

Manager Adoption / Defensiveness Lens

Evidence summary, critical implications, and project-useful recommendations for Kashi

Prepared for	Kashi project team
Purpose	Turn the manager-adoption / defensiveness research into decision-ready guidance for product, governance, and pitch design
Scope	External research + Kashi internal docs; focused on manager adoption, surveillance backlash, coaching efficacy, accountability boundaries, and metric-gaming risk

Bottom line. The research does not support the lazy story that nicer wording or “coaching framing” is enough to make monitoring-like systems acceptable. Developmental framing helps at the margin, but the decisive variables are architecture, role boundaries, procedural fairness, escalation rules, and whether managers believe the system is a hidden disciplinary file. For Kashi, that means the Manager Mirror should exist, but only as a bounded lane inside a broader accountability system - not as the product’s primary identity.

Executive summary

- **Manager resistance is structural, not cosmetic.** Research on algorithmic management and monitoring shows that managers worry about accountability, explainability, and worker-health implications, while staff resistance is a material reason for non-adoption in Japan and several European countries. This means manager defensiveness is not just a messaging problem; it is a deployment and governance problem. [1][2]
- **Developmental framing helps less than people want to believe.** A 2025 study found that electronic monitoring increased worker stress and that developmental purpose and participation did not produce the expected moderator effects. A 2022 meta-analysis similarly found lower job satisfaction, higher stress, no performance gain, and slightly higher counterproductive work behavior under monitoring. So 'this is for development' is necessary, but not sufficient. [3][4]
- **Raw feedback does not reliably change manager behavior.** In one upward-feedback field experiment, only about half of leaders improved after feedback. By contrast, future-oriented feedforward interviews outperformed traditional appraisal interviews on later job performance. Kashi therefore needs behavior-change mechanics, not just dashboards. [5][6]
- **Manager receptivity itself is a variable.** Managers who value feedback are perceived as better coaches, and leader feedback-sharing - openly discussing criticism they have received - can improve team psychological safety more durably than simply asking for feedback. Kashi should assume uneven manager readiness and scaffold it. [7][8]

- **Separation between developmental and evaluative use is real, but separation alone is not enough.**

The appraisal literature shows that development and evaluation pull in different directions. Yet one study found that simply separating the two did not improve attitudes and even reduced reported intention to use development in the future. Kashi should separate lanes, but must also define what happens when self-correction fails. [9]

- **Any measured signal can be gamed.** Goodhart's law applies directly here: once managers know the score, some will improve behavior, while others will improve the metric. Kashi should treat measured improvement as a useful but incomplete signal, not as proof that the underlying power asymmetry is gone. [10]

1. What this memo is trying to answer

The immediate question is not whether managers "like" Kashi. The real question is whether Kashi's manager-facing logic can survive the political and psychological reactions that arise when a system makes power-asymmetry patterns visible. This memo therefore asks five narrower questions:

- What does the literature say about manager trust, surveillance backlash, and adoption barriers?
- Does developmental or coaching framing materially reduce those risks?
- What kind of feedback actually changes manager behavior, and how often does it fail?
- How should developmental and accountability functions be separated?
- What do those findings imply for Kashi's current architecture, pitch, and next-step decisions?

Important scope note. The research summarized below does not directly test Kashi itself. Most of it comes from adjacent literatures: algorithmic management, electronic monitoring, performance appraisal, feedback/coaching, and psychological safety. That means the evidence is good enough to constrain design and messaging, but not good enough to justify strong causal claims such as 'Kashi will reduce harassment incidents' or 'Manager Mirror will improve behavior on its own.'

2. What the external research actually says

2.1 Manager resistance is a governance problem, not a tone problem

OECD work on algorithmic management in workplaces is useful here because it shows two things at once. First, managers themselves report trustworthiness concerns with these systems. In a 2025 OECD brief, the most reported concerns were unclear accountability in the case of a wrong decision (28%), inability to follow the logic of algorithmic decisions or recommendations (27%), and inadequate protection of workers' physical and mental health (27%). [2] Second, in the larger 2025 OECD report, staff resistance was the second-most cited reason for non-adoption in France, Germany, Italy, Japan, and Spain, and concern for employees was also a common non-adoption reason. [1]

What this means for Kashi. Interpretation for Kashi: the adoption barrier is not merely that managers dislike criticism. The barrier is that organizations read behavior-analytics systems through the lenses of accountability risk, employee backlash, legal defensibility, and political fallout. A product that feels like covert evidence accumulation will therefore be attacked at the rollout stage, not just grumbled about after deployment.

2.2 Developmental purpose does not neutralize the harms of monitoring

The strongest evidence against naive optimism comes from monitoring research. A 2025 study of 365 German blue-collar workers examined whether developmental purpose, controlling purpose, and participation would moderate the well-being effects of electronic monitoring. It found detrimental effects of monitoring on worker stress and could not find the proposed moderator effects. [3] A 2022 meta-analysis of 70 independent samples and 233 effect sizes found that electronic monitoring slightly decreased job satisfaction ($r = -0.10$), slightly increased stress ($r = .11$), had no relationship with performance ($r = -0.01$), and had a small positive relationship with counterproductive work behavior ($r = .09$). The same paper notes that performance targets and feedback may exacerbate negative effects. [4]

What this means for Kashi. Interpretation for Kashi: 'this is developmental, not punitive' is worth saying, but it is not a magic solvent. If the architecture still feels like surveillance, or if the measured signals appear high-stakes, developmental wording will not erase the deeper reaction. This is especially relevant because Kashi works on meeting interaction data, which sits closer to identity, status, and interpersonal control than ordinary productivity telemetry.

2.3 Feedback alone is a weak and uneven intervention

Manager feedback is not useless, but the literature does not support treating it as a reliable self-correction machine. In Atwater et al.'s upward-feedback field experiment, only 50% of leaders improved after feedback. The same study also found that cynicism about organizational change was associated with less improvement, suggesting that motivation and interpretation matter. [5] A different intervention, the feedforward interview, performed better: in a field test, employees who had a feedforward interview with their manager performed significantly better four months later than employees who received the firm's traditional performance appraisal interview. [6]

What this means for Kashi. Interpretation for Kashi: a manager-facing dashboard is not enough. If Kashi wants actual correction rather than symbolic reflection, it needs future-oriented behavior prompts, repeated review cycles, and a path from detection to coaching action. Otherwise Manager Mirror risks becoming a sophisticated mirror that people glance at once and then ignore.

2.4 Receptivity and leader vulnerability change whether feedback lands

Two adjacent findings matter here. First, Steelman and Wolfeld found that managers with higher feedback orientation - i.e., managers who value feedback for themselves - were perceived as better coaches and created better coaching relationships and feedback environments. [8] Second, research on leader feedback-sharing found that leaders openly sharing criticism they had received improved team psychological safety one year later, whereas simply seeking feedback did not. The qualitative follow-up suggested that feedback-seeking could dissolve into defensiveness and inaction, while feedback-sharing normalized vulnerability and accountability. [7]

What this means for Kashi. Interpretation for Kashi: the system should assume that some managers are coachable and some are not. It should also distinguish between private consumption of feedback and visible commitment to change. If Kashi ever adds manager workflows beyond the dashboard, a strong pattern would be 'share your change commitment' rather than 'please click that you understand.'

2.5 Development and evaluation conflict with each other - but separation alone is not a cure

The performance-appraisal literature has long argued that developmental and evaluative uses pull in different directions. Boswell and Boudreau explicitly studied the effects of separating developmental and evaluative performance-appraisal uses in time and by person. Their results showed no effect on employee attitudes, and employees in the separated group reported they were less likely to use development in the future. [9] This does not mean separation is pointless. It means the relationship is more complicated: separating lanes can avoid immediate contamination, but it does not automatically create trust or developmental uptake.

What this means for Kashi. Interpretation for Kashi: you still need separation, because a single interface cannot credibly say 'this is just for coaching' while also serving as a quiet evidentiary feeder into discipline. But Kashi should not oversell separation as sufficient. The system also needs explicit rules for when self-correction remains private, when it becomes governed remediation, and when it becomes a formal review problem.

2.6 Metric-gaming risk is not optional; it comes with measurement itself

Goodhart's law is the cleanest theoretical frame here: when a measure becomes a target, it ceases to be a good measure. The 2022 CNA report explains that incentives to optimize measured performance can produce manipulation of the measure rather than genuine improvement in the underlying system. [10] Applied to Kashi, this means a manager could learn not to interrupt in visible meetings while shifting the same dominance into agenda control, 1:1 interactions, async channels, or more polished but equally suppressive forms of turn-taking.

What this means for Kashi. Interpretation for Kashi: measured improvement is still useful. It is just not dispositive. The product should never imply that cleaner structural scores equal solved power problems. They only show that one observable slice of behavior improved or changed.

2.7 Evidence-to-decision matrix

The table below compresses the literature into the product decisions it most directly supports.

Research claim	Best evidence	How strong?	Product consequence
Managers and organizations resist	OECD 2025 [1][2]	Strong direction; not	Design and pitch for contestability, bounded

monitoring-like tools for accountability and worker-impact reasons.		Kashi-specific	access, and political survivability.
Developmental framing does not erase stress effects.	Siegel 2025; Siegel et al. 2022 [3][4]	Strong enough to reject naive claims	Do not treat 'coaching' language as sufficient governance.
Feedback alone produces mixed change.	Atwater 2000 [5]	Strong cautionary signal	Mirror needs behavior prompts and escalation logic, not just self-view.
Future-oriented feedback can outperform traditional appraisal.	Budworth et al. 2015 [6]	Promising but context-bounded	Favor feedforward prompts over backward moralizing.
Manager readiness changes coaching quality.	Steelman & Wolfeld 2016 [8]	Useful directional evidence	Assume uneven coachability; scaffold engagement.
Public feedback-sharing can strengthen psychological safety.	Constant et al. 2021 [7]	Useful but not central	Optional future feature: manager change commitments rather than mere acknowledgment.
Separating development and evaluation matters, but separation alone is insufficient.	Boswell & Boudreau 2002 [9]	Nuanced / mixed	Use lane separation, but define crossings and consequences.
Metrics can be gamed.	Goodhart / CNA 2022 [10]	General but highly relevant	Never equate score improvement with harm resolution.

3. Tailoring the research to Kashi

This section translates the evidence into Kashi-specific design judgments by reading the research against the project's current posture in the progress overview and concept note. Kashi already has a stronger foundation than most adjacent products because it is structural rather than affective, deterministic rather than fully inferential, tiered in data retention, and explicit about refusing HR-decision use. [11][12] The

question is not whether Kashi is already better than generic bossware. It is where Kashi is still vulnerable despite that strength.

3.1 What Kashi already gets right

- Kashi's 'mirrors, not microscopes' principle is aligned with the literature's warning against broad visibility and casual browsing of employee data. [11][12]
- The refusal to infer affect or emotion is aligned both with trust concerns and with the legal/regulatory red lines Kashi itself highlights. [11]
- The no-HR-decisions stance is directionally correct. It reduces immediate contamination between developmental and evaluative use. [11][12]
- The refusal to expose named-individual behavioral telemetry to managers is one of Kashi's strongest protective design choices. It directly addresses the retaliation and coercion risk that a manager-facing system could otherwise create. [11]
- The v2 idea of a victim-owned encrypted evidence vault is unusually strong. It reduces the chance that the organization turns the system into a generalized transcript archive while still preserving a path for user-driven escalation. [11]

3.2 Where Kashi is still strategically vulnerable

The 'CEO instrument' framing is commercially sharp but politically risky.

Internally, that line helps anchor buyer economics. Externally, it risks making Kashi sound like top-down oversight technology. The manager may hear 'I am being watched for executive risk management.' The employee may hear 'this ultimately serves executive visibility first.' That does not mean drop the CEO story; it means demote it from product identity to buyer-economics logic. [11]

Manager Mirror is conceptually right but procedurally underdefined.

Kashi says managers see their own behavior as data, not moral judgment, and the concept note says manager feedback should be corrective and reflective, not accusatory. Good. But there is still no sharp answer to what happens after a manager sees the signal, what the protected self-correction window is, and what happens if the pattern persists. [11][12]

The bounded-specificity problem is unresolved.

Kashi correctly refuses named-individual telemetry to managers. But too much abstraction will make feedback useless. The system therefore needs a rule for how specific manager-facing outputs can be while remaining safe: e.g., self-baseline deltas, directional concentration by category, lower-participation-speaker concentrations, or mixed-seniority asymmetry - but not direct subordinate naming by default. [11]

The anti-laundering rule is still implicit rather than explicit.

A company could use Kashi to say 'we gave the manager feedback' and then quietly stop there. If persistent asymmetry after notice does not trigger governed remediation, Kashi can become a managerial-laundering device rather than accountability infrastructure. The current materials imply escalation support, but they need the explicit sentence that coaching is not the endpoint when repeated asymmetry persists. [11][12]

The success narrative can still be misread as adoption-heavy rather than outcome-heavy.

The concept note already includes outcome measures such as reduced concentrated interruption and improved escalation quality. But Kashi should actively subordinate vanity metrics such as dashboard review rates to harder questions: did post-notice behavior change, did escalation quality improve, and did non-improvement become visible instead of being absorbed? [12]

3.3 The most useful synthesis for the project

Question	Research answer	Kashi decision
Will softer wording remove resistance?	No. It may reduce surface reactance, but architecture still dominates perceived legitimacy. [3][4]	Keep non-accusatory language, but prioritize boundaries and procedural design.
Should Kashi keep a manager-facing layer?	Yes. The feedback/coaching literature supports developmental mirrors more than pure punishment. [5][6][8]	Keep Manager Mirror, but treat it as a bounded lane, not the product identity.
Can self-correction stay private forever?	No. Research on feedback failure and mixed developmental/evaluative dynamics argues against endless private coaching. [5][9]	Define protected self-correction window, then governed remediation, then formal review.
Can structural scores prove harm is solved?	No. Goodhart risk means metrics can improve while the underlying pattern mutates. [10]	Treat score movement as evidence, not closure.
Can Kashi claim coaching framing makes it non-surveillant?	No. Monitoring research rejects that simplification. [3][4]	Describe Kashi as governance infrastructure with bounded manager self-correction, not as 'friendly monitoring'.

4. Project-ready recommendations

These are the changes most likely to strengthen Kashi's manager logic without hollowing out the employee-protective thesis.

Priority 1 - Reframe the manager story as a bounded developmental lane, not a core identity.

Keep the manager-facing value proposition, but phrase it as early behavioral correction before patterns harden - not as executive behavior analytics. Move the 'CEO instrument' language into buyer-economics sections rather than the moral center of the product. [11]

Priority 2 - Define a 3-lane model explicitly.

Lane A: private self-correction. Lane B: governed remediation after repeated pattern persistence or post-notice non-improvement. Lane C: formal review with strict entitlement and audit. This is the cleanest way to reconcile coaching with accountability. [5][9][11][12]

Priority 3 - Resolve the specificity rule for Manager Mirror.

Manager-facing outputs should stay structural and self-comparative by default. Examples: 'intrusive interruption concentration toward lower-participation speakers is 2.3x your 90-day baseline' or 'mixed-seniority asymmetry increased over the last 30 days.' Avoid subordinate naming by default. [11]

Priority 4 - Add an explicit anti-laundering rule to the deck and governance page.

State clearly that manager feedback is an early correction mechanism, not proof that the organization has adequately responded. If directional asymmetry persists after protected review, the issue leaves the purely developmental lane. [11][12]

Priority 5 - Build feedforward, not just score display.

Whenever Kashi shows a manager a pattern, pair it with one or two future-oriented behavior commitments. Example: 'In your next two mixed-seniority meetings, wait for sentence completion before redirecting, and ask one clarifying question before stating a decision.' This is more evidence-aligned than backward-looking moral commentary. [6]

Priority 6 - Design success metrics around post-notice change and escalation quality.

Do not let dashboard open rates become the headline. Measure whether signals lead to better remediation, whether recurrence falls after notice, whether affected employees get clearer support options, and whether organizational response happens earlier. [12]

5. Drop-in wording that is safer and more coherent for Kashi

This text is written to be directly reusable in the deck, concept note, or governance page.

Managers are a necessary adoption constituency, but they are not Kashi's primary moral center. Kashi's manager-facing layer exists to support early behavioral correction before harmful patterns become entrenched, not to shield managers from accountability. The Manager Mirror should therefore function as a bounded developmental lane within a broader governance architecture: private self-correction first, governed remediation if patterns persist, and formal review only under explicit threshold and access rules. Manager-facing outputs should remain structural, self-comparative, and non-accusatory, while avoiding named-individual employee telemetry by default. Kashi is not designed to tell managers what they morally are. It is designed to make repeated directional asymmetries visible early enough that they can either be corrected or, if not corrected, no longer remain institutionally deniable.

6. Claims Kashi can and cannot responsibly make

Responsible claims	Claims to avoid
--------------------	-----------------

Kashi surfaces repeated structural interaction asymmetries for human review. [11][12]	Kashi detects harassment, intent, or illegality.
Kashi can support earlier behavioral correction and more structured escalation if the surrounding governance process is sound. [11][12]	Developmental framing makes monitoring backlash disappear. [3][4]
Manager-facing feedback may help some managers change, especially when paired with future-oriented prompts and visible commitments. [5][6][7][8]	Manager Mirror alone will reliably improve behavior.
Kashi reduces one important institutional failure mode: plausible deniability about repeated meeting-level patterns. [11][12]	A cleaner score proves the underlying power problem is solved. [10]

7. Limits of the evidence

- There is no large direct literature on 'meeting-governance tools that privately mirror manager behavior while withholding named-employee telemetry.' Kashi is operating in a newer design space.
- Much of the evidence comes from adjacent systems: electronic monitoring, appraisal systems, coaching, and algorithmic management. These constrain design, but they do not directly measure Kashi-specific effect sizes.
- The evidence supports directional product principles much more strongly than precise business claims. In practice, that means Kashi should be careful with marketing promises and strong about design boundaries.

Sources consulted

- [1] OECD. Algorithmic management in the workplace. 2025. OECD employer survey findings on adoption, monitoring tools, and trustworthiness concerns.
- [2] OECD. How widespread is algorithmic management in workplaces? 2025. Short OECD brief reporting specific manager trustworthiness concerns including accountability, explainability, and health.
- [3] Siegel, R. An expanded view on electronic monitoring: Testing moderating effects of participation and perceived purposes of electronic monitoring. 2025.
- [4] Siegel, R., Konig, C. J., & Lazar, V. The impact of electronic monitoring on employees' job satisfaction, stress, performance, and counterproductive work behavior: A meta-analysis. *Computers in Human Behavior Reports*, 8, 100227. 2022.
- [5] Atwater, L. E., Waldman, D. A., Atwater, D., & Cartier, P. An upward feedback field experiment: supervisors' cynicism, reactions, and commitment to subordinates. *Personnel Psychology*. 2000.
- [6] Budworth, M.-H., Latham, G. P., & Manroop, L. Looking Forward to Performance Improvement: A Field Test of the Feedforward Interview for Performance Management. *Human Resource Management*. 2015.
- [7] Constant, C. S., Lee, N. J., He, Y., & Grossman, A. Taking Your Team Behind the Curtain: The Effects of Leader Feedback-Sharing and Feedback-Seeking on Team Psychological Safety. *Organization Science*. 2021.
- [8] Steelman, L. A., & Wolfeld, L. The Manager as Coach: The Role of Feedback Orientation. *Journal of Business and Psychology*. 2016.
- [9] Boswell, W. R., & Boudreau, J. W. Separating the Developmental and Evaluative Performance Appraisal Uses. *Journal of Business and Psychology*, 16, 391-412. 2002.

- [10] CNA. Goodhart's Law: Recognizing and Mitigating Manipulation Measures in Analysis. 2022.
- [11] Kashi - Progress & Project Overview (2026-04-21). Internal project document.
- [12] meeting_governance_ai_concept_note.docx. Internal project document.