Future Technology
Open Access Journal
ISSN 2832-0379

February 2026| Volume 05 | Issue 01 | Pages 47-54

Journal homepage: https://fupubco.com/futech

https://doi.org/10.55670/fpll.futech.5.1.5

Future Publishing LLC

Article

Reconstruction of knowledge worker performance evaluation system in the ChatGPT era: an exploratory study based on human-AI collaborative work model

Zhixin Yu, Zhicheng Yu*

University College Dublin, National University of Ireland, Dublin, Ireland, D04V1W8

ARTICLE INFO

Article history:
Received 21 August 2025
Received in revised form
29 September 2025
Accepted 16 October 2025

Keywords:
Performance evaluation, ChatGPT, Human-AI collaboration, Knowledge workers, AI skills

*Corresponding author
Email address:
20020721yuzhicheng@gmail.com

DOI: 10.55670/fpll.futech.5.1.5

ABSTRACT

The emergence of ChatGPT in November 2022 disrupted practice in knowledge work and defied performance-measurement systems in human-exclusive task accomplishment under unprecedented comparability. This current study fills the gap in the literature between traditional models of appraisal and AI-enabled workspaces through the development of an evidence-based model of measuring performance in human-AI collaborative settings. Drawing on systematic analysis of 5,000 LinkedIn job adverts and 2,000 Indeed salary information between 2022-2024, the present study examined the shift in performance needs and skill needs in knowledge sectors following the release of ChatGPT. The study's findings indicated that AI skills are especially needed in 27.8% of knowledge workers' jobs, with a growth rate of 376% since the release of ChatGPT. AI-trained staff are rewarded with a 17.7% overall premium for their wages, and occupational competence varies from 43.2% in high-tech to 9.7% in the public sector. Systematic skill differences cannot be captured by conventional measuring systems, according to the results. The study discovers a three-dimensional model for measuring performance, including AI Tool Mastery, Collaborative Work Quality, and Human-AI Synergy to measure hybrid skills developed through human-machine collaboration. The research establishes the theory of performance management by developing operational measurement solutions for companies going through workplace redesign due to AI.

1. Introduction

The introduction of ChatGPT in November 2022 transformed conventional knowledge work processes, and never-before-seen problems were unveiled for human-exclusive task accomplishment-oriented performance measurement systems [1]. Generative AI-supported knowledge workers achieve tangible gains in writing, programming, and analytical work [2], and extensive evidence establishes revolutionary impacts on value creation processes in knowledge-intensive environments [3]. Institutions have seen greater practice and administrative procedure adoption of AI [4], with stakeholder research showing significant effects on traditional evaluation processes [5]. Studies of student writing demonstrate mature interactions among AI support and skill development, suggesting that traditional measures do not capture real capabilities well in AI-supported contexts [6]. Workplace assimilation research reveals ChatGPT exerts a significant influence on the process of knowledge workers searching for, processing, and making use of information sources [7]. Such influence extends far beyond micro productivity effects at the level, qualitatively transforming organizational processes and decision-making processes. Developing evidence suggests generative AI adoption is highly heterogeneous across organizational levels, with knowledge workers at various points in their careers presenting differential adoption patterns of AI into work. Early adopters indicate that they spend as much as 30% of their working hours working with AI tools, raising the question of how performance management systems need to account for such a dramatic work process change. Organizations increasingly have to redefine performance metrics since conventional output measures are no longer able to capture value created in AI-aided processes. Seventy years of performance management research are still held back by models that are designed for human-alone performance [8]. Evaluating systems have been radically transformed, but cannot support AI-driven work processes [9]. The literature of organizational behavior can determine AI multi-dimensional workplace effects, but is not sufficiently prepared with systems to measure hybrid performance [10]. Methodological controversy still suggests AI evaluation processes are task-dependent and not universally applicable [11]. Empirical research indicates AI collaboration improves organizational performance through improved resource coordination [12]. Human-AI collaboration evidence varies on employee performance, particularly in safety-sensitive domains [13]. Experiments indicate that explainable AI improves group task performance, but traditional frameworks fail to reflect these interactions [14]. Human-centered AI teaming frameworks emphasize maintaining human agency in combination with enhanced technological capability [15]. Generative AI has uncovered inbuilt constraints in existing performance evaluation processes reliant on human mental effort as the sole source of organizational value production. Contemporary evaluation systems measure human potential against established standards, assuming evident human performance at tasks. However, for AI scenarios, performance becomes increasingly dependent on the metacompetence of staff to successfully enable human-AI collaboration – a function foreign to classic models and unmeasured by them. Misalignments such as these between measurement and actual practice produce systemically blind company talent management. Trust machinery exerts compelling forces on human-AI quality of decision-making but lies beyond existing systems [16]. Despite these constraints, performance appraisal systems are a long-standing organizational value as much as quality management perspectives are concerned [17]. Systematic reviews confirm global trends in building skills, with evidence noting growing upskilling and reskilling requirements as AI penetration is getting deeper roots in industries [18, 19].

Despite growing academic interest in AI employment effects, the literature inadequately addresses performance appraisal reconstruction. While research addresses operational and safety dimensions of human-AI collaboration, it abstains from considering primary measurement and evaluation concerns, leaving the literature ungrounded in adequate theory incorporating classical evaluation cultures and AI-enabled workplaces. To address these gaps, this study pursues three specific objectives: (1) quantify the transformation of knowledge worker skill requirements following ChatGPT's release through systematic analysis of job market data; (2) assess the economic value of AI competencies by examining compensation differentials across industries and organizational levels; and (3) develop a three-dimensional performance evaluation framework specifically designed for human-AI collaborative work environments. This research makes three distinct contributions that differentiate it from existing literature. First, it provides the first large-scale empirical analysis (5,000 job postings, 2,000 salary records) of post-ChatGPT performance requirements, moving beyond theoretical discussions to market-driven evidence. Second, while prior studies examine human-AI collaboration in isolated tasks, this framework systematically measures hybrid competencies across organizational contexts through three integrated dimensions. Third, the documented 376% growth in AI skills demand and 17.7% salary premium establishes economic validation absent in existing performance management literature, demonstrating that traditional evaluation systems systematically undervalue emerging workplace capabilities.

2. Research design and methodology

2.1 Research design

This study adopts a quantitative research design using systematic analysis of publicly available employment market data to investigate performance evaluation transformation in the ChatGPT era. The conceptual framework (Figure 1) illustrates the paradigmatic shift from individual-focused assessment to collaborative human-AI effectiveness evaluation [20]. The research employs a secondary data analysis approach to capture the evolutionary trajectory of knowledge worker performance requirements [21]. The study leverages job posting data from LinkedIn (5,000 positions) and salary information from Indeed (2,000 entries), spanning 2022-2024, to provide an empirical foundation for understanding the transition from output-based metrics to AI-collaborative competencies. Key operational definitions are provided below. Human-AI collaborative work denotes task completion wherein knowledge workers utilize AI tools and validate outputs through human judgment. Hybrid competencies represent skills emerging from human-machine interaction that transcend individual human or AI capabilities. Collaborative effectiveness measures human-AI joint performance quality through three dimensions: AI Tool Mastery, Collaborative Work Quality, and Human-AI Synergy.

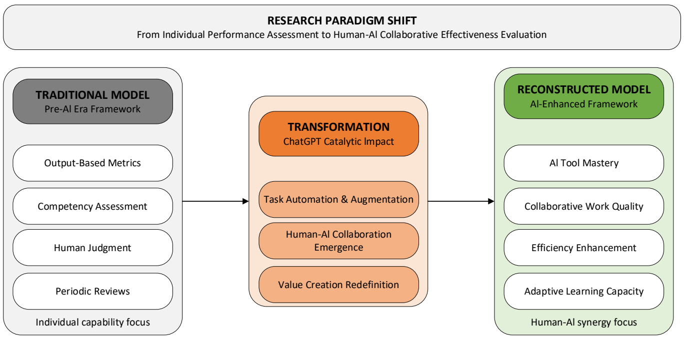

Figure 1. Conceptual framework for performance evaluation transformation

2.2 Data sources

The empirical foundation of this research rests on a multi-source data collection strategy designed to capture the transformation of performance evaluation requirements in knowledge-intensive sectors. The data architecture encompasses three complementary sources providing triangulated evidence of market-driven changes (Table 1). LinkedIn job posting data (5,000 positions) enables systematic tracking of skill requirement evolution, particularly AI-related competencies across organizational contexts. Indeed, salary records (2,000 entries) facilitate quantitative assessment of compensation differentials associated with AI proficiency, establishing economic validation of evolving performance criteria. Corporate case studies (10 organizations) supplement these market indicators with organizational implementation evidence. The temporal scope spanning 2022-2024 captures both pre-ChatGPT baseline conditions and subsequent transformational patterns, enabling comparative analysis of requirement shifts [22]. This study acknowledges several data limitations. LinkedIn and Indeed platforms may exhibit demographic and industry biases toward technology-intensive sectors. Job postings potentially represent idealized rather than actual skill requirements. The U.S.-focused sample limits international generalizability, as AI adoption patterns vary across regulatory environments. The 2022-2024 timeframe captures immediate responses rather than long-term trends.

Table 1. Data sources and analytical framework

Data Source	Data Type	Sample Size	Time Frame	Key Variables	Analytical Purpose
LinkedIn	Job postings	5,000 positions	2022-2024	AI skill requirements, job titles, and industry sectors	Skill demand evolution analysis
Indeed	Salary records	2,000 entries	2022-2024	Compensation levels, AI skill premiums	Wage differential analysis
Corporate Reports	Case studies	10 organizations	2023-2024	AI implementation strategies, performance metrics	Framework validation

Note: Corporate case studies encompass ten organizations across five sectors: technology (n=3), financial services (n=2), management consulting (n=2), healthcare (n=2), and manufacturing (n=1). Data derived from publicly available annual reports and HR white papers (2023-2024), providing organizational validation of framework applicability across diverse industry contexts.

2.3 Analysis methods

The analytical method follows a multi-method quantitative approach to systematically examine performance requirement evolution in knowledge-intensive industries. Descriptive statistical analysis provides baseline information on skill requirement trends, detecting significant differences in AI-related competency demands across the study period. Regression analysis estimates the economic impact of AI proficiency on compensation levels, establishing empirical evidence for market valuation of emerging skills by examining salary differentials between AI-skilled and traditionally-skilled positions. Text mining methods extract and categorize performance-driven terms through a three-stage process. Python NLP libraries (NLTK, spaCy) identified AI-related keywords from job postings. Two researchers independently coded 200 postings across eight skill categories, achieving inter-rater reliability (Cohen's Kappa = 0.84). The validated coding scheme was applied to the full 5,000-position dataset, distinguishing emerging AI-related requirements from traditional skill mentions. This comprehensive analytical strategy aligns with rigorous methodological guidelines for performance measurement research [23], ensuring total coverage of market-driven changes while maintaining analytical rigor.

3. Findings

3.1 AI skills demand growth

The empirical analysis reveals a pronounced transformation in AI skills demand across knowledge-intensive sectors (Figure 2). During the pre-ChatGPT period (2022-01 to 2022-10), AI skills mentions remained relatively stable, fluctuating between 8.1% and 9.5% of total job postings. This baseline pattern indicates nascent generative AI integration in workplace requirements. The release of ChatGPT in November 2022 marked a decisive inflection point, triggering an immediate acceleration in demand from 9.5% to 14.7% within the subsequent quarter. This study documents a sustained upward trajectory throughout 2023-2024, culminating in 27.8% of knowledge worker positions explicitly requiring AI competencies by April 2024. The overall growth rate of 376% represents a fundamental shift in skill requirements. Absolute mentions increased from 42 to 200 instances across the 5,000-position sample. The consistent post-ChatGPT acceleration suggests that generative AI capabilities have become integral to organizational performance expectations.

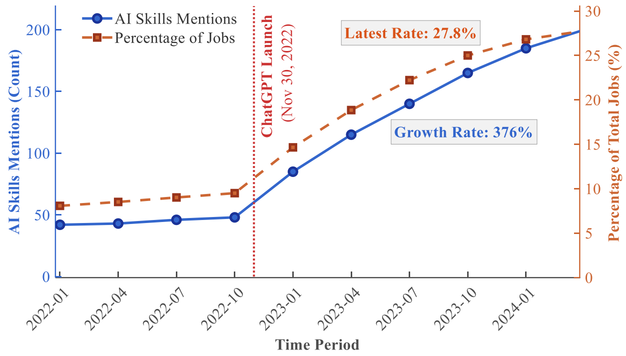

Figure 2. Temporal evolution of AI skills demand in knowledge work (2022-2024)

3.2 Salary premium

The salary premium analysis reveals substantial economic value associated with AI skills across all organizational levels (Figure 3). This analysis demonstrates that AI-required positions command significant compensation advantages. Mid-level roles exhibit the highest premium at 21.5%, followed by senior-level positions at 16.7% and entry-level positions at 15.0%. Normalized against traditional entry-level positions, salary index data indicate an average premium of 17.7% for AI-competent workers. This compensation differential reflects market recognition of AI skills as valuable organizational assets. Mid-level professionals capture the greatest premium due to their optimal combination of technical proficiency and practical experience. The economic validation supports the demand growth patterns observed in job posting analysis.

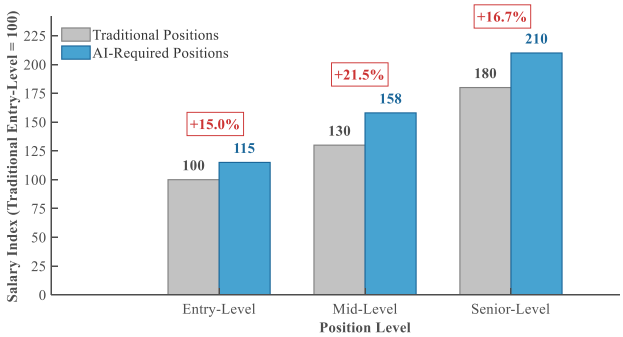

Figure 3. AI skills salary premium analysis (2024 data)

3.3 New Performance Indicators

The detailed skill requirement analysis reveals systematic transformation in the competencies demanded by knowledge-intensive organizations (Table 2). This investigation documents substantial shifts across eight core performance domains, with technical proficiency demonstrating the highest adoption rate at 16.8% of job postings explicitly requiring AI-related capabilities by 2024. Traditional skill categories have been fundamentally supplemented by AI-collaborative competencies, with growth rates ranging from 216.7% to 300.0% across all domains. Analytical capabilities and communication skills emerge as particularly significant transformation areas, reflecting organizational recognition that knowledge work increasingly involves human-AI collaboration rather than purely individual performance.

Table 2. Emerging vs. traditional skill requirements in knowledge work job postings (2022-2024)

Skill Category	Traditional Requirements	Emerging AI-Related Requirements	2022 Frequency (%)	2024 Frequency (%)	Change (%)
Technical Proficiency	Excel proficiency, Data entry, Software operation	ChatGPT experience, AI tools familiarity, Generative AI usage	4.2	16.8	+300.0
Analytical Capabilities	Statistical analysis, Research skills, Report writing	AI-assisted analysis, Data interpretation with AI, Automated reporting	3.3	12.1	+266.7
Communication Skills	Written communication, Presentation skills, Client interaction	AI content creation, Enhanced writing with AI, Digital collaboration	2.5	8.9	+256.0
Problem-Solving	Critical thinking, Independent analysis, Solution development	AI-supported reasoning, Enhanced problem solving, Technology integration	1.9	6.2	+226.3
Project Management	Timeline management, Resource allocation, Team coordination	Digital workflow optimization, AI-assisted planning, Process automation	1.3	4.8	+269.2
Creative Tasks	Design thinking, Content creation, Innovation	AI-enhanced creativity, Content generation, Creative collaboration	1.6	5.6	+250.0
Quality Assurance	Manual review, Error checking, Compliance monitoring	Output validation, Quality control, Review processes	1.0	3.2	+220.0
Learning & Adaptation	Professional development, Skill updating, Training completion	Technology adaptation, Platform learning, Continuous upskilling	1.2	3.8	+216.7

Note: Comparison of traditional and emerging AI-related skill requirements in knowledge work job postings based on 5,000 LinkedIn positions (2022-2024). Frequencies represent the percentage of job postings explicitly mentioning these requirements within each skill category.

However, these hybrid competencies present fundamental measurement challenges for traditional performance evaluation systems designed to assess discrete individual capabilities rather than collaborative human-AI effectiveness. Organizations currently lack standardized metrics to evaluate how effectively employees leverage AI assistance, validate AI-generated outputs, or integrate artificial intelligence into complex decision-making processes.

3.4 Industry differences

The cross-sectional analysis reveals substantial heterogeneity in AI skills demand across knowledge-intensive sectors (Figure 4). Technology and software organizations demonstrate the highest adoption rates at 43.2% of job postings, while government and public sector positions exhibit the lowest demand at 9.7%, creating a differential of 33.5 percentage points between sectors. Financial services and management consulting sectors position themselves above the cross-industry average of 27.8%, reflecting their strategic emphasis on data-driven decision-making and analytical capabilities. Healthcare and pharmaceuticals organizations approach near-average adoption levels at 28.6%, indicating moderate integration of AI competencies into clinical and research workflows. Marketing and advertising sectors exceed education and training sectors at 21.5% versus 19.2% respectively, suggesting commercial applications drive faster AI adoption than institutional educational contexts. Manufacturing and engineering sectors remain below average at 16.8%, potentially reflecting traditional operational structures and regulatory constraints.

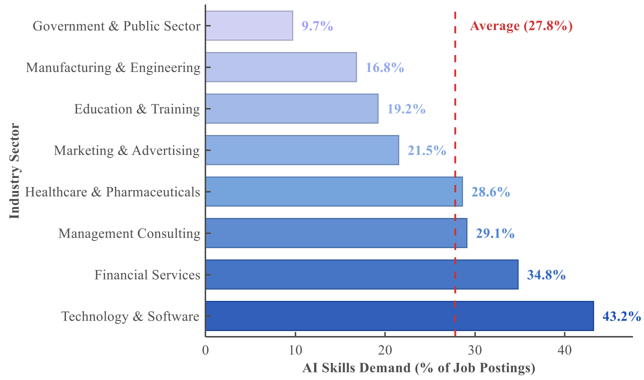

Figure 4. AI skills demand across industry sectors (2024)

4. Proposed framework

4.1 Three-dimensional model

The proposed framework reconceptualizes performance evaluation through three integrated dimensions that address the empirical skill transformation documented in Section 3 (Figure 5). AI Tool Mastery encompasses technical proficiency requirements, ranging from basic generative AI familiarity to advanced tool integration capabilities. Collaborative Work Quality captures the effectiveness of human-AI cooperative processes, including output validation, quality control, and enhanced decision-making workflows. Human-AI Synergy represents the emergent capability to optimize human cognitive strengths alongside artificial intelligence assistance, creating value through complementary task allocation. These dimensions intersect to form performance zones reflecting integrated competency levels rather than isolated skill assessments. The framework addresses traditional evaluation limitations by enabling organizations to measure hybrid capabilities emerging from human-AI collaboration. Performance assessment occurs within three-dimensional space, where employee effectiveness depends on balanced development across all dimensions rather than excellence in single competencies. This approach accommodates the finding that 27.8% of knowledge worker positions now require AI-related skills, while providing flexibility for industry-specific weighting adjustments.

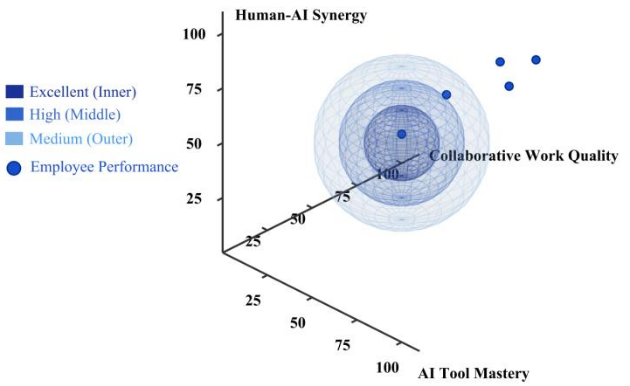

Figure 5. Three-dimensional AI performance framework

4.2 Simple metrics

The operationalization of the three-dimensional framework requires practical measurement indicators that enable organizations to assess human-AI collaborative performance systematically (Table 3). This study proposes streamlined metrics that address the documented skill transformation where 27.8% of knowledge worker positions now demand AI competencies. AI Tool Mastery encompasses technical skill assessment through ChatGPT experience evaluation and AI tools familiarity testing, providing quantifiable measures of technological proficiency. Collaborative Work Quality focuses on output validation capabilities and AI-assisted analysis competencies, establishing benchmarks for quality control effectiveness and analytical accuracy in human-AI collaborative environments. Human-AI Synergy measures AI-supported reasoning abilities and technology integration effectiveness through problem-solving evaluation and integration capability assessment. These indicators derive from the empirical analysis of emerging skill requirements, enabling organizations to evaluate employee performance within AI-augmented work contexts rather than traditional individual-focused metrics. The framework addresses measurement challenges identified in conventional performance systems by capturing collaborative effectiveness between knowledge workers and artificial intelligence tools, providing organizations with actionable assessment criteria for the evolving workplace landscape.

Table 3. Three-dimensional performance evaluation framework

Dimension	Key Indicators	Measurement Methods
AI Tool Mastery	ChatGPT experience and AI tools familiarity	Technical skill assessment, Usage proficiency evaluation
Collaborative Work Quality	Output validation and AI-assisted analysis capabilities	Quality control metrics, Analytical accuracy assessment
Human-AI Synergy	AI-supported reasoning and technology integration	Problem-solving effectiveness evaluation, Integration capability assessment

Note: Framework addresses the documented transformation where 27.8% of knowledge worker positions now require AI competencies. Indicators derived from empirical analysis of emerging skill requirements in human-AI collaborative work environments.

4.3 Framework application guidelines

The transformation from traditional performance evaluation systems to AI-enhanced collaborative assessment requires systematic integration of the three-dimensional framework developed through this research (Figure 6). Organizations must transition beyond individual performance focus and output-based metrics toward a comprehensive evaluation of human-AI collaborative effectiveness. The framework incorporates AI Tool Mastery, Collaborative Work Quality, and Human-AI Synergy as interconnected dimensions that collectively address the documented skill transformation where 27.8% of knowledge worker positions now require AI competencies. The empirical evidence supporting this paradigm shift includes the 376% growth in AI skills demand and 17.7% average salary premium for AI-competent workers, establishing market validation for the proposed evaluation approach. Application of this framework enables organizations to measure hybrid competencies emerging from human-AI collaboration rather than discrete individual capabilities. Organizations implementing this framework can expect enhanced alignment with evolving market demands, improved talent attraction and retention capabilities, and more accurate assessment of knowledge worker productivity in AI-augmented environments.

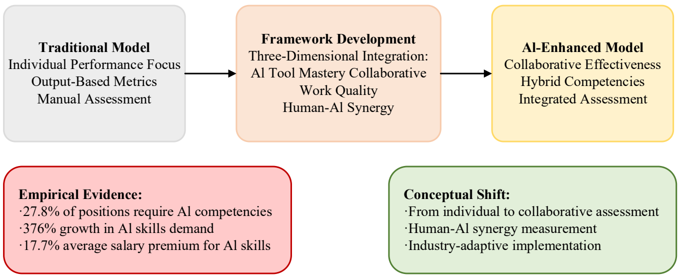

Figure 6. Performance evaluation framework transformation

5. Discussion

From a theoretical perspective, this research advances performance management theory by integrating the resource-based view and task-technology fit frameworks to explain AI's role in organizational effectiveness. The three-dimensional framework reconceptualizes AI as a complementary organizational resource rather than a substitutive technology, addressing fundamental gaps in traditional models such as Kaplan and Norton's Balanced Scorecard, which assume individual-based value creation. The empirical finding that 27.8% of knowledge worker positions now require AI competencies supports the theoretical proposition that technology integration transforms the nature of performance, necessitating evaluation systems that measure collaborative effectiveness rather than isolated individual contributions. The 376% increase in demand for AI capability is a paradigm shift in knowledge organization value creation [24]. The shift contradicts conventional assumptions regarding individual capability measurement because current models are incapable of fully explaining collective efficacy at the point where human cognitive capacity meets artificial intelligence capacity [25]. Generative AI has revolutionized knowledge work itself, developing hybrid task environments where traditional productivity metrics no longer function as valid predictors of real performance contribution [26]. Companies that use pre-AI assessment models risk systematically underestimating workers who possess valuable human-AI collaboration skills [27]. Scholarship on worker-AI coexistence points to the imperative of recalibrating workplace assessment systems to fit new collaboration paradigms [28].

Aside from compensation equity issues, AI-driven performance measurement brings with it serious organizational change management issues. Studies on technology adoption in performance management contexts identify employee resistance as being due to fairness and transparency [29]. As the criteria of evaluation change to incorporate AI expert knowledge, employees who had mastered previous paradigms would lose out, and this would trigger serious organizational conflict. Change management best practice equals rollout success, founded on large-scale stakeholder communication, pilot rollout by business units, and phased rollout timelines for competency development step-wise [30]. Organizations need to balance the necessity for appraisal system change with employee interest and faith in change. The average 17.7% salary increase for AI-skilled professionals is a market demand for composite skills that cannot be measured by current appraisal systems [31]. This imbalance presages organizational hazards since the lack of know-how in performance management systems for AI promise has the potential to misallocate human capital structurally [32]. Our proposed three-dimensional approach addresses the burning human-AI collaborative measurement gaps that the literature has posited, but could not address effectively [29]. Human-AI collaboration research is discovered to lean towards maintenance of human agency with greater technical expertise, but provides little advice on how to gauge such complex interactions [33]. This paper's articulation of AI Tool Mastery, Collaborative Work Quality, and Human-AI Synergy variables offers systematic solutions to the measurement of hybrid skills as a product of human-machine collaboration [34], in accordance with integrative job performance measurement frameworks that call for multidimensional measurement [35]. The real-world trade-offs between standardization and contextualization are uncovered through the implementation experience of early adopter organizations. Since three-dimensional structure aids conceptual integrity, organizations face high heterogeneity in the realization of AI Tool Mastery across functional areas [36]. Whereas programming and API integration capabilities are valued by technical functions, timely engineering and output verification capabilities are valued by administrative functions. Such functional diversity demands a competency library's mapping framework dimensions to role-based behavioral indicators, creating implementation complexity but facilitating evaluation relevance in varied organizational settings [37]. Industry variation implies powerful low-level contextual determinants calling for framework adaptation. There is a 33.5 percentage point disparity in underlying organizational readiness levels required for AI adoption [36]. Performance measurement systems should have the ability to capture flexibility in meeting different integration levels without compromising the consistency of measurements [30].

Further research should investigate usage issues in the operations of AI-based performance appraisal systems and organizational and employee development consequences [37]. Interdependencies between AI-based performance systems, motivation, skill learning channels, and employee career development should be researched [38,39]. The development and testing of AI-augmented skills standardization are key requirements for building theory-driven skills and everyday practice. AI measurement is also ethically incorrect as it raises issues such as algorithmic control, overwork, and data concealment. The businesses need to implement this system in the open world and disclose the assessment parameters to employees without diminishing human judgment in the conclusion. Standard audits shouldn't be biased in favor of algorithmic bias, which favors specific groups, and must provide equal evaluation outcomes.

6. Conclusion

This study presumes that the arrival of ChatGPT brought scale-level requirements of knowledge work performance, and 27.8% officially working tasks require AI competency, and 376% growth rates require AI competency as a prerequisite for organizational capital. The greatest contribution of this study is to construct a three-dimensional performance measurement model on the basis of human-AI collaborative workplace environments. Positioning itself at the intersection of AI Tool Mastery drivers, Collaborative Work Quality drivers, and Human-AI Synergy drivers, the model makes it possible for the company to measure hybrid skills hidden from the previous individual-based systems. The recent 17.7% wage gap between workers augmented by AI only reflects marketplace trial runs of such fledgling necessity skills, which initiated performance management system redesign. Such an institution that fails to update its system of evaluation risks becoming a victim of structural misallocation of human capital and an inability to capitalize on talent in an increasingly knowledge-based economy that is more competitive. The study enables the development of theory since the transition to team output-based measure and collaborative effectiveness measure constructs can be created, completing the research gaps in performance management. While methodological constraints are opposite to generalizability, conclusions provide pragmatic recommendations to organizations that are undertaking AI-facilitated workplace change. Future research should focus on empirical validation of the proposed framework through longitudinal organizational studies and the development of standardized assessment tools for human-AI collaborative competencies. The study ultimately contributes to performance management theory by providing evidence-based solutions for measuring value creation in the evolving landscape of AI-enhanced knowledge work.

Ethical issue

The authors are aware of and comply with best practices in publication ethics, specifically with regard to authorship (avoidance of guest authorship), dual submission, manipulation of figures, competing interests, and compliance with policies on research ethics. The authors adhere to publication requirements that the submitted work is original and has not been published elsewhere.

Data availability statement

The manuscript contains all the data. However, more data will be available upon request from the authors.

Conflict of interest

The authors declare no potential conflict of interest.

References

[1] Y. K. Dwivedi et al., "Opinion Paper:"So what if ChatGPT wrote it?" Multidisciplinary perspectives on opportunities, challenges and implications of generative conversational AI for research, practice and policy," International journal of information management, vol. 71, p. 102642, 2023. https://doi.org/10.1016/j.ijinfomgt.2023.102642

[2] S. Noy and W. Zhang, "Experimental evidence on the productivity effects of generative artificial intelligence," Science, vol. 381, no. 6654, pp. 187-192, Jul 14 2023, doi: 10.1126/science.adh2586. DOI:10.1126/science.adh2586

[3] E. Brynjolfsson, D. Li, and L. Raymond, "Generative AI at work," The Quarterly Journal of Economics, vol. 140, no. 2, pp. 889-942, 2025. https://doi.org/10.1093/qje/qjae044

[4] A. M. Bettayeb, M. Abu Talib, A. Z. Sobhe Altayasinah, and F. Dakalbab, "Exploring the impact of ChatGPT: conversational AI in education," in Frontiers in Education, 2024, vol. 9: Frontiers Media SA, p. 1379796. https://doi.org/10.3389/feduc.2024.1379796

[5] A. M. Hasanein and A. E. E. Sobaih, "Drivers and consequences of ChatGPT use in higher education: Key stakeholder perspectives," European journal of investigation in health, psychology and education, vol. 13, no. 11, pp. 2599-2614, 2023. https://doi.org/10.3390/ejihpe13110181

[6] S. Mahapatra, "Impact of ChatGPT on ESL students' academic writing skills: A mixed methods intervention study," Smart Learning Environments, vol. 11, no. 1, p. 9, 2024. https://doi.org/10.1186/s40561-024-00295-9

[7] H. Jo and D.-H. Park, "AI in the workplace: Examining the effects of ChatGPT on information support and knowledge acquisition," International Journal of Human–Computer Interaction, vol. 40, no. 23, pp. 8091-8106, 2024. https://doi.org/10.1080/10447318.2023.2278283

[8] T. C. Brown, P. O'Kane, B. Mazumdar, and M. McCracken, "Performance management: A scoping review of the literature and an agenda for future research," Human Resource Development Review, vol. 18, no. 1, pp. 47-82, 2019. DOI:10.1177/1534484318798533

[9] A. S. DeNisi and K. R. Murphy, "Performance appraisal and performance management: 100 years of progress?," J Appl Psychol, vol. 102, no. 3, pp. 421-433, Mar 2017, doi: 10.1037/apl0000085. https://doi.org/10.1037/apl0000085

[10] S. Bankins, A. C. Ocampo, M. Marrone, S. L. D. Restubog, and S. E. Woo, "A multilevel review of artificial intelligence in organizations: Implications for organizational behavior research and practice," Journal of organizational behavior, vol. 45, no. 2, pp. 159-182, 2024. https://doi.org/10.1002/job.2735

[11] G. Fragiadakis, C. Diou, G. Kousiouris, and M. Nikolaidou, "Evaluating human-ai collaboration: A review and methodological framework," arXiv preprint arXiv:2407.19098, 2024. https://doi.org/10.48550/arXiv.2407.19098

[12] A. Przegalinska, T. Triantoro, A. Kovbasiuk, L. Ciechanowski, R. B. Freeman, and K. Sowa, "Collaborative AI in the workplace: Enhancing organizational performance through resource-based and task-technology fit perspectives," International Journal of Information Management, vol. 81, p. 102853, 2025.https://doi.org/10.1016/j.ijinfomgt.2024.102853

[13] Y. Liu and Y. Li, "Does human-AI collaboration promote or hinder employees' safety performance? A job demands-resources perspective," Safety Science, vol. 188, p. 106872, 2025. https://doi.org/10.1016/j.ssci.2025.106872

[14] J. Senoner, S. Schallmoser, B. Kratzwald, S. Feuerriegel, and T. Netland, "Explainable AI improves task performance in human-AI collaboration," Sci Rep, vol. 14, no. 1, p. 31150, Dec 28 2024. https://doi.org/10.1038/s41598-024-82501-9

[15] S. Berretta, A. Tausch, G. Ontrup, B. Gilles, C. Peifer, and A. Kluge, "Defining human-AI teaming the human-centered way: a scoping review and network analysis," Frontiers in Artificial Intelligence, vol. 6, p. 1250725, 2023. https://doi.org/10.3389/frai.2023.1250725

[16] Y. Wen, J. Wang, and X. Chen, "Trust and AI weight: human-AI collaboration in organizational management decision-making," Frontiers in Organizational Psychology, vol. 3, p. 1419403, 2025. https://doi.org/10.3389/forgp.2025.1419403

[17] Y. Y. Liao, E. Soltani, A. Iqbal, and R. van der Meer, "The utility of performance review systems: a total quality management perspective," Strategic Change, vol. 33, no. 4, pp. 287-310, 2024. https://doi.org/10.1002/jsc.2580

[18] L. Babashahi et al., "AI in the workplace: A systematic review of skill transformation in the industry," Administrative Sciences, vol. 14, no. 6, p. 127, 2024. https://doi.org/10.3390/admsci14060127

[19] S. Morandini, F. Fraboni, M. De Angelis, G. Puzzo, D. Giusino, and L. Pietrantoni, "The impact of artificial intelligence on workers' skills: Upskilling and reskilling in organisations," Informing Science, vol. 26, pp. 39-68, 2023. https://doi.org/10.28945/5078

[20] D. Tranfield, D. Denyer, and P. Smart, "Towards a methodology for developing evidence‐informed management knowledge by means of systematic review," British journal of management, vol. 14, no. 3, pp. 207-222, 2003. https://doi.org/10.1111/1467-8551.00375

[21] S. Zhang, A. H. A. Hamid, and B. S. Alias, "A Study on Research Performance Evaluation: A Systematic Literature Review," Journal of Scientometric Research, vol. 14, no. 2, pp. 383-398, 2025. https://doi.org/10.5530/jscires.20251910

[22] P. Varsha, A. Chakraborty, and A. K. Kar, "How to undertake an impactful literature review: Understanding review approaches and guidelines for high-impact systematic literature reviews," South Asian Journal of Business and Management Cases, vol.

13, no. 1, pp. 18-35, 2024. DOI:10.1177/22779779241227654

[23] C. Mio, A. Costantini, and S. Panfilo, "Performance measurement tools for sustainable business: A systematic literature review on the sustainability balanced scorecard use," Corporate social responsibility and environmental management, vol. 29, no. 2, pp. 367-384, 2022. https://doi.org/10.1002/csr.2206

[24] A. Cebulla and Z. Szpak, "Workplace Relations with AI in Mind: What Is Likely to Change?," in The Future of Work and Technology: Chapman and Hall/CRC, 2023, pp. 151-172. https://doi.org/10.1201/9781003393757-8

[25] M. Mitchell, "Why AI is harder than we think," arXiv preprint arXiv:2104.12871, 2021. https://doi.org/10.48550/arXiv.2104.12871

[26] Woodruff, Allison, et al. "How knowledge workers think generative ai will (not) transform their industries." Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems. 2024.https://doi.org/10.1145/3613904.364270

[27] J. Howard, "Artificial intelligence: Implications for the future of work," Am J Ind Med, vol. 62, no. 11, pp. 917-926, Nov 2019, doi: 10.1002/ajim.23037. https://doi.org/10.1002/ajim.23037

[28] A. Zirar, S. I. Ali, and N. Islam, "Worker and workplace Artificial Intelligence (AI) coexistence: Emerging themes and research agenda," Technovation, vol. 124, p. 102747, 2023. https://doi.org/10.1016/j.technovation.2023.102747

[29] M. I. Biswas, M. S. Talukder, and A. R. Khan, "Who do you choose? Employees' perceptions of artificial intelligence versus humans in performance feedback," China Accounting and Finance Review, vol. 26, no. 4, pp. 512-532, 2024. https://doi.org/10.1108/CAFR-08-2023-0095

[30] I. do Carmo Lameque, M. J. P. Velez, and C. M. D. A. Botelho, "Purposes of performance appraisal: A systematic review and agenda for future research," International Journal of Professional Business Review: Int. J. Prof. Bus. Rev., vol. 8, no. 7, p. 40, 2023. DOI:10.26668/businessreview/2023.v8i7.2274

[31] R. Bukartaite and D. Hooper, "Automation, artificial intelligence and future skills needs: an Irish perspective," European Journal of Training and Development, vol. 47, no. 10, pp. 163-185, 2023. https://doi.org/10.1108/EJTD-03-2023-0045

[32] K. R. Murphy, "Performance evaluation will not die, but it should," Human Resource Management Journal, vol. 30, no. 1, pp. 13-31, 2020. https://doi.org/10.1111/1748-8583.12259

[33] H. G. Óskarsdóttir, G. V. Oddsson, J. Þ. Sturluson, and R. J. Sæmundsson, "Towards a holistic framework of knowledge worker productivity," Administrative sciences, vol. 12, no. 2, p. 50, 2022. https://doi.org/10.3390/admsci12020050

[34] M. Cosa and R. Torelli, "Digital transformation and flexible performance management: A systematic literature review of the evolution of performance measurement systems," Global Journal of Flexible Systems Management, vol. 25, no. 3, pp. 445-466, 2024. https://doi.org/10.1007/s40171-024-00409-9

[35] H. Sandall, L. M. C. e. Silva, and F. Queiroga, "A comprehensive approach to job performance in the service sector: A systematic literature review," BAR-Brazilian Administration Review, vol. 19, no. 02, p. e210046, 2022. DOI:10.1590/1807-7692bar2022210046

[36] H. M. Alzoubi, M. In'airat, and G. Ahmed, "Investigating the impact of total quality management practices and Six Sigma processes to enhance the quality and reduce the cost of quality: the case of Dubai," International journal of business excellence, vol. 27, no. 1, pp. 94-109, 2022. DOI:10.1504/IJBEX.2020.10039342

[37] N. Khan, Z. Khan, A. Koubaa, M. K. Khan, and R. B. Salleh, "Global insights and the impact of generative AI-ChatGPT on multidisciplinary: a systematic review and bibliometric analysis," Connection science, vol. 36, no. 1, p. 2353630, 2024. DOI:10.36227/techrxiv.171466626.67294030/v1

[38] J. Dempere, K. Modugu, A. Hesham, and L. K. Ramasamy, "The impact of ChatGPT on higher education," in Frontiers in Education, 2023, vol. 8: Frontiers Media SA, p. 1206936. https://doi.org/10.3389/feduc.2023.1206936

[39] L. Isiaku, A. S. Muhammad, H. I. Kefas, and F. C. Ukaegbu, "Enhancing technological sustainability in academia: leveraging ChatGPT for teaching, learning and evaluation," Quality Education for All, vol. 1, no. 1, pp. 385-416, 2024. https://doi.org/10.1108/QEA-07-2024-0055

This article is an open-access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).