, and repository deposits. A citation to a projected book does not establish that the edition was published. This distinction matters for the earlier Management Bits project discussed in Chapter 31.

Hypothetical examples and the worked investigation are constructed illustrations, not observations from an employer. Connections to software analysis patterns are our analytical comparisons. They suggest questions to investigate; the evidence must establish whether a proposed mechanism applies.

Neighboring practices and their questions

Practice	Main question	What the result alone cannot establish
Monitoring	Has a selected measure changed?	Why it changed or whether its definition remained stable
Diagnosis	Which explanations fit the condition and evidence?	That the proposed remedy will work
Audit	Is a specified assertion or procedure supported?	That the organization achieves every relevant purpose
Evaluation	What consequences followed an intervention?	That the result transfers unchanged to another setting
Debugging	Which specified defect was corrected and checked?	That every disagreement or future failure has disappeared
Organizational learning	Did subsequent action incorporate what was learned?	That publishing a lesson changed the operating routine

Contents

Seven parts and 35 chapters

SECTION	PAGE
Abstract	2
How to read this history	2
PART I. RECORDS AND THE EMERGENCE OF MANAGERIAL INQUIRY	6
1. Vocabulary and the object of diagnosis	6
2. Administrative records and the persistence of an event	7
3. Double entry and the examination of accounts	7
4. Industrial work and scientific management	8
5. Administration and formal authority	9
PART II. WORK OBSERVATION AND IMPROVEMENT	11
6. Motion studies and the visibility of sequence	11
7. Hawthorne and the diagnosis of observation	12
8. Cooperation and the authority of the situation	12
9. Variation and statistical quality control	13
10. Toyota and the visibility of abnormal work	14
PART III. DECISIONS AND THE FORMULATION OF PROBLEMS	16
11. Operations research and the examination of a model	16
12. Bounded rationality and organized choice	17
13. Objectives and the meaning of performance	17
14. Structured problem analysis and comparative evidence	18
15. Problem formulation and contested purposes	19
PART IV. ORGANIZATIONAL SYSTEMS AND LEARNING	21
16. Action research and sociotechnical work	21
17. Organizational diagnosis as an explicit practice	22
18. Cybernetics and distributed regulation	22
19. System dynamics and delayed consequences	23
20. Learning and the correction of governing assumptions	24

SECTION	PAGE
PART V. MEASUREMENT AND THE EXAMINATION OF FAILURE	26
21. Management information and the work of attention	26
22. Scorecards and different uses of control	27
23. Incentives and the transformation of a measure	27
24. Audit and administrative visibility	28
25. Accidents and the organization of unwelcome evidence	29
PART VI. DIGITAL PROCESSES AND EVIDENCE BASED INTERVENTION	31
26. Process redesign and event records	31
27. Software projects and operational learning	31
28. Evidence based management and field interventions	32
29. Narratives and the reconstruction of organizational events	33
30. Algorithmic management and the diagnosis of automated judgment	34
PART VII. MANAGERIAL REASON AND PATTERN ORIENTED DIAGNOSTICS	36
31. Critique of management reason and the rational manager	36
32. Designing organizational observability	37
33. Connections to trace and log analysis patterns	38
34. A worked investigation of apparent improvement	39
35. Continuing developments and the scope of this synthesis	40
Conclusion	42
Appendix A Selected chronology	43
Appendix B Glossary	46
Appendix C Evidence taxonomy and blind spots	48
References	50

PART I

Records and the emergence of managerial inquiry

1. Vocabulary and the object of diagnosis

Let us consider a service whose monthly report shows an improvement while customers report longer waits. Before looking for a responsible person, we need to identify what the report measures. It may count completed cases rather than received requests. It may exclude cases waiting for another department. Its average may conceal a small group of customers whose waiting time increased considerably. The discrepancy is already diagnostic evidence, although it does not yet identify a cause.

The object of diagnosis can be the work, its representation, or the managerial explanation. A delayed delivery belongs to the first. An incorrectly calculated delivery measure belongs to the second. The claim that all delays result from insufficient effort belongs to the third. These objects interact, but correcting one does not automatically correct the others. A sound investigation keeps a record of which object is being examined at each step.

This distinction also separates a symptom from a mechanism. A missed deadline is an outcome relative to an agreed date. It can result from an unrealistic estimate, an unavailable input, a change in scope, a hidden dependency, or a combination of conditions. A familiar label such as poor communication may merely rename the symptom unless we identify the particular message, recipient, timing, and decision affected by it.

Organizational diagnosis developed explicit frameworks for investigating such relationships. Weisbord's 1976 model, for example, organized attention around purposes, structure, relationships, rewards, leadership, and helpful mechanisms. Its purpose was to make inquiry manageable, not to provide a universal causal answer from six headings. The existence of this literature matters historically: management did not need to borrow the entire idea of diagnosis from software engineering. [1]

We can nevertheless bring a useful distinction from software diagnostics. An analysis pattern describes a recurrent arrangement in evidence, while a defect mechanism explains how a particular failure can occur. Seeing a repeated escalation pattern does not establish why escalation repeats. The next investigation may concern an authority boundary, a missing skill, an incentive, or a customer requirement that the process cannot satisfy. Our pattern vocabulary should preserve these alternatives. [2]

There is also a question of purpose. An arrangement may satisfy a budget while imposing unreasonable waiting or unsafe work. A manager's preferred objective is evidence about that manager's decision, but it does not settle the legitimacy of the objective. This history therefore includes disputes about what counts as failure. Debugging becomes meaningful only after we state the defect precisely enough to distinguish an observable correction from a change in the story told about it.

2. Administrative records and the persistence of an event

An event becomes available for later inquiry only if something of it persists. A delivery may leave a receipt. An allocation may leave an account. A verbal instruction may leave no independent record at all. Administrative writing changed the temporal reach of coordination by allowing a later reader to compare obligations and recorded actions without having witnessed every transaction.

Surviving Mesopotamian tablets provide concrete evidence of this capacity. The Metropolitan Museum's Ur III tablet described as the balanced account of Dugga records transactions involving sheep, ewes, and goats and dates to about 2039 BCE. The Library of Congress cuneiform collection spans a wider period and also contains accounting records, including a balanced barley account. These artifacts document particular administrative practices; they do not justify attributing a complete modern management system to their producers. [3, 4]

For our purposes, the diagnostic operation is comparison under a shared representation. Once quantities, persons, and dates can be related, a reader can ask whether recorded receipts and distributions agree. The record can preserve an obligation beyond the memory of one participant. It can also permit someone distant from the work to exercise authority over it. Persistence and control develop together, although they are not the same function.

The surviving account is selective. A quantity of grain says little by itself about the circumstances of its production, the fairness of its allocation, or the experience of the people involved. A discrepancy may concern arithmetic, physical loss, a different unit, or an omitted transaction. The diagnostic problem is therefore partly semantic even when the marks represent numbers. We need to know what was counted and what relationship the account intended to preserve.

Consider a hypothetical store record that contains an opening quantity and several issues but no final physical count. The arithmetic can be internally consistent while the physical stock is different. Conversely, a surviving fragment can appear inconsistent because another part of the record has been lost. These are different limitations: the first concerns verification at the time of administration; the second concerns the evidence available to a later investigator.

We can call the account an early form of organizational memory, provided that we keep the comparison modest. It preserves selected states and transitions. It does not preserve the whole organization. The historical importance for diagnostics lies in the possibility of returning to an artifact, relating it to another artifact, and challenging a claim with evidence that outlasts the immediate exchange.

3. Double entry and the examination of accounts

Accounting gave administrative comparison increasingly explicit structures. The diagnostic significance of double entry lies in the relationships it requires among entries. A transaction is represented through

corresponding effects rather than left as an isolated statement. Those relationships can expose certain inconsistencies and support the reconstruction of a financial history.

Luca Pacioli's Summa, printed in Venice in 1494, included a description of the bookkeeping practices associated with Venetian merchants. The ICAEW historical collection identifies it as the first printed exposition of double-entry bookkeeping. Pacioli described an existing practice; calling him its inventor would confuse publication with origin. His account helped make a method transmissible through print. [5]

The distinction matters beyond accounting history. A published management method can standardize language, training, and review without having created every practice it describes. Publication also creates a reference against which later users can compare their own arrangements. It makes some deviations visible and may encourage others to disappear from the official representation.

A balanced account is a constrained representation, not a certificate that every underlying event occurred as described. A transaction can be omitted in a way that preserves balance. A false transaction can be recorded consistently. A correct quantity can be assigned to the wrong period or category. Arithmetic agreement narrows an investigation but cannot replace examination of supporting documents and the circumstances of the transaction.

In a hypothetical purchasing process, the invoice, approval, delivery record, and ledger entry provide different views of one commercial activity. An invoice alone establishes a demand for payment. An approval establishes that someone authorized an action. A delivery record supports a claim about receipt. Their agreement is informative only when their origins and dependencies are understood. Four documents copied from one incorrect source are not four independent confirmations.

The enduring contribution is an inspectable relationship among artifacts. Later audits, reconciliations, and management accounts extend this operation under more elaborate institutional arrangements. We can recognize a general diagnostic principle here: an invariant is useful because it constrains possible explanations, but its scope must be stated. A ledger can balance while a business fails, and an organization can comply with its recording procedure while misunderstanding the condition that the procedure was intended to reveal.

4. Industrial work and scientific management

Industrial production made the organization of repeated work an explicit object of measurement. The factory brought together machines, materials, skills, schedules, and wage arrangements whose interactions could be compared across tasks and periods. The resulting records supported practical improvements, but they also extended managerial authority over the definition of a task and the evaluation of the person performing it.

Charles Babbage's On the Economy of Machinery and Manufactures, published in 1832, examined machinery, division of labor, costs, and the organization of production. His work connected technical

particulars with economic arrangements. A machine's operation could be studied together with the distribution of skills and the conditions under which its use became economical. This was a substantial extension beyond merely counting finished goods. [6]

Frederick Winslow Taylor's The Principles of Scientific Management appeared in 1911. Taylor advocated systematic study of work, selection and training, cooperation, and a redistribution of responsibilities between management and workers. His historical claims and examples belong to an argument for a particular management program. We should distinguish that program from the broader claim that all organizational problems have a single scientifically discoverable solution. [7]

From a diagnostic perspective, the measured task becomes a representation. The investigator decides where the task starts, where it ends, which interruptions count, and what conditions are held comparable. A measured difference may reflect a better method, a different material, prior learning, selection of workers, or a workload transferred elsewhere. The stopwatch does not decide which explanation is correct.

Suppose a hypothetical assembly operation becomes faster after tools are repositioned. This supports a local improvement under the observed conditions. It does not establish that the same arrangement is suitable for every worker or that the resulting production rate is sustainable throughout a shift. The diagnostic record needs the changed arrangement, the work conditions, and the effects that the local timing omitted, including fatigue or later rework where relevant.

Scientific management made the analysis of work visible as a managerial responsibility. It also raised a question that remains with digital management: who controls the representation through which work is judged? If workers can challenge task definitions and contribute knowledge of exceptional conditions, measurement can support inquiry. If the recorded standard is treated as an unquestionable description of reality, the same instrument can prevent the organization from seeing why its own expectations are wrong.

5. Administration and formal authority

As organizations grew, managerial inquiry extended from individual operations to the arrangements connecting them. A production difficulty might result from an unclear responsibility, conflicting instructions, or a planning failure. The organization itself could be examined as a set of offices, relationships, and procedures rather than as a collection of separately measured tasks.

Henri Fayol's administrative work, serialized in French in 1916, issued as a book in 1917, and widely known in Constance Storrs's 1949 English translation, considered planning, organization, command, coordination, and control. His principles supplied a vocabulary for examining managerial arrangements across an enterprise. They were not software interfaces, but they made responsibilities and relationships discussable in a way that supports our present comparison. [8]

Max Weber's analysis of bureaucracy examined authority, offices, rules, expertise, and records within a broader sociology of domination. Economy and Society appeared posthumously in German in 1922. Its treatment of bureaucracy is an analytical account of a form of organization, not simply a recommendation to make every activity more bureaucratic. An ideal type is a conceptual construction used for comparison; it is not an assurance that an actual organization possesses ideal performance. [9]

Formalization can help diagnosis by preserving who was authorized to do what and which procedure applied. It also allows comparison between an intended process and actual conduct. However, the existence of a rule does not show that the rule was understood, workable, or followed. An organizational chart records a formal arrangement, while coordination may depend on relationships that the chart does not display.

Consider a hypothetical exception that requires two departments to approve a customer request. Each department follows its own rule, yet neither accepts responsibility for resolving disagreement. An inquiry confined to compliance may find no local violation. An inquiry into the overall service reveals an unresolved dependency. The relevant defect may be in the relation between rules rather than in the execution of either rule separately.

The historical development of formal administration therefore gives us two objects of inspection. We can examine behavior against a rule, and we can examine the rule against the purpose and conditions of the work. This second operation becomes increasingly explicit in later organizational learning. A procedure that repeatedly generates the same blockage should itself enter the diagnostic evidence, even when every participant can demonstrate compliance with it.

PART II

Work observation and improvement

6. Motion studies and the visibility of sequence

A total duration can indicate a problem without showing where the time went. Motion studies and graphical planning made the internal sequence of work more inspectable. They represented movements, activities, and progress so that an investigator could return to particulars that would otherwise disappear during performance.

Frank and Lillian Gilbreth's work used observation and recording to study how tasks were performed. Purdue's archives preserve original motion-study films alongside later compilations and related research materials. The distinction between an original film and a later compilation is itself important: the date and purpose of the surviving presentation may differ from those of the recorded work. [10]

Graphical schedules provided another representation. Wallace Clark's The Gantt Chart: A Working Tool of Management, published in 1922, explained charts associated with Henry Gantt for making plans and progress visible. We should distinguish these historical charts from every later software product called a Gantt chart. Their diagnostic value lies in allowing a reader to compare expected and recorded work within a common visual arrangement. [11]

These instruments bring time into organizational evidence in different ways. A film records a sequence for later inspection. A progress chart relates activities to a plan. Neither automatically explains a delay. A repeated pause on film may be necessary checking, waiting for material, or an interruption outside the camera's view. A late bar on a chart does not reveal whether the original estimate was credible.

Let us consider a hypothetical repair operation. The recording shows that a technician repeatedly leaves the bench. A narrow interpretation classifies the movement as wasted effort. A wider investigation finds that the technician is obtaining information that the work instruction omits. Moving the information to the bench may reduce delay; forbidding movement may instead make the repair less reliable. The same observation supports different interventions until the surrounding activity is understood.

The lasting contribution is a recoverable sequence that can be compared with an expected one. This is close to the practical value of a trace, although the mechanisms of recording and the meaning of the events differ. The investigator should preserve the units, boundaries, and omitted regions of the representation. A detailed sequence can still be incomplete, and greater visual detail can increase confidence without increasing the range of explanations actually tested.

7. Hawthorne and the diagnosis of observation

The Hawthorne studies made the relation between work, social conditions, and observation a persistent management question. They also provide a case in which the history of a result became simpler than the surviving evidence. A diagnostic history must examine that simplification as well as the original experiments.

The studies at Western Electric's Hawthorne Works extended from 1924 to 1933. The illumination experiments began in 1924, and the relay assembly investigations began in 1927. Harvard's historical exhibition records several possible influences on the observed productivity changes, including rest periods, hours, incentives, group conditions, and special attention. Treating the entire program as one experiment with one unambiguous result removes these differences. [12, 13]

The familiar claim that productivity improved whenever lighting changed, including when it became worse, requires particular care. Steven Levitt and John List reanalyzed surviving illumination data in their 2011 paper. They did not confirm the conventional account of the supposedly remarkable patterns, although they reported indications compatible with more subtle observation effects. Their analysis concerns those illumination data, not a universal demonstration that observation can never affect behavior. [14]

For management diagnostics, two lessons follow. First, the observer can become part of the conditions being studied. Second, later accounts of an investigation can acquire a stability that the evidence does not warrant. A memorable management lesson may survive because it is easy to teach, while the qualifications and competing explanations are gradually omitted.

Imagine a hypothetical team whose output increases during a pilot. It receives a new tool, additional attention, faster access to specialists, and temporary relief from other duties. Calling the increase a tool effect is premature. Calling it an attention effect is also premature. We need a comparison capable of separating these changes, or a narrower conclusion that the combined arrangement was associated with improvement during the pilot.

The diagnostic object can therefore be the research narrative itself. We compare the published lesson with the original design, the available records, and later reanalysis. The history of management contains many such opportunities. A method's popularity, the confidence of its advocates, and the repetition of its success story are evidence about its reception. They are not substitutes for evidence about the mechanism or the conditions under which a reported result can be reproduced.

8. Cooperation and the authority of the situation

An organization can possess clear instructions and still fail to coordinate. The person issuing an order may lack relevant local knowledge. The recipient may interpret the instruction differently or find that

the situation has changed. The problem cannot always be diagnosed by checking whether a message traveled from a superior to a subordinate.

Mary Parker Follett's discussion of the giving of orders in 1925 proposed joint examination of the situation. Her account linked authority to what the developing situation required and emphasized that acting on an instruction also changes the situation. This gives us a historically grounded alternative to treating an order as a fixed command whose only possible defect is disobedience. [15]

Chester Barnard's The Functions of the Executive, published in 1938, examined formal organization as cooperation sustained through communication, contributions, and purpose. His account of authority gave an important role to acceptance by those receiving a communication. The executive's problem was therefore broader than issuing commands through a formal hierarchy. [16]

These approaches enlarge the evidence available for diagnosis. We need to understand how an instruction was received, what knowledge participants possessed, and which conditions made cooperation possible. A formal acknowledgment may show that a message arrived. It does not establish that the recipient could carry it out or that the instruction remained suitable after circumstances changed.

Consider a hypothetical maintenance team asked to restore service quickly while preserving evidence of an intermittent fault. Immediate replacement of the suspect unit may satisfy the first objective and destroy an opportunity to investigate the second. The conflict is real even if every participant behaves competently. Joint examination can make the competing purposes explicit and determine which evidence must be collected before intervention.

For pattern-oriented inquiry, this is a warning about the interpretation of a blocked activity. A pause may indicate indecision, an unresolved authority boundary, or a deliberate attempt to protect another objective. A useful diagnosis identifies the relation that prevents progress. It should also preserve the participants' account of that relation. Without it, a later report can transform reasonable local action into an apparent error merely because the report assumes knowledge or authority that the participants did not have.

9. Variation and statistical quality control

Repeated measurements vary even when no identifiable local event has changed. A management system that treats every movement as a new failure can produce unnecessary intervention. Statistical quality control made the distinction between routine variation and evidence of a changed process central to practical investigation.

Walter Shewhart's work in the 1920s introduced the control chart, and his Economic Control of Quality of Manufactured Product appeared in 1931. The historical date of the early chart is 1924. Its importance was the organization of observations so that variation could be considered in relation to the behavior of a process, rather than only through inspection of individual finished items. [17, 18]

Control limits and specification limits answer different questions. The former describe a statistical basis for judging process behavior under the chart's assumptions. The latter express requirements. A process can behave consistently while producing unacceptable results. It can also produce an item within specification while showing evidence that the process has changed. Confusing these questions can turn a chart into an attractive but misleading management display. [18]

Deming's later work connected statistical thinking with management and learning. The Plan–Do–Study–Act cycle, associated with his development of Shewhart's ideas, emphasizes a prediction, a change, and examination of what was learned. It should not be presented as a sequence invented in one step under every familiar acronym. The Deming Institute specifically explains PDSA as a learning cycle. [19] Deming's The New Economics for Industry, Government, Education (1993) develops the connection between knowledge, variation, and management. [20]

Quality management also widened beyond the statistical technique. Joseph Juran's Quality Control Handbook first appeared in 1951, and his work connected quality with managerial responsibility. Kaoru Ishikawa helped develop Japanese company-wide approaches and quality circles through which people close to the work could participate in improvement. These contributions enlarged both the organizational scope of quality and the population able to contribute diagnostic knowledge. [21, 22]

Suppose a hypothetical department responds to every weekly rise in backlog by changing the assignment rule. If weekly fluctuations are expected under the existing workload, the changing rules may add variation instead of removing a cause. A chart can help determine whether further investigation is warranted, but the chart cannot establish whether the underlying service design is satisfactory. That requires another question and additional evidence.

The diagnostic advance is disciplined restraint as well as detection. We can avoid assigning a special explanation to every difference while remaining open to evidence of a real change. The associated managerial advance is to examine the process that repeatedly produces a condition. A remedy aimed at one worker or one week may leave that process intact. Statistical evidence helps locate the level at which the next investigation should proceed, provided that measurement definitions and sampling conditions remain available for review.

10. Toyota and the visibility of abnormal work

A production system can expose a problem while work is still in progress or conceal it until final inspection. The location of detection changes the possible response. A defect recognized immediately may permit examination of the operation that produced it. A defect discovered much later may have passed through many activities and accumulated further consequences.

Toyota's account of its production system identifies jidoka and just-in-time as its two central concepts. Its explanation of jidoka includes stopping when an abnormality occurs, and its plant materials describe andon signals through which workers can call attention to a problem. The historical roots include Sakichi

Toyoda's loom mechanisms and the later development of automobile production. These are related developments, not one invention at one moment. [23]

The diagnostic relation is between detection, a visible signal, and an available response. A signal without someone able to investigate it can become background noise. A formal right to stop work can be ineffective if using that right produces punishment. Conversely, a small visible signal can be useful when it brings appropriate help while the evidence is still present.

Consider a hypothetical assembly station whose error is discovered before the part moves onward. The inquiry can compare the part, tool setting, work instruction, and immediately preceding action. If the part is corrected without retaining any record, the next recurrence may be treated as a new isolated event. The signal has helped containment but has not yet created organizational memory.

Toyota's published descriptions also connect flow with the supply of what is needed at the appropriate time and quantity. For our comparison, reducing hidden queues can make dependencies more visible. It does not follow that removing every buffer is always safe. A buffer may absorb variation or protect a critical service. Its purpose and the consequences of its removal require examination in the particular system. [23]

The contribution to management history is an arrangement in which abnormalities can become occasions for inquiry close to the work. Copying the visible board or terminology does not reproduce that arrangement by itself. The organization must connect the signal with authority, technical understanding, and a record of what changed. Otherwise the display can show that a problem was closed while leaving the conditions of its recurrence unchanged.

PART III

Decisions and the formulation of problems

11. Operations research and the examination of a model

Some management problems involve too many interacting choices for unaided comparison. Operations research developed methods for representing such choices, comparing consequences, and allocating scarce resources. Its diagnostic relevance concerns both the organization under study and the model through which the organization is studied.

Wartime operations research brought scientific investigation to operational questions, including search, deployment, and the use of equipment. Its postwar development extended these methods into civilian organizations. INFORMS's historical account places these developments in relation to the emergence of a professional field rather than attributing every analytical method to a single wartime discovery. [24]

George Dantzig's work on linear programming and the simplex method in 1947 provided a means of solving a large class of allocation problems. In his oral history, he described the effort to mechanize planning and the practical difficulties of computation before reliable electronic machines became widely available. The historical problem included formulation and execution of the calculation as well as the mathematical method. [25]

An optimal solution is conditional on a model. Its objective, constraints, variables, and data determine what the calculation can establish. A mathematically correct result can be operationally unsuitable because an important dependency is missing or a supposedly available resource is unavailable in practice. Debugging the implementation of an optimization routine and diagnosing the adequacy of its organizational model are separate investigations.

For example, a hypothetical schedule minimizes overtime by assigning every qualified employee to available work. The model omits the need for experienced staff to supervise unfamiliar tasks. Its solution is feasible within its written constraints but creates delays on the floor. The useful response is to identify the omitted relationship and examine its evidence, rather than conclude that calculation itself is useless or that workers failed to follow an optimal plan.

Operations research also makes counterfactual inquiry more explicit. We can change an assumption and examine the resulting solution before changing the organization. Such analysis is informative when the model preserves the relationships relevant to the question. It is less informative when a small numerical change hides a large institutional change, such as assuming that authority, skills, or cooperation can be transferred instantly. The history of management analytics is therefore also a history of discovering what must remain inside the model and what must be checked outside it.

12. Bounded rationality and organized choice

A decision record often makes a choice look more orderly than the activity that produced it. The record may list alternatives and reasons without showing which alternatives were never considered, which information arrived late, or which participants were absent. The diagnosis of a decision needs more than its final justification.

Herbert Simon's work challenged the assumption that actual decision makers consistently optimize over a complete set of known alternatives. His 1955 article, A Behavioral Model of Rational Choice, examined choice under limits of knowledge and calculation. The associated idea of satisfactory search directs attention to aspiration levels, available alternatives, and stopping conditions. It does not mean that any convenient decision should be accepted as rational. [26]

The implication for observability is practical. We need evidence of the search process, including what the decision maker knew at the time. A later analyst who possesses the final outcome and a complete incident chronology occupies a different position. Judging the earlier choice as though that later evidence had always been available introduces a defect into the diagnosis itself.

Cohen, March, and Olsen's 1972 garbage can model examined organizational choice under problematic preferences, unclear technology, and fluid participation. Problems, solutions, participants, and opportunities for choice can meet without following an orderly sequence from a clearly defined problem to a designed solution. The model is a particular account of organized anarchy, not a claim that every decision is random. [27]

Consider a hypothetical organization that purchases a reporting platform before agreeing which operational question it needs to answer. Later, a backlog problem is framed as a reporting problem because the platform is already available. The sequence can be reconstructed from proposals, meeting records, and procurement dates. That reconstruction may explain why an apparently reasonable solution repeatedly fails to address the experienced difficulty.

We can therefore investigate a decision as a temporally distributed activity. The relevant evidence includes discarded alternatives and changes in participation where those records exist. Missing records should remain missing in the analysis. A coherent explanation created afterward must not be silently substituted for an undocumented earlier reasoning process. Bounded rationality helps us ask what was possible to know and do, while organized-choice models help us examine how the problem itself became attached to a particular action.

13. Objectives and the meaning of performance

Before a deviation can be diagnosed, someone must establish the reference against which it is a deviation. Management objectives supply such references, but they also introduce a problem: an objective can clarify purpose while a narrow measure gradually replaces that purpose.

Peter Drucker's The Practice of Management, published in 1954, included management by objectives and self-control. Its treatment connects a manager's contribution with the larger enterprise and gives a role to the manager's own assessment. The historical idea should not be reduced to the later practice of assigning a numerical quota and calling it an objective. The wording of the original chapter matters because it preserves the relation between purpose, responsibility, and self-direction. [28]

A performance measure represents an aspect of work under a definition. It may be useful even when incomplete, provided that the organization can return to the work behind it. Trouble begins when the measure becomes the only admissible account. A manager can then improve recorded performance by changing classification, shifting work to another period, or transferring difficult cases without improving the outcome that justified the measure.

Suppose a hypothetical support team is asked to reduce the time taken to resolve requests. It can work on technical capability, assignment delays, and the quality of initial information. It can also close requests before customers regard them as resolved. The same reported average may arise from these different mechanisms. Diagnosis requires a stable definition and evidence of the transitions represented by the number.

Objectives can also conflict. A purchasing team may minimize unit price while an operations team needs reliable delivery and a service team needs replaceable parts. These differences are not necessarily signs that one department misunderstood its instructions. They may reveal that the organization has distributed objectives without specifying how their consequences should be reconciled.

The historical development of performance management thus creates an additional object for diagnostic inquiry: the objective system. We examine how objectives were selected, how local measures relate to overall outcomes, and who can challenge the relation. A discrepancy between local success and overall failure is especially informative. It suggests that the next correction may concern the definition or coordination of objectives, rather than increased pressure to achieve the existing measures more efficiently.

14. Structured problem analysis and comparative evidence

A problem description often arrives with an explanation already embedded in it. A statement such as the new team is slow contains an observed condition, a comparison, and an attribution, although none may have been made explicit. Structured problem analysis attempts to separate these elements before selecting a remedy.

Charles Kepner and Benjamin Tregoe published The Rational Manager in 1965. Their subsequent work distinguished the assessment of a situation, analysis of a problem, decision analysis, and anticipation of potential problems. This separation matters because explaining an existing deviation and choosing among future actions require different evidence. A preferred action does not become a cause merely because it is available. [29]

The authorized Problem Analysis chapter from The New Rational Manager describes a deviation through what, where, when, and extent, together with relevant cases where the problem could occur but does not. Differences and changes can then support possible causes, which are checked against the specification. The chapter also cautions that human performance requires adaptations to the questions and the use of additional reasoning processes. [30]

The comparative operation is useful because it constrains explanation. If a proposed cause applies equally to affected and unaffected cases, it needs an additional condition to explain the difference. This is not a mechanical disproof: the cases may differ in an unrecorded interaction, or the supposedly unaffected group may have incomplete reporting. The comparison initiates further inquiry into those possibilities.

Let us consider a hypothetical invoice delay that appears only after orders cross a particular value. General claims about motivation do little to explain the boundary. An approval threshold, a different document requirement, or a change in customer mix fits the observed distinction more closely. We would inspect those relationships before changing staffing or buying another workflow tool.

Structured analysis cannot decide every disagreement about organizational purpose. It is strongest when a deviation can be described and useful comparisons can be established. A dispute over the proper distribution of work may include facts that can be checked and values that cannot be settled by an IS and IS NOT comparison alone. Keeping that boundary visible preserves the value of the method. It prevents a disciplined analytical procedure from becoming a claim that every managerial question has one answer waiting to be mechanically extracted from a completed form.

15. Problem formulation and contested purposes

Not every management problem begins with an agreed desired state. A hospital, public agency, university, or company may contain several defensible accounts of what should improve. A remedy for one participant can create a problem for another. The difficulty is then partly the formulation of the problem itself.

Horst Rittel and Melvin Webber’s 1973 paper on planning described wicked problems whose formulation and resolution cannot be separated in the manner assumed for many well-defined technical tasks. Their discussion emphasized contested judgments and the consequences of intervention. It did not propose that evidence is irrelevant or that careful analysis should be abandoned. It challenged the transfer of a particular problem-solving model into situations whose goals and boundaries remain disputed. [31]

Russell Ackoff’s 1979 critique of operational research similarly objected to treating interrelated organizational problems as if they could always be solved separately. His argument directs attention to systems of problems and to the consequences of optimizing a part while leaving its relations unexamined. We can use this criticism without assuming that mathematical models are inherently incapable of representing interaction. [32]

Peter Checkland's Systems Thinking, Systems Practice, published in 1981, developed soft systems methodology for inquiry into problematic situations. Models can be used to structure learning and debate about purposeful activity rather than treated as literal descriptions of the organization. This gives participants a way to compare perspectives and explore changes without assuming that the initial problem formulation is already agreed. [33]

For diagnostics, the practical question is which problem statement a proposed explanation actually addresses. A reorganization may improve a department's financial result by moving costs elsewhere. A service change may reduce queue length by making entry more difficult. In each case the chosen boundary determines whether the change appears successful.

Consider a hypothetical public service with a waiting list. One proposal adds appointment capacity. Another changes eligibility. Both may reduce the recorded list, but they produce different consequences for people seeking service. Calling the list length the defect obscures the distinction. The investigation needs to preserve arrivals, eligibility decisions, unmet demand, and the experiences of those who no longer appear in the list.

The historical importance of problem formulation is that managerial correction becomes open to examination at the level of purpose. We can ask whether the proposed intervention addresses the stated difficulty, whether the statement excludes affected participants, and which consequences remain outside the measure. Debugging is still possible for a specified fault, such as an incorrect eligibility calculation. It should not absorb the larger political or ethical question of which eligibility rule the organization ought to adopt.

PART IV

Organizational systems and learning

16. Action research and sociotechnical work

An investigator who changes an organization also changes the conditions under which further evidence will be produced. Action research made this relation between inquiry and intervention explicit. Instead of treating practical action as something that begins only after research has ended, it connected investigation with cycles of action and examination.

Kurt Lewin's 1946 paper Action Research and Minority Problems placed research in relation to social action and the evaluation of its effects. Its historical setting concerned social problems and intergroup relations, not merely a convenient corporate improvement cycle. The wider setting reminds us that participants and consequences matter to the definition of the inquiry. [34]

Eric Trist and Ken Bamforth's 1951 study of the longwall method of coal-getting examined psychological and social consequences in relation to the technical organization of work. It became a central reference for sociotechnical thinking. The task arrangement, work group, and technology could be studied together instead of treating social difficulties as residual problems after an engineering design had been completed. [35]

This changes what counts as a relevant cause. A technically efficient arrangement may fragment the knowledge or cooperation needed to make it work. A worker's apparent failure to follow a procedure may reveal a conflict between that procedure and the actual sequence of dependencies. Conversely, a locally successful workaround may protect production while hiding a design defect from those able to correct it.

In a hypothetical service center, work is divided into narrow specialist stages. Each stage reports good utilization, but cases wait repeatedly between stages because no one has enough context to resolve an exception. Adding a tracking screen may expose the queues. Correcting the difficulty may require changes to team responsibilities, access to information, and authority to make decisions. The technical and social changes belong to the same intervention.

Action research adds an evidential responsibility: preserve what was changed, what was predicted, and what participants subsequently experienced. If the intervention changes several conditions, the report should not attribute the result to one favored component without support. Participation can improve access to local knowledge, but participation alone does not guarantee agreement or validity. The diagnostic contribution is a more explicit relation between inquiry, action, and the next state of the work.

17. Organizational diagnosis as an explicit practice

By the later twentieth century, organizational diagnosis had a substantial vocabulary of models and interventions. These models attempted to make a complex organization investigable without reducing every problem to one individual or one measure. They also created a new risk: the categories of a model could become substitutes for the evidence they were intended to organize.

Weisbord's six-box framework of 1976 joined purposes, structure, relationships, rewards, leadership, and helpful mechanisms. Nadler and Tushman's 1980 model directed attention to organizational functioning and the congruence of related elements. Both can be used to ask whether arrangements that appear reasonable separately fit together in operation. Their categories support a search; filling the categories does not complete the diagnosis. [1, 36]

Burke and Litwin's 1992 model proposed causal relationships among organizational conditions, distinguishing transformational and transactional factors. Their article explicitly presented hypothesized links supported through theory, research, and practice. A diagram of proposed causality should retain that status. It is not a measured causal graph for every organization to which its labels are applied. [37]

The models differ in their emphasis and level of detail, but they share an attempt to broaden the diagnostic field. An incentive may conflict with a stated purpose. A formal structure may depend on informal coordination. A change in the external environment may make an established arrangement unsuitable even though no internal participant has made a new mistake.

Let us consider a hypothetical product group whose members are asked to collaborate while rewards are assigned exclusively to individual output. A survey may record dissatisfaction, but the dissatisfaction is only one part of the evidence. We would also inspect assignment decisions, reward criteria, dependencies, and examples of work that benefits the group while reducing an individual's score. The diagnosis concerns the relationship among those arrangements.

A model is most useful when it makes an omitted relationship available for examination. It becomes less useful when it converts every observation into confirmation of its own categories. The historical development of organizational diagnosis therefore invites a second-level inquiry into the diagnostic instrument. Which questions does it make easy to ask? Which conditions does it leave unnamed? Can participants challenge the interpretation? A framework supports disciplined attention when its use remains accountable to the organization's actual work and evidence.

18. Cybernetics and distributed regulation

Feedback makes action dependent on information about its consequences. Management cybernetics examined how organizations could regulate their activities while facing changing conditions. The diagnostic question extends beyond whether information exists to whether it reaches a place where an appropriate response can occur.

W. Ross Ashby's An Introduction to Cybernetics, published in 1956, developed concepts of regulation, disturbance, and variety. Requisite variety concerns the relation between the range of disturbances and the regulatory capacity available to deal with them. In organizational comparison, it cautions against expecting one rigid response to handle many materially different situations. The formal theory and its managerial application should nevertheless be distinguished. [38]

Stafford Beer developed management cybernetics, including the analysis of viable organization and relations between local autonomy and wider coordination. His Brain of the Firm first appeared in 1972, with an expanded second edition in 1981. [39] Project Cybersyn in Chile, developed during 1971–1973, attempted to connect production information and economic management through a cybernetic design. Eden Medina's history examines its technical, organizational, and political dimensions. The famous operations room was one element of a larger project, not a completed autonomous controller of the national economy. [40, 41]

The distinction between a display and a functioning system is crucial. Information may be late, incomplete, or interpreted through the wrong model. A central view can remove details needed for local action. Local autonomy can also leave dependencies unresolved if no arrangement connects activities that affect one another. More centralization and more decentralization are not universal remedies.

Suppose a hypothetical regional service receives a summary showing that demand remains within capacity. Local offices nevertheless report severe delays because demand is concentrated in one specialist activity. Aggregation has removed the distribution that matters. A useful response may preserve local discretion while exposing the shared specialist constraint to those coordinating the whole service.

Cybernetics contributes a history of relationships between information and possible action. It encourages us to inspect delays, amplification, filtering, and the location of decision rights. The corresponding diagnostic limit is equally valuable. A regulator acts through a model and within a chosen boundary. The existence of feedback does not establish that the objective is appropriate, that every affected person is represented, or that the resulting stability is desirable.

19. System dynamics and delayed consequences

An intervention can improve a measure immediately and worsen the same measure later. This is difficult to understand when reports are organized as isolated monthly snapshots. System dynamics made feedback, accumulation, and delay explicit objects of managerial modeling.

Jay Forrester's Industrial Dynamics, published in 1961, examined industrial behavior through models linking flows, accumulated quantities, policies, and delays. Its production and distribution examples showed how organizational behavior could arise from interactions within a system. The analytical question moved from locating one disturbing event to examining structures capable of generating a recurring pattern over time. [42]

A queue illustrates the difference between a level and a flow. The number of waiting cases is an accumulated condition. Arrivals and completions change it. An increase in completions can coexist with a growing queue if arrivals increase faster. Reporting only completed work can therefore support a mistaken impression of improving service. This reasoning does not require a complex simulation, but simulation can help examine more elaborate dependencies.

Consider a hypothetical team that repeatedly hires temporary staff to clear a backlog. Experienced staff must train them, reducing short-term capacity. If management interprets the temporary reduction as evidence that hiring failed and changes the arrangement again, the organization can remain in a cycle of disruption. A model can represent the training delay and generate predictions to compare with subsequent evidence.

The model must remain inspectable. A plausible curve does not prove that the model contains the actual mechanism. Several structures can produce similar visible behavior. The investigator should examine parameter choices, the origin of data, and alternative explanations, then use the model to identify observations that would distinguish among them. Sensitivity analysis can show which conclusions depend heavily on uncertain assumptions.

The contribution to diagnostics is a longer temporal field. We inspect consequences beyond the first reporting period and compare them with the sequence expected from the proposed mechanism. A remedy should have an observation window appropriate to its delays. Closing an improvement action immediately after a favorable local result can prevent the organization from seeing whether the action merely transferred work into a later queue or another department.

20. Learning and the correction of governing assumptions

An organization can repeatedly correct a local error while preserving the rule that generates it. Each case is closed, yet the same type of case returns. Organizational learning asks whether inquiry changes subsequent action and whether the assumptions governing that action can themselves be examined.

Chris Argyris and Donald Schön’s Organizational Learning, published in 1978, distinguished correction within existing governing arrangements from learning that questions those arrangements. Single-loop learning can improve the way an established objective is pursued. Double-loop learning can examine the objectives, norms, or policies involved. The distinction concerns the object of correction, not a simple ranking in which every local adjustment is inferior. [43]

A faulty calculation can be corrected without reconsidering an organization’s purpose. Conversely, a recurring conflict between a target and the actual service may require examination of the target. The diagnostic task is to determine which level the evidence supports. Calling a discussion double-loop learning does not establish that any governing assumption changed.

Suppose a hypothetical department repeatedly trains staff after cases are assigned to the wrong queue. Further investigation shows that two categories overlap and that different managers interpret them

differently. Another training session may improve consistency temporarily. Revising the categories and recording how borderline cases are handled addresses a different level of the problem. The effect should still be checked after the change.

Organizational memory complicates the result. A lesson in a report may remain separate from the assignment system, induction material, or performance criteria that guide daily work. Personnel changes can remove knowledge that was never incorporated into those arrangements. A recurring incident may therefore reveal a failure of retention or transmission rather than a failure to discover the explanation originally.

For our comparison with debugging, the equivalent of saving a corrected version is not merely publishing the lesson. We need an identifiable change in the operating arrangement and evidence that it remains in use. We also need a way to reconsider that arrangement when conditions change. The history of organizational learning makes management itself part of the diagnostic field: the procedure used to identify and correct problems can become one of the procedures that requires correction.

PART V

Measurement and the examination of failure

21. Management information and the work of attention

Management information systems promised to make relevant organizational conditions available to decision makers. Their history includes the increasing speed of collection and calculation, but speed alone does not explain whether a report helps. The information has to enter an activity in which someone can interpret it and decide what to examine next.

Henry Mintzberg's studies of managerial work challenged the picture of the manager as a detached planner proceeding through an orderly sequence. His 1975 article The Manager's Job: Folklore and Fact summarized roles and characteristics of managerial activity, including its fragmented and interpersonal character. This directs attention to the actual use of information, not merely to its formal availability in a reporting system. [44]

A report may be accurate and still fail as an instrument of inquiry. It can arrive after the relevant decision, aggregate away the affected group, or use categories that do not correspond to the manager's practical question. Informal conversation may supply missing context, although it can also introduce selective recollection and unequal access. The diagnostic task is to relate these forms of evidence without assuming that either is inherently complete.

Consider a hypothetical executive summary that reports no significant delivery problems. A customer manager knows of several serious exceptions, but they are excluded because the summary counts only a standard product category. The disagreement is initially about the reporting boundary. An investigation of customer service should preserve the excluded cases instead of treating them as irrelevant because they do not fit the existing report.

Attention is also scarce. A system that sends every anomaly to the same manager may reduce the probability that any one anomaly is examined. A threshold can reduce noise, but it can also hide a slowly developing condition that never crosses the threshold in one period. The choice of aggregation, timing, and escalation therefore belongs to the design of organizational observability.

We can inspect that design through concrete records. Which event generated the report? Who received it? Was a response possible within the available time? Which information was retained when the report was summarized? These questions connect management information with the earlier history of records and authority. The main historical change is not that organizations finally possess information. It is that increasingly elaborate transformations stand between the activity and the person who acts on its representation.

22. Scorecards and different uses of control

Financial results are important evidence, but they can arrive after the conditions producing them have developed. Performance systems therefore expanded the range of measures used to describe an organization. The diagnostic problem became how to relate those measures without assuming that a larger collection automatically gives a more complete explanation.

Robert Kaplan and David Norton introduced the balanced scorecard in 1992. The original article named financial, customer, internal business, and innovation and learning perspectives. Their later account uses the familiar internal process and learning and growth labels. This widened the representation of organizational performance and supported discussion of relationships between current activities and longer-term results. The perspectives are a framework for inquiry, not proof of a causal chain in a particular enterprise. [45]

Robert Simons's Levers of Control, published in 1995, distinguished beliefs, boundaries, diagnostic control, and interactive control. For our purposes, the difference between monitoring selected variables and using information to examine strategic uncertainty is especially relevant. A measure can participate in different managerial activities even when its calculation remains unchanged. [46]

A dashboard used to confirm that targets have been met may discourage attention to an unexpected pattern. The same dashboard used in a discussion of assumptions may support further inquiry. We should therefore examine the meeting, decisions, and follow-up around the display. The interface alone cannot establish how the control system functions.

Suppose a hypothetical organization shows improving customer satisfaction and declining repeat purchases. The result can reflect a changed customer sample, a delayed effect, a product category shift, or a genuine divergence between satisfaction and continued use. A scorecard makes the discrepancy visible, but an explanatory arrow between satisfaction and revenue does not resolve it. The organization needs evidence capable of distinguishing the alternatives.

The history of measurement systems consequently includes a history of their use. Adding another indicator may be helpful when it reveals a previously hidden condition. It can be counterproductive when it creates another target detached from its purpose. Diagnostic value depends on traceability back to the work, stability of definitions, and the possibility of questioning the assumed relationship among measures. A balanced display should remain open to evidence that its own balance excludes an important part of the organization's activity.

23. Incentives and the transformation of a measure

Once a measure affects rewards or sanctions, it becomes part of the system it describes. Participants may change useful behavior, change the recorded representation, or move activity beyond the

measure's boundary. A history of managerial observation must therefore include the consequences of being observed through a target.

Steven Kerr's 1975 paper On the Folly of Rewarding A, While Hoping for B examined arrangements in which rewards support behavior different from the stated intention. The discrepancy can persist without anyone misunderstanding the formal objective. Participants may respond quite intelligibly to the consequences attached to their actions. [47]

Donald Campbell's discussion of planned social change, issued in 1976 and later reprinted, warned about pressures on social indicators when they are used for consequential decisions. Charles Goodhart's monetary-policy argument in the mid-1970s concerned the instability of statistical relationships under attempts to use them for control. Their settings and formulations differ. [48, 49] The later formulation about a target ceasing to be a good measure appears in Marilyn Strathern's 1997 discussion of university audit. It should not be presented as a verbatim statement from both earlier sources. [50]

For diagnosis, this means that a change in a metric can have at least two broad origins. The underlying work may have changed, or the relationship between the work and its recorded representation may have changed. Both can occur together. We need evidence about classification rules, incentives, and the treatment of exceptional cases in addition to the metric itself.

In a hypothetical support organization, a reward for rapid closure may encourage earlier closure and more subsequent contacts under new case identifiers. The measured closure time improves, while customers experience the same unresolved problem across several records. A case-level view hides the continuity that a customer-level or problem-level view reveals. The investigation concerns the unit of analysis as well as the incentive.

The appropriate correction need not be abandonment of measurement. It may involve a different unit, a check of downstream consequences, a more suitable reward arrangement, or qualitative examination of selected cases. However, every revised measure remains a representation with a purpose and a boundary. The historical lesson is to make those relationships inspectable and to retain independent evidence of the outcome that the measure was introduced to improve.

24. Audit and administrative visibility

Organizations often need to demonstrate that specified procedures were followed or that particular assertions are supported. Audit can make this demonstration inspectable. It also shapes what organizations record and what they regard as worth doing, because activities that can be documented are easier to present as evidence of compliance.

Michael Power's The Audit Society, first published in 1997, examined the expansion of auditing and the organizational consequences of making activities auditable. The diagnostic question is whether the verification procedure remains connected to the substantive purpose it is meant to support. A well-documented process can be easier to audit without necessarily producing a better service. [51]

The COSO Internal Control—Integrated Framework, updated in 2013, provides a different kind of contribution: a formal framework connecting the control environment, risk assessment, control activities, information and communication, and monitoring activities. It helps organize questions about whether an organization's controls work together. Listing the components is not evidence that they are effective in a particular institution; their operation still requires examination. [52]

James Scott's Seeing Like a State, published in 1998, investigated administrative simplification and large schemes for social improvement. His account draws attention to local knowledge and to what can disappear when diverse practices are made legible through standardized categories. This is relevant to public management and to enterprises whose central reports must compress heterogeneous local activities. [53]

The distinction between visibility and understanding is practical. Standard categories permit comparison, but they can remove a condition essential to explaining a particular case. A local workaround may appear as noncompliance until the investigator learns which failure it prevents. Conversely, the local description may defend an arrangement that transfers costs to another group. Both the central representation and the local account require examination.

Suppose a hypothetical organization audits whether every project has a completed risk register. All projects pass. Later, several fail because entries were copied from a template and never connected to decisions. The audit established the presence of an artifact. It did not establish the quality of the inquiry or the effect of the recorded risks on action. A better investigation asks when the entries changed, what evidence prompted them, and which decisions followed.

This does not make documentation dispensable. It makes the relation between documentation and work an object of diagnosis. Records can protect people from arbitrary claims, preserve disagreement, and support later reconstruction. They can also create an appearance of control. The historical task is to examine which of these functions a particular arrangement performs, for whom, and under what conditions. Auditable evidence is valuable when its limits remain visible alongside the assertion it supports.

25. Accidents and the organization of unwelcome evidence

Major accident investigations show why management diagnostics cannot stop at the immediate physical failure. They can reconstruct how information was interpreted, transmitted, and acted upon before the event. The resulting history often concerns a sequence of locally understandable decisions whose combined consequences were not adequately recognized.

The 1986 Presidential Commission on the Space Shuttle Challenger accident identified serious defects in the launch decision process. Its chapter on contributing causes described incomplete or misleading information, communication failures, and a structure through which safety concerns bypassed key decision makers. This is a documented organizational finding, distinct from the physical mechanism involving the solid rocket motor joint. [54]

Diane Vaughan's The Challenger Launch Decision, published in 1996, reconstructed the development of organizational judgments preceding the accident. Her analysis associated the normalization of deviance with processes through which departures came to be accepted in routine decision making. This adds a longer history of interpretation to the immediate sequence before launch. It also warns against explaining a complex organizational failure solely through a few visibly consequential decisions at the end. [55]

The Columbia Accident Investigation Board's 2003 report also examined management practices as causal contributors alongside the physical damage initiated by foam striking the wing. The two investigations should not be collapsed into the claim that NASA simply ignored an identical warning twice. Their events, technical evidence, and decision histories differ. What supports comparison is the need to reconstruct how technical evidence acquired or lost force within an organization. [56]

Amy Edmondson's 1999 study of work teams found an association between psychological safety and learning behavior. This concerns the conditions under which people can raise questions, discuss errors, and participate in inquiry. The study does not imply that supportive language alone establishes safety or that its measured association proves every proposed intervention will succeed. [57]

For management observability, the important distinction is between a signal being recorded somewhere and its consequences becoming available to the relevant decision. An objection can survive in a document while disappearing from the summary used for authorization. A concern can be acknowledged without anyone possessing the authority or resources to investigate it. Counting reports is therefore an incomplete account of whether the organization can learn from them.

A useful diagnostic reconstruction follows the evidence through these transformations. It preserves what participants knew, how uncertainty was expressed, which alternatives were considered, and where responsibilities changed. This also guards against hindsight. The outcome can make an earlier warning look decisive even when its interpretation was uncertain at the time. A rigorous investigation can identify avoidable failures without pretending that every participant already possessed the completed investigation report.

PART VI

Digital processes and evidence based intervention

26. Process redesign and event records

Computerized workflows can preserve many transitions that earlier organizations recorded only intermittently. A request can have an identifier, an arrival time, a sequence of assignments, and a final state. These records make new forms of inquiry possible, but only if their relationship to the work is understood.

Hammer discussed process redesign in 1990. [58] The diagnostic question for our purposes is whether a new system reproduces an avoidable dependency or allows the organization to reconsider it. A faster implementation of an unsuitable sequence can reproduce its consequences at greater scale.

Process mining later connected event logs with the discovery and examination of process models. The 2004 paper by van der Aalst, Weijters, and Maruster presented a workflow-mining algorithm that extracts a model from logged events. Its formal assumptions matter: a log and a discovered model do not give unrestricted access to everything that happened in an organization. [59]

Several evidential choices determine what can be reconstructed. The investigator needs a meaningful case identifier, activity definitions, and an account of ordering. A timestamp can represent entry into a system rather than performance of the work. Batch recording can make sequential work appear simultaneous. Missing events can make an actual handoff disappear, while a changed identifier can divide one continuing problem into several apparent cases.

Consider a hypothetical approval process whose log shows rapid transitions between departments. Staff explain that requests wait in email before anyone creates the formal case. The logged process is accurate within its boundary, but the customer’s waiting time begins earlier. A process-mining result may therefore identify a well-behaved administrative subprocess inside a poorly performing overall service.

The historical advance is the ability to examine repeated sequences at scale and compare them with an intended model. The corresponding responsibility is to preserve the event semantics and collection history. Conformance to a recorded model establishes a relationship between a log and that model. It does not establish that the model expresses the right purpose, that unlogged work is irrelevant, or that an observed difference is a defect rather than a necessary adaptation.

27. Software projects and operational learning

Software work brought management into contact with systems whose behavior can change quickly and whose internal dependencies are difficult to infer from a completed plan. Iterative development and

operational incident analysis made feedback available at shorter intervals, although the managerial value still depended on how the organization used it.

The 2001 Manifesto for Agile Software Development and its principles emphasized working software, collaboration, response to change, and regular reflection. The 2020 Scrum Guide describes an empirical approach organized around transparency, inspection, and adaptation. These sources describe principles and a framework; they do not establish that adopting particular meetings or labels guarantees successful delivery. [60, 61]

In this setting, a demonstration or increment can serve as evidence about an earlier assumption. It may expose a misunderstood requirement or an integration difficulty before the full project is complete. A plan remains useful as a statement of expectations, but its diagnostic role depends on preserving the differences between prediction, observation, and revised commitment.

Operational practice adds another evidential sequence. The Site Reliability Engineering book published in 2016 describes postmortems that preserve incident impact, contributing causes, response, and follow-up actions. Its account of blameless inquiry aims to make the conditions of action available for learning rather than ending the investigation with an individual's mistake. Follow-up work is part of the practice, not an optional decoration on the report. [62]

Suppose a hypothetical release fails because an undocumented dependency was omitted. A corrective action that merely tells the engineer to be more careful does not specify how the next engineer will discover the dependency. Recording it in the deployment process, checking that the check runs, and observing subsequent releases creates a more inspectable correction. The action can still fail if the dependency changes and its documentation remains static.

The management lesson concerns continuity across planning, execution, and learning. A retrospective can identify a problem that no one has time or authority to fix. A postmortem action can be marked complete without testing the changed behavior. A new reporting system can hide the earlier incident history. We therefore examine the whole sequence by which evidence produces a change and the change remains open to further evidence.

28. Evidence based management and field interventions

A plausible management explanation is not the same as evidence that an intervention improves outcomes. Success stories often combine several changes and observe one organization after the event. Evidence-based management asks how decisions can make better use of research while retaining attention to the conditions of the local problem.

Denise Rousseau's 2006 article Is There Such a Thing as Evidence-Based Management? argued for connecting managerial decisions with scientific evidence. The issue is not whether managers possess experience, but how claims based on experience can be examined alongside other evidence. A remembered success can generate a useful hypothesis without becoming a universal prescription. [63]

Bloom, Eifert, Mahajan, McKenzie, and Roberts's study Does Management Matter? Evidence from India, published in the 2013 issue of The Quarterly Journal of Economics, investigated a management intervention in textile firms. The study used experimental variation in consulting support and reported improvements associated with the adopted practices. Its setting and bundled intervention are essential to interpretation; it does not establish that every management technique has the same effect in every industry. [64]

For diagnostics, there is a distinction between investigating a local mechanism and estimating an intervention's effect. A detailed case can reveal how a queue formed without estimating the average effect of a policy across organizations. An experiment can estimate an effect in its studied setting without explaining every pathway or guaranteeing transfer to another institution. These forms of evidence can support one another when their questions remain explicit.

Consider a hypothetical intervention that adds a daily review and changes the assignment rule. Backlog subsequently falls. To attribute the improvement, we need to consider arrivals, staffing, case complexity, and other concurrent changes. A comparison group or a carefully planned sequence may strengthen the inference. If those conditions are unavailable, the report should state the narrower result that the combined change preceded the observed improvement.

The history of evidence-based management makes a useful demand on managerial claims: specify what evidence would justify the decision and what result would lead to revision. Evidence includes consequences for those affected, not only the measure chosen by the sponsor. The manager's practical judgment remains necessary, but its assumptions become available for examination. That is a diagnostic improvement even when the available evidence supports a conditional decision rather than certainty.

29. Narratives and the reconstruction of organizational events

Organizations explain themselves through stories. A project succeeded because a leader acted decisively. A department failed because people resisted change. A customer problem was exceptional. Such narratives select events and connect them into an explanation. The selection can help inquiry, but it can also conceal the relationships needed to test the explanation.

Karl Weick's Sensemaking in Organizations, published in 1995, examined how people make situations intelligible through ongoing organizational activity. This shifts attention from a finished decision to the processes by which participants establish what is happening. Retrospective interpretation matters because an event's significance can change as later events become known. [65]

Gareth Morgan's Images of Organization, first published in 1986, used multiple metaphors to examine organizational life. Seeing an organization as a machine, a culture, or a political system directs attention toward different features. The practical contribution is the possibility of changing viewpoints; the danger is allowing a successful metaphor to become an exhaustive account. [66]

For diagnostic purposes, a narrative should remain connected to identifiable artifacts. The statement that a requirement changed late can be examined through versions, messages, meeting records, and acceptance decisions. These artifacts may disagree because they represent different stages or purposes. Reconstructing the sequence should preserve that disagreement instead of selecting whichever source supports the preferred story.

Imagine a hypothetical project report that attributes a missed date to an external supplier. The supplier's delay is real, but earlier records show that internal specifications arrived after the agreed date. A fuller narrative changes the relation between the events. It may still assign responsibility to the supplier for part of the outcome, while removing the claim that the organization's own sequence was unaffected.

This is where software narratology and management inquiry can meet without becoming identical disciplines. A trace is also read through a selected viewpoint and reconstructed sequence. In management, statements have additional social purposes: they can justify decisions, allocate credit, protect participants, or support a claim. Diagnosing an organizational narrative therefore involves both its evidential structure and its role in the organization that produces and receives it.

30. Algorithmic management and the diagnosis of automated judgment

Digital management systems can assign work, calculate performance, rank alternatives, and recommend interventions. Some employ machine learning; others use explicit rules. Treating all of them as one kind of artificial intelligence hides differences that matter when a decision must be explained or corrected.

The 2022 paper by Sara Baiocco, Enrique Fernández-Macías, Uma Rani, and Annarosa Pesole, published by the International Labour Organization and the European Commission's Joint Research Centre, describes algorithmic management as the use of programmed procedures in coordinating labor and relates it to both platforms and conventional workplaces. It emphasizes continuities with earlier forms of organization as well as new consequences of digital coordination. Automated assignment does not begin the history of managerial control, but it can change its scale, speed, and inspectability. [67]

An OECD employer survey reported in 2025 examined adoption and managerial perceptions in six countries. Managers reported benefits alongside concerns about unclear accountability, difficulty following the tools' logic, and insufficient protection of workers' health. These are survey findings about reported use and perceptions. They should not be converted into a causal estimate that algorithmic management improves every organization using it. [68]

NIST's AI Risk Management Framework 1.0, released in 2023, organizes voluntary risk-management activity through Govern, Map, Measure, and Manage. In our context, this directs attention to the system's purpose, use, evaluation, and response to identified risks. The framework is not a certification that a particular model makes appropriate management decisions. [69]

Consider a hypothetical assignment model that ranks employees by recorded completion time. The training data may omit time spent helping colleagues, handling difficult exceptions, or preventing rework. The model can predict the recorded label accurately while rewarding an incomplete representation of contribution. Debugging its code would not correct the definition of the label. That requires organizational diagnosis and examination of the measurement process.

Generative systems add another transformation: they can produce a coherent managerial explanation from incomplete records. The explanation may introduce a causal connection that no source establishes. A useful arrangement retains links to the underlying evidence, distinguishes recorded facts from generated hypotheses, and permits review of the data and model version used. The historical continuity is clear. The instrument becomes more elaborate, while the old question remains: what relationship connects the visible report to the work and people it claims to describe?

PART VII

Managerial reason and pattern oriented diagnostics

31. Critique of management reason and the rational manager

My book Critique of Management Reason: Management Bits and Tips from Can't provides a particular connection between management reflection and the diagnostic viewpoint used here. It developed from Management Bits and Tips, a blog concerned with software engineering and software technical support management. Its publication history should be kept separate from the history of the ideas and posts it collects. [70]

The anthology centers on posts from 2008–2010, with later additions. The projected title was Management Bits: An Anthology from Reductionist Manager. An announcement in June 2008 already listed that project and its ISBN among planned OpenTask books. [71] A further announcement on July 10, 2010, named September 2010 as the intended publication date. [72] The later Leanpub description recalls the publication plan in connection with my return to engineering in 2009; the 2008 announcement documents an earlier intention. Around 2015, I changed the title to Critique of Management Reason, and additions continued as late as 2017. An announced date for an earlier project is not evidence of a completed publication. [70]

The published contents place project failure analysis patterns alongside management philosophy, management poetics, process parasites, and reflections on organizational memory and language. This mixture permits us to examine management through several forms of inquiry. A concise observation can identify a contradiction before a formal model has been developed. Humor can expose an assumption, although it does not by itself establish the frequency or cause of the behavior described. [70]

Jerzy Kociatkiewicz and Monika Kostera's 2012 paper, Sherlock Holmes and the Adventure of the Rational Manager: Organizational Reason and Its Discontents, mentions my earlier project when considering links between software design and managerial problems. Its bibliography cites Management Bits: An Anthology from Reductionist Manager (OpenTask, 2010). [73]

The projected edition was still unpublished. The bibliography's 2010 date is consistent with the announced schedule, but the citation does not establish publication or identify which announcement the authors used. [72] It should not be read as the publication date of Critique of Management Reason. The Leanpub sample's copyright page identifies Version 1.01, November 2025; the OpenTask table-of-contents PDF carries Version 1.0 for the same month. Both explicitly assign ISBN 978-1-906717-13-1 to the paperback. [70] The references identify both the journal article and the accessible manuscript, including the relevant manuscript pages.

The authors' interest in a rationality that includes imagination provides a useful point of comparison. For our purposes, imaginative hypothesis construction and evidential checking are distinct operations. An

analogy can propose an explanation; examining the artifacts determines how far that explanation applies. [73]

We can use that distinction to examine our own analogies. A project may resemble a blocked computation, but the resemblance does not establish the same mechanism. The analogy becomes useful when it suggests a concrete question: which dependency is unresolved, which event is missing, or which representation hides continuing work? Critique then applies to the diagnostic model as well as to management. We retain the power of a comparison while checking what it permits us to infer.

32. Designing organizational observability

Organizational observability begins with a question that can be answered or narrowed through evidence. It does not begin with a requirement to collect everything. An organization can accumulate a large volume of records while remaining unable to reconstruct one consequential handoff.

For a particular activity, we need to identify the entities, transitions, and relationships that matter. A request may pass through several systems and acquire a new identifier at each boundary. A person may participate in several activities without any one activity representing the whole of their work. If the question concerns customer experience, a local departmental identifier may be insufficient. If it concerns a decision, the relevant evidence may include an assumption and an authorization rather than a transaction count.

Time also requires design. Event time, entry time, processing time, and reporting time can differ. A weekly report may include records created this week about work done earlier. Comparing it with another report based on completion time can produce an apparent disagreement even when both calculations are correct. The observation system should preserve enough metadata to make these differences inspectable.

We can express the design as a chain of evidential responsibilities. The activity produces an artifact. A collection procedure selects and transforms it. A representation presents a result. An analyst constructs an explanation. A decision changes the activity or its recording. At every transition we can ask what was preserved, what was omitted, and which version of a rule was applied. This is our synthesis of the historical practices, not a claim that they all used the same architecture.

Participation is part of this design. Workers can explain unrecorded dependencies, customers can identify outcomes that closure codes miss, and specialists can identify conditions hidden by aggregation. Their accounts should be preserved with their context and examined alongside other evidence. The aim is an inspectable relationship among accounts, rather than a forced agreement that removes the disagreement before it has been understood.

Observability also has a scope. Recording information that cannot answer the diagnostic question adds burden and can alter behavior without improving the investigation. A more useful design states why a record is needed, how long its meaning remains valid, and who can challenge an interpretation built

from it. The result is not complete visibility of an organization. It is a better ability to investigate specified questions while knowing which regions of the work remain outside the available evidence.

33. Connections to trace and log analysis patterns

Our trace and log analysis pattern language was developed for software evidence and later extended to texts, narratives, and data more generally. Its reference text explicitly discusses this widening scope. Management artifacts provide a further field in which recurrent arrangements can suggest useful analysis operations, provided that we preserve differences in meaning and mechanism. [74]

The most immediate connection is a thread of activity. In software, a computational thread gives one way to group messages. An Adjoint Thread of Activity groups them by another selected attribute. In management, a customer, order, asset, policy version, or unresolved question can provide a comparable grouping. These groupings are proposed analytical uses of the pattern, not claims that an employee literally functions as an operating-system thread.

For example, departmental views may each show completed work while a customer view shows an unresolved sequence of transfers. The choice of grouping changes the visible continuity. Likewise, a repeated message can represent repeated failure, repeated reporting of one failure, or a template copied into several records. Message Context and Message Invariant help us separate the stable wording from the changing conditions of its use.

Existing analysis pattern	Proposed managerial reading	Check before inferring a cause
Adjoint Thread of Activity	Follow one customer or issue across systems	Confirm identity links and distinguish one continuing issue from several issues
Use Case Trail	Reconstruct a request from entry to actual outcome	Include handoffs and work outside the formal workflow
Message Context	Read an escalation with its preceding and following records	Check what participants knew at that time
Message Invariant	Separate repeated report wording from variable values	Determine whether the wording is templated or independently composed
Discontinuity	Locate a gap in an activity's recorded sequence	Test missing collection, changed identifiers, and genuine inactivity
No Activity	Examine an absence of expected records	Establish that the activity would normally have produced a retained record
Activity Disruption	Inspect interruptions within a continuing activity	Distinguish containment, planned pauses, and actual loss of function
Inter-Correlation	Compare related records from separate sources	Check clocks, definitions, dependencies, and possible common origins
Intra-Correlation	Compare regions or attributes within one record set	Avoid treating repeated copies as independent observations
False Positive Error	Reconsider warnings that do not explain the reported problem	Preserve their possible relevance to a different problem

Existing analysis pattern	Proposed managerial reading	Check before inferring a cause
Trace Viewpoints	Compare departmental and customer interpretations	State the unit, boundary, and purpose of each view
Visibility Limit	Mark what the available records cannot distinguish	Do not convert absence of evidence into a categorical verdict

The table proposes ways to read evidence. It does not diagnose an organization from a pattern name. A discontinuity can be caused by a reporting change, a stopped activity, or missing retention. A recurring escalation can indicate a persistent dependency or an appropriate response to repeated independent events. Further evidence must distinguish these possibilities.

Pattern orientation can nevertheless reduce avoidable reinvention. Analysts can name a recurrent evidential arrangement, remember which alternatives it admits, and identify the next artifact to inspect. This is especially useful when technical and managerial investigations meet. A service incident may require both a software trace and a record of the decisions that changed its operating conditions. The shared vocabulary supports coordination while leaving each discipline responsible for its own claims.

34. A worked investigation of apparent improvement

Let us consider a constructed service investigation. An organization reports that its new workflow closes requests more quickly. Customers nevertheless describe more repeat contacts and unresolved handoffs. All quantities and records in this chapter are hypothetical. They illustrate an analysis; they do not describe an actual company.

We first define comparable cohorts. Each contains 100 requests received during an equally long intake period and followed for 30 days. The counts below refer to those original requests, even when the system later creates additional case identifiers. This prevents a repeated contact from silently becoming a new independent success.

Observation within each 100 request cohort	Earlier workflow	New workflow
Original requests with a recorded closure by day 30	90	90
Original requests with a recorded closure within two days	42	77
Original requests with customer-confirmed resolution by day 30	82	67
Original requests followed by a repeat contact about the same issue	12	28
Original requests involving work in a separate specialist queue	9	24

The report’s statement about faster recorded closure is supported within this constructed dataset. The broader statement that the service improved is not established by that measure. Customer-confirmed resolution decreased, and repeat contact increased. These observations justify further inquiry without yet proving that the workflow change caused all the differences. In a real investigation, demand, case mix, staffing, and follow-up completeness would also require examination.

We then reconstruct one of the new-workflow cases. The organizational thread is the original customer issue. Its records are distributed across two systems and a later case identifier.

Time	Artifact	Recorded event
Day 1 at 09:00	Service case R204	Original request received
Day 1 at 09:04	Service case R204	Initial assessment completed
Day 1 at 09:09	Specialist case X77	Request transferred for specialist work
Day 1 at 09:10	Service case R204	Local case closed after transfer
Day 3 at 14:00	Service case R981	Customer contacts the service about the continuing issue
Day 5 at 11:20	Specialist case X77	Specialist records completion of the requested work

The local service trace ends after ten minutes. The issue continues for several days. An Adjoint Thread of Activity formed around the original issue makes that continuity visible. A departmental view and a customer view are therefore different representations of related evidence, rather than mutually exclusive accounts of what happened.

Next we inspect the hypothetical workflow definitions. The earlier rule closed a case after the requested work was completed. The new rule closes the local case after transfer to a specialist. This change establishes a concrete mechanism for part of the metric improvement: the meaning of the recorded endpoint changed. It does not establish why specialist work took longer or why every repeat contact occurred. Those remain separate questions.

We now have several possible corrections. The service can preserve transfer as a distinct state, maintain the link between the original request and specialist work, and report customer-level resolution separately from local handling. It can also investigate the specialist queue to determine whether the change created additional work, exposed existing work, or coincided with another capacity problem. Deleting the old closure records would destroy useful evidence of the earlier reporting rule.

Verification must follow the corrected definition. We would check new cohorts, inspect unresolved cases, and examine whether work has moved into another unreported queue. Cases still unresolved at the observation boundary cannot simply be omitted from a duration average and then treated as successes. Their incomplete observation needs explicit treatment. We would also sample identity links to ensure that unrelated issues have not been merged.

The diagnosis has changed the object of correction. The first suspicion concerned staff performance. The evidence identified a changed recording rule and a continuing cross-system activity. It also left a capacity question unresolved. A useful report preserves all three results: what was corrected, what the evidence supports, and what the next investigation still needs to establish.

35. Continuing developments and the scope of this synthesis

The practices examined here do not form a sequence in which each new method makes the previous one obsolete. An investigation may combine accounting reconciliation, comparative problem analysis,

feedback analysis, and examination of a decision narrative. Digital tools can connect these operations while their evidential requirements remain distinct.

Our proposed contribution is a pattern-oriented historical synthesis. It relates records, representations, explanations, interventions, and subsequent checks across established management traditions. The framework distinguishes failure in the work from failure in its measurement or interpretation. Organizational diagnosis, feedback, learning, and process analysis have histories that precede this comparison.

A continuing problem is diagnosis of the diagnostic system. A dashboard becomes stale, a classification conceals recurrence, or an automated summary omits consequential disagreement. Preserving versions of rules and transformations allows these failures to be investigated. Another problem concerns the boundary of correction: improved departmental reports may conceal additional work imposed on customers. Verification must follow the relevant activity across that boundary.

Automation can help search records and propose recurring arrangements. A retrieved incident may nevertheless differ in the feature that matters. A persuasive explanation of missing events may overlook a change in collection. Pattern recognition remains useful when the analyst examines the conditions supporting an inference.

Management concerns people who interpret, respond, and disagree. Their ability to challenge a representation contributes to organizational learning. A specified defect can be corrected while a disagreement about purpose remains. The continuing task is to make evidence and the limits of explanation inspectable, then connect inquiry with changes whose consequences can be observed and reconsidered.

Conclusion

We began with records that preserve an event beyond its immediate occurrence. Accounting organized relations among entries. Industrial studies made work sequences and variation available for comparison. Administrative theory examined responsibilities and coordination. Later approaches expanded the diagnostic field to cooperation, bounded choice, organizational systems, learning, and contested purposes. Digital systems now preserve and transform evidence at a scale that earlier investigators could not readily examine.

Across these developments, the same analytical distinctions remain useful. An observed condition is different from its representation. A representation is different from an explanation. An explanation is different from evidence that a remedy works. The transitions among them require records and arguments that another investigator can inspect.

Management also changes the conditions of its own observation. A target changes behavior. An audit changes documentation. A new workflow changes what counts as an event. An automated model can make an incomplete label consequential. Diagnostics must therefore include the system that produces the managerial evidence, together with the work that evidence is intended to describe.

Our pattern-oriented comparison provides a vocabulary for navigating these relationships. We can follow an activity across boundaries, compare viewpoints, inspect discontinuities, and identify visibility limits. We can use an analogy to generate a question and then return to evidence to examine it. The connection with Critique of Management Reason belongs here: managerial reason, including our own explanatory models, remains an object of inquiry.

A correction is most useful when its scope is clear and its consequences remain observable. The resulting organizational memory should preserve the problem, the evidence, the changed arrangement, and the conditions under which the change was checked. It should also preserve what remains unresolved. This makes the next investigation a continuation of knowledge rather than another attempt to reconstruct why the same problem has returned.

Appendix A Selected chronology

The chronology identifies selected artifacts, publications, and events discussed in the chapters. A publication date marks a documented formulation, not necessarily the first appearance of every practice associated with it. Later digital editions and repository dates are distinguished in the references.

Date	Development	Diagnostic significance
Late third millennium BCE	Ur III administrative accounts [3]	Selected transactions persist for comparison and accountability
1494	Pacioli's printed account of double-entry bookkeeping [5]	Relations among entries become widely teachable through print
1832	Babbage's On the Economy of Machinery and Manufactures [6]	Technical operations are examined with costs and division of labor
1911	Taylor's The Principles of Scientific Management [7]	Systematic examination of work becomes a managerial program
Early twentieth century	Gilbreth motion studies [10]	Recorded sequences permit repeated inspection of work
1916–1917	Fayol's French administrative work, serialized and then issued as a book [8]	Coordination and control receive an explicit general vocabulary
1922	Weber's posthumous Economy and Society [9]	Formal authority and bureaucracy become objects of systematic analysis
1922	Clark's The Gantt Chart [11]	Graphical comparison connects planned and recorded work
1924	Shewhart's early control-chart work [17]	Process variation becomes an object of disciplined interpretation
1924–1933	Hawthorne research program [12]	Work observation expands to social conditions and the research setting
1925	Follett's discussion of giving orders [15]	Joint examination of an evolving situation informs authority
1931	Shewhart's major quality-control book [17]	Statistical inquiry supports investigation of process behavior
1938	Barnard's The Functions of the Executive [16]	Cooperation and communication enter the analysis of executive work
1940s	Wartime operations research and postwar expansion [24]	Models support investigation of operational choices
1946	Lewin's action-research paper [34]	Inquiry is connected with action and examination of consequences
1947	Dantzig's linear programming and simplex work [25]	Allocation choices become computationally investigable
1951	Trist and Bamforth's coal-mining study [35]	Technical and social arrangements are examined together
1951 and after	Juran's handbook and later Japanese quality initiatives [21, 22]	Quality inquiry expands through managerial responsibility and participation

Date	Development	Diagnostic significance
1954	Drucker's The Practice of Management [28]	Objectives are related to managerial contribution and self-control
1955	Simon's behavioral model of rational choice [26]	Limits of knowledge and calculation become central to choice
1956	Ashby's An Introduction to Cybernetics [38]	Regulation is examined in relation to disturbance and variety
1961	Forrester's Industrial Dynamics [42]	Feedback and delays become explicit in managerial models
1965	Kepner and Tregoe's The Rational Manager [29]	Structured analysis separates explanation from choice of action
1971–1973	Project Cybersyn in Chile [40]	Economic information and coordination are explored through cybernetics
1972	The garbage can model [27]	Choice is examined through changing combinations of organizational streams
1972 / 1981	Beer's Brain of the Firm and its expanded second edition [39]	Viability, local autonomy, and wider coordination receive a cybernetic account
1973	Rittel and Webber's planning critique [31]	Problem formulation and contested judgment receive explicit attention
1975	Mintzberg's account of managerial work and Kerr's incentive critique [44, 47]	Actual activity and reward consequences challenge official descriptions
1975–1976	Goodhart's and Campbell's indicator critiques [49, 48]	Control and consequential use can alter evidential relationships
1976	Weisbord's six-box diagnosis [1]	A practical framework broadens the search for organizational trouble
1978	Argyris and Schön's Organizational Learning [43]	Governing assumptions become possible objects of correction
1979–1980	Ackoff's critique and Nadler–Tushman diagnosis [32, 36]	Interdependence and fit challenge isolated problem treatment
1981	Checkland's Systems Thinking, Systems Practice [33]	Models support learning about differently understood problematic situations
1986	Challenger investigation and Morgan's Images of Organization [54, 66]	Decision structures and interpretive viewpoints enter different diagnostic inquiries
1990	Hammer's reengineering article [58]	Process redesign questions inherited work before automating it
1992	Balanced scorecard and Burke–Litwin model [45, 37]	Multiple performance perspectives and proposed change relationships gain formal expression
1993	Deming's The New Economics [20]	Knowledge, variation, and learning are connected with management
1995	Simons's Levers of Control and Weick's Sensemaking in Organizations [46, 65]	Information use and construction of meaning are examined explicitly
1996	Vaughan's The Challenger Launch Decision [55]	Routine interpretation is connected with the normalization of departures
1997	Strathern's discussion of university audit [50]	Target setting is examined as a transformation of measurement

Date	Development	Diagnostic significance
1997–1999	Power, Scott, and Edmondson publications [51, 53, 57]	Auditability, administrative simplification, and learning conditions receive scrutiny
2001	Agile Manifesto [60]	Software development principles emphasize feedback and reflection
2003	Columbia investigation report [56]	Organizational practices are examined alongside physical failure
2004	Workflow-mining paper [59]	Event logs support discovery of formal process models
2006	Rousseau's evidence-based management paper [63]	Research evidence is connected explicitly with managerial decisions
2008–2010	Management Bits posts and anthology announcements [71, 72, 70]	Announcements preserve publication intentions, including the unfulfilled September 2010 schedule
2011	Levitt and List's Hawthorne reanalysis [14]	An inherited management account is checked against surviving data
2012	Kociatkiewicz and Kostera's Holmes paper [73]	The earlier Management Bits project appears in a discussion of rationality
2013	Bloom and colleagues' India study [64]	A field intervention examines consequences of management practices
2013	Updated COSO internal-control framework [52]	Control components are related within an integrated organizational framework
Around 2015	Change to the Critique of Management Reason title [70]	A later title must be distinguished from the earlier projected edition
2016	Site Reliability Engineering book [62]	Operational postmortems and follow-up are documented as organizational practice
2020	Revised Scrum Guide [61]	Transparency, inspection, and adaptation receive a compact framework statement
2022–2025	ILO, NIST, and OECD publications [67, 69, 68]	Algorithmic coordination and AI use become explicit objects of organizational inquiry
2025	Critique of Management Reason, November version [70]	The published title and paperback ISBN are distinguished from the earlier projected anthology

Appendix B Glossary

Activity. A sequence or grouping of actions relevant to a question. Its boundary may differ from a department's reporting boundary.

Adjoint Thread of Activity. In our trace analysis vocabulary, messages grouped by a selected attribute beyond a computational thread identifier. A customer or issue can provide an analogous grouping in managerial evidence.

Anomaly. A condition that differs from an expectation or comparison set. Its significance depends on how that expectation was established.

Artifact. A record or object available for investigation, including a ledger, message, film, model, report, or saved configuration.

Audit. Examination of specified assertions, records, or procedures against criteria. The scope of the assurance must be distinguished from overall organizational success.

Baseline. A reference condition used for comparison. A valid comparison needs an account of relevant differences between the baseline and the later condition.

Bounded rationality. Decision making under limits of knowledge, attention, and calculation. It directs inquiry toward the available search and choice process.

Case identifier. A value used to connect events attributed to one case. A change of identifier can hide continuity; reuse can create false continuity.

Causal mechanism. An account of how interacting conditions produce a result. It requires more than naming a symptom or finding a correlation.

Common-cause variation. Variation associated with the continuing process system in statistical quality-control usage. A stable process can still be unacceptable for its purpose.

Conformance. A specified relation between observed or recorded behavior and a rule or model. Conformance alone does not establish the suitability of the rule.

Control limit. A statistical reference used in interpreting a control chart. It is different from a specification or contractual requirement.

Debugging. In this book's comparative usage, locating and correcting a specified defect and examining the result of the correction.

Defect. A specified failure relative to a requirement or justified expectation. A disagreement about the requirement is a different question.

Diagnosis. Investigation and interpretation of a condition through evidence, possible explanations, and checks of those explanations.

Diagnostic control. In Simons's terminology, use of formal information systems to monitor selected performance variables against expectations. This is narrower than all diagnostic inquiry.

Double-loop learning. Learning that can question or change the governing arrangements within which an error was identified.

Event time. The time attributed to an event's occurrence. It may differ from the time at which the event was entered, processed, or reported.

Evidence. An artifact or observation considered in relation to a claim. The same artifact can support one claim while failing to support another.

Feedback. Information about consequences used in subsequent regulation or action. Its effects depend on delay, interpretation, and available responses.

Handoff. A transfer of work, information, or responsibility. These transfers may occur at different times even when a report presents them as one event.

Hypothesis. A proposed explanation that can be examined against evidence. Its usefulness includes the observations it suggests looking for next.

Indicator. A selected measure used to represent a condition. Its relationship to the underlying activity may change when its use changes.

Invariant. A relationship or feature expected to remain stable under specified conditions. Its failure can indicate a problem but does not identify every possible cause.

Jidoka. Toyota's concept of building the detection of abnormalities and appropriate stopping into production. Its organizational operation includes the response to the signal.

Model. A selective representation used for reasoning. Its omissions and assumptions determine which conclusions it can support.

Observability. Here, the arrangements that make evidence relevant to an organizational question obtainable and inspectable. This usage makes no general mathematical state-reconstruction claim.

Organizational memory. Records and operating arrangements through which earlier experience can affect later action. A stored report is only one possible component.

Pattern. A named recurrent arrangement useful for analysis in a specified context. Recognizing it does not by itself establish the causal mechanism.

Postmortem. A retrospective investigation of an incident intended to preserve evidence, explain contributing conditions, and support subsequent correction.

Provenance. Information about an artifact's origin, transformations, and custody relevant to assessing its meaning and use.

Psychological safety. A shared belief concerning the interpersonal consequences of speaking, questioning, or acknowledging error within a team.

Reconciliation. Comparison of related representations to identify and explain disagreement. Agreement may remain insufficient if the sources share an error.

Right censoring. An observation ends before the event of interest has occurred. An unresolved request at the end of follow-up provides incomplete duration information.

Single-loop learning. Detection and correction within existing governing arrangements. It can be appropriate when the governing arrangement is suitable and the defect is local.

Sociotechnical system. Work considered through the interaction of its social and technical arrangements rather than through either in isolation.

Special-cause variation. Variation indicating a departure from the continuing process system under the relevant statistical interpretation. It invites investigation of particular conditions.

Trace. An ordered representation of recorded activity. Its order, boundaries, and event meanings depend on how it was produced.

Triangulation. Examination through different kinds or sources of evidence. The sources' independence and distinct limitations still require checking.

Visibility limit. A boundary beyond which the available evidence cannot distinguish relevant alternatives. More analysis of the same evidence may not remove it.

Wicked problem. A planning problem whose formulation and acceptable response are bound up with contested judgments and consequences, as discussed by Rittel and Webber.

Appendix C Evidence taxonomy and blind spots

The taxonomy classifies artifacts by the questions they can help investigate. It is not a ranking in which a numerical record automatically has greater authority than a participant's account. Evidential strength depends on the claim, the method of collection, and the alternatives that the artifact can distinguish.

Evidence class	Questions it can help answer	Characteristic blind spots
Transaction records	What was recorded as received, allocated, paid, or completed?	Omitted transactions, misleading categories, and internally consistent false entries
Direct observation and recordings	What sequence or local condition was visible?	Framing, observation effects, unrecorded context, and selection of occasions
Plans and schedules	What was expected, by whom, and under which version?	Optimistic assumptions, informal changes, and later replacement of earlier plans
Workflow event logs	Which recorded transitions can be linked into a case?	Missing events, reused identifiers, batch entry, and work outside the system
Measures and dashboards	Which selected quantities changed under a definition?	Aggregation, denominator changes, target responses, and changed definitions
Decision records	What was authorized and what reasons were recorded?	Omitted alternatives, undocumented uncertainty, and retrospective justification
Interviews and narratives	How did participants understand conditions and events?	Recall, context, power relationships, and unequal opportunities to speak
Surveys	How did sampled respondents answer specified questions?	Nonresponse, question interpretation, sampling limits, and changed composition
Experiments and pilots	What happened under an intentionally changed arrangement?	Concurrent changes, spillovers, limited follow-up, and restricted transferability
Simulations and analytical models	What follows from stated assumptions and structure?	Omitted relationships, uncertain parameters, and several mechanisms fitting one pattern
Audit and investigation reports	What did a defined inquiry establish?	Scope limitations, dependence on supplied records, and confusion between assurance and success
Generated summaries and predictions	What does an automated transformation suggest?	Lost provenance, unsupported causal links, label defects, and changing model behavior

Relationships among artifacts

An artifact should be classified by its actual relation to the event as well as its format. A database field can contain a direct instrument reading, a person's judgment, or a value copied from another system. These have different provenance even when they appear in the same table. Likewise, a narrative can preserve an exact contemporaneous statement or a later reconstruction.

When two artifacts agree, we examine whether one copied the other or whether both depend on a common source. When they disagree, we examine differences in event boundary, clock, category, and version before deciding that one is false. A disagreement can reveal a useful relationship that a merged dataset would conceal.

Absence and collection conditions

An absent record becomes evidence only in relation to an expectation about collection. If every completed handoff normally generates a retained event, a missing event is diagnostically useful. If recording was optional or retention changed, the same absence supports a weaker conclusion. Collection failure and activity failure should remain separate hypotheses until further evidence connects them.

Sampling also matters. A collection of escalated cases describes the cases that reached escalation. It does not directly describe all unsuccessful cases, all customers, or all workers. A comparison of two reporting periods can be distorted by a change in who enters the collection even when the measurement rule remains unchanged.

Intervention and preservation

An intervention can remove evidence. Replacing a workflow, overwriting a classification, or merging case identifiers can make the earlier behavior difficult to reconstruct. A useful diagnostic record preserves the relevant earlier definitions and relationships alongside the changed arrangement. This allows later analysts to understand whether a reported improvement concerns the work, the representation, or both.

The final record of an investigation should state the question, evidence, supported explanation, intervention, and observation period. It should identify unresolved alternatives and the next evidence that could distinguish them. This is a practical continuation of the historical methods examined throughout the book: preserve enough of the analysis that a future correction does not have to begin from organizational amnesia.

References

Numbered references follow their first appearance in the text. Web links were checked during preparation in October 2026. Where a link identifies a later translation, reissue, repository manuscript, or explanatory resource, the entry states that relationship. Historical claims concern the work or event named, not the date on which its web page was created.

[1] Weisbord, M. R. (1976). Organizational Diagnosis: Six Places to Look for Trouble with or without a Theory. Group & Organization Studies, 1(4), 430–447. https://doi.org/10.1177/105960117600100405

[2] Vostokov, D. Theoretical Software Diagnostics. Fourth edition. OpenTask. Author’s reference text and associated Software Diagnostics Library. https://www.dumpanalysis.org/

[3] The Metropolitan Museum of Art. Cuneiform tablet: balanced account of Dugga. Neo-Sumerian, Ur III period, ca. 2039 BCE, Drehem. Accession 11.217.3; collection record 322447. https://www.metmuseum.org/art/collection/search/322447

[4] Library of Congress. Cuneiform Tablets: From the Reign of Gudea of Lagash to Shalmanassar III. About this collection. https://www.loc.gov/collections/cuneiform-tablets/about-this-collection/

[5] Institute of Chartered Accountants in England and Wales. The Earliest Books on Bookkeeping 1494 to 1683. Discussion of Luca Pacioli’s Summa de arithmetica, geometria, proportioni et proportionalita, Venice, 1494. https://www.icaew.com/library/library-collection/historical-accounting-literature/earliest-books-on-bookkeeping

[6] Babbage, C. (1832). On the Economy of Machinery and Manufactures. Charles Knight. Digital text, Project Gutenberg. https://www.gutenberg.org/ebooks/4238

[7] Taylor, F. W. (1911). The Principles of Scientific Management. Harper & Brothers. Digital text, Project Gutenberg. https://www.gutenberg.org/ebooks/6435

[8] Fayol, H. (1949). General and Industrial Management. Translated by C. Storrs. Pitman. French original serialized in 1916; book edition, Dunod, 1917. https://search.worldcat.org/title/825227

[9] Weber, M. (1978). Economy and Society: An Outline of Interpretive Sociology. G. Roth and C. Wittich, eds. University of California Press. Two volumes. Original German publication, 1922; linked catalog record identifies the 1978 English edition. https://search.worldcat.org/title/Economy-and-society-:-an-outline-of-interpretive-sociology/oclc/925682795

[10] Purdue University Archives and Special Collections. Frank and Lillian Gilbreth papers, MSP 7. Series 7, Audio Visual. Frank and Lillian Gilbreth Motion Study Films, ca. 1918–1968, including undated material. Original films and later compilations. Finding aid, revised November 2, 2023, pp. 29–31. https://lib.purdue.edu/archives/wp-content/uploads/sites/3/2026/01/msp7-gilbreth_papers.pdf

[11] Clark, W. (1922). The Gantt Chart: A Working Tool of Management. With appendices by W. N. Polakov and F. W. Trabold. Ronald Press. Digitized copy, Library of Congress. https://www.loc.gov/item/22009947/

[12] Baker Library, Harvard Business School. The Human Relations Movement: Harvard Business School and the Hawthorne Experiments, 1924–1933. Historical exhibition. https://www.library.hbs.edu/hc/hawthorne/

[13] Baker Library, Harvard Business School. Illumination Studies and Relay Assembly Test Room. The Human Relations Movement exhibition. https://www.library.hbs.edu/hc/hawthorne/03.html

[14] Levitt, S. D., and List, J. A. (2011). Was There Really a Hawthorne Effect at the Hawthorne Plant? An Analysis of the Original Illumination Experiments. American Economic Journal: Applied Economics, 3(1), 224–238. https://doi.org/10.1257/app.3.1.224

[15] Follett, M. P. The Giving of Orders. Extract reproduced as The Law of the Situation, dated 1925 by the Panarchy text archive. https://panarchy.org/follett/lawsituation.html

[16] Barnard, C. I. (1938). The Functions of the Executive. Harvard University Press. https://search.worldcat.org/title/555075

[17] American Society for Quality. Walter A. Shewhart. Biographical account including Economic Control of Quality of Manufactured Product (1931). https://asq.org/about-asq/honorary-members/shewhart

[18] American Society for Quality. Control Chart. Statistical process control guidance and interpretation. https://asq.org/quality-resources/control-chart

[19] The W. Edwards Deming Institute. PDSA Cycle. Account of Plan–Do–Study–Act and its relationship to Shewhart's work. https://deming.org/explore/pdsa/

[20] Deming, W. E. (1993). The New Economics for Industry, Government, Education. Massachusetts Institute of Technology, Center for Advanced Engineering Study. First-edition catalog record; publisher link identifies the third edition (2018). https://books.google.com/books?id=WA_tAAAAMAAJ https://mitpress.mit.edu/9780262535939/the-new-economics-for-industry-government-education/

[21] American Society for Quality. Joseph M. Juran. Biographical account of his quality-management work and the Quality Control Handbook, first published in 1951. https://asq.org/about-asq/honorary-members/juran

[22] American Society for Quality. Kaoru Ishikawa. Biographical account of company-wide quality control and quality circles. https://asq.org/about-asq/honorary-members/ishikawa

[23] Toyota Motor Corporation. Toyota Production System. Toyota Virtual Plant Tour. Jidoka, andon, mistake prevention, and just-in-time production. https://global.toyota/en/company/plant-tours/production-system/

[24] INFORMS (2002). Time Line for Operations Research, On the Occasion of OR's 50th Anniversary. March 22. https://www.informs.org/News-Room/INFORMS-Releases/News-Releases/Time-Line-for-Operations-Research-On-the-Occasion-of-OR-s-50th-Anniversary-March-22-2002

[25] Dantzig, G. B. Oral history interview. INFORMS History of O.R. Excellence. Interviewed by Irv Lustig. Recollection of linear programming and the simplex method in 1947. https://www.informs.org/Explore/History-of-O.R.-Excellence/Oral-Histories/George-Dantzig

[26] Simon, H. A. (1955). A Behavioral Model of Rational Choice. The Quarterly Journal of Economics, 69(1), 99–118. https://doi.org/10.2307/1884852

[27] Cohen, M. D., March, J. G., and Olsen, J. P. (1972). A Garbage Can Model of Organizational Choice. Administrative Science Quarterly, 17(1), 1–25. https://doi.org/10.2307/2392088

[28] Drucker, P. F. (1954). The Practice of Management. Harper & Brothers. Chapter 11, Management by Objectives and Self-Control. Publisher link identifies the 2007 reissue. https://www.routledge.com/The-Practice-of-Management/Drucker/p/book/9780750685047

[29] Kepner, C. H., and Tregoe, B. B. (1965). The Rational Manager: A Systematic Approach to Problem Solving and Decision Making. McGraw-Hill. https://books.google.com/books?id=Qn8qAAAAMAAJ

[30] Kepner, C. H., and Tregoe, B. B. The New Rational Manager. Chapter 2, Problem Analysis. Chapter extract hosted by Kepner-Tregoe. https://kepner-tregoe.com/wp-content/uploads/2021/09/Chapter-02_New-Rational-Manager-1.pdf

[31] Rittel, H. W. J., and Webber, M. M. (1973). Dilemmas in a General Theory of Planning. Policy Sciences, 4(2), 155–169. https://doi.org/10.1007/BF01405730

[32] Ackoff, R. L. (1979). The Future of Operational Research Is Past. Journal of the Operational Research Society, 30(2), 93–104. https://doi.org/10.1057/jors.1979.22

[33] Checkland, P. (1981). Systems Thinking, Systems Practice. Wiley. Later reissue includes a retrospective. https://search.worldcat.org/title/1036880420

[34] Lewin, K. (1946). Action Research and Minority Problems. Journal of Social Issues, 2(4), 34–46. https://doi.org/10.1111/j.1540-4560.1946.tb02295.x

[35] Trist, E. L., and Bamforth, K. W. (1951). Some Social and Psychological Consequences of the Longwall Method of Coal-Getting: An Examination of the Psychological Situation and Defences of a Work Group in Relation to the Social Structure and Technological Content of the Work System. Human Relations, 4(1), 3–38. https://doi.org/10.1177/001872675100400101

[36] Nadler, D. A., and Tushman, M. L. (1980). A Model for Diagnosing Organizational Behavior. Organizational Dynamics, 9(2), 35–51. https://doi.org/10.1016/0090-2616(80)90039-X

[37] Burke, W. W., and Litwin, G. H. (1992). A Causal Model of Organizational Performance and Change. Journal of Management, 18(3), 523–545. https://doi.org/10.1177/014920639201800306

[38] Ashby, W. R. (1956). An Introduction to Cybernetics. Chapman & Hall. Electronic reproduction permitted for nonprofit educational use; scanned title page is the second impression, 1957. https://www.ashby.info/Ashby-Introduction-to-Cybernetics.pdf

[39] Beer, S. (1981). Brain of the Firm: The Managerial Cybernetics of Organization. Second edition. Wiley. First edition published by Allen Lane in 1972; publisher link identifies the 1994 paperback reissue of the second edition. https://www.wiley-vch.de/de/fachgebiete/finanzen-wirtschaft-recht/brain-of-the-firm-978-0-471-94839-1

[40] Medina, E. (2011). Cybernetic Revolutionaries: Technology and Politics in Allende's Chile. MIT Press. https://mitpress.mit.edu/9780262016490/cybernetic-revolutionaries/

[41] Medina, E. (2023). Project Cybersyn: Chile's Radical Experiment in Cybernetic Socialism. The MIT Press Reader. Extract discussing the project begun in 1971. https://thereader.mitpress.mit.edu/project-cybersyn-chiles-radical-experiment-in-cybernetic-socialism/

[42] Forrester, J. W. (1961). Industrial Dynamics. MIT Press. https://books.google.com/books?id=KRzHAAAAIAAJ

[43] Argyris, C., and Schön, D. A. (1978). Organizational Learning: A Theory of Action Perspective. Addison-Wesley. https://search.worldcat.org/title/895549588

[44] Mintzberg, H. (1975). The Manager's Job: Folklore and Fact. Harvard Business Review, 53(4), 49–61. Author's publication record. https://mintzberg.org/articles/managers-job-folklore-and-fact

[45] Kaplan, R. S., and Norton, D. P. (1992). The Balanced Scorecard—Measures that Drive Performance. Harvard Business Review, 70(1), 71–79. The 1992 article and the related author account, Partnering and the Balanced Scorecard (2002), are linked below. https://hbr.org/1992/01/the-balanced-scorecard-measures-that-drive-performance-2 https://www.library.hbs.edu/working-knowledge/partnering-and-the-balanced-scorecard

[46] Simons, R. (1995). Levers of Control: How Managers Use Innovative Control Systems to Drive Strategic Renewal. Harvard Business School Press. Author's Harvard Business School publication record. https://www.hbs.edu/faculty/Pages/item.aspx?num=257

[47] Kerr, S. (1975). On the Folly of Rewarding A, While Hoping for B. Academy of Management Journal, 18(4), 769–783. https://doi.org/10.2307/255378

[48] Campbell, D. T. (1976). Assessing the Impact of Planned Social Change. Occasional Paper 8. Reprinted in Journal of MultiDisciplinary Evaluation, 7(15), 3–43 (2011).
https://doi.org/10.56645/jmde.v7i15.297

[49] Goodhart, C. A. E. (1975). Problems of Monetary Management: The U.K. Experience. Paper associated with the Reserve Bank of Australia's July 1975 conference. The RBA's 1990 bibliography dates the revised conference volume 1976. Reprinted in Monetary Theory and Practice: The UK Experience (Macmillan, 1984), pp. 91–121.
https://www.rba.gov.au/publications/rdp/1990/9013/conference-volumes.html
https://doi.org/10.1007/978-1-349-17295-5_4

[50] Strathern, M. (1997). "Improving ratings": Audit in the British University System. European Review, 5(3), 305–321. The formulation about a measure becoming a target appears on p. 308.
https://doi.org/10.1002/(SICI)1234-981X(199707)5:3<305::AID-EURO184>3.0.CO;2-4

[51] Power, M. (1997). The Audit Society: Rituals of Verification. Oxford University Press. Linked edition is the 1999 paperback.
https://doi.org/10.1093/acprof:oso/9780198296034.001.0001

[52] Committee of Sponsoring Organizations of the Treadway Commission (2013). Internal Control—Integrated Framework: Executive Summary. May.
https://www.coso.org/_files/ugd/3059fc_1df7d5dd38074006bce8fdf621a942cf.pdf

[53] Scott, J. C. (1998). Seeing Like a State: How Certain Schemes to Improve the Human Condition Have Failed. Yale University Press. Linked edition is the 1999 paperback.
https://yalebooks.yale.edu/book/9780300078152/seeing-like-a-state/

[54] Presidential Commission on the Space Shuttle Challenger Accident (1986). Report, Volume I, Chapter V, The Contributing Cause of the Accident. NASA historical archive.
https://www.nasa.gov/history/rogersrep/v1ch5.htm

[55] Vaughan, D. (1996). The Challenger Launch Decision: Risky Technology, Culture, and Deviance at NASA. University of Chicago Press. Publisher link identifies the enlarged edition with a new preface.
https://press.uchicago.edu/ucp/books/book/chicago/C/bo22781921.html

[56] Columbia Accident Investigation Board (2003). Columbia Accident Investigation Board Report, Volume I. August 1. NASA Technical Reports Server record 20030066167.
https://ntrs.nasa.gov/citations/20030066167

[57] Edmondson, A. (1999). Psychological Safety and Learning Behavior in Work Teams. Administrative Science Quarterly, 44(2), 350–383.
https://doi.org/10.2307/2666999

[58] Hammer, M. (1990). Reengineering Work: Don't Automate, Obliterate. Harvard Business Review, 68(4), 104–112.
https://hbr.org/1990/07/reengineering-work-dont-automate-obliterate

[59] van der Aalst, W. M. P., Weijters, A. J. M. M., and Măruşter, L. (2004). Workflow Mining: Discovering Process Models from Event Logs. IEEE Transactions on Knowledge and Data Engineering, 16(9), 1128–1142.
https://doi.org/10.1109/TKDE.2004.47

[60] Beck, K., and sixteen co-signatories (2001). Manifesto for Agile Software Development, and Principles behind the Agile Manifesto.
https://agilemanifesto.org/
https://agilemanifesto.org/principles.html

[61] Schwaber, K., and Sutherland, J. (2020). The Scrum Guide. November.
https://scrumguides.org/scrum-guide.html

[62] Beyer, B., Jones, C., Petoff, J., and Murphy, N. R., eds. (2016). Site Reliability Engineering: How Google Runs Production Systems. O'Reilly Media. Chapter 15, Postmortem Culture: Learning from Failure, by John Lunney and Sue Lueder.
https://sre.google/sre-book/postmortem-culture/

[63] Rousseau, D. M. (2006). Is There Such a Thing as "Evidence-Based Management"? Academy of Management Review, 31(2), 256–269.
https://doi.org/10.5465/amr.2006.20208679

[64] Bloom, N., Eifert, B., Mahajan, A., McKenzie, D., and Roberts, J. (2013). Does Management Matter? Evidence from India. The Quarterly Journal of Economics, 128(1), 1–51. Published online in November 2012. https://doi.org/10.1093/qje/qjs044

[65] Weick, K. E. (1995). Sensemaking in Organizations. SAGE. https://www.sagepub.com/shop/buy-a-book/sensemaking-in-organizations-1-4988

[66] Morgan, G. (1986). Images of Organization. SAGE. https://search.worldcat.org/title/13424067

[67] Baiocco, S., Fernández-Macías, E., Rani, U., and Pesole, A. (2022). The Algorithmic Management of Work and Its Implications in Different Contexts. ILO background paper and JRC Working Papers on Labour, Education and Technology, 2022-02. June 21. https://www.ilo.org/publications/algorithmic-management-work-and-its-implications-different-contexts

[68] Milanez, A., Lemmens, A., and Ruggiu, C. (2025). Algorithmic Management in the Workplace: New Evidence from an OECD Employer Survey. OECD Artificial Intelligence Papers, No. 31. OECD Publishing. https://doi.org/10.1787/287c13c4-en

[69] National Institute of Standards and Technology (2023). Artificial Intelligence Risk Management Framework (AI RMF 1.0). NIST AI 100-1. January. https://doi.org/10.6028/NIST.AI.100-1

[70] Vostokov, D. Critique of Management Reason: Management Bits and Tips from Can't. OpenTask. Copyright pages: © 2010, 2025; Leanpub sample, Version 1.01; OpenTask table-of-contents PDF, Version 1.0. Both dated November 2025 and identifying ISBN 978-1-906717-13-1 as the paperback ISBN. Leanpub digital edition marked updated January 16, 2026. https://leanpub.com/bits https://leanpub.com/bits/sample.pdf https://www.dumpanalysis.org/CMR/Critique-of-Management-Reason-TOC.pdf

[71] Vostokov, D. (2008). Opentask Publishing Plans. Software Diagnostics Library. June 26. Announcement includes the projected Management Bits anthology and its ISBN. https://www.dumpanalysis.org/blog/index.php/2008/06/26/opentask-publishing-plans/

[72] Vostokov, D. (2010). Forthcoming Management Bits: An Anthology from Reductionist Manager. Software Diagnostics Library. July 10. Announcement gives September 2010 as the intended publication date and ISBN 978-1906717131. https://www.dumpanalysis.org/blog/index.php/2010/07/10/forthcoming-management-bits-an-anthology-from-reductionist-manager/

[73] Kociatkiewicz, J., and Kostera, M. (2012). Sherlock Holmes and the Adventure of the Rational Manager: Organizational Reason and Its Discontents. Scandinavian Journal of Management, 28(2), 162–172. Author manuscript: discussion on manuscript page 8 and Vostokov reference on manuscript page 15; these are PDF pages 9 and 16 in the White Rose copy, not journal pagination. https://doi.org/10.1016/j.scaman.2012.01.003 https://hal.science/hal-02423778v1 https://eprints.whiterose.ac.uk/101001/1/Kociatkiewicz%20and%20Kostera%20-%20Holmes%20and%20the%20rational%20manager%20%28unformatted%29.pdf

[74] Vostokov, D. Trace, Log, Text, Narrative, Data: An Analysis Pattern Reference for Information Mining, Diagnostics, Anomaly Detection. Fifth edition. OpenTask. Author's reference text and associated online trace analysis pattern catalog. https://www.dumpanalysis.org/blog/index.php/category/trace-analysis-patterns/

