Int. j. adv. multidisc. res. stud. 2026; 6(3):1266-1271

Received: 21-04-2026
Accepted: 01-06-2026

International Journal of Advanced Multidisciplinary Research and Studies

ISSN: 2583-049X

The Dark Side of AI: When Automation Reduces Employee Engagement and Output

Vo Phuoc Tai
Ho Chi Minh City University of Economics and Finance (UEF), Vietnam

DOI: https://doi.org/10.62225/2583049X.2026.6.3.6417
Corresponding Author: **Vo Phuoc Tai**

Abstract
Artificial intelligence (AI) and automation are increasingly adopted by organizations to improve efficiency, reduce operational costs, accelerate decision-making, and enhance productivity. However, a growing body of research suggests that AI does not always generate positive outcomes for employees or organizations. This study examines the dark side of AI by exploring the conditions under which automation may reduce employee engagement and output. Using a desk-based qualitative research design, the paper synthesizes prior studies on AI adoption, technostress, job insecurity, emotional exhaustion, psychological capital, perceived organizational support, work engagement, and individual productivity. Drawing on Conservation of Resources Theory, Job Demands–Resources Theory, Self-Determination Theory, and socio-technical systems theory, the study proposes that AI automation can become harmful when it is perceived as a threat to job security, autonomy, competence, professional identity, and psychological resources. The thematic analysis indicates that AI may reduce employee engagement and sustainable output through several mechanisms: increased technostress, emotional exhaustion, job insecurity, role ambiguity, work intensification, reduced autonomy, over-reliance on algorithmic recommendations, and depletion of psychological capital. Evidence from recent studies further suggests that organizational support, employee involvement, digital training, psychological capital, and transparent AI implementation can mitigate these negative effects. The findings contribute to the AI and work literature by shifting the discussion from a purely productivity-oriented view of automation to a more balanced understanding of its psychological and organizational consequences. Practically, the study suggests that organizations should not evaluate AI adoption only through cost reduction, speed, or output volume, but also through employee engagement, well-being, autonomy, learning, and long-term work quality. AI creates sustainable value only when it is implemented as a human-augmentation tool rather than merely as a labor-substitution mechanism.

Keywords: Artificial Intelligence, Automation, Technostress, Employee Engagement, Employee Output, Job Insecurity, Emotional Exhaustion, Psychological Capital, Perceived Organizational Support, Human-AI Collaboration

1. Introduction: Background and Problem Statement

Artificial intelligence (AI) has become one of the most influential technologies shaping contemporary organizations. In service, hospitality, finance, healthcare, manufacturing, logistics, education, and professional work, AI is increasingly used to automate routine tasks, support decision-making, predict customer behavior, optimize workflows, monitor performance, and generate text, images, code, and recommendations. Much of the dominant discourse presents AI as a productivity-enhancing technology. Empirical studies indeed show that AI can increase task efficiency in some contexts. For example, Brynjolfsson, Li, and Raymond (2025) [1] found that generative AI assistance increased customer-support agents' productivity by approximately 15%, with larger benefits for less experienced workers. However, the same literature also suggests that AI's effects are uneven, conditional, and potentially harmful when automation increases stress, weakens autonomy, creates job insecurity, reduces psychological safety, or pushes employees beyond sustainable performance thresholds.

This study focuses on the "dark side" of AI: situations in which automation reduces employee engagement and output rather than improving them. This problem is increasingly important because organizations often implement AI with the expectation that it will reduce costs and increase efficiency, while paying insufficient attention to employees' psychological, social, and work-design conditions. Earlier technostress research showed that information and communication technologies can create overload, invasion, complexity, insecurity, and uncertainty, which may reduce productivity and job satisfaction (Tarafdar *et*

al., 2007; Ragu-Nathan *et al*., 2008) [12, 9]. More recent AI-focused studies suggest similar but intensified mechanisms: AI awareness may increase emotional exhaustion and depression (Xu *et al*., 2023) [14], AI-induced stress may reduce work engagement through depleted psychological capital (Hou & Fan, 2024) [7], and AI use may influence employee outcomes through job insecurity and technostress (Sharif *et al*., 2025) [11].

The research problem is therefore not whether AI is generally good or bad. Rather, the key question is: **Under what conditions does AI automation reduce employee engagement and output?** This question matters because engagement and output are central to organizational sustainability. Work engagement reflects vigor, dedication, and absorption in work. Output reflects the quantity and quality of work produced. When AI is implemented poorly, it can generate "productivity loss through productivity pressure": employees may be expected to produce more with fewer resources, monitor more systems, learn new tools rapidly, respond faster, and remain accountable for AI-generated errors. In such cases, automation does not remove work; it redistributes work in ways that may increase cognitive, emotional, and ethical burden.

The purpose of this desk-based qualitative study is to synthesize previous research on AI, automation, technostress, job insecurity, psychological resource depletion, engagement, and productivity. The study aims to explain why AI can reduce engagement and output, identify the main mechanisms through which this occurs, and propose a conceptual interpretation useful for future empirical research and managerial practice.

2. Literature Review

The literature on AI and work can be divided into two broad streams. The first emphasizes AI's positive role in productivity, efficiency, and augmentation. The second emphasizes AI's negative effects on stress, insecurity, work intensification, autonomy, and well-being. This study integrates both streams but focuses on the second.

Early evidence on generative AI shows that AI can improve task performance when the technology is well-matched to the work. Noy and Zhang (2023) [8] found productivity gains in professional writing tasks, while Brynjolfsson *et al*. (2025) [1] found that AI support improved customer-service productivity and accelerated learning among novice employees. However, these findings should not be interpreted as universal evidence that AI always improves work outcomes. Brynjolfsson *et al*. (2025) [1] also found that AI had little effect on experienced high-skill workers and may slightly reduce conversation quality among the most skilled agents when they over-follow AI recommendations.

The "jagged technological frontier" literature further complicates the productivity narrative. Dell'Acqua *et al*. (2023) [3] argue that AI performs very well on some tasks but poorly on others that may appear superficially similar. Their experimental work with consultants shows that AI can raise productivity and quality for tasks within its capability frontier, but it can reduce performance when users rely on AI outside that frontier. This is central to the present study because reduced employee output may occur not because employees resist AI, but because they over-trust AI in ambiguous, complex, context-sensitive, or high-judgment tasks.

Technostress research provides a major foundation for understanding AI's dark side. Tarafdar *et al*. (2007) [12] demonstrated that technostress is negatively related to individual productivity and positively related to role stress. Their work shows that technology can reduce output when it generates role overload, role conflict, or ambiguity. Ragu-Nathan *et al*. (2008) [9] further developed the distinction between technostress creators and technostress inhibitors. Technostress creators reduce job satisfaction and organizational commitment, while organizational mechanisms such as technical support, training, and involvement reduce the negative effects of technology stress. Salanova *et al*. (2012/2013) added a psychological perspective by identifying technostrain and technoaddiction as two distinct experiences of technology-related stress. Technostrain involves anxiety, fatigue, skepticism, and inefficacy; technoaddiction involves compulsive overuse of technology.

Recent AI-specific studies show that these technology-stress mechanisms are becoming more complex. Xu *et al*. (2023) [14] examined AI awareness and employee depression using Conservation of Resources theory. Their findings show that AI awareness is positively associated with depression, emotional exhaustion mediates this relationship, and perceived organizational support weakens the harmful indirect effect. Hou and Fan (2024) [7] found that AI-induced stress negatively affects hotel employees' work engagement, psychological capital mediates the relationship, and perceived organizational support reduces the negative effect of AI stress on psychological capital. Sharif *et al*. (2025) [11] found that AI use in hospitality organizations affects employee well-being and career outcomes through job insecurity, while technostress acts as a boundary condition.

The labor-market literature also shows that AI exposure is not evenly distributed. The International Labour Organization (ILO) argues that generative AI is more likely to augment many occupations than fully automate them, but clerical occupations face particularly high exposure. The ILO's 2025 refined index further reports that clerical occupations remain the most exposed group, while some digitized professional and technical occupations show increasing exposure due to expanding generative AI capabilities. The World Economic Forum (2025) [13] similarly emphasizes that technology, AI, demographic shifts, economic uncertainty, and green transition trends will reshape jobs and skills through 2030.

The emerging literature therefore suggests a paradox: AI can increase productivity at the task level while reducing engagement or sustainable output at the worker level. Chuang, Chiang, and Lin (2025) [2], using a Job Demands–Resources perspective, found that generative AI and AI efficacy may increase productivity, but AI technostress induces exhaustion, work–family conflict, and lower job satisfaction. Högemann (2025) [6] similarly notes that generative AI is widely expected to improve productivity, but workplace adoption may introduce new forms of technostress, especially among young professionals. These studies support the argument that AI's dark side emerges when automation increases demand faster than it increases resources.

3. Theoretical Foundation

This study integrates four theoretical perspectives: Conservation of Resources theory, Job Demands–Resources theory, Self-Determination Theory, and socio-technical

systems theory.

Conservation of Resources theory explains why AI can reduce engagement. According to COR theory, individuals seek to obtain, retain, and protect valued resources. These resources include time, energy, skills, status, job security, confidence, psychological capital, and social support. AI threatens these resources when employees perceive that their skills are becoming obsolete, their jobs are replaceable, their professional identity is weakened, or their work is being intensified. Xu *et al*. (2023) [14] show that AI awareness can deplete emotional resources, leading to emotional exhaustion and depression. Hou and Fan (2024) [7] show that AI-induced stress reduces psychological capital, thereby reducing work engagement.

Job Demands–Resources theory explains the work-design mechanism. AI can function as a resource when it reduces repetitive work, supports decision-making, and improves task efficiency. However, it becomes demand when employees must learn new systems, monitor AI outputs, correct errors, handle customers' dissatisfaction with automated systems, or meet higher performance targets because AI is assumed to make work faster. Chuang *et al*. (2025) [2] show this duality: AI can increase productivity, but AI technostress can increase exhaustion and reduce job satisfaction.

Self-Determination Theory helps explain why automation may reduce engagement even when it improves efficiency. SDT argues that motivation depends on autonomy, competence, and relatedness. AI may undermine autonomy when employees feel controlled by algorithmic systems, scripts, dashboards, surveillance tools, or automated performance evaluations. It may undermine competence when employees feel that AI performs core tasks better than they do or when they are reduced to monitoring machine outputs. It may undermine relatedness when work becomes more mediated by systems and less grounded in human interaction. Sharif *et al*. (2025) [11] use SDT to show how AI use may influence well-being and career outcomes through job insecurity and technostress.

Socio-technical systems theory explains why AI implementation cannot be treated as a purely technical intervention. AI changes not only tools but also roles, routines, authority, accountability, incentives, and social relationships. Tarafdar *et al*. (2007) [12] show that ICT-related stress arises from both technical changes in tasks and social changes in roles, reward systems, and authority structures. Thus, automation reduces engagement and output when the technical system is optimized at the expense of the social system.

Together, these theories suggest the following conceptual logic: **AI automation → increased work demands/resource threat → technostress/job insecurity/emotional exhaustion → reduced engagement → reduced sustainable output**. Organizational support, psychological capital, AI self-efficacy, participatory implementation, and job redesign may weaken this negative pathway.

4. Methodology: Desk-Based Qualitative Research Design

This paper applies a desk-based qualitative research design. It synthesizes secondary academic and institutional sources rather than collecting primary data. This approach is appropriate because the research question requires integration across multiple research streams: AI adoption, technostress, employee engagement, productivity, job insecurity, automation, and organizational psychology.

The review prioritized peer-reviewed journal articles, working papers, and institutional reports published mainly between 2021 and 2026. However, several foundational studies published before 2021 were included because they established core concepts necessary for interpreting recent AI research, especially technostress creators, technostress inhibitors, role stress, and technostrain. Key foundational sources include Tarafdar *et al*. (2007) [12], Ragu-Nathan *et al*. (2008) [9], and Salanova *et al*. (2012/2013). Key recent sources include Xu *et al*. (2023) [14], Hou and Fan (2024) [7], Brynjolfsson *et al*. (2025) [1], Sharif *et al*. (2025) [11], Chuang *et al*. (2025) [2], Högemann (2025) [6], and ILO and WEF reports from 2023–2025.

The analysis followed a thematic synthesis procedure. First, studies were screened for relevance to AI, automation, technostress, job insecurity, engagement, output, productivity, or well-being. Second, findings were coded into conceptual categories: AI-induced stress, automation anxiety, job insecurity, emotional exhaustion, work engagement, reduced autonomy, skill obsolescence, algorithmic control, and productivity loss. Third, categories were grouped into broader themes. Fourth, the themes were interpreted using COR theory, JD-R theory, SDT, and socio-technical systems theory.

The purpose of this method is not to estimate a statistical effect size. Instead, the goal is to build a theoretically grounded explanation of when and why AI automation may reduce employee engagement and output.

5. Thematic Analysis: Synthesis of Previous Studies

Theme 1: AI creates loss of engagement when it is perceived as a threat to job security

A central theme across recent studies is that AI can reduce engagement when employees perceive automation as a threat to job continuity or career development. Job insecurity is not merely an economic concern; it is a psychological stressor that affects motivation, identity, and future orientation. Xu *et al*. (2023) [14] show that AI awareness can increase depression through emotional exhaustion, suggesting that perceived AI threat depletes emotional resources. Sharif *et al*. (2025) [11] similarly show that job insecurity mediates the link between AI use and employee outcomes in hospitality.

When employees fear replacement, they may reduce discretionary effort, avoid experimentation, hide knowledge, or psychologically withdraw from work. Engagement declines because the employment relationship becomes less secure. Instead of seeing AI as a tool for support, employees may view it as a signal that the organization is preparing to reduce labor dependence.

Theme 2: AI reduces output when technostress exceeds technological support

Technostress is one of the most consistent explanations for reduced employee output. Tarafdar *et al*. (2007) [12] found that technostress is inversely related to individual productivity. Ragu-Nathan *et al*. (2008) [9] found that technostress creators reduce job satisfaction and commitment, while technostress inhibitors improve employee outcomes. These findings remain relevant to AI because AI tools often increase technological complexity,

uncertainty, workload, and invasion.

AI technostress may arise from several sources. Employees may feel overloaded by too many AI systems. They may feel invaded when AI extends work expectations beyond normal boundaries. They may feel insecure if AI makes their skills appear obsolete. They may feel uncertain because AI tools evolve rapidly and require continuous learning. They may feel complexified when they must understand, verify, and explain AI outputs. When these demands exceed available support, output declines because employees spend more time coping with the technology than producing meaningful work.

Theme 3: AI weakens engagement when it reduces autonomy and meaningfulness

AI automation can reduce engagement when employees feel that their work has become controlled, standardized, or deskilled. Engagement depends not only on workload but also on autonomy, purpose, and ownership. When AI prescribes how employees should respond, what decisions they should make, or how their performance should be evaluated, employees may experience reduced self-determination.

This mechanism is visible in AI-assisted customer service, algorithmic management, and automated decision-support systems. Brynjolfsson *et al.* (2025) [1] found that AI assistance improved average productivity but also raised concerns about reduced original contributions from highly skilled workers and possible long-term effects on model learning and expertise. The implication is that AI may improve standardization while weakening employee agency, especially among experts whose value lies in judgment, creativity, and tacit knowledge.

Theme 4: AI reduces output when employees over-rely on it outside its capability frontier

The "jagged technological frontier" explains why AI can reduce output quality. Dell'Acqua *et al*. (2023) [3] show that AI produces strong gains when used within its capability frontier but can reduce performance when used outside it. This finding is critical because many employees do not know where AI is reliable and where it is not.

Over-reliance may lead to inaccurate reports, shallow analysis, biased decisions, generic customer responses, poor problem diagnosis, and reduced critical thinking. In high-stakes work, the cost of such errors can be significant. Output reduction therefore appears not only as fewer tasks completed, but also as lower quality, more rework, higher error correction costs, reduced trust, and weakened professional accountability.

Theme 5: AI creates hidden work and work intensification

AI is often introduced to reduce workload, but in practice it may create hidden work. Employees may need to prompt AI systems, verify outputs, correct hallucinations, explain automated decisions, handle customer frustration, update data, manage system errors, and learn continuously changing tools. These activities are often invisible in formal productivity metrics.

Chuang *et al*. (2025) [2] show that AI can increase productivity while AI technostress produces exhaustion, work–family conflict, and lower job satisfaction. Högemann (2025) [6] also reports that although managers expect productivity gains from generative AI, many workers experience increased workload and reduced productivity perceptions. This suggests that AI may shift work from visible production to invisible coordination and correction.

Theme 6: AI reduces sustainable output by depleting psychological capital

Work engagement requires energy, hope, resilience, efficacy, and optimism. Hou and Fan (2024) [7] show that AI-induced stress reduces work engagement partly by reducing psychological capital. This is highly relevant because employees may initially comply with AI adoption while gradually losing confidence and motivation.

Psychological capital depletion occurs when employees repeatedly feel unable to master new tools, fear comparison with AI, receive insufficient training, or experience AI as a sign that human skills are less valued. When psychological capital declines, employees may still complete required tasks, but they are less likely to show initiative, creativity, persistence, and service-oriented behavior. Thus, output may remain temporarily stable while engagement declines, creating delayed productivity loss.

6. Discussion: Interpretation of Findings

The evidence suggests that AI reduces employee engagement and output when automation is implemented as a labor-substitution strategy rather than a human-augmentation strategy. The dark side of AI does not arise from technology alone; it emerges from the interaction between technology, work design, employee resources, managerial expectations, and organizational support.

The first major interpretation is that AI changes the psychological contract between employee and organization. When workers believe that AI is used to support them, engagement may increase. When they believe, AI is used to monitor, replace, accelerate, or deskill them, engagement declines. This distinction explains why the same technology may produce positive outcomes in one organization and negative outcomes in another.

The second interpretation is that AI may improve short-term output while damaging long-term productive capacity. For example, an AI tool may help employees produce more reports, answer more tickets, or process more transactions. However, if employees become exhausted, lose autonomy, stop developing expertise, or disengage from problem-solving, the organization may suffer from declining creativity, quality, learning, and resilience. This is the hidden productivity cost of automation.

The third interpretation is that AI creates a new form of role stress. Employees are expected to collaborate with AI, supervise AI, correct AI, explain AI, and remain accountable for AI-assisted outputs. Yet many organizations do not formally redesign jobs to recognize these new responsibilities. This creates role ambiguity and role overload, which Tarafdar *et al*. (2007) [12] identified as pathways through which technology reduces productivity.

The fourth interpretation is that organizational support is a central protective factor. Xu *et al*. (2023) [14] and Hou and Fan (2024) [7] both show that perceived organizational support reduces the harmful psychological effects of AI-related stress. Support must go beyond technical training. It should include transparent communication, job-security assurances where possible, employee participation in AI implementation, clear accountability rules, mental health

resources, and redesign of performance metrics.

The fifth interpretation is that AI implementation should be evaluated using both productivity and well-being indicators. If managers only measure speed, cost, or output volume, they may conclude that AI is successful while missing early warning signs of burnout, disengagement, quality decline, and knowledge loss. A more balanced evaluation should include engagement, psychological safety, perceived autonomy, workload, learning, error rates, customer satisfaction, and employee retention.

7. Limitations and Future Research

This study has several limitations. First, it is based on secondary sources and does not collect primary empirical data. Therefore, its conclusions are interpretive rather than statistically causal. Second, the reviewed literature covers multiple sectors and technologies, including ICT, generative AI, algorithmic management, and hospitality AI systems. Although these technologies overlap, their effects may differ across industries, occupations, and national contexts. Third, many AI studies are recent and short-term. Longitudinal evidence on how AI affects engagement, expertise, output quality, and career development remains limited.

Future research should conduct longitudinal studies examining whether AI-induced engagement loss accumulates over time. Researchers should also distinguish between output quantity and output quality, because AI may increase the former while reducing the latter. More sector-specific studies are needed in healthcare, banking, education, logistics, manufacturing, hospitality, and public administration. Future research should also examine moderating variables such as AI self-efficacy, digital literacy, leadership style, psychological safety, participatory technology design, and organizational justice.

Another important direction is to study "AI over-reliance" and "skill atrophy." As employees use AI more frequently, organizations need to know whether human judgment, creativity, and problem-solving capabilities are strengthened or weakened. Future research should also explore how AI affects vulnerable groups, including older workers, clerical workers, frontline service employees, and employees in routine cognitive occupations.

8. Conclusion

This study examined the dark side of AI by analyzing when automation reduces employee engagement and output. The review shows that AI can improve productivity under favorable conditions, but it can also reduce engagement and sustainable output when it creates job insecurity, technostress, emotional exhaustion, reduced autonomy, over-reliance, hidden work, and psychological resource depletion.

The main conclusion is that AI is not inherently engagement-enhancing or engagement-destroying. Its effects depend on implementation. AI becomes harmful when organizations use it to intensify work, reduce human agency, obscure accountability, or signal labor replacement. It becomes beneficial when organizations use it to support employees, reduce low-value work, strengthen skills, and redesign jobs around human–AI complementarity.

For managers, the key lesson is clear: AI adoption should not be measured only by speed, cost reduction, or output volume. It must also be assessed by its effects on employee engagement, psychological safety, autonomy, learning, and long-term output quality. The central managerial question is not "How many tasks can AI automate?" but "How can AI be implemented without destroying the human energy, judgment, and commitment that make organizations productive?"

9. References

1. Brynjolfsson E, Li D, Raymond LR. Generative AI at work. The Quarterly Journal of Economics. 2025; 140(2):889-942. Doi: 10.1093/qje/qjae044
2. Chuang YT, Chiang HL, Lin AP. Insights from the Job Demands-Resources Model: AI's dual impact on employees' work and life well-being. International Journal of Information Management. 2025; 83, Article 102887. Doi: 10.1016/j.ijinfomgt.2025.102887
3. Dell'Acqua F, McFowland E, Mollick ER, Lifshitz-Assaf H, Kellogg KC, Rajendran S, *et al*. Navigating the jagged technological frontier: Field experimental evidence of the effects of AI on knowledge worker productivity and quality. Harvard Business School Working Paper, 2023.
4. Gmyrek P, Berg J, Bescond D. Generative AI and jobs: A global analysis of potential effects on job quantity and quality. International Labour Organization, 2023.
5. Gmyrek P, Berg J, Kamiński K, Konopczyński F, Ładna A, Nafradi B. Generative AI and jobs: A refined global index of occupational exposure. International Labour Organization, 2025.
6. Högemann M. Technostress and generative AI in the workplace: A qualitative analysis of young professionals. Frontiers in Artificial Intelligence, 2025.
7. Hou Y, Fan L. Working with AI: The effect of job stress on hotel employees' work engagement. Behavioral Sciences. 2024; 14(11):1076. Doi: 10.3390/bs14111076
8. Noy S, Zhang W. Experimental evidence on the productivity effects of generative artificial intelligence. Science. 2023; 381(6654):187-192. Doi: 10.1126/science.adh2586
9. Ragu-Nathan TS, Tarafdar M, Ragu-Nathan BS, Tu Q. The consequences of technostress for end users in organizations: Conceptual development and empirical validation. Information Systems Research. 2008; 19(4):417-433. Doi: 10.1287/isre.1070.0165
10. Salanova M, Llorens S, Cifre E. The dark side of technologies: Technostress among users of information and communication technologies. International Journal of Psychology. 2013; 48(3):422-436. Doi: 10.1080/00207594.2012.680460
11. Sharif MN, Zhang L, Asif M, Alshdaifat SM, Hanaysha JR. Artificial intelligence and employee outcomes: Investigating the role of job insecurity and technostress in the hospitality industry. Acta Psychologica. 2025; 253, Article 104733. Doi: 10.1016/j.actpsy.2025.104733
12. Tarafdar M, Tu Q, Ragu-Nathan BS, Ragu-Nathan TS. The impact of technostress on role stress and productivity. Journal of Management Information Systems. 2007; 24(1):301-328. Doi: 10.2753/MIS0742-1222240109
13. World Economic Forum. The Future of Jobs Report 2025. World Economic Forum, 2025.
14. Xu G, Xue M, Zhao J. The association between artificial intelligence awareness and employee

depression: The mediating role of emotional exhaustion and the moderating role of perceived organizational support. International Journal of Environmental Research and Public Health. 2023; 20(6):5147. Doi: 10.3390/ijerph20065147