Evaluating the Effect of AI-Assisted Tools on Engineering Team Productivity and Decision Quality

Harsimran Kaur

Master of Science in Engineering Management

Abstract: ***The integration of artificial intelligence (AI) tools into engineering teams has rapidly shifted from an experimental curiosity to a mainstream operational reality. As organizations adopt AI-assisted platforms for coding, scheduling, risk assessment, and project decision-making, a critical question emerges: do these tools genuinely enhance team productivity and the quality of engineering decisions, or do they introduce new forms of dependency and risk? This paper conducts a conceptual evaluation of the effects of AI-assisted tools on engineering team performance, drawing on recent empirical studies and organizational behavior frameworks. It proposes a Human-AI Collaboration (HAC) model that distinguishes between individual-level output gains and team-level decision quality outcomes, argues that the net effect of AI adoption is highly context-dependent, and offers a set of managerial recommendations for engineering leaders seeking to deploy AI tools responsibly and effectively.***

Keywords: Human-AI Collaboration, Engineering Management, AI Tools, Team Productivity, Decision Quality, Industry 4.0, Organizational Behavior

1. Introduction

Engineering teams today operate in environments of increasing complexity. Projects span multiple disciplines, geographies, and technology stacks, while timelines grow shorter and stakeholder expectations climb higher. In response, organizations across sectors- from software development to civil infrastructure to manufacturing systems- have turned to AI-assisted tools to help engineers work faster, smarter, and with greater accuracy.

The appeal is intuitive: if AI can handle routine pattern recognition, data aggregation, scheduling optimization, and document synthesis, human engineers are freed to focus on higher-order judgment, creative problem framing, and collaborative decision-making. Yet this optimistic narrative is increasingly complicated by empirical evidence showing that AI adoption does not uniformly improve team outcomes. Benefits tend to be uneven across experience levels, disciplines, and organizational cultures. In some cases, productivity gains at the individual level have failed to translate into measurable improvements at the team or project level- a phenomenon sometimes termed the AI Productivity Paradox.

This paper aims to examine both the promise and the complexity of human-AI collaboration in engineering teams. It reviews current evidence, proposes a conceptual framework for understanding the mechanisms at play, and offers practical guidance for engineering managers navigating this transition.

2. Literature Review

A growing body of research confirms that AI tools can meaningfully accelerate individual engineering task completion. A multi-company randomized controlled trial spanning nearly 5,000 developers found that access to AI code assistance correlated with approximately a 26 percent increase in individual output, with participants completing more tasks per unit of time without a measurable decline in code quality. A separate longitudinal study tracking 300 engineers over twelve months found that adoption of AI-assisted development environments was initially slow—around 4 percent active usage in the first month- but accelerated to over 80 percent engagement by month six before stabilizing near 60 percent.

However, team-level outcomes tell a more nuanced story. The 2025 DevOps Research and Assessment (DORA) report, drawing on nearly 5,000 technology professionals and more than 100 hours of qualitative interviews, concluded that AI does not automatically improve software delivery performance. The report's central finding is that the success of AI in engineering contexts depends less on tool sophistication and more on the organizational systems surrounding deployment- including engineering culture, platform capability, and internal knowledge-sharing structures.

From an engineering management perspective, the distinction between individual-level productivity and team-level decision quality is critical. Prior research on group decision-making suggests that team decisions are influenced not only by the information available but by the processes through which that information is surfaced, discussed, and evaluated. AI tools that accelerate information retrieval and synthesis may paradoxically reduce the deliberative richness of team discussions if members defer to algorithmic recommendations without critical interrogation—a risk scholars have termed automation bias.

Literature on human-computer interaction further highlights the role of explainability. When AI systems present recommendations without interpretable rationale, engineers with less domain experience may accept outputs uncritically, while experienced engineers may override valid recommendations based on intuition. Neither extreme optimizes decision quality, underscoring the importance of Explainable AI (XAI) design principles in team deployment contexts.

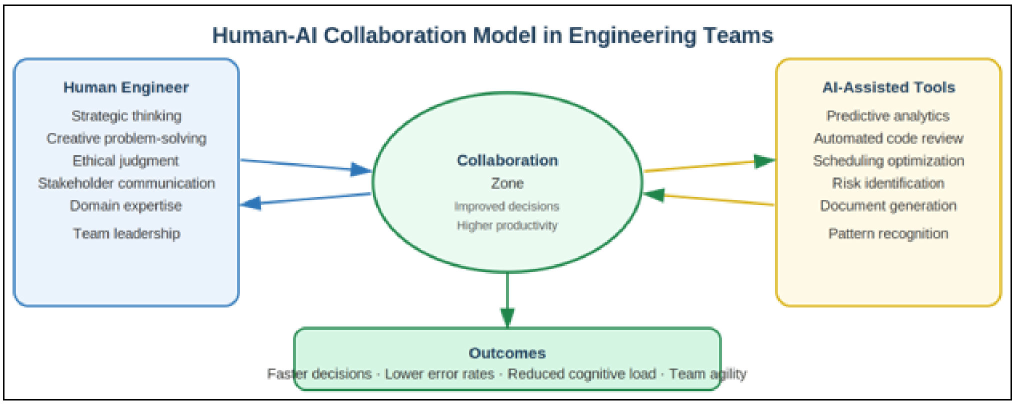

Figure 1: Human-AI Collaboration Model: the interaction between human engineers and AI tools, and the outcomes this collaboration produces

3. A Framework for Human-AI Collaboration in Engineering Teams

To systematically evaluate the effect of AI-assisted tools on team performance, this paper proposes the Human-AI Collaboration (HAC) Framework, structured around three dimensions: (1) capability complementarity, (2) trust calibration, and (3) organizational readiness.

Capability Complementarity refers to the degree to which AI tools address genuine skill gaps or cognitive bottlenecks in the team rather than simply automating tasks engineers already perform well. Teams achieve the greatest gains when AI handles high-volume, pattern-based work- such as anomaly detection in sensor data, automated scheduling, or standards compliance checking- while human engineers retain ownership of contextual judgment, stakeholder negotiation, and novel problem formulation. Misalignment between AI capabilities and team needs is a leading cause of underperformance and disengagement.
Trust Calibration addresses the challenge of ensuring that team members neither over-rely on AI outputs nor reflexively dismiss them. Optimal trust is calibrated- meaning engineers understand when to defer to AI recommendations and when to override them. Engineering managers can foster calibrated trust through structured workflows that require engineers to document their rationale when overriding AI suggestions, creating accountability and enabling organizational learning about edge cases where human judgment consistently outperforms algorithmic prediction.

Organizational Readiness encompasses the data infrastructure, training programs, leadership alignment, and cultural norms necessary to support effective human-AI collaboration. Organizations that score high on readiness- with clean, well-governed data pipelines, cross-functional AI literacy training, and psychological safety for experimentation- consistently achieve stronger and more equitable productivity outcomes across team members of varying experience levels.

4. Effects on Engineering Decision Quality

While productivity metrics such as tasks completed per sprint or time-to-resolution are relatively tractable to measure, decision quality is inherently harder to operationalize. In engineering management, a high-quality decision is one that optimally balances technical feasibility, resource constraints, stakeholder alignment, and long-term maintainability- criteria that resist simple quantification.

Available evidence suggests that AI tools improve certain dimensions of decision quality while introducing risks in others. On the positive side, AI-driven risk identification systems have been shown to surface potential failure modes earlier in project cycles, reducing costly late-stage corrections. Automated scheduling tools that draw on historical project data can generate more accurate timeline estimates than manual planning, reducing schedule overruns. AI-assisted documentation generation has been associated with improved clarity and consistency in engineering specifications.

However, decision quality risks are real and must be actively managed. First, when team members attribute authority to AI outputs by virtue of their algorithmic origin rather than their substantive merit, groupthink can emerge at the human-AI interface rather than within the human team alone. Second, AI models trained on historical data may perpetuate outdated design assumptions or risk tolerances that no longer reflect current project contexts. Third, the accelerated pace that AI enables can compress the deliberation time that complex engineering decisions genuinely require, substituting speed for thoroughness.

Paper ID: SR26901083036 DOI: https://dx.doi.org/10.21275/SR26901083036

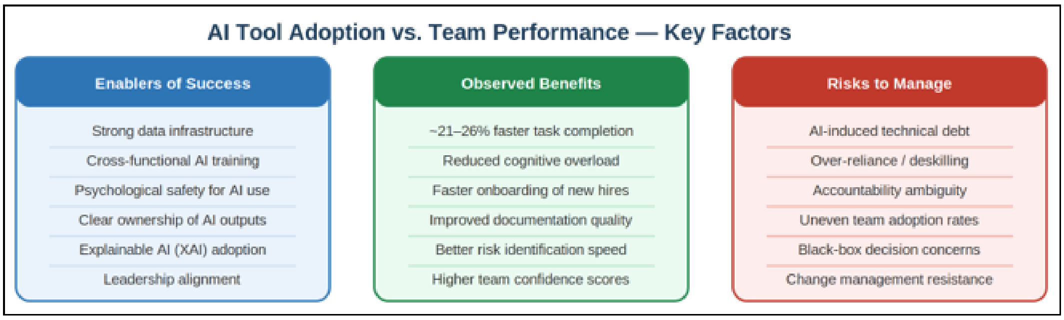

Figure 2: Key enablers, observed benefits, and risks associated with AI tool adoption in engineering teams

5. Managerial Recommendations

Based on the HAC framework and the evidence reviewed, this paper offers the following recommendations for engineering managers overseeing AI tool adoption:

First, audit capability gaps before selecting AI tools. Productivity gains are maximized when AI targets genuine bottlenecks. A pre-adoption assessment of where engineers spend the most time on low-judgment, high-volume tasks will identify the highest-value deployment opportunities.

Second, invest in XAI and structured review practices. Requiring teams to engage critically with AI recommendations- through structured review protocols and decision logs—builds trust calibration over time and creates institutional knowledge about where AI performs reliably versus where human judgment should dominate.

Third, measure team-level outcomes, not just individual output. Organizations should complement individual productivity metrics with measures of decision quality, rework rates, and team satisfaction to ensure that individual gains are translating into project-level value.

Fourth, phase adoption to build literacy progressively. The longitudinal data on developer adoption rates suggests that gradual rollout with embedded training outperforms forced adoption mandates in sustaining long-term engagement and performance improvement.

6. Conclusion

The integration of AI-assisted tools into engineering teams represents one of the most consequential organizational challenges of the current decade. The evidence is clear that these tools can meaningfully accelerate individual task completion and improve certain quality dimensions when deployed thoughtfully. However, productivity at the individual level does not automatically translate into better team decisions or stronger project outcomes. Engineering managers who treat AI adoption as a purely technical implementation challenge- ignoring the cultural, organizational, and human factors that govern how teams actually use these tools- are unlikely to realize the full potential of their investments.

This paper has proposed the Human-AI Collaboration framework as a structured lens through which engineering leaders can evaluate and guide AI adoption in their teams. Future empirical research should develop validated instruments for measuring trust calibration and decision quality in human-AI team contexts, and longitudinal case studies across engineering disciplines would substantially deepen our understanding of how these dynamics evolve over time.

References

[1] DORA Research Program. (2025). State of AI-Assisted Software Development. Google Cloud DevOps Research and Assessment.
[2] Faros.ai. (2025). The AI Productivity Paradox Report. Faros AI Inc.
[3] Parasuraman, R., & Riley, V. (1997). Humans and automation: Use, misuse, disuse, abuse. Human Factors, 39(2), 230–253.
[4] Dafoe, A. (2018). AI Governance: A Research Agenda. Future of Humanity Institute, University of Oxford.
[5] GitHub & Accenture. (2024). The Impact of AI Coding Tools on Developer Productivity: A Multi-Company RCT. SSRN Working Paper.
[6] DX Engineering Intelligence. (2025). AI-Assisted Engineering: Q4 Impact Report. DX Inc.
[7] Gigerenzer, G., & Gaissmaier, W. (2011). Heuristic decision making. Annual Review of Psychology, 62, 451–482.
[8] Brynjolfsson, E., Li, D., & Raymond, L. R. (2024). Generative AI at Work. NBER Working Paper No. 31161.