ORIGINAL RESEARCH Open access

# The Degradation of Decision Evidence When AI Optimization Reshapes the Business Metrics and Data Distributions Used to Evaluate Its Own Performance

David Chen[1*], Julia Weber[2], Mei Lin[3], Daniel Okoro[4]

## Abstract

Artificial intelligence increasingly participates in the organizational processes through which performance is measured, interpreted, and improved. This creates a theoretical problem when an AI system does more than predict outcomes: its recommendations, rankings, controls, or allocations can alter the activities and observations that later constitute evidence of its own performance. This article develops a systems and feedback-theoretic account of that problem. It distinguishes decision evidence from passive data, treats AI optimization as an intervention in the data-generating process, and explains how target adaptation, selective visibility, distributional feedback, and organizational learning can make post-deployment evaluation progressively self-referential. The central claim is conditional rather than universal. Evidence degradation becomes plausible when optimization changes behavior or case selection, those changes reshape measured outcomes or observable distributions, and evaluation reuses the altered evidence without a sufficiently independent reference channel. Reinforcing loops can then make locally improved metrics increasingly poor indicators of system-level performance, whereas balancing mechanisms such as independent validation, preserved counterfactual observations, and organizational challenge can interrupt the process. The resulting account separates beneficial adaptive feedback from evidence-corrupting feedback and identifies the conditions under which identical apparent performance trajectories can imply different evidentiary states.

**Keywords** Algorithmic management, Distribution shift, Feedback systems, Endogenous measurement, Decision evidence, Performative prediction

*Correspondence:
David Chen
david.chen@nus.edu.sg
[1] Department of Information Systems, National University of Singapore, Singapore, Singapore
[2] Department of Information Systems, Free University of Berlin, Berlin, Germany
[3] Department of Analytics and Operations, National University of Singapore, Singapore, Singapore
[4] Department of Business Administration, University of Lagos, Lagos, Nigeria

## Introduction

Artificial intelligence is increasingly embedded in performance measurement and management systems, where it supports monitoring, prediction, prioritization, and prescriptive decision support [1]. Recent work similarly connects AI-enabled performance systems with human capital and organizational decision-making, emphasizing that measurement and managerial interpretation are becoming more tightly coupled [2]. In management accounting, digitalization and AI are also altering information production, professional roles, and the organizational interfaces through which performance is interpreted [3]. These developments make a familiar assumption less secure: that the evidence used to evaluate an analytical system is generated independently of the system being evaluated.

This is an open access article distributed under the terms of the Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License (CC BY-NC-SA 4.0), which permits non-commercial use, distribution, reproduction, and adaptation in any medium or format, provided that appropriate credit is given to the original author(s) and source, a link to the license is provided, and any changes made are indicated. Adapted material must be distributed under the same license.

© 2026 The Author(s). | Published by Institute for Management, Business, and Accounting Studies

The theoretical difficulty becomes clearer once AI is treated as heterogeneous rather than as a single technological intervention. Predictive, generative, agentic, and embodied systems differ in the kinds of organizational actions they can initiate or reshape [4]. AI management is further complicated by the interaction of autonomy, learning, and inscrutability [5]. Automation and augmentation also coexist dynamically rather than occupying mutually exclusive organizational states [6]. A system-level perspective is therefore needed because AI performance can emerge from changing relations among algorithms, workers, managers, data infrastructures, and decision rights rather than from model properties alone [7].

This article examines a specific consequence of that interdependence: the possibility that AI optimization changes the business metrics and data distributions subsequently used to assess whether the optimization worked. The problem is not simply model drift, biased data, incentive gaming, or inaccurate KPIs. It arises when evaluation becomes recursively linked to an intervention that has already altered the environment from which evaluation evidence is drawn.

Systems and feedback theory provides the appropriate analytical form because the phenomenon depends on reciprocal causation, delays, reinforcing and balancing responses, and divergence between local and system objectives. The central question is therefore conditional: under what configurations does adaptive AI improve the evidence available for decision-making, and under what configurations does it progressively reduce the independence and diagnostic value of that evidence?

# Decision Evidence and Endogenous Measurement

Organizational evidence is rarely produced by observation alone. Rules, classifications, routines, and decision procedures participate in determining what organizations record and subsequently treat as evidence. The organization-as-algorithm perspective makes this constitutive role visible by showing how algorithmic rules can become part of the organizing process rather than remaining external analytical aids [8]. AI deployment similarly depends on continuing forms of data work, knowledge work, and values work, meaning that the information supplied to analytical systems reflects organizational choices about collection, interpretation, correction, and relevance [9].

For the present analysis, decision evidence denotes information used to justify, compare, revise, or continue an organizational decision or decision-support regime. Such evidence becomes endogenous when its content, availability, composition, or meaning is partly produced by actions taken under the regime being evaluated. Endogeneity is therefore a matter of dependence between intervention and observation, not merely the existence of measurement error.

Algorithmic management demonstrates why that distinction matters. Validated measures of algorithmic management separate functions such as monitoring, goal setting, scheduling, performance rating, and compensation rather than treating managerial automation as a single exposure [10]. These functions differ in how strongly they can influence the behavior later recorded by the same systems. Longitudinal evidence on algorithmic HR systems further shows that employees respond differently when automated performance systems are interpreted as mechanisms of managerial control versus constructive feedback; those interpretations can influence motivation and task prioritization [11]. Recorded performance therefore may partly reflect adaptation to the measurement regime itself.

Learning algorithms deepen this endogeneity because organizational practices can adjust as algorithms learn, while algorithms subsequently learn from the outcomes of those adjusted practices [12]. The relevant analytical object is consequently not a static dataset feeding a static model. It is a coupled system in which measurement, organizational response, and model updating can repeatedly alter one another.

# AI Optimization as an Intervention in the Data-Generating Process

© 2026 The Author(s). | Published by Institute for Management, Business, and Accounting Studies

An AI system becomes consequential for future evidence when its outputs influence what people do, which cases they pursue, what gets monitored, and which outcomes become organizationally important. Research on algorithmic management in scientific work demonstrates that AI can participate directly in task division, allocation, direction, coordination, motivation, and learning support [13]. Once such functions are coupled to operational activity, the model is no longer merely estimating a pre-existing environment; it is helping create the environment from which later observations are drawn.

This intervention can occur through more than one layer. Experimental work on algorithmic control distinguishes the design of the control mechanism from the data layer and from the broader organizational embedding of that control [14]. The distinction is important for feedback theory because different layers can alter the data-generating process through different mechanisms. A scheduling recommendation changes who encounters which cases. Monitoring changes what becomes observable. Target setting changes which actions are rewarded. Automated evaluation changes which performances are classified as successful.

Broader syntheses of algorithmic management show that these interventions are organizationally heterogeneous rather than mechanically uniform [15]. Workers and managers can comply, reinterpret, resist, circumvent, or strategically adapt to algorithmic direction. Algorithmic control is therefore a contested organizational process in which monitoring and evaluation can generate responses that modify the conditions under which subsequent monitoring and evaluation occur [16].

The resulting feedback system contains at least two distinct objective levels. A local optimization objective may reward increased conversion, throughput, ranking accuracy, utilization, or another operational metric. The wider system may instead depend on durable customer value, learning, safety, fairness, coverage of difficult cases, or information about alternatives that were not chosen. When AI intervention changes the data-generating process, local improvement can occur simultaneously with declining visibility into the wider system objective.

The transition from a passive measurement regime to an intervention-coupled regime therefore changes the meaning of subsequent evidence: observations recorded after deployment are partly outcomes of the very optimization policy whose performance they are later asked to assess.

The relationships discussed here are shown in **Figure 1.**

© 2026 The Author(s). | Published by Institute for Management, Business, and Accounting Studies

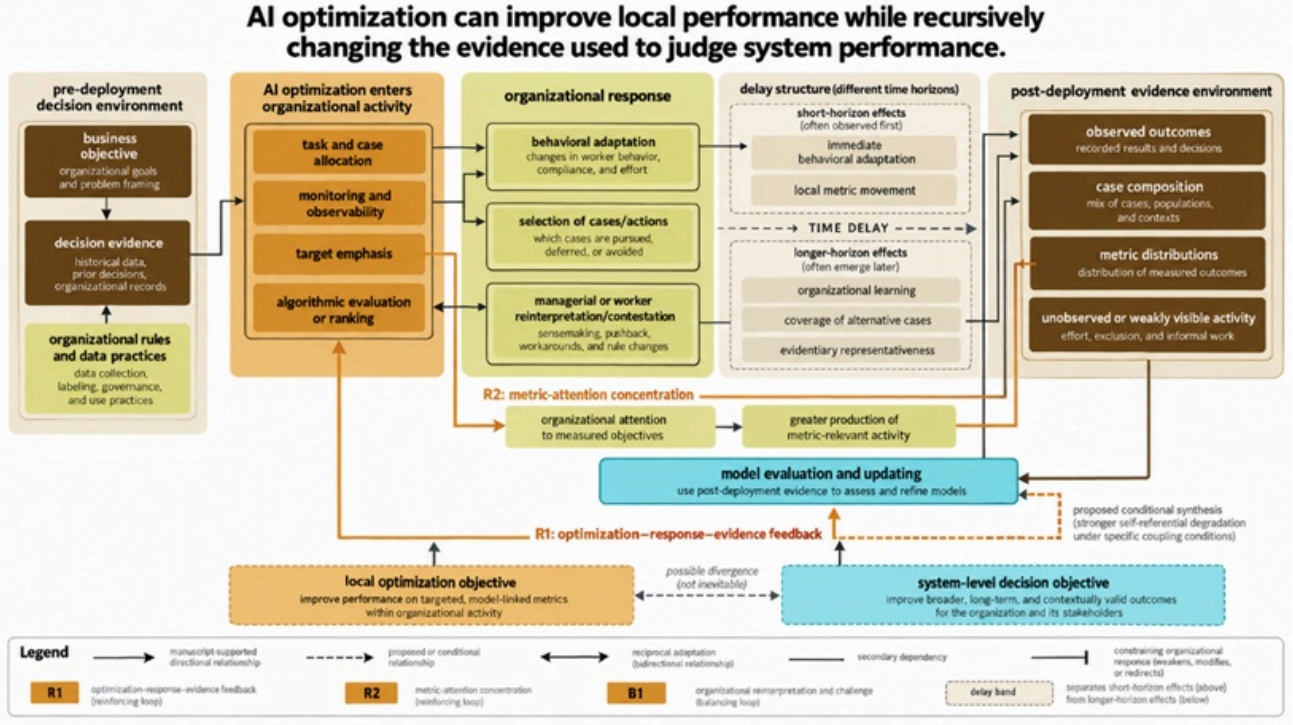


**Figure 1.** AI optimization can improve local performance while recursively changing the evidence used to judge system performance.

# Metric Reshaping

Intervention in activity does not automatically degrade decision evidence. The stronger problem begins when organizational actors or processes adapt to the metric itself. Formal analysis of Goodhart-type effects in machine-learning settings shows how strategic modification of measured covariates can generate bias when predictive systems are coupled to incentives and optimization [17]. The implication for organizational AI is conditional: a metric can lose diagnostic value when it becomes sufficiently important to the allocation of rewards, attention, access, or sanctions that behavior is reorganized around the measure.

The same adaptation need not take the form of deliberate gaming. Systematic research on algorithmic management in gig work shows substantial heterogeneity in how workers experience and respond to algorithmic management [18]. Some responses can represent attempts to preserve autonomy, cope with opaque control, or reallocate effort rather than purposeful manipulation of the system. Empirical work nevertheless documents adverse behavioral pathways under algorithmic control [19], while other evidence shows that control can generate countervailing effects instead of a uniformly negative outcome [20].

Metric reshaping should therefore be understood as a change in the relationship between an indicator and the underlying organizational objective. That relationship can weaken because actors optimize the indicator directly, because AI changes which activities occur, or because the organization shifts attention toward dimensions that the system measures well. It can also remain stable when incentives are weak, measurement captures the substantive objective closely, or independent evidence constrains adaptation.

The crucial feedback-theoretic distinction is between improving the measured state and improving the system state that the measure was intended to represent. The two can move together, but systems that repeatedly optimize and evaluate using the same measurement channel create conditions under which they may separate.

© 2026 The Author(s). | Published by Institute for Management, Business, and Accounting Studies

The intervention mechanisms developed above imply distinct forms of metric reshaping rather than a single generic Goodhart effect.

**Table 1** presents the analytical distinctions discussed here.

**Table 1.** How AI interventions can reshape the informational meaning of the metrics they optimize

| Intervention locus | How AI enters the data-generating process | What becomes selectively optimized | How the recorded metric can change meaning | Condition that can preserve diagnostic value |
|---|---|---|---|---|
| Task or case allocation | Directs actors toward some cases and away from others | Cases predicted to yield preferred outcomes or efficient completion | Aggregate performance increasingly reflects the selected case mix | Evaluation retains comparable unselected or independently sampled cases |
| Monitoring intensity | Expands observation of selected behaviors or stages | Activities that are digitally visible and auditable | Improvement can partly reflect greater observability rather than broader system improvement | Measurement coverage includes weakly observed or noninstrumented activity |
| Target emphasis | Reweights organizational attention around explicit performance objectives | Actions that improve the targeted indicator | Indicator becomes progressively more behavior-shaping and less purely descriptive | Ultimate outcome remains independently measured and difficult to proxy-game |
| Algorithmic ranking or evaluation | Converts observed behavior into comparative scores or decisions | Activities that improve position under the scoring rule | Score distributions increasingly reflect adaptation to the evaluator | Scoring rules are periodically challenged against external outcomes |
| Automated feedback and updating | Uses newly observed outcomes to revise future recommendations | Patterns reinforced by prior model-directed activity | Future training/evaluation data become increasingly policy-dependent | Updating includes data not generated under the current policy |

# Distributional Feedback and Selective Visibility

Metric reshaping concerns the changing relationship between a measure and an objective. Distributional feedback concerns a different process: the model or decision regime changes which observations subsequently exist or become visible. Performative-prediction research formalizes this possibility by analyzing settings in which predictions affect the distribution from which future predictions are evaluated [21]. Applied organizationally, the implication is not that every deployed model is performative, but that deployment becomes performative when its outputs influence behavior, selection, exposure, or resource allocation strongly enough to change later data.

A sales-prioritization model, for example, may concentrate contact on customers predicted to convert. A fraud model may intensify scrutiny of transactions already classified as suspicious. A staffing model may direct experienced personnel toward cases classified as difficult. Subsequent outcome data then contain information about performance under the intervention, not simply about the pre-existing environment. Cases that were never selected, contacted, inspected, or attempted may become weakly represented precisely because the optimization policy made them less likely to occur.

© 2026 The Author(s). | Published by Institute for Management, Business, and Accounting Studies

This distinction matters because an observed performance decline or improvement can originate from multiple kinds of distributional change. Methods for diagnosing model performance under distribution shift separate changes in population frequency, conditional relationships, and poorly represented or newly encountered observations rather than attributing all degradation to one source [22]. Robust-validation methods similarly show that future predictive validity can be evaluated under specified forms of distribution shift, although the protection obtained depends on explicit statistical assumptions about the relationship between training and target distributions [23].

These methodological developments suggest a balancing principle for organizational evaluation: the more strongly a model shapes the evidence stream it later encounters, the more important it becomes to retain comparison evidence whose generation is less dependent on the current optimization policy. That evidence may involve independent sampling, delayed outcomes, parallel operational channels, externally defined targets, or other reference observations. Such mechanisms do not eliminate endogeneity automatically, but they provide a way to diagnose whether apparent performance reflects environmental learning or self-produced evidence.

Feedback also has a constructive configuration. Data-network-effect theory explains how repeated use can generate additional data, improve an AI-enabled service, and create further user value when stewardship, legitimacy, and system design support the loop [24]. Endogenous data production is therefore not synonymous with degradation. The critical theoretical problem is whether recursive data generation expands valid knowledge about the organizational objective or progressively concentrates evidence around behavior already induced by the optimizer.

# Self-Referential Performance Evaluation

Evaluation becomes self-referential when evidence generated under an AI-guided decision regime is interpreted as though it were independent evidence about the quality of that regime. The problem is especially visible when predictive performance is used as a proxy for organizational decision performance. Field evidence on algorithm-supported inspections shows that improvements in predictive sophistication do not automatically translate into better realized organizational decisions; decision authority and the way algorithmic information enters the organizational process matter [25]. A system can therefore improve one local analytical metric without establishing that the wider decision system has improved.

A second form of self-reference appears over longer horizons. Theory on substituting human decision-making with machine learning argues that automation can change the richness of organizational learning because it changes which actions are attempted and which outcomes become available for experience-based learning [26]. When future evaluation relies on observations generated after those opportunities have already been narrowed or redirected, evidence can become increasingly dependent on prior model choices.

The resulting self-reference has three analytically separate components. First, selection dependence occurs when the model influences which cases enter the observed sample. Second, behavioral dependence occurs when people adapt to monitored or rewarded variables. Third, learning dependence occurs when prior model-directed actions shape which relationships the organization can subsequently discover. None establishes degradation by itself. The stronger condition arises when these dependencies accumulate while the organization continues to use the resulting evidence as its primary basis for evaluating the same optimization regime.

This explains why apparent stability can be misleading. A model may remain well calibrated on the cases it continues to generate, and a KPI may improve among the activities the organization continues to measure, while the organization becomes progressively less informed about excluded alternatives, counterfactual actions, or longer-horizon consequences. Conversely, self-reference can remain limited when important outcome evidence originates outside the optimizer’s influence or when governance preserves meaningful opportunities to challenge the operational evidence stream.

© 2026 The Author(s). | Published by Institute for Management, Business, and Accounting Studies

The evidence stream therefore needs to be characterized by how it was generated, not only by its statistical properties.

Table 2 presents the analytical distinctions discussed here.

**Table 2.** Evidence-stream changes that can make post-deployment AI evaluation increasingly self-referential

| Evidence-stream change | What changes in the observed environment | Why apparent AI performance can change | Form of self-reference created | Evidence needed to discriminate adaptation from degradation |
|---|---|---|---|---|
| Case-composition shift | The system repeatedly encounters cases favored by prior allocation decisions | Performance may rise because the evaluated population becomes easier or more homogeneous | Current policy determines part of the future evaluation sample | Independent or randomized sampling of cases outside the normal allocation policy |
| Selective visibility | Measured activity becomes more observable than unmeasured or weakly instrumented activity | Recorded performance can improve while poorly observed dimensions disappear from evaluation | The evaluator increasingly sees what the optimization regime makes visible | Parallel measures of outcomes not tied to the optimizer's monitoring channel |
| Behavioral adaptation | Workers, customers, or managers change actions in response to targets, monitoring, rankings, or recommendations | Metric trajectories reflect response to the evaluator as well as underlying capability | Evaluation partly measures adaptation to the measurement rule | External outcome measures and repeated checks after rule or incentive changes |
| Conditional-relation shift | Relationships between predictors, decisions, and outcomes change after deployment | Previously valid performance estimates can cease to describe current decision consequences | The model helps create the conditions under which its old associations are retested | Shift decomposition and temporally separated validation |
| Learning-opportunity compression | Model-directed choices reduce exposure to alternative actions or difficult cases | Stable short-run performance can coexist with weaker long-run knowledge about alternatives | Prior optimization constrains what future evidence can reveal | Preserved exploration, counterfactual evidence, or deliberately protected learning channels |

# Decision-Evidence Degradation Framework

Decision-evidence degradation is best understood as a system configuration rather than as a property of a dataset or metric considered in isolation. Three conditions must become coupled. First, optimization must materially influence organizational behavior, case selection, or attention. Second, those responses must change the evidence later available for assessment. Third, the changed evidence must be reused to evaluate or update the optimization regime without an adequately independent reference channel. Strategic adaptation to measured variables provides one route through which the first two conditions can become linked [17]. Performative prediction provides a complementary mechanism: predictions can alter the distribution on which subsequent prediction and evaluation occur [21].

© 2026 The Author(s). | Published by Institute for Management, Business, and Accounting Studies

The resulting feedback can reinforce itself without producing an obvious failure signal. If a model directs activity toward cases on which its recommendations work well, subsequent performance may improve partly because the observed population has become increasingly compatible with the policy. If organizational attention concentrates on measured objectives, the measured objective may become easier to optimize while progressively representing a narrower portion of organizational performance. Repeated updating can then strengthen the apparent empirical foundation of the policy because each cycle contains more data produced under prior cycles.

This process becomes decision-evidence degradation when the evidence stream loses information needed to discriminate between genuine system improvement and improvement generated by the optimization regime's own influence on measurement. Degradation is therefore relational: an operational dataset may remain accurate about what happened while becoming less informative about what would have happened under relevant alternatives. The problem concerns evidentiary independence and counterfactual reach, not necessarily data quality in the conventional sense.

The mechanism also admits balancing configurations. Distribution-shift diagnosis can identify whether changes originate in population composition, conditional relationships, or weakly represented observations [22]. Robust validation can preserve inferential protection against specified shifts when its assumptions are met [23]. Data feedback can also be productive: recursive use–data–learning relationships may increase user value when stewardship and system design preserve informative learning [24]. Feedback becomes theoretically problematic only when it narrows the evidence needed to evaluate the objective that justified optimization.

The framework therefore distinguishes adaptive evidence accumulation from self-referential evidence concentration. Adaptive accumulation expands or improves knowledge about the system objective. Self-referential concentration increasingly documents performance within conditions created by the optimizer itself. The distinction is not observable from a rising or falling KPI alone; it requires evidence about how observations entered the evaluation stream.

# Detection and Governance Conditions

Detection begins by separating changes in the analytical model from changes in the system that generates its inputs and outcomes. Algorithmic control can be decomposed into control mechanisms, data practices, and organizational embedding, which cautions against interpreting an aggregate performance change as evidence of a single causal mechanism [14]. Statistical diagnostics provide a complementary distinction by decomposing distributional change rather than treating any performance deterioration as homogeneous shift [22].

A useful diagnostic question is therefore whether an observed trajectory persists when evaluation is moved away from evidence generated under the current optimization policy. Where feasible, independent sampling, delayed outcomes, protected exploration, or externally defined outcome measures can provide comparison channels. Robust-validation methods demonstrate that inference can explicitly accommodate specified future distribution shifts [23], although such methods cannot by themselves establish why the shift occurred. Organizational diagnosis must therefore accompany statistical diagnosis.

Governance also requires separating decision authority from predictive capability. Evidence that more sophisticated algorithmic predictions do not automatically generate better organizational decisions shows why model-centric evaluation can be insufficient [25]. Authority arrangements influence whether algorithmic recommendations become consequential enough to reshape the environment. An advisory system with weak decision influence may generate little endogenous feedback; a system whose output determines allocation, ranking, pricing, inspection, or access may produce a much stronger intervention in its own future evaluation environment.

These distinctions imply that governance should focus on maintaining the capacity to challenge operational evidence rather than merely monitoring whether model metrics remain within expected ranges. The article therefore proposes four analytically different governance responses: preserving evidence not generated under the dominant policy, monitoring

© 2026 The Author(s). | Published by Institute for Management, Business, and Accounting Studies

changes in case composition and visibility, separating predictive success from decision-system success, and periodically testing whether the organization is still observing meaningful alternatives. Their relevance depends on the strength of optimization–environment coupling and should not be treated as universal requirements.

The configurations developed in the sections on decision-evidence degradation and detection and governance conditions can be organized according to the type of evidence dependence they create and the observation required to distinguish them.

Table 3 presents the analytical distinctions discussed here.

**Table 3.** Distinguishing evidence degradation from ordinary adaptation in AI-optimized decision systems

| System configuration | Primary feedback mechanism | What can remain deceptively stable | Discriminating observation | Governance condition implied by the framework |
|---|---|---|---|---|
| Policy-shaped case exposure | Prior recommendations determine which cases are subsequently encountered | Accuracy or outcome rates within the selected population | Compare exposed cases with independently sampled, protected, or previously unselected cases | Preserve an evidence channel not governed by the dominant allocation policy |
| Metric-attention concentration | Targets redirect effort toward activities represented in the KPI | Target metric and compliance indicators | Examine whether unmeasured or weakly measured outcomes change differently | Maintain outcome measures that are not reducible to the optimization target |
| Behavioral response to evaluation | Actors alter effort or strategy because measurement affects rewards, ranking, or control | Apparent productivity or score distributions | Test performance after changes in incentives, monitoring, or evaluation rules | Treat behavioral response as part of model evaluation rather than as background noise |
| Distributional adaptation with retained external validity | Deployment changes data but the new data continue to improve knowledge of the underlying objective | Both local and external performance | Compare operational learning with independent system-level outcomes | Permit adaptation while retaining external validation |
| Self-referential evidence concentration | Altered evidence is repeatedly reused to judge and update the same optimization regime | Internal validation, calibration, or operational KPI | Evaluate against temporally, institutionally, or operationally less-coupled evidence | Introduce sufficient evidentiary separation to permit falsification of apparent improvement |
| Learning-opportunity narrowing | Repeated optimization reduces exposure to alternative actions or difficult cases | Short-horizon operational consistency | Measure diversity of attempted actions and information gained about alternatives | Preserve bounded exploration or other channels for organizational learning |

Similar apparent AI performance can consequently correspond to different evidentiary states depending on whether post-deployment evidence remains capable of challenging the regime that helped produce it.

© 2026 The Author(s). | Published by Institute for Management, Business, and Accounting Studies

The relationships discussed here are shown in **Figure 2.**

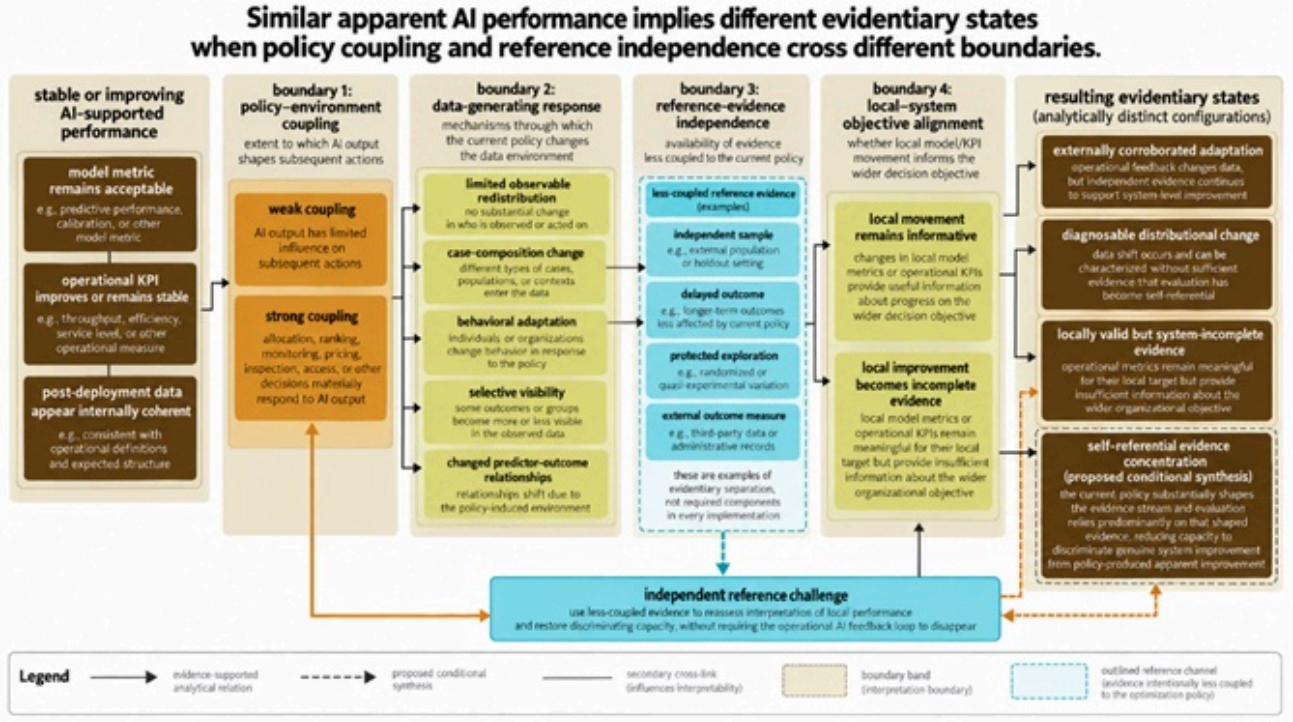


**Figure 2.** Similar apparent AI performance implies different evidentiary states when policy coupling and reference independence cross different boundaries.

# Analytics Implications

The principal implication for analytics is that post-deployment evaluation should treat the data-generating process as potentially endogenous to the analytical intervention. AI-enabled performance systems already combine measurement with prediction and decision support [1]. Once model outputs influence organizational activity, however, conventional validation against later operational data can answer a narrower question than analysts intend. Performative-prediction theory makes this problem explicit because model deployment can alter the future distribution on which the model is judged [21].

Analytics practice should therefore distinguish three evaluative levels. Predictive evaluation asks whether estimates remain accurate for observed cases. Decision evaluation asks whether using those estimates improves organizational choices. System-learning evaluation asks whether repeated use preserves or improves the organization's ability to learn about relevant alternatives over time. Evidence from organizational deployment shows why the first cannot automatically substitute for the second [25], while organizational-learning theory shows why neither necessarily captures the third [26].

Distribution-shift diagnostics can help identify whether post-deployment change reflects altered population composition, changed conditional relationships, or observations that were previously weakly represented [22]. The stronger empirical test of the argument developed here would compare evidence streams exposed differently to the optimization policy. If apparent improvement persists in an independent or less-coupled reference stream, adaptive learning becomes more plausible. If improvement is concentrated in the policy-shaped operational stream while external or protected evidence diverges, the self-referential explanation becomes more credible.

© 2026 The Author(s). | Published by Institute for Management, Business, and Accounting Studies

Future research can consequently test feedback structure rather than merely correlate AI adoption with organizational outcomes. Designs can exploit staggered deployment, protected exploration, independently sampled cases, delayed outcomes, or changes in decision authority to identify whether optimization alters its own evidence environment. Such tests would establish when endogenous measurement is productive adaptation and when it removes precisely the variation required to evaluate whether the decision system is improving.

# Conclusion

AI optimization changes the epistemic conditions of performance evaluation when it materially changes the activities, cases, behaviors, or distributions from which later evaluation evidence is drawn. Under those conditions, post-deployment data cannot automatically be treated as independent observations of system quality.

The resulting problem is conditional. Feedback becomes evidence-degrading when optimization shapes the data-generating process, the altered evidence is repeatedly reused to assess the same regime, and the organization lacks sufficiently independent information with which to challenge apparent improvement. Reinforcing loops can then make operational metrics increasingly coherent while reducing information about alternatives, unobserved cases, or longer-horizon system consequences.

The opposite configuration is also possible. Recursive data generation can support genuine learning when interaction improves coverage, external outcomes continue to corroborate local performance, and balancing mechanisms preserve opportunities for falsification and comparison. The relevant distinction is therefore not between feedback and no feedback, or between AI and human decision-making. It is between feedback that expands valid knowledge about the organizational objective and feedback that progressively concentrates evaluation inside conditions created by the optimizer itself.

For decision analytics, the central requirement is consequently evidentiary separation sufficient to determine whether observed improvement survives outside the environment that produced it. AI-supported performance should be interpreted most confidently when local gains remain visible through evidence channels that the optimization regime does not substantially control.

# Acknowledgements

None

# Conflict of interest

None

# Financial support

None

# Ethics statement

None

Received: 29 Dec 2024 Revised: 05 Mar 2025 Accepted: 21 Mar 2025

© 2026 The Author(s). | Published by Institute for Management, Business, and Accounting Studies

Published online: 18 September 2026

## Rights and permissions

**Open Access** The author(s) retain copyright. This article is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License. It may be shared and adapted for non-commercial purposes with appropriate attribution, an indication of changes, and distribution of adaptations under the same license. Third-party material may be subject to separate terms identified in its credit line. View the license at https://creativecommons.org/licenses/by-nc-sa/4.0/.

# References

1. Sardi A, Modarelli G. Integrating artificial intelligence into performance measurement and management: A systematic literature review. Int J Product Perform Manag. 2026;75(6):1979-2022. https://doi.org/10.1108/IJPPM-05-2025-0362.

2. Modarelli G, Amelio S, Gazzola P, Sardi A. Unlocking human power: How artificial intelligence can transform performance measurement and management systems enhancing organizational decision-making. Int Entrep Manag J. 2026;22(2):100. https://doi.org/10.1007/s11365-026-01234-5.

3. Abbas K. Management accounting and artificial intelligence: A comprehensive literature review and recommendations for future research. Br Account Rev. 2026;58(2):101551. https://doi.org/10.1016/j.bar.2025.101551.

4. Chalmers D, Hunt R, Pachidi S, Potočnik K, Townsend D. The acceleration of artificial intelligence: Rethinking organization and work in an era of rapid technological change. J Manag Stud. 2026;63(2):285-314. https://doi.org/10.1111/joms.70063.

5. Berente N, Gu B, Recker J, Santhanam R. Managing artificial intelligence. MIS Q. 2021;45(3):1433-50.

6. Raisch S, Krakowski S. Artificial intelligence and management: The automation-augmentation paradox. Acad Manage Rev. 2021;46(1):192-210. https://doi.org/10.5465/amr.2018.0072.

7. Anthony C, Bechky BA, Fayard AL. “Collaborating” with AI: Taking a system view to explore the future of work. Organ Sci. 2023;34(5):1672-94. https://doi.org/10.1287/orsc.2022.1651.

8. Glaser VL, Sloan J, Gehman J. Organizations as algorithms: A new metaphor for advancing management theory. J Manag Stud. 2024;61(6):2748-69. https://doi.org/10.1111/joms.13033.

9. van den Broek E. Unpacking AI at work: Data work, knowledge work, and values work. Inf Organ. 2025;35(3):100584. https://doi.org/10.1016/j.infoandorg.2025.100584.

10. Parent-Rocheleau X, Parker SK, Bujold A, Gaudet MC. Creation of the algorithmic management questionnaire: A six-phase scale development process. Hum Resour Manage. 2024;63(1):25-44. https://doi.org/10.1002/hrm.22185.

11. Edwards MR, Zubielevitch E, Okimoto T, Parker SK, Anseel F. Managerial control or feedback provision: How perceptions of algorithmic HR systems shape employee motivation, behavior, and well-being. Hum Resour Manage. 2024;63(4):691-710. https://doi.org/10.1002/hrm.22218.

12. Faraj S, Pachidi S, Sayegh K. Working and organizing in the age of the learning algorithm. Inf Organ. 2018;28(1):62-70. https://doi.org/10.1016/j.infoandorg.2018.02.005.

13. Koehler M, Sauermann H. Algorithmic management in scientific research. Res Policy. 2024;53(4):104985. https://doi.org/10.1016/j.respol.2024.104985.

14. Alizadeh A, Wiener M, Benlian A. A multi-layered perspective on algorithmic control systems: Managerial design choices and worker legitimacy judgments. Eur J Inf Syst. 2026;35(3):404-29. https://doi.org/10.1080/0960085X.2025.2576230.

15. Zhang MM, Cooke FL, Ahlstrom D, McNeil N. The rise of algorithmic management and implications for work and organisations. New Technol Work Employ. 2025;40(3):659-71. https://doi.org/10.1111/ntwe.12343.

16. Kellogg KC, Valentine MA, Christin A. Algorithms at work: The new contested terrain of control. Acad Manage Ann. 2020;14(1):366-410. https://doi.org/10.5465/annals.2018.0174.

17. Hennessy CA, Goodhart CAE. Goodhart's law and machine learning: A structural perspective. Int Econ Rev. 2023;64(3):1075-86. https://doi.org/10.1111/iere.12633.

18. Kadolkar I, Kepes S, Subramony M. Algorithmic management in the gig economy: A systematic review and research integration. J Organ Behav. 2025;46(7):1057-80. https://doi.org/10.1002/job.2831.

19. Yang Y, Chi M, Bi X, Xu Y. Dark sides of algorithmic control in app-based gig work: An objectification perspective. J Bus Res. 2025;195:115407. https://doi.org/10.1016/j.jbusres.2025.115407.

20. Chen RR, Gao J, Chen X, Huang Q. Blessing or curse? The two-sided effects of algorithmic control on ego-depletion and safety performance of gig workers. Comput Human Behav. 2025;162:108461. https://doi.org/10.1016/j.chb.2024.108461.

21. Zhong G, Liu Y, Liu R, Liu J. Performative prediction in the wild: Adapting to arbitrary data distribution maps. Mach Learn. 2026;115(3):42. https://doi.org/10.1007/s10994-025-06980-1.

22. Cai T, Namkoong H, Yadlowsky S. Diagnosing model performance under distribution shift. Oper Res. 2026;74(2):898-916. https://doi.org/10.1287/opre.2023.0217.

23. Cauchois M, Gupta S, Ali A, Duchi JC. Robust validation: Confident predictions even when distributions shift. J Am Stat Assoc. 2024;119(548):3033-44. https://doi.org/10.1080/01621459.2023.2298037.

24. Gregory RW, Henfridsson O, Kaganer E, Kyriakou H. The role of artificial intelligence and data network effects for creating user value. Acad Manage Rev. 2021;46(3):534-51. https://doi.org/10.5465/amr.2019.0178.

25. Kim H, Glaeser EL, Hillis A, Kominers SD, Luca M. Decision authority and the returns to algorithms. Strateg Manage J. 2024;45(4):619-48. https://doi.org/10.1002/smj.3569.

26. Balasubramanian N, Ye Y, Xu M. Substituting human decision-making with machine learning: Implications for organizational learning. Acad Manage Rev. 2022;47(3):448-65.

© 2026 The Author(s). | Published by Institute for Management, Business, and Accounting Studies

https://doi.org/10.5465/amr.2019.0470.

© 2026 The Author(s). | Published by Institute for Management, Business, and Accounting Studies