

# Beyond the Playbook

## The Human and Organizational Work AI Cannot Do

Aurorae Group, LLC • AI Resource Library • 2025

### Overview

---

A substantial body of research now documents that generative AI tools produce measurable improvements in productivity, output quality, and worker experience under controlled conditions. The results are compelling: professional writing tasks completed 40% faster. Consulting deliverables rated more than 40% higher in quality. Customer support issues resolved 14% more efficiently. For organizations navigating the question of whether and how to adopt AI, this evidence is encouraging.

But a second, equally important body of research tells a different story. When researchers shift from controlled experiments to real-world organizational settings, the gains documented in the lab largely disappear. AI chatbots deployed across thousands of actual workplaces produce average time savings of less than 3% of work hours. Productivity tools rolled out to large knowledge worker populations change which tasks people do, but not how many hours they work or what they earn. The technology is the same. The workers are real. The gap between what AI can do and what organizations actually gain from it is striking.

This brief examines that gap, what research reveals about why it exists, and what it means for leaders making decisions about AI adoption.

### What the Research Shows

---

Two recent studies are particularly instructive because they study AI in real organizational settings, not controlled experiments.

**Shifting Work Patterns with Generative AI** (Dillon, Jaffe, Peng & Cambon, Microsoft Research / NBER, 2025) followed knowledge workers across a large organization over 13 months as Microsoft Copilot was rolled out. This was not a simulated task in a research setting. It was the actual, messy, longitudinal experience of a major AI tool entering a real workforce.

The findings were nuanced and important. Workers did change how they spent their time: some tasks shifted, certain activities became easier, and communication patterns evolved. But the headline finding is that these changes did not translate into the broader organizational transformation that productivity research might predict. The study identifies a consistent pattern: broader shifts in how organizations work require more than tool access. They require time, institutional investment, and deliberate redesign of workflows, roles, and expectations.

**Large Language Models, Small Labor Market Effects** (Humlum & Vestergaard, University of Chicago / University of Copenhagen, 2025) is the most rigorous real-world counterpoint to AI productivity optimism available. The researchers linked large-scale adoption surveys to actual administrative labor records across 25,000 workers in 7,000 workplaces in Denmark, covering 11 occupations including HR professionals, legal professionals, teachers, accountants, marketing professionals, and software developers.

Their finding is unambiguous: AI chatbots produced no significant impact on earnings or recorded hours in any occupation studied. Not among intensive users. Not among early adopters. Not in workplaces that had made substantial investments in the technology. Workers reported using AI

tools and, in surveys, described productivity gains. But those gains did not materialize in any measurable change in output, hours, or compensation as recorded in actual employment data.

Average time savings reported by workers: 2.8% of work hours. This is not nothing, but it is far below the 15% to 50% productivity gains documented in controlled studies of the same occupations.

*The gap between AI's demonstrated potential in research settings and its actual impact in organizations is not a measurement error. It is a signal about what is missing from most AI implementation strategies.*

## Why the Gap Exists

---

The researchers in both studies are careful not to conclude that AI is ineffective. The controlled experiments documenting significant productivity gains are rigorous and replicable. The gap between those findings and real-world outcomes points to something specific: the conditions that make AI effective in a controlled setting are rarely present in actual organizations.

Several factors consistently emerge across the research:

- **Task clarity and workflow fit.** Workers in controlled experiments are assigned clear, bounded tasks and given structured access to AI tools. Workers in real organizations must first determine which of their many responsibilities benefit from AI, learn to use the tools effectively for each one, and integrate that use into workflows that were not designed with AI in mind. This takes time that most implementation plans do not account for.
- **Organizational friction.** Controlled experiments eliminate organizational friction. Real workplaces do not. Institutional norms, approval processes, tool access policies, and the informal social dynamics of teams all shape whether workers actually use AI tools and how. The Microsoft Copilot study found that broader shifts in how organizations work require coordinated institutional effort, not just individual tool adoption.
- **Motivation and adoption depth.** In controlled experiments, workers are motivated to use the tool and perform well on the assigned task. In real workplaces, motivation is more complex. Workers who are uncertain about AI's implications for their roles may use tools cautiously or selectively. The Danish study found that even in workplaces with substantial employer investment, impacts on recorded hours remained negligible.
- **The difference between task productivity and organizational productivity.** The productivity gains in controlled experiments accrue to individual task performance. Organizational productivity depends on more than individual tasks performed more quickly. It depends on how work is coordinated, how decisions are made, how knowledge moves through a team, and how the organization learns. AI tools affect individual task performance. They do not automatically improve the systems those tasks are embedded in.

---

## What This Means for Leaders

---

The research does not argue against AI adoption. It argues for a more honest account of what AI adoption actually requires. Several implications are consistent across the evidence.

- **Implementation readiness precedes tool selection.** The question is not whether AI tools work. The controlled evidence is clear that they do, under the right conditions. The question is whether your organization has created those conditions. That means examining workflows before deploying tools, not after. It means identifying where task boundaries are clear enough for AI to be reliably helpful, and where they are not.
- **AI changes what people do, not whether they work.** Workers across both studies changed what tasks they performed when given access to AI. In some cases, lower-value tasks were reduced. In others, AI created new tasks, such as evaluating AI outputs, that replaced rather than eliminated the time previously spent. The work is not disappearing. It is shifting. Leaders need to understand what work is shifting to, not just what is being automated.
- **Individual efficiency gains do not automatically become organizational gains.** The Danish study found that despite widespread AI adoption and self-reported productivity gains, no measurable changes appeared in hours worked or compensation. This gap between workers' experience of AI and its organizational impact suggests that the gains from individual use are being absorbed into existing workloads rather than translating into organizational capacity. Realizing organizational value from AI requires deliberate choices about how freed capacity is redirected.
- **The equity dimension is a strategic dimension.** The research consistently documents that newer and lower-skilled workers benefit most from AI assistance. This is one of the most consistent and meaningful findings across the evidence. Organizations that deploy AI with an equity lens, specifically designing for the workers with the most to gain rather than the workers already performing at the highest level, are likely to see more durable organizational impact.
- **The investment is human before it is technological.** Both studies point to the same conclusion: realizing the organizational potential of AI requires investment in the human and structural conditions that make tools effective. That means time for workers to develop genuine fluency, not just tool access. It means redesigning workflows, not just adding a new step. It means leadership that treats AI adoption as an organizational change initiative, not a software rollout.

*The playbooks, prompt libraries, and implementation guides available to leaders today document what AI can do. This research documents what organizations must do for AI to deliver on that potential.*

---

## A Note on These Sources

---

This brief draws on two peer-reviewed or working-paper-stage studies from independent academic and research institutions with no commercial interest in AI adoption outcomes. Both studies are freely available and linked below. They represent a small but growing body of research examining AI in real organizational settings rather than controlled experiments, and they are among the most rigorous sources currently available for leaders making practical decisions about AI.

Readers are encouraged to examine the primary sources directly. The findings are more nuanced than any summary can capture, and the methodological choices in each study shape what can and cannot be concluded from them.

---

## Source Studies

---

- **Shifting Work Patterns with Generative AI:** Dillon, Jaffe, Peng & Cambon. Microsoft Research / NBER, 2025. [https://www.nber.org/system/files/working\\_papers/w33795/w33795.pdf](https://www.nber.org/system/files/working_papers/w33795/w33795.pdf)
- **Large Language Models, Small Labor Market Effects:** Humlum & Vestergaard. University of Chicago / University of Copenhagen, 2025. [https://bfi.uchicago.edu/wp-content/uploads/2025/04/BFI\\_WP\\_2025-56-2.pdf](https://bfi.uchicago.edu/wp-content/uploads/2025/04/BFI_WP_2025-56-2.pdf)

*This brief was prepared by Aurorae Group, LLC as part of the Aurorae Group AI Resource Library. It is intended for leadership audiences navigating AI adoption decisions and does not constitute consulting or implementation advice. © 2025 Aurorae Group, LLC.*