Future of Work and Digital Management Journal

Article type:
Original Research

Article history:
Received 05 May 2026
Revised 26 June 2026
Accepted 03 August 2026
Initial Publish 27 August 2026
Published online 01 July 2027

Designing a Performance Evaluation Model for Knowledge Workers in the Atomic Energy Organization of Iran

Mehrdad. Faraghat[1], bahram. Alishiri[2*], Seyed Ahmad. Hosseini Golafshani[3]

1 PhD Student, Department of Public Administration, ST.C., Islamic Azad University, Tehran, Iran.
2 Department of Public Administration, NT.C., Islamic Azad University, Tehran, Iran.
3 Department of Public Administration, ST.C., Islamic Azad University, Tehran, Iran.

Corresponding author email address:
bahram.alishiri@iau.ac.ir

How to cite this article:
Faraghat, M., Alishiri, B., & Hosseini Golafshani, S. A. (2027). Designing a Performance Evaluation Model for Knowledge Workers in the Atomic Energy Organization of Iran. *Future of Work and Digital Management Journal, 5*(4), 1-25. https://doi.org/10.61838/fwdmj.307

ABSTRACT

This study aimed to design and validate a comprehensive performance evaluation model for knowledge workers in the Atomic Energy Organization of Iran. An exploratory sequential mixed-methods design was employed. In the qualitative phase, semi-structured interviews were conducted with 12 academic and organizational experts selected purposively until theoretical saturation was reached. Interview data were analyzed through thematic analysis, thematic-network analysis, and the fuzzy Delphi technique using MAXQDA. In the quantitative phase, the statistical population consisted of 110 knowledge workers, managers, and supervisors, of whom 85 were selected through stratified simple random sampling based on Cochran's formula. Data were collected using a researcher-made questionnaire whose content validity was confirmed by 10 experts. A pilot study was conducted with 30 employees, and the questionnaire demonstrated high internal consistency. Quantitative data were analyzed using SPSS and partial least squares structural equation modeling. Qualitative coding generated 60 indicators classified into eight dimensions: perceived supervisory support, team empowerment, perceived support for innovation, psychological safety, social cohesion, information sharing, internal communication, and external communication. Inter-coder reliability was confirmed by a significant Cohen's kappa coefficient of 0.758 ($p < .001$). All indicators were retained after two fuzzy Delphi rounds, with inter-round differences below 0.80. Exploratory factor analysis was supported by KMO = 0.914 and a significant Bartlett's test, $\chi^2(1770) = 11{,}876.408$, $p < .001$. Eight factors with eigenvalues greater than one explained 57.585% of the variance. All first-order factor loadings exceeded 0.40 and were significant at $p < .001$. Second-order loadings ranged from 0.730 to 0.867. Cronbach's alpha values ranged from 0.733 to 0.947, composite reliability ranged from 0.849 to 0.952, and all AVE values exceeded 0.50. Discriminant validity was confirmed through the Fornell–Larcker criterion, cross-loadings, and HTMT values below 0.90. Model fit was acceptable, with SRMR = 0.071 and NFI = 0.800. Expert validation also confirmed the model's internal, external, and overall validity ($p < .01$). The validated model provides a reliable and multidimensional framework for evaluating knowledge-worker performance by integrating managerial, team-based, innovative, psychological, social, and communicational dimensions.

Keywords: Knowledge workers, performance evaluation, perceived supervisory support, team empowerment, knowledge management, Atomic Energy Organization of Iran.

Introduction

The transition from industrial economies to knowledge-based economies has fundamentally altered the nature of work, the sources of organizational value, and the criteria through which employee performance should be understood and assessed. In traditional organizational settings, performance evaluation frequently focused on observable outputs, compliance with predetermined procedures, attendance, task completion, and direct productivity. Such criteria remain relevant in many operational roles, but they are insufficient for evaluating employees whose principal contribution is the production, interpretation, integration, and application of specialized knowledge. Knowledge workers generate value through intellectual judgment, problem solving, technical expertise, creativity, learning, collaboration, and the capacity to transform information into decisions or innovations [1]. Consequently, their performance cannot be evaluated adequately through systems designed primarily for routine and standardized work. Contemporary performance-evaluation models must therefore capture not only individual outputs but also the organizational, managerial, team-based, communicational, and knowledge-related conditions through which specialized performance is produced.

The growing strategic importance of knowledge has made knowledge management an essential component of organizational performance. Organizations increasingly depend on their ability to acquire, preserve, retrieve, exchange, and apply knowledge across individuals, teams, and organizational units. Evidence indicates that knowledge-management infrastructure can strengthen sustainable organizational performance, particularly when organizations can simultaneously exploit existing capabilities and explore new opportunities through organizational ambidexterity [2]. Likewise, data-driven knowledge frameworks are increasingly used to predict risks, optimize decisions, and improve organizational performance in complex and multinational environments [3]. These developments demonstrate that employee performance is no longer produced solely through individual competence. It also depends on the systems through which knowledge is governed, circulated, converted into action, and aligned with strategic objectives.

Knowledge workers occupy a particularly important position in this process because their work involves a high degree of cognitive complexity, specialization, discretion, and dependence on tacit as well as explicit knowledge. Tacit knowledge is often embedded in professional experience, personal judgment, technical intuition, and repeated participation in complex work processes. It is difficult to document fully and cannot always be transferred through formal instructions. Research on tacit knowledge-management processes has shown that the effective exchange and application of tacit knowledge can contribute to product innovation and organizational performance, especially when affective trust and task efficiency support the knowledge-transfer process [4]. Accordingly, a valid performance-evaluation model for knowledge workers should not be restricted to measurable end products. It should also recognize behaviors that facilitate knowledge creation, knowledge exchange, collaborative problem solving, professional learning, and the development of innovative solutions.

The relationship between knowledge management and performance is also shaped by leadership and organizational culture. Knowledge-management systems do not operate independently of the social and managerial context in which employees work. Leadership determines whether employees are encouraged to share expertise, question established procedures, seek assistance, experiment with new approaches, and participate in decisions. Organizational culture similarly influences whether knowledge is treated as a private source of individual power or a collective organizational resource. Studies have demonstrated that knowledge management, digital leadership, and organizational culture jointly influence employee performance, suggesting that performance improvement requires an integrated rather than exclusively individual approach [5]. Effective leadership of knowledge management can also improve organizational performance by aligning knowledge-related activities with organizational priorities and ensuring that knowledge resources are used systematically rather than incidentally [6]. Therefore, evaluating knowledge workers requires attention to the leadership environment that enables or restricts the productive use of professional knowledge.

Supervisory support is especially important in knowledge-intensive work. Knowledge workers often require autonomy, but autonomy does not eliminate the need for managerial guidance, access, trust, and support. Rather, the supervisor’s role changes from direct control of routine tasks to facilitation of professional performance. Supervisors can contribute to knowledge-worker performance by clarifying strategic expectations, providing developmental feedback, removing organizational barriers, supporting employee judgment, protecting legitimate experimentation, and connecting specialized teams with senior management. Leadership models developed for knowledge-based companies emphasize that leaders must create conditions that support expertise, learning, innovation, and employee participation rather than relying exclusively on formal authority and hierarchical control [7]. This implies that perceived supervisory support should be treated as a potentially central component of knowledge-worker performance evaluation because it reflects both the quality of direct supervision and the organization's willingness to empower supervisors to support specialized employees effectively.

Team empowerment constitutes another important dimension of knowledge-intensive performance. Many complex organizational problems cannot be solved by isolated individuals, regardless of their expertise. They require interdisciplinary collaboration, integration of different forms of knowledge, and coordinated decision-making. Empowered teams are able to choose appropriate methods, organize their work, use members' expertise, and adapt their activities to changing circumstances. Such teams are also more likely to regard work as an opportunity for development and to assume responsibility for achieving organizational objectives. Knowledge governance research suggests that the organizational regulation and coordination of knowledge processes can enhance innovative employee performance when employees have access to appropriate structures, authority, and knowledge resources [8]. Team empowerment is therefore not merely a motivational practice; it is a mechanism for mobilizing distributed knowledge and transforming specialized competence into collective performance.

Innovation is similarly inseparable from the performance of knowledge workers. In knowledge-intensive organizations, superior performance often involves identifying new opportunities, developing alternative methods, improving technical processes, and creating solutions for problems that have no predetermined answer. Innovation capability depends partly on how organizations manage knowledge and whether employees are encouraged to apply it in novel ways. Research has shown that knowledge-management practices can strengthen innovation capability and, through innovation, contribute to business performance [9]. Moreover, the relationship between artificial intelligence and employee performance may operate through knowledge sharing, knowledge retrieval, and organizational meaning systems, indicating that technological tools improve performance when they are embedded in effective knowledge processes [10]. A performance-evaluation model for knowledge workers should therefore recognize innovation-supportive behaviors and conditions, including encouragement of new ideas, willingness to test alternative approaches, openness to employee suggestions, and attention to opportunities for continuous improvement.

Technological change has intensified the need to reconsider employee performance evaluation. Artificial intelligence, big data, digital platforms, and smart human resource systems have expanded the range of information available for recruitment, development, monitoring, and evaluation. Artificial intelligence can improve the speed and analytical capacity of human resource processes, but it also introduces challenges concerning transparency, interpretability, fairness, privacy, and the risk of reducing complex human performance to incomplete digital indicators [11]. In knowledge-intensive settings, these concerns are especially significant because many valuable contributions are qualitative, collaborative, delayed, or difficult to attribute to a single employee. Smart human resource governance models in Iran's public sector consequently emphasize the need to align digital tools, governance arrangements, organizational accountability, and human resource practices [12]. Thus, the development of an appropriate evaluation model requires more than the adoption of new technology; it requires a conceptually valid framework capable of distinguishing meaningful indicators from data that are merely easy to collect.

The work environment is another determinant of knowledge-worker performance. Employees' capacity to use knowledge depends on access to information, supportive interpersonal relations, appropriate communication channels, adequate resources, and opportunities to apply professional judgment. The combined influence of work environment and knowledge management on employee performance indicates that performance emerges from the interaction between employee capability and the organizational conditions surrounding work [1]. This interaction is particularly important in highly specialized organizations, where the consequences of poor communication, restricted information flow, or inadequate coordination may extend beyond individual inefficiency and affect organizational reliability, learning, and strategic decision-making.

Psychological safety is a central part of such an environment. Knowledge workers must often communicate uncertainty, identify errors, raise difficult questions, and challenge existing assumptions. These behaviors involve interpersonal risk, especially in hierarchical organizations where disagreement may be interpreted as disloyalty or incompetence. When employees fear embarrassment, blame, or retaliation, they may withhold concerns and remain silent even when they possess knowledge that could prevent mistakes or improve decisions. A psychologically safe environment enables employees to request assistance, acknowledge limitations, express different perspectives, and use their specialized knowledge without fear of interpersonal harm. For performance evaluation purposes, psychological safety is relevant because the visible absence of errors or disagreement does not necessarily indicate high performance; it may instead indicate silence, avoidance, or concealment. An effective model should therefore consider whether the organizational environment supports the open use of knowledge rather than merely rewarding apparent conformity.

Social cohesion also influences knowledge-based performance. Knowledge exchange depends not only on formal systems but also on trust, belonging, interpersonal comfort, and the willingness of employees to cooperate. Cohesive teams are more likely to exchange experience, coordinate effort, and provide mutual support. However, cohesion should not be interpreted simply as social harmony. Its performance value lies in its capacity to facilitate collaboration while preserving the expression of different professional perspectives. A cohesive knowledge team combines a sense of collective identity with respect for specialized contributions. Such cohesion may strengthen the effects of empowerment, psychological safety, and knowledge sharing by reducing interpersonal barriers to communication and helping employees interpret their success as part of a shared organizational accomplishment.

Information sharing and communication are among the most visible behavioral expressions of knowledge-worker performance. Knowledge can contribute to organizational outcomes only when it reaches the individuals and units capable of using it. Internal information sharing includes the exchange of work reports, technical documents, professional experiences, planning information, and decision-relevant knowledge among team members. Broader internal communication connects teams with other departments and organizational levels, while external communication allows employees to acquire technical ideas, coordinate with outside groups, and remain informed about developments beyond the organization. The effectiveness of these processes depends on both communication quality and the organizational willingness to treat information as a shared resource. Research on employee performance has increasingly identified knowledge sharing and retrieval as mechanisms connecting technology, management systems, and employee outcomes [10]. Therefore, a model that ignores internal and external communication would overlook a major pathway through which knowledge workers create organizational value.

Despite the growing recognition of these multidimensional influences, many performance-evaluation systems continue to emphasize general indicators that do not distinguish knowledge work from other occupational categories. Public-sector systems may be especially vulnerable to standardization because uniform procedures are often adopted to promote administrative consistency and accountability. However, excessive uniformity can weaken construct validity when the same criteria are applied to routine, managerial, scientific, and technical roles. Recent research in Iranian government organizations has identified several drivers for improving human-capital performance-evaluation systems, demonstrating the need for more responsive, evidence-based, and context-sensitive approaches [13]. Similarly, the development of a performance-evaluation and accountability model for Iran’s Audit Organization shows that evaluation systems in specialized public institutions must reflect their professional missions, accountability structures, and expert judgments [14]. These studies underscore the necessity of designing evaluation systems that are institutionally specific without sacrificing methodological rigor.

Performance evaluation also has a sustainability dimension. Sustainable human resource performance evaluation extends beyond short-term output and considers whether employee performance contributes to long-term organizational capability, learning, resilience, and responsible resource use. A sustainable evaluation model should support employee development, reduce counterproductive pressures, preserve critical expertise, and align individual contributions with organizational continuity [15]. This perspective is particularly relevant for knowledge workers because the loss, underutilization, or mismanagement of specialized knowledge can have long-lasting consequences. Evaluation systems that reward only immediate, easily quantifiable outputs may discourage mentoring, documentation, collaborative learning, and innovation, even though these activities are essential for maintaining organizational knowledge over time.

Methodologically, the complexity of performance evaluation has encouraged the use of multidimensional and hybrid decision-making approaches. Models based on combinations of DEMATEL, ELECTRE, and intuitionistic fuzzy techniques illustrate how interdependent indicators can be identified, weighted, and prioritized when evaluating organizational units [16]. Such approaches are valuable because performance indicators rarely operate independently. Supervisory support may affect innovation, empowerment may influence information sharing, psychological safety may shape communication, and social cohesion may facilitate knowledge exchange. A valid model must therefore identify the relevant components and empirically examine whether they form a coherent structure rather than treating them as an arbitrary list of criteria.

The Atomic Energy Organization of Iran represents a setting in which these issues are especially consequential. Its work depends heavily on specialized scientific, technical, engineering, managerial, and regulatory expertise. Employees in such an organization are required to operate within a context characterized by technical complexity, strategic sensitivity, extensive interdependence, and a strong need for accuracy, reliability, organizational learning, and controlled innovation. Conventional evaluation criteria may capture compliance and task completion but fail to represent the quality of professional judgment, collaborative knowledge use, supervisory facilitation, team autonomy, psychological safety, or communication across specialized boundaries. A context-specific performance-evaluation model is therefore necessary to reflect how knowledge-worker performance is actually generated and sustained in this organizational environment.

Existing Iranian studies have contributed important insights into sustainable performance evaluation, smart human resource governance, public-sector evaluation systems, organizational accountability, leadership in knowledge-based firms, and the drivers of human-capital performance improvement [7, 12-15]. International studies have likewise clarified the relationships among knowledge management, digital leadership, work environment, innovation, tacit knowledge, knowledge governance, and organizational performance [1, 3-6, 8, 9]. Nevertheless, an important gap remains. These contributions do not provide an integrated, empirically validated model specifically designed to evaluate knowledge workers in the Atomic Energy Organization of Iran by combining qualitative exploration, expert consensus, exploratory factor analysis, and confirmatory structural modeling.

Accordingly, the aim of this study was to design and validate a multidimensional performance-evaluation model for knowledge workers in the Atomic Energy Organization of Iran.

Methodology

This study employed an exploratory sequential mixed-methods design integrating qualitative and quantitative approaches. The qualitative phase was conducted first to explore the principal dimensions, components, variables, and indicators required for evaluating the performance of knowledge workers in the Atomic Energy Organization of Iran. The findings of this phase were subsequently used to develop a researcher-made questionnaire and construct the preliminary performance evaluation model. In the quantitative phase, a survey was conducted among knowledge workers to examine the identified components, assess the relationships among the study variables, and validate the proposed model. Before collecting empirical data, an extensive review of the theoretical and empirical literature was undertaken through library-based research and systematic note-taking. Relevant Persian and international books, journal articles, dissertations, and theses addressing knowledge workers, employee performance, performance evaluation systems, human resource management, and knowledge-intensive organizations were reviewed. The results of this literature review provided the theoretical foundation for developing the interview protocol and interpreting the empirical findings.

The participants in the qualitative phase were academic specialists and experienced organizational experts with substantial knowledge of human resource management, performance evaluation, knowledge workers, and the organizational characteristics of the Atomic Energy Organization of Iran. Participants were selected purposively based on their professional experience, academic expertise, familiarity with knowledge-intensive work, and ability to provide detailed information relevant to the research problem. Semi-structured interviews continued concurrently with data analysis until theoretical saturation was achieved and additional interviews no longer produced substantially new concepts, categories, or indicators. Theoretical saturation was reached after 12 interviews. This sample size was considered sufficient because the collected data demonstrated conceptual repetition and adequately represented the principal dimensions of the phenomenon under investigation.

The statistical population of the quantitative phase consisted of 110 knowledge workers employed by the Atomic Energy Organization of Iran. Knowledge workers were distinguished from other employees according to two operational criteria. The first criterion included employees working in technical and engineering disciplines who held university qualifications, and the second included managers and supervisors who worked directly with knowledge workers and were involved in supervising, coordinating, or evaluating their specialized activities. Based on these criteria, eligible individuals were identified and included in the sampling frame. Using Cochran's sample-size formula and a stratified random sampling procedure, 85 participants were selected for the quantitative survey. Stratification was used to ensure adequate representation of the relevant occupational and managerial categories, while eligible individuals within each stratum were selected randomly.

Several procedures were applied to strengthen the trustworthiness and confirmability of the qualitative findings. Data were collected from multiple sources, and triangulation was used to compare the evidence obtained from the literature, interviews, expert judgments, and organizational experiences. Negative-case analysis was also performed to identify and resolve contradictory explanations or interpretations emerging from the interview data. The interview transcripts were read repeatedly, and the identified concepts, categories, and interpretive processes were continually reviewed and reassessed. Methodological flexibility was maintained throughout the analysis so that interpretations, findings, and proposed modifications could be revised in response to new evidence. Peer review was undertaken at different stages of the research by presenting the emerging findings to specialists and incorporating their critical observations into the analytical process. After the preliminary theoretical model had been developed, it was returned to a selected group of participants and experts for member checking. Their views concerning the modification, elimination, clarification, or addition of model components were examined, and relevant recommendations were incorporated into the final model.

The principal data collection instrument in the qualitative phase was a semi-structured interview protocol developed on the basis of the research objectives and the literature review. The interviews were designed to elicit experts' perceptions regarding the defining characteristics of knowledge work, the individual and organizational determinants of knowledge-worker performance, appropriate performance indicators, contextual factors influencing performance evaluation, and the mechanisms required for implementing an effective evaluation system within the Atomic Energy Organization of Iran. The semi-structured format ensured that the principal research domains were addressed consistently while allowing participants to elaborate on their professional experiences, introduce additional concepts, and explain context-specific considerations. The interviews were documented and transcribed for subsequent coding and thematic analysis.

The reliability of the qualitative coding was evaluated using Cohen's kappa coefficient. An independent expert who had not been informed of the researcher's coding procedures or the concepts previously developed was asked to code a portion of the interview data independently. The coding results produced by the researcher and the independent expert were then compared. The resulting agreement coefficient was 0.74, indicating an acceptable level of inter-coder reliability and providing evidence that the coding process had been conducted consistently. In addition to inter-coder agreement, repeated transcript review, peer examination, member checking, triangulation, and negative-case analysis were used to enhance the credibility and dependability of the qualitative results.

The quantitative data collection instrument was a researcher-made questionnaire developed from the themes, categories, dimensions, and indicators identified through the literature review, expert interviews, thematic analysis, and fuzzy Delphi process. The questionnaire was designed to operationalize the proposed dimensions of knowledge-worker performance and obtain standardized responses from the selected organizational sample. To assess the content and face validity of the instrument, the preliminary questionnaire was reviewed by 10 experts. Each expert evaluated the questionnaire items in terms of comprehensibility, relevance to the dimensions of the research problem, appropriateness of wording, and potential ambiguity or likelihood of misinterpretation. Based on their comments, unclear, difficult, repetitive, or insufficiently relevant items were revised, clarified, or removed.

A pilot study was subsequently conducted with 30 employees who possessed characteristics similar to those of the main study population. Participants in the pilot phase evaluated the clarity and comprehensibility of the items and reported any uncertainty or ambiguity they experienced while completing the questionnaire. The preliminary data and participants' corrective comments were examined, and the identified deficiencies were resolved before the final version of the questionnaire was administered. The internal consistency of the final questionnaire was assessed using Cronbach's alpha. The overall Cronbach's alpha coefficient was 0.94, demonstrating excellent internal consistency and indicating that the questionnaire items measured the proposed constructs with an acceptable degree of reliability.

Qualitative and quantitative data were analyzed through complementary analytical procedures corresponding to the sequential mixed-methods design. The qualitative interview data were organized and analyzed using MAXQDA software. The analysis began with repeated reading of the interview transcripts to achieve immersion in the data and identify meaningful statements related to knowledge-worker performance evaluation. Relevant statements were coded, similar codes were compared and grouped, and progressively broader categories and themes were developed. Thematic analysis was used to identify the basic, organizing, and overarching themes represented in the interview data. The analysis was further developed through a thematic network approach, which clarified the relationships among the extracted indicators, categories, and higher-order dimensions and contributed to the construction of the preliminary conceptual model.

Following thematic analysis, the fuzzy Delphi method was used to confirm, prioritize, and screen the indicators identified from the 12 expert interviews. This method made it possible to transform the linguistic judgments of experts into quantitative values while accounting for uncertainty and variation in their assessments. Triangular fuzzy numbers were used to represent expert opinions regarding the relevance and importance of each proposed indicator. The aggregated expert judgments were examined to determine the degree of consensus concerning the retention, modification, or elimination of the indicators. Indicators that achieved an acceptable level of expert agreement were retained in the model, whereas indicators with insufficient support were reconsidered or excluded. The combination of thematic analysis and the fuzzy Delphi method ensured that the proposed model was grounded both in empirical interview data and in systematic expert consensus.

The quantitative survey data were analyzed using SPSS and partial least squares structural equation modeling software. SPSS was employed for data preparation, descriptive analysis, preliminary reliability assessment, and examination of the general characteristics of the collected responses. Partial least squares structural equation modeling was then used to evaluate the measurement model, examine the structural relationships among the identified dimensions, and validate the proposed performance evaluation model. Indicator reliability was assessed through factor loadings, Cronbach's alpha coefficients, and composite reliability coefficients. Convergent validity was examined using the average variance extracted, with acceptable values indicating that the observed indicators adequately represented their corresponding latent constructs.

The structural model was assessed by examining the significance of the estimated path coefficients through bootstrapped t statistics. A t value greater than 1.96 was considered indicative of statistical significance at the 0.05 level. The explanatory capacity of the endogenous constructs was evaluated using the coefficient of determination, or R^2. Values of approximately 0.32 and 0.67 were interpreted as indicating moderate and strong explanatory power, respectively. The overall predictive and explanatory adequacy of the proposed model was additionally evaluated through the goodness-of-fit index. The combined assessment of factor loadings, internal consistency coefficients, composite reliability, average variance extracted, path significance, R^2 values, and the goodness-of-fit index provided the basis for determining the reliability, validity, and empirical adequacy of the final performance evaluation model for knowledge workers in the Atomic Energy Organization of Iran.

Findings and Results

Analysis of the 12 semi-structured interviews was conducted through open and axial coding in MAXQDA. Repeated comparison of the interview statements, initial codes, and conceptually related categories generated 60 primary indicators. During axial coding, these indicators were organized into eight overarching categories: perceived supervisory support, team empowerment, perceived support for innovation, psychological safety, social cohesion, information sharing, internal communication, and external communication. The MAXQDA code-system visualization also confirmed that the 60 initial codes formed eight conceptually distinguishable clusters. The resulting dimensions and their associated indicators are presented in Table 1.

Table 1

Axial coding results and the indicators constituting the proposed model

Axial category	Number of indicators	Open codes and performance indicators
Perceived supervisory support	21	The supervisor's efforts to promote employees' occupational interests; provision of an environment for open organizational discussion; provision of specific developmental guidance by the supervisor; the supervisor's trust and confidence in employees; supervisory guidance when problems arise; ease of access to the supervisor as a team member; ease of access to the supervisor during problems; the supervisor's willingness to listen to employee requests; encouragement to approach the supervisor when unexpected or emerging problems arise; employees' comfort in discussing unexpected or emerging problems with the supervisor; organizational respect for the supervisor; the supervisor's authority to make important decisions; organizational appreciation of the supervisor's contributions; organizational permission for the supervisor to experiment with new practices; organizational support for supervisory decisions; the supervisor's participation in major organizational decisions; the supervisor's influence on senior-management decisions; organizational permission for the supervisor to manage affairs according to professional judgment; organizational consultation with the supervisor regarding new policies and procedures; permission for the direct supervisor to select behavior appropriate to individual employees; and the importance attributed by the organization to continued cooperation with the supervisor.
Team empowerment	10	The team's ability to select different methods of performing work; implementation of activities selected by the team; freedom to decide what the team performs; confidence in the team's ability to accomplish its work; belief in the team's influence and effectiveness; belief that the team empowers organizational members; possession of a strong knowledge base within the team; belief that work provides an opportunity for growth; access to necessary information concerning the organization's future goals and expectations; and provision of information to employees regarding how organizational goals can be achieved.
Perceived support for innovation	7	Supervisory encouragement to develop ideas; encouragement to test new ways of performing work; encouragement to identify new solutions rather than relying on older approaches; willingness of the supervisor to listen to new ideas; supervisory attention to new opportunities for work improvement; willingness to discuss desired objectives and new methods of achieving them; and supervisory encouragement of innovative ideas.
Psychological safety	6	The ability of team members to raise difficult problems and issues; avoidance of undermining other team members; the possibility of taking interpersonal risks without harmful consequences; the ability to request assistance from other team members; acceptance of individual differences among team members; and recognition and use of members' unique skills and talents.
Social cohesion	5	A sense of belonging to the team as an important social group; team members spending time together outside working hours; interpersonal comfort among team members; happiness associated with team membership; and enjoyment of the team's social activities.
Information sharing	5	Sharing work reports and documents with other team members; sharing team experience or knowledge with other teams; freely sharing information required for key decisions; sharing useful information for planning; and sharing information useful for organizational development.
Internal communication	3	Sharing formal work reports and documents with other team members; sharing personal experience or knowledge with other team members; and sharing information required for key decisions among team members.
External communication	3	Sharing information about the team's activities with other groups in the organization; coordinating activities with groups outside the organization; and collecting technical information or ideas from sources outside the organization.

The largest axial category was perceived supervisory support, which contained 21 indicators and represented both direct leader–employee interactions and the broader organizational authority, legitimacy, and support afforded to supervisors. Team empowerment contained 10 indicators related to autonomy, team competence, influence, access to strategic information, and developmental opportunity. Perceived support for innovation included seven indicators reflecting supervisory encouragement of experimentation and idea development. Psychological safety, social cohesion, and information sharing comprised six, five, and five indicators, respectively, while internal and external communication each reflected distinct patterns of knowledge exchange within and beyond the immediate team.

The stability of the qualitative categorization was examined using Cohen's kappa. An independent expert in the subject area, who had not been informed about the researcher's integration of the initial codes or the resulting conceptual categories, independently classified the codes. The two sets of classifications were then compared. The kappa coefficient was 0.758, with a standard error of 0.105 and a test statistic of 7.219. The result was statistically significant at $p < .001$.

Table 2

Inter-coder agreement and interpretation of the kappa coefficient

Evaluation component	Value
Number of axial categories compared	8
Cohen's kappa	0.758
Standard error	0.105
Test statistic	7.219
Significance level	< .001
Interpretation	Valid/substantial agreement

For interpretation, kappa values below zero were considered indicative of poor agreement, values from 0.00 to 0.20 indicated negligible agreement, values from 0.21 to 0.40 indicated moderate agreement, values from 0.41 to 0.60 indicated appropriate agreement, values from 0.61 to 0.80 indicated valid or substantial agreement, and values from 0.81 to 1.00 indicated excellent agreement. Accordingly, the coefficient of 0.758 was located in the valid-agreement range and supported the reliability of the coding and categorization process.

The fuzzy Delphi technique was used to screen the 60 indicators and determine whether the indicators identified through the qualitative phase should be retained in the model. Twelve experts assessed the importance of each indicator. Because conventional numerical judgments cannot fully represent the ambiguity and linguistic nature of expert evaluations, the experts' judgments were transformed into triangular fuzzy numbers.

The nine-point linguistic scale ranged from "very unimportant," represented by the triangular fuzzy number (1, 1, 1), to "very important," represented by (8, 9, 9). The intermediate categories were "very unimportant to unimportant," represented by (1, 2, 3); "unimportant," represented by (2, 3, 4); "unimportant to moderately important," represented by (3, 4, 5); "moderately important," represented by (4, 5, 6); "moderately important to important," represented by (5, 6, 7); "important," represented by (6, 7, 8); and "important to very important," represented by (7, 8, 9). The corresponding definitive scores ranged from 1 to 9.

The first-round panel matrix included 720 judgments, comprising the assessment of 60 indicators by 12 experts. Most judgments were concentrated in the three upper fuzzy categories, namely (6, 7, 8), (7, 8, 9), and (8, 9, 9). A limited number of lower ratings, including (1, 1, 1), (1, 2, 3), (2, 3, 4), (3, 4, 5), and (4, 5, 6), were observed for some indicators; however, these ratings did not reduce the aggregate score of any indicator below the acceptance threshold. An indicator was accepted when its defuzzified score exceeded 7.00. In the second round, expert consensus was considered achieved when the absolute difference between the first- and second-round definitive values was less than 0.80.

Table 3

First- and second-round fuzzy Delphi results for the 60 indicators

Indicator	Round 1 crisp score	Round 1 decision	Round 2 crisp score	Difference	Consensus decision
C1	7.47	Accepted	7.57	0.10	Consensus
C2	7.42	Accepted	7.39	−0.03	Consensus
C3	7.42	Accepted	7.54	0.12	Consensus
C4	7.36	Accepted	7.32	−0.04	Consensus
C5	7.34	Accepted	7.38	0.04	Consensus
C6	7.28	Accepted	7.27	−0.01	Consensus
C7	7.17	Accepted	7.24	0.07	Consensus
C8	7.75	Accepted	7.67	−0.08	Consensus
C9	7.36	Accepted	7.28	−0.08	Consensus
C10	7.36	Accepted	7.27	−0.09	Consensus
C11	7.22	Accepted	7.15	−0.07	Consensus
C12	7.06	Accepted	7.11	0.05	Consensus
C13	7.08	Accepted	7.55	0.47	Consensus
C14	7.69	Accepted	7.84	0.15	Consensus
C15	7.86	Accepted	7.36	−0.50	Consensus
C16	7.34	Accepted	7.34	0.00	Consensus
C17	7.14	Accepted	7.44	0.30	Consensus
C18	7.47	Accepted	7.27	−0.20	Consensus
C19	7.14	Accepted	7.14	0.00	Consensus
C20	7.34	Accepted	7.84	0.50	Consensus
C21	7.89	Accepted	7.59	−0.30	Consensus
C22	7.59	Accepted	7.59	0.00	Consensus
C23	7.58	Accepted	7.88	0.30	Consensus
C24	7.84	Accepted	7.54	−0.30	Consensus
C25	7.47	Accepted	7.62	0.15	Consensus
C26	7.58	Accepted	7.97	0.39	Consensus
C27	7.89	Accepted	7.17	−0.72	Consensus
C28	7.11	Accepted	7.22	0.11	Consensus
C29	7.36	Accepted	7.22	−0.14	Consensus
C30	7.05	Accepted	7.49	0.44	Consensus
C31	7.50	Accepted	7.37	−0.13	Consensus
C32	7.31	Accepted	7.36	0.05	Consensus
C33	7.36	Accepted	7.41	0.05	Consensus
C34	7.36	Accepted	7.27	−0.09	Consensus
C35	7.25	Accepted	7.64	0.39	Consensus
C36	7.56	Accepted	7.43	−0.13	Consensus
C37	7.39	Accepted	7.38	−0.01	Consensus
C38	7.31	Accepted	7.42	0.11	Consensus
C39	7.50	Accepted	7.50	0.00	Consensus
C40	7.36	Accepted	7.36	0.00	Consensus
C41	7.36	Accepted	7.83	0.47	Consensus
C42	7.44	Accepted	7.44	0.00	Consensus
C43	7.72	Accepted	7.72	0.00	Consensus
C44	7.28	Accepted	7.28	0.00	Consensus
C45	7.53	Accepted	7.53	0.00	Consensus
C46	7.47	Accepted	7.47	0.00	Consensus
C47	7.53	Accepted	7.33	−0.20	Consensus
C48	7.50	Accepted	7.50	0.00	Consensus
C49	7.25	Accepted	7.85	0.60	Consensus
C50	7.61	Accepted	7.61	0.00	Consensus
C51	7.94	Accepted	7.64	−0.30	Consensus
C52	7.36	Accepted	7.24	−0.12	Consensus
C53	7.42	Accepted	7.42	0.00	Consensus
C54	7.61	Accepted	7.61	0.00	Consensus
C55	7.22	Accepted	7.72	0.50	Consensus
C56	7.47	Accepted	7.47	0.00	Consensus
C57	7.67	Accepted	7.67	0.00	Consensus
C58	7.66	Accepted	7.06	−0.60	Consensus
C59	7.42	Accepted	7.42	0.00	Consensus
C60	7.66	Accepted	7.97	0.31	Consensus

All 60 indicators obtained a first-round score above 7.00 and were therefore retained for the second Delphi round. The highest first-round score was observed for C51 at 7.94, followed by C21 and C27 at 7.89, C15 at 7.86, and C24 at 7.84. The lowest first-round scores were 7.05 for C30 and 7.06 for C12, both of which remained above the acceptance threshold.

No indicator was eliminated in the second round. The reported absolute differences between rounds ranged from 0.00 to 0.72 and were therefore below the stopping threshold of 0.80. The largest change was observed for C27, with an absolute difference of 0.72, followed by C49 and C58, each with an absolute difference of 0.60, and C15, C20, and C55, each with an absolute difference of 0.50. Because all second-round indicators remained accepted and the round-to-round differences satisfied the consensus criterion, the fuzzy Delphi process was terminated after the second round.

Exploratory factor analysis was performed to statistically evaluate whether the categorization obtained from the thematic analysis could be reproduced in the quantitative data. Principal component analysis was used to extract the underlying factors. Sampling adequacy was evaluated with the Kaiser–Meyer–Olkin index, and the suitability of the correlation matrix was evaluated with Bartlett's test of sphericity.

The KMO value was 0.914, substantially exceeding the minimum recommended value of 0.70 and indicating excellent sampling adequacy. Bartlett's test was statistically significant, $\chi^2(1770) = 11{,}876.408$, $p < .001$, demonstrating that the correlation matrix was not an identity matrix and that the correlations among the indicators were sufficiently strong for factor extraction.

Communality values after extraction ranged from 0.415 to 0.735. Most indicators had extraction communalities above 0.50. The indicators with values below 0.50 were Q20 = 0.450, Q24 = 0.485, Q25 = 0.432, Q26 = 0.459, Q27 = 0.485, Q30 = 0.472, Q33 = 0.472, Q36 = 0.441, Q37 = 0.426, Q40 = 0.480, and Q41 = 0.415. Although these items had comparatively weaker communalities, they were retained because of their qualitative relevance, fuzzy Delphi approval, and acceptable performance in the subsequent confirmatory measurement model. The highest communalities were observed for Q52 = 0.735, Q11 = 0.716, Q44 = 0.703, Q59 = 0.673, Q8 = 0.662, Q43 = 0.656, and Q32 = 0.649.

Eight components had initial eigenvalues greater than 1.00 and were retained. Together, the eight factors explained 57.585% of the total variance. The scree plot also displayed eight points above the eigenvalue threshold, supporting the retention of an eight-factor solution.

Table 4

Sampling adequacy, Bartlett's test, eigenvalues, and variance explained by the extracted factors

Analysis	Factor	Initial eigenvalue	Initial variance explained (%)	Initial cumulative variance (%)	Rotated eigenvalue	Rotated variance explained (%)	Rotated cumulative variance (%)
KMO index	—	0.914	—	—	—	—	—
Bartlett's test	—	χ^2 = 11,876.408	df = 1770	p < .001	—	—	—
Factor extraction	1	22.050	36.750	36.750	6.956	11.594	11.594
	2	2.699	4.498	41.248	4.833	8.055	19.649
	3	2.265	3.776	45.023	4.570	7.617	27.266
	4	1.742	2.903	47.927	4.464	7.441	34.707
	5	1.570	2.616	50.543	4.191	6.985	41.692
	6	1.468	2.446	52.989	3.512	5.853	47.545
	7	1.414	2.357	55.346	3.359	5.599	53.144
	8	1.343	2.239	57.585	2.665	4.442	57.585

Before rotation, the first factor accounted for 36.750% of the total variance. After rotation, the explained variance was distributed more evenly across the eight factors. The first rotated factor accounted for 11.594% of the variance, while the eighth factor accounted for 4.442%. The other rotated factors explained between 5.599% and 8.055% of the total variance.

Inspection of the rotated component matrix indicated that the items could be interpreted consistently with the eight dimensions produced by the qualitative analysis. Each item was assigned to the factor on which it had the highest substantively meaningful loading. Although several items exhibited secondary loadings, their dominant loadings and conceptual meanings supported the intended classification.

Table 5

Eight factors identified through exploratory factor analysis

Factor	Interpreted dimension	Number of indicators	Approximate range of primary loadings	Principal content represented
1	Perceived supervisory support	21	0.431–0.675	Supervisor accessibility, guidance, listening, trust, authority, organizational support, participation in decision-making, and influence on senior management
2	Perceived support for innovation	7	0.452–0.617	Encouragement of idea generation, experimentation, new solutions, and discussion of innovative methods
3	Team empowerment	10	0.429–0.620	Team discretion, autonomy, competence, influence, growth opportunity, and access to strategic information
4	Psychological safety	6	0.422–0.589	Ability to discuss difficult issues, seek assistance, take interpersonal risks, and express individual differences
5	Social cohesion	5	0.426–0.660	Team belonging, interpersonal comfort, happiness, social interaction, and enjoyment of collective activities
6	Information sharing	5	0.514–0.681	Exchange of work documents, experience, decision information, planning information, and development-related knowledge
7	Internal communication	3	0.569–0.677	Formal and informal exchange of documents, experience, and decision-relevant information within the team
8	External communication	3	0.561–0.697	Intergroup information flow, external coordination, and acquisition of technical knowledge from outside the organization

Within perceived supervisory support, particularly strong exploratory loadings were observed for participation of the supervisor in major organizational decisions, organizational support for supervisory decisions, supervisory authority to make important decisions, encouragement to approach the supervisor during unexpected problems, and organizational respect for the supervisor. In the innovation dimension, supervisory encouragement of innovative ideas and idea development displayed the strongest loadings. Team empowerment was principally characterized by freedom to decide team activities, the team's ability to empower members, implementation of team-selected activities, and provision of information concerning how organizational objectives could be achieved.

Psychological safety was most strongly represented by the absence of interpersonal undermining, the ability to request assistance, and the opportunity to raise difficult issues. Social cohesion was primarily reflected in comfort among team members, spending time together, enjoyment of social activities, and happiness associated with membership. Information sharing was characterized by sharing reports and documents, planning information, experience, and knowledge across teams. Internal communication was represented by the exchange of formal reports, experience, and key decision information, whereas external communication was represented by coordination with external groups, collection of technical ideas from outside the organization, and communication of team activities to other organizational groups.

The questionnaire collected demographic information concerning sex, age group, length of service, and educational level. Analysis of variance and relevant group comparisons were conducted in SPSS to determine whether perceptions of the knowledge-worker performance-evaluation components differed across demographic categories.

The null hypotheses proposed that the mean responses were equal across male and female participants, across different age groups, across employees with different lengths of service, and across different educational levels. The alternative hypotheses proposed that at least one group mean differed. The null hypothesis was retained in all four analyses, and the corresponding alternative hypothesis was rejected. No statistically significant differences were therefore detected according to sex, age group, service history, or educational level. The supplementary comparisons also did not indicate significant differences according to job type or academic-field group.

These findings indicated a relatively homogeneous understanding of the performance-evaluation system among the participating knowledge workers, managers, and supervisors. The homogeneity may reflect the presence of a broadly communicated organizational evaluation system and a common understanding of its objectives and components. Most knowledge workers had backgrounds in mathematics, physics, engineering, and related technical disciplines, whereas approximately one-third of managers and supervisors had backgrounds in the humanities. Despite these educational differences, their evaluations of the identified performance components were statistically comparable.

Confirmatory factor analysis using partial least squares structural equation modeling was conducted to determine whether the eight-factor measurement structure identified in the qualitative and exploratory phases was supported by the observed data. A hierarchical measurement model was tested in which the 60 observed indicators loaded on eight first-order constructs, and the eight first-order constructs loaded on the second-order construct of knowledge-worker performance evaluation.

All first-order factor loadings exceeded 0.40. The lowest first-order loading was 0.519 for Q14, while the highest was 0.906 for Q56. All associated t statistics exceeded 2.58 and all p values were below .001. The relationship between every observed item and its intended latent construct was therefore supported at the 99% confidence level.

The eight second-order loadings ranged from 0.730 to 0.867. Perceived support for innovation and social cohesion had the highest second-order loadings, both equal to 0.867. Psychological safety had a second-order loading of 0.834, information sharing had a loading of 0.815, perceived supervisory support had a loading of 0.803, team empowerment had a loading of 0.787, internal communication had a loading of 0.764, and external communication had a loading of 0.730. Every second-order loading was statistically significant at $p < .001$.

Table 6

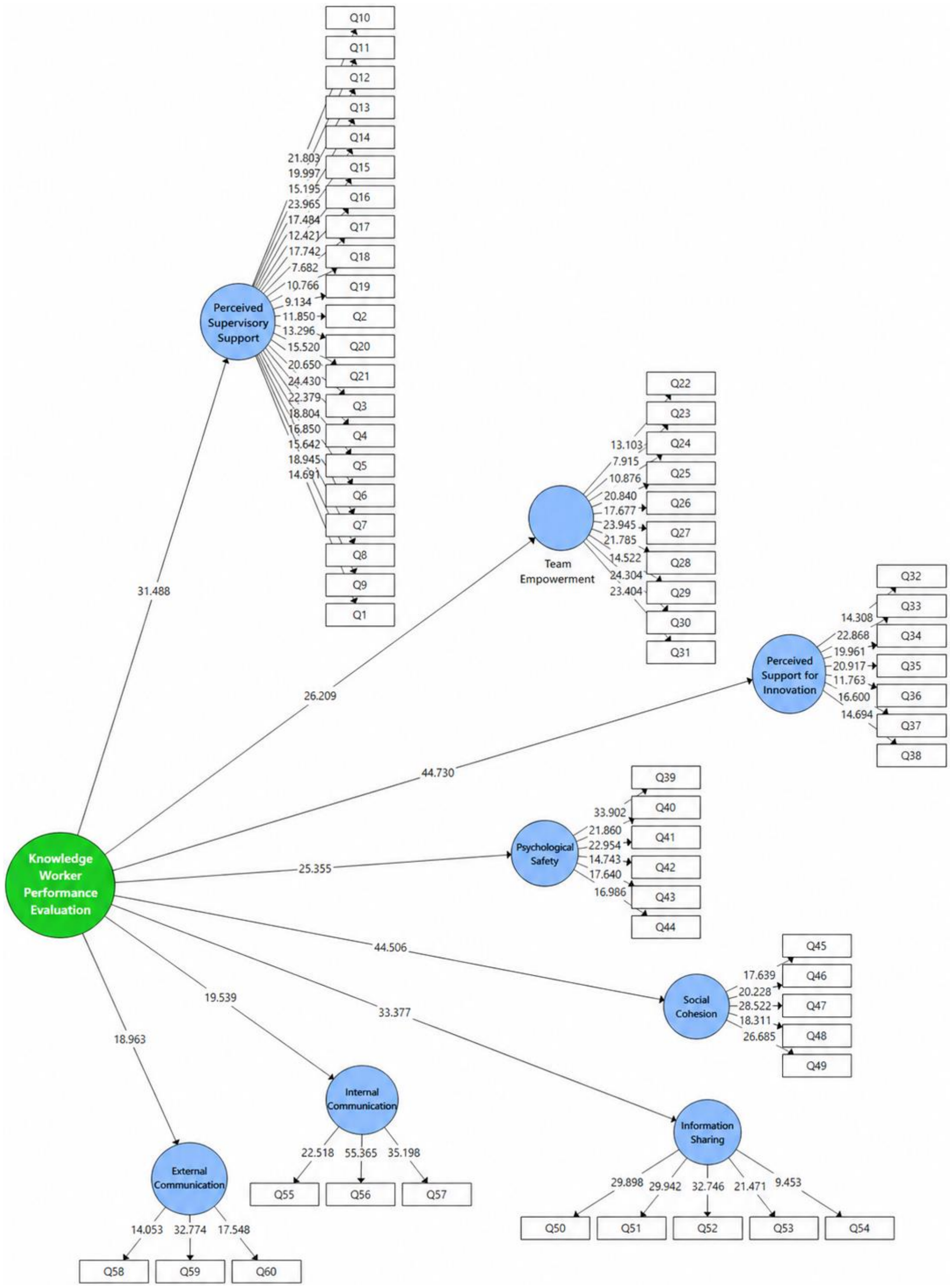
First- and second-order confirmatory factor-analysis results

First-order construct	Item	First-order loading	t statistic	p value	Second-order loading	Second-order t	Second-order p
Perceived supervisory support	Q1	0.643	14.697	< .001	0.803	31.488	< .001
	Q2	0.717	11.850	< .001	—	—	—
	Q3	0.752	20.650	< .001	—	—	—
	Q4	0.769	24.430	< .001	—	—	—
	Q5	0.788	22.379	< .001	—	—	—
	Q6	0.608	8.804	< .001	—	—	—
	Q7	0.767	16.850	< .001	—	—	—
	Q8	0.763	15.642	< .001	—	—	—
	Q9	0.758	18.945	< .001	—	—	—
	Q10	0.782	21.803	< .001	—	—	—
	Q11	0.745	19.997	< .001	—	—	—
	Q12	0.720	15.185	< .001	—	—	—
	Q13	0.793	23.965	< .001	—	—	—
	Q14	0.519	7.484	< .001	—	—	—
	Q15	0.693	12.421	< .001	—	—	—
	Q16	0.738	17.742	< .001	—	—	—

	Q17	0.579	7.682	< .001	—	—	—
	Q18	0.609	10.766	< .001	—	—	—
	Q19	0.602	9.134	< .001	—	—	—
	Q20	0.632	13.296	< .001	—	—	—
	Q21	0.624	15.520	< .001	—	—	—
Team empowerment	Q22	0.628	13.103	< .001	0.787	26.209	< .001
	Q23	0.655	7.915	< .001	—	—	—
	Q24	0.638	10.876	< .001	—	—	—
	Q25	0.771	20.840	< .001	—	—	—
	Q26	0.761	17.677	< .001	—	—	—
	Q27	0.771	23.945	< .001	—	—	—
	Q28	0.755	21.785	< .001	—	—	—
	Q29	0.701	14.522	< .001	—	—	—
	Q30	0.780	24.304	< .001	—	—	—
	Q31	0.742	23.404	< .001	—	—	—
Perceived support for innovation	Q32	0.737	14.308	< .001	0.867	44.730	< .001
	Q33	0.789	22.868	< .001	—	—	—
	Q34	0.797	19.961	< .001	—	—	—
	Q35	0.787	20.917	< .001	—	—	—
	Q36	0.745	11.763	< .001	—	—	—
	Q37	0.791	16.600	< .001	—	—	—
	Q38	0.738	14.694	< .001	—	—	—
Psychological safety	Q39	0.854	33.902	< .001	0.834	25.355	< .001
	Q40	0.802	21.860	< .001	—	—	—
	Q41	0.813	22.954	< .001	—	—	—
	Q42	0.770	14.743	< .001	—	—	—
	Q43	0.784	17.640	< .001	—	—	—
	Q44	0.777	16.986	< .001	—	—	—
Social cohesion	Q45	0.810	17.639	< .001	0.867	44.506	< .001
	Q46	0.800	20.228	< .001	—	—	—
	Q47	0.862	28.522	< .001	—	—	—
	Q48	0.829	18.311	< .001	—	—	—
	Q49	0.807	26.685	< .001	—	—	—
Information sharing	Q50	0.797	29.898	< .001	0.815	33.377	< .001
	Q51	0.830	29.942	< .001	—	—	—
	Q52	0.768	20.746	< .001	—	—	—
	Q53	0.780	21.471	< .001	—	—	—
	Q54	0.572	9.453	< .001	—	—	—
Internal communication	Q55	0.817	22.518	< .001	0.764	19.539	< .001
	Q56	0.906	55.365	< .001	—	—	—
	Q57	0.865	35.198	< .001	—	—	—
External communication	Q58	0.746	14.053	< .001	0.730	18.963	< .001
	Q59	0.851	32.774	< .001	—	—	—
	Q60	0.822	17.548	< .001	—	—	—

Figure 1

Final Model with Factor Loadings



The findings confirmed that the observed variables were sufficiently homogeneous within their respective constructs. The second-order results further demonstrated that all eight dimensions were meaningful manifestations of the broader knowledge-worker performance-evaluation construct.

Cronbach's alpha, rho_A, and composite reliability were calculated to evaluate internal consistency. All construct-level coefficients exceeded 0.70. Cronbach's alpha ranged from 0.733 for external communication to 0.947 for perceived supervisory support. Composite reliability ranged from 0.849 for external communication to 0.952 for perceived supervisory support. At the level of the overall knowledge-worker performance-evaluation construct, Cronbach's alpha was 0.917, rho_A was 0.920, and composite reliability was 0.928. These results supported the internal consistency and reliability of the measurement instrument.

Convergent validity was evaluated using the average variance extracted. All AVE values exceeded 0.50, ranging from 0.522 for team empowerment to 0.745 for internal communication. Each construct therefore explained more than 50% of the variance in its observed indicators.

Discriminant validity was initially evaluated with the Fornell–Larcker criterion. The square root of the AVE for every construct exceeded its largest correlation with another construct. The square roots of the AVE were 0.807 for external communication, 0.863 for internal communication, 0.822 for social cohesion, 0.800 for psychological safety, 0.755 for information sharing, 0.722 for team empowerment, 0.770 for perceived support for innovation, and 0.760 for perceived supervisory support.

The largest interconstruct correlations were 0.759 between social cohesion and perceived support for innovation, 0.758 between social cohesion and psychological safety, 0.753 between perceived supervisory support and perceived support for innovation, and 0.717 between team empowerment and perceived supervisory support. In each instance, the square root of the relevant AVE exceeded the corresponding correlation.

Cross-loading analysis provided further evidence of discriminant validity. For each of the 60 indicators, the loading on the intended latent variable exceeded its loadings on the other seven latent variables. For example, Q1–Q21 loaded most strongly on perceived supervisory support, Q22–Q31 loaded most strongly on team empowerment, Q32–Q38 loaded most strongly on perceived support for innovation, Q39–Q44 loaded most strongly on psychological safety, Q45–Q49 loaded most strongly on social cohesion, Q50–Q54 loaded most strongly on information sharing, Q55–Q57 loaded most strongly on internal communication, and Q58–Q60 loaded most strongly on external communication.

The heterotrait–monotrait ratio was also examined. All HTMT coefficients were below the critical value of 0.90. The coefficients ranged from 0.644 to 0.856. The highest value was 0.856 between social cohesion and perceived support for innovation, followed by 0.853 between social cohesion and psychological safety and 0.847 between team empowerment and perceived supervisory support. Thus, discriminant validity was supported by the Fornell–Larcker criterion, cross-loadings, and HTMT.

Table 7

Reliability, convergent validity, and discriminant-validity evidence

Construct	Cronbach's alpha	rho_A	Composite reliability	AVE	Square root of AVE	Largest correlation with another construct	Highest associated HTMT
External communication	0.733	0.746	0.849	0.652	0.807	0.622	0.817
Internal communication	0.828	0.833	0.898	0.745	0.863	0.688	0.817
Social cohesion	0.880	0.880	0.912	0.675	0.822	0.759	0.856
Psychological safety	0.888	0.892	0.914	0.641	0.800	0.758	0.853
Information sharing	0.809	0.832	0.867	0.570	0.755	0.710	0.826
Team empowerment	0.897	0.898	0.916	0.522	0.722	0.717	0.847
Perceived support for innovation	0.885	0.887	0.910	0.592	0.770	0.759	0.856
Perceived supervisory support	0.947	0.949	0.952	0.590	0.760	0.753	0.847
Overall knowledge-worker performance-evaluation construct	0.917	0.920	0.928	—	—	—	—

The standardized root mean square residual was 0.071 for both the saturated and estimated models. This value was below the conventional threshold of 0.10 and supported an acceptable fit between the empirical correlation matrix and the model-implied matrix. The squared Euclidean distance, d_ULS, was 0.688, and the geodesic distance, d_G, was 0.414. The chi-square value was 735.865, and the normed fit index was 0.800. The equality of the saturated and estimated model values indicated that the estimated structural specification did not produce meaningful deterioration in the measurement-model fit.

The final model was returned to participating experts for member checking and expert validation. A six-item, five-point Likert-type model-evaluation questionnaire was initially prepared to assess the correctness and practical appropriateness of the proposed framework. Following expert review of the validation instrument, the model was evaluated according to internal-validity criteria, external-validity criteria, and the acceptability of its overall form.

The internal-validation criteria were fit, comprehensibility, generalizability, and control. The external-validation criteria were philosophy and objectives, theoretical foundations, implementation stages, and evaluation and reengineering. The descriptive results showed that the modal response for all criteria was 4. The descriptive mean reported for each criterion was approximately 4, indicating initial expert approval. Standard deviations ranged from 0.374 for evaluation and reengineering to 0.615 for control, and variances ranged from 0.140 to 0.331. Minimum ratings were generally 3, except for theoretical foundations, for which the minimum was 2; the maximum rating was 5 for every component.

Because descriptive indices alone were not considered sufficient for decision-making, one-sample t tests were performed against a theoretical midpoint of 3. All nine tests were statistically significant. The empirical means ranged from 3.567 to 4.121 and were consistently above the theoretical midpoint.

Table 8

Partial least squares model fit and expert validation of the final model

Evidence type	Criterion or index	Empirical value/mean	Reference value	Mean difference	95% confidence interval	t	df	p	Decision
PLS model fit	SRMR, saturated model	0.071	< 0.10	—	—	—	—	—	Supported
PLS model fit	SRMR, estimated model	0.071	< 0.10	—	—	—	—	—	Supported
PLS model fit	d_ULS	0.688	—	—	—	—	—	—	Supported
PLS model fit	d_G	0.414	—	—	—	—	—	—	Supported
PLS model fit	Chi-square	735.865	—	—	—	—	—	—	Reported
PLS model fit	NFI	0.800	—	—	—	—	—	—	Acceptable
Internal validation	Fit	3.872	3.00	0.872	[0.6907, 1.1093]	8.121	11	< .001	Confirmed
Internal validation	Comprehensibility	3.976	3.00	0.976	[0.4327, 0.9673]	5.480	11	< .001	Confirmed
Internal validation	Generalizability	3.895	3.00	0.895	[0.5552, 1.0448]	6.839	11	< .001	Confirmed
Internal validation	Control	4.121	3.00	1.121	[0.6963, 1.3037]	6.892	11	< .001	Confirmed
External validation	Philosophy and objectives	3.567	3.00	0.567	[0.5119, 1.0881]	5.812	11	< .001	Confirmed
External validation	Theoretical foundations	3.637	3.00	0.637	[0.1636, 0.9364]	2.979	11	.008	Confirmed
External validation	Model implementation stages	3.987	3.00	0.987	[0.4925, 1.0075]	6.097	11	< .001	Confirmed
External validation	Evaluation and reengineering	3.792	3.00	0.792	[0.4511, 1.0489]	5.252	11	< .001	Confirmed
Overall model validation	Overall form of the model	3.784	3.00	0.784	[0.6907, 1.1093]	9.000	11	< .001	Confirmed

For the fit criterion, the calculated statistic was $t = 8.121$ and the empirical mean was 3.872. Because the statistic exceeded the critical value of 2.58 and the mean exceeded 3, the fit of the proposed model was confirmed at the 99% confidence level. Comprehensibility was also supported, with $M = 3.976$ and $t = 5.480$, $p < .001$. Generalizability obtained a mean of 3.895 and a statistically significant test statistic of 6.839, while control obtained the highest empirical mean, $M = 4.121$, with $t = 6.892$, $p < .001$.

The philosophy and objectives of the model were supported with $M = 3.567$ and $t = 5.812$, $p < .001$. Theoretical foundations obtained $M = 3.637$ and $t = 2.979$, $p = .008$. Although this was the smallest t statistic among the validation criteria, it remained statistically significant at the .01 level and exceeded the critical value of 2.58. The model implementation stages were supported with $M = 3.987$ and $t = 6.097$, $p < .001$. Evaluation and reengineering were supported with $M = 3.792$ and $t = 5.252$, $p < .001$. Finally, the overall form of the model obtained a mean of 3.784 and a t statistic of 9.000, $p < .001$, confirming the overall acceptability of the model at the 99% confidence level.

Collectively, the qualitative findings, inter-coder agreement, two-round fuzzy Delphi process, exploratory factor analysis, confirmatory factor analysis, reliability coefficients, convergent-validity evidence, three complementary assessments of discriminant validity, partial least squares model-fit indices, and expert-validation results supported the final eight-dimensional model. The proposed knowledge-worker performance-evaluation model for the Atomic Energy Organization of Iran therefore comprised perceived supervisory support, team empowerment, perceived support for innovation, psychological safety, social cohesion, information sharing, internal communication, and external communication, operationalized through 60 empirically retained indicators.

Discussion and Conclusion

The present study aimed to design and validate a multidimensional model for evaluating the performance of knowledge workers in the Atomic Energy Organization of Iran. The qualitative findings identified 60 indicators that were organized into eight dimensions: perceived supervisory support, team empowerment, perceived support for innovation, psychological safety, social cohesion, information sharing, internal communication, and external communication. The reliability of the qualitative categorization was supported by a Cohen's kappa coefficient of 0.758, indicating substantial agreement between the researcher and the independent coder. The fuzzy Delphi analysis further showed that all 60 indicators exceeded the acceptance threshold in the first round and remained in the model after the second round. The differences between the two rounds were below the predetermined stopping criterion, indicating expert consensus regarding the relevance of the identified indicators. These results suggest that the performance of knowledge workers in the studied organization is not adequately represented by individual productivity or technical output alone. Rather, it emerges from an integrated system of managerial support, team functioning, innovation, psychological conditions, and knowledge-related communication.

The quantitative results confirmed the structure identified in the qualitative phase. The Kaiser–Meyer–Olkin coefficient was 0.914, and Bartlett's test of sphericity was statistically significant, demonstrating that the correlation matrix was suitable for factor analysis and that the sample contained sufficient common variance for extracting underlying dimensions. Exploratory factor analysis retained eight factors with eigenvalues greater than one, collectively accounting for 57.585% of the variance in the 60 indicators. The alignment between the eight statistically extracted factors and the eight qualitative categories strengthens the credibility of the proposed model because the structure was supported independently through expert interviews and quantitative analysis. This correspondence also indicates that the identified dimensions are not merely conceptually plausible but reflect distinguishable patterns in the responses of knowledge workers.

Confirmatory factor analysis provided additional support for the proposed measurement model. All first-order factor loadings were greater than 0.40 and statistically significant at the 99% confidence level. The second-order loadings of the eight dimensions on the general construct of knowledge-worker performance evaluation ranged from 0.730 to 0.867. Perceived support for innovation and social cohesion had the strongest second-order loadings, followed by psychological safety, information sharing, perceived supervisory support, team empowerment, internal communication, and external communication. These results indicate that all eight dimensions contribute meaningfully to the overarching construct, while also showing that innovation support and social cohesion occupy particularly influential positions in the model. The overall measurement structure therefore reflects both task-related and relational aspects of knowledge work.

The centrality of perceived supervisory support is consistent with the nature of leadership in knowledge-intensive organizations. Knowledge workers require professional autonomy, but they also need supervisors who provide guidance, facilitate access to organizational resources, listen to emerging concerns, support experimentation, and represent specialized teams in higher-level decision-making. The present findings showed that perceived supervisory support included not only direct interpersonal assistance but also the extent to which the organization respected, empowered, and consulted supervisors. This broader conceptualization is consistent with leadership models for knowledge-based companies, in which leaders are expected to facilitate expertise, learning, participation, and innovation rather than merely exercise formal control [7]. The result also aligns with the argument that effective leadership of knowledge management improves organizational performance by directing knowledge activities toward strategic priorities and creating the conditions necessary for productive knowledge use [6]. Accordingly, supervisory support should be understood as a structural mechanism connecting individual expertise with organizational authority and resources.

The identification of team empowerment as a major dimension is also theoretically meaningful. Complex scientific and technical work is rarely performed entirely by isolated individuals. It depends on the integration of multiple forms of expertise, collective problem solving, and the capacity of teams to adjust their methods to changing circumstances. The indicators retained in this dimension reflected autonomy in selecting work methods, confidence in team competence, belief in collective influence, developmental opportunity, and access to organizational information. These findings support the view that knowledge governance can improve innovative employee performance when employees and teams possess suitable authority, resources, and knowledge structures [8]. They are also consistent with evidence that knowledge management, leadership, and organizational culture interact in shaping employee performance [5]. Team empowerment therefore appears to be an essential bridge between individual knowledge and coordinated organizational performance.

Perceived support for innovation emerged as one of the strongest dimensions in the second-order model. This dimension included supervisory encouragement of new ideas, experimentation with alternative work methods, replacement of outdated solutions, attention to improvement opportunities, and willingness to discuss new approaches. Its strong loading reflects the importance of innovation in evaluating knowledge workers, particularly in organizations whose work depends on scientific and technical advancement. Previous research has shown that knowledge-management processes can strengthen innovation capability and subsequently improve organizational performance [9]. Similarly, effective management of tacit knowledge contributes to innovation and performance when supported by trust and task efficiency [4]. The present results therefore reinforce the conclusion that the value of knowledge workers is expressed not only through the accurate execution of existing procedures but also through their capacity to generate, communicate, and apply new ideas.

The findings concerning psychological safety show that knowledge-worker performance is partly dependent on whether employees can raise difficult issues, request assistance, acknowledge uncertainty, take interpersonal risks, and express different viewpoints without fear of humiliation or retaliation. In highly specialized environments, employees may possess critical knowledge that is not immediately visible to supervisors or colleagues. If the organizational climate discourages speaking up, errors and weaknesses may remain hidden. Psychological safety is therefore not simply an indicator of employee comfort; it is a condition for organizational learning, technical reliability, and responsible decision-making. This interpretation is compatible with research showing that work environment and knowledge-management conditions influence employee performance jointly rather than independently [1]. It is also consistent with evidence that affective trust facilitates the transfer and effective use of tacit knowledge [4]. The strong factor loading obtained for psychological safety suggests that a valid evaluation system should consider whether employees actively contribute to problem identification and learning rather than rewarding only the absence of visible mistakes.

Social cohesion had one of the highest second-order loadings, indicating that belonging, interpersonal comfort, happiness with team membership, and enjoyment of collective interaction are highly relevant to knowledge-worker performance. Knowledge exchange often depends on informal relationships and mutual trust in addition to formal reporting systems. Employees may be more willing to share incomplete ideas, experiential insights, or professional concerns when they feel accepted within the team. The importance of social cohesion is consistent with studies emphasizing the role of organizational culture in employee performance [5]. It also complements findings that corporate social responsibility and knowledge-management infrastructure contribute to sustainable performance through organizational ambidexterity, because collaboration and cohesion can facilitate simultaneous exploitation of established capabilities and exploration of new opportunities [2]. Nevertheless, the value of cohesion should not be interpreted as a preference for conformity. Its organizational benefit arises when belonging supports open communication, constructive disagreement, and the combination of diverse expertise.

Information sharing, internal communication, and external communication formed three related but empirically distinguishable dimensions. Information sharing concerned the exchange of work documents, professional experience, planning information, and decision-relevant knowledge. Internal communication referred to the circulation of formal and informal information among organizational members, while external communication involved coordination with outside groups and acquisition of technical ideas beyond the organization. The separation of these dimensions is important because knowledge may be available within an organization but still remain inaccessible across team or departmental boundaries. The results align with research indicating that knowledge sharing and knowledge retrieval mediate the relationship between artificial intelligence and employee performance [10]. They also correspond with big-data-driven approaches in which organizational performance depends on collecting, integrating, and applying information across complex systems [3]. In the context of the Atomic Energy Organization of Iran, such communication processes are especially significant because specialized work requires coordination among technical units while also depending on access to external scientific and technological developments.

The reliability and validity findings provide strong evidence for the psychometric quality of the proposed model. Cronbach's alpha values ranged from 0.733 to 0.947, while composite reliability values ranged from 0.849 to 0.952. The overall construct also demonstrated high internal consistency. All average variance extracted values exceeded 0.50, confirming convergent validity. The Fornell–Larcker criterion, cross-loadings, and HTMT values below 0.90 provided complementary evidence of discriminant validity. These findings show that the eight dimensions were internally coherent while remaining sufficiently distinct from one another. The results are consistent with the broader movement toward multidimensional performance-evaluation models that combine several interdependent criteria rather than relying on a single productivity measure. Hybrid models using techniques such as DEMATEL, ELECTRE, and intuitionistic fuzzy analysis similarly recognize that organizational performance consists of related yet distinguishable dimensions that must be evaluated systematically [16].

The model-fit results further supported the adequacy of the proposed structure. The SRMR value of 0.071 for both the saturated and estimated models was below the acceptable threshold, indicating limited discrepancy between the observed and model-implied correlation matrices. The NFI value of 0.800 also indicated an acceptable fit for the partial least squares model. Expert validation confirmed both the internal and external credibility of the model. Fit, comprehensibility, generalizability, control, philosophy and objectives, theoretical foundations, implementation stages, evaluation and reengineering, and the overall form of the model all obtained means above the theoretical midpoint and statistically significant one-sample t values. These findings indicate that the model was not only statistically defensible but was also viewed by specialists as understandable, applicable, controllable, and theoretically justified.

The absence of statistically significant differences according to sex, age, length of service, educational level, job type, and academic-field group suggests that the identified dimensions were understood relatively consistently across the participating groups. This homogeneity may indicate that the proposed indicators represent broadly shared features of knowledge work rather than concerns limited to a particular demographic or occupational subgroup. It may also reflect common exposure to the organization's performance-evaluation system and shared awareness of its requirements. This result is relevant to public-sector performance evaluation because models must often be applied across heterogeneous groups while maintaining fairness and interpretive consistency. Studies of performance-evaluation systems in Iranian government organizations similarly emphasize the importance of coherent drivers, institutional alignment, and transparent criteria [13]. Research on performance evaluation and accountability in specialized Iranian public organizations likewise supports the need for systems that are professionally relevant while remaining understandable across organizational levels [14].

The findings also have implications for the increasing digitization of human resource management. Smart human resource governance and artificial intelligence can strengthen performance monitoring, data analysis, and decision support, but digital systems require valid conceptual models to determine what should be measured. Without such models, technology may privilege easily quantifiable indicators while overlooking collaboration, psychological safety, tacit knowledge, and innovation. This concern is consistent with research on the applications and challenges of artificial intelligence in human resource processes [11] and with models of smart human resource governance in Iran's public sector [12]. The present model can therefore provide a conceptual foundation for future digital evaluation systems by defining a broader set of meaningful performance dimensions.

Finally, the proposed model is compatible with a sustainable approach to human resource performance evaluation. Knowledge-intensive organizations must preserve expertise, support learning, develop future capability, and avoid evaluation practices that sacrifice long-term knowledge resources for short-term numerical targets. Sustainable performance-evaluation models emphasize the integration of organizational outcomes with employee development and continuity [15]. The present model contributes to this perspective by including empowerment, innovation support, knowledge exchange, communication, and relational conditions alongside direct supervisory factors. It therefore evaluates not only what knowledge workers produce but also how they contribute to the organization's capacity to learn, adapt, and maintain specialized performance over time.

This study was conducted within one highly specialized public organization, and the distinctive strategic, technical, regulatory, and security-related characteristics of that setting may limit the direct generalization of the findings to other organizations. The quantitative population and sample were also relatively small, although they represented a substantial proportion of the eligible knowledge workers. The study relied primarily on expert judgments and self-reported questionnaire data, which may be affected by social desirability, organizational caution, common-method variance, or differences between perceived and objectively observed performance. The cross-sectional design did not permit assessment of changes in the model across time or determination of whether the identified dimensions predict subsequent individual, team, or organizational outcomes. In addition, although the study retained all 60 indicators on qualitative and statistical grounds, some exploratory communalities were comparatively low, and the large number of indicators may increase the administrative burden of practical implementation.

Future studies should test the model in other knowledge-intensive public and private organizations, including research institutes, energy companies, universities, technology firms, hospitals, and engineering organizations, to examine its contextual stability and generalizability. Multi-group analyses could determine whether the factor structure remains invariant across occupational disciplines, managerial levels, organizational units, and different types of knowledge workers. Longitudinal research should investigate whether changes in supervisory support, empowerment, psychological safety, cohesion, and communication predict subsequent innovation, technical quality, learning, retention, and organizational reliability. Researchers could also compare the full 60-item instrument with a psychometrically developed short form and incorporate objective indicators, supervisor ratings, peer assessments, and digital work-process data. Further research should examine causal relationships among the eight dimensions and determine whether some dimensions mediate or moderate the influence of others on knowledge-worker outcomes.

The Atomic Energy Organization of Iran should implement the model as a developmental and strategic evaluation system rather than as a solely administrative scoring mechanism. Evaluators should receive standardized training in interpreting the eight dimensions, collecting evidence, reducing rating bias, and providing constructive feedback. Performance reviews should combine employee self-assessment, supervisor evaluation, peer input, and verifiable work evidence while protecting confidentiality and preventing punitive use of knowledge-sharing or problem-reporting behaviors. Managers should translate evaluation results into individualized development plans, team-learning initiatives, supervisory coaching, innovation-support mechanisms, and improvements in internal and external communication. The organization should periodically review indicator relevance, monitor the fairness of scores across units and demographic groups, simplify the instrument where necessary, and integrate the validated dimensions into digital human resource systems without reducing complex knowledge contributions to automated numerical surveillance.

Acknowledgments

We would like to express our appreciation and gratitude to all those who cooperated in carrying out this study.

Authors' Contributions

All authors equally contributed to this study.

Declaration of Interest

The authors of this article declared no conflict of interest.

Ethical Considerations

The study protocol adhered to the principles outlined in the Helsinki Declaration, which provides guidelines for ethical research involving human participants. Written consent was obtained from all participants in the study.

Transparency of Data

In accordance with the principles of transparency and open research, we declare that all data and materials used in this study are available upon request.

Funding

This research was carried out independently with personal funding and without the financial support of any governmental or private institution or organization.

References

[1] C. H. I. Xinjia and W. Thongsiri, "The Influence of Work Environment and Knowledge Management on Employee Performance: A Case of Sunshine Asset Management CO., LTD," *Procedia of Multidisciplinary Research,* vol. 3, no. 4, p. 5, 2025. [Online]. Available: https://so09.tci-thaijo.org/index.php/PMR/article/view/6252.

[2] A. Pazhoohan, P. Akbari, and M. A. Nouri, "Analyzing the Effect of Corporate Social Responsibility and Knowledge Management Infrastructure on Sustainable Performance Considering the Mediating Role of Organizational Ambidexterity," *Quarterly Journal of New Approaches in Management and Marketing,* vol. 5, no. 1, 2026.

[3] D. Gao, "A Big Data-Driven Knowledge Framework for Supply Chain Risk Prediction and Performance Optimization in Multinational Corporations," *International Journal of Knowledge Management,* vol. 22, no. 1, pp. 1-25, 2026, doi: 10.4018/ijkm.402393.

[4] J. Zhang, U. Zia, M. U. Shehzad, and Sherani, "Tacit knowledge management process, product innovation and organizational performance: exploring the role of affective trust and task efficiency," *Business Process Management Journal,* vol. 31, no. 1, pp. 267-297, 2025, doi: 10.1108/BPMJ-11-2023-0873.

[5] I. Muis, "The Effects of Knowledge Management, Digital Leadership and Organizational Culture on Employee Performance," *Sjam,* vol. 3, no. 1, pp. 1-16, 2025, doi: 10.38035/sjam.v3i1.411.

[6] S. Nah and G. D. Saxton, "How effective leadership of knowledge management impacts organizational performance," *Business Information Review,* vol. 36, pp. 30-38, 2025, doi: 10.1177/0266382119829643.

[7] S. Salem and H. Teymouri, "Designing a Leadership Model for Knowledge-Based Companies," *Strategic Management Studies,* vol. 15, no. 58, pp. 231-249, 2024.

[8] X. Ye, N. Wang, and Y. Han, "Knowledge Governance and Innovative Employee Performance," *Management Decision,* 2025, doi: 10.1108/md-05-2024-1015.

[9] A. Zahrawi, S. Khashashneh, and A. Althunibat, "The impact of customer knowledge management on innovation capability and business performance: Evidence from Jordanian financial services," *International Journal of Innovative Research and Scientific Studies,* 2025, doi: 10.53894/ijirss.v8i3.7401.

[10] S. Andayesh and Z. Kianzad, *Examining the Relationship between Artificial Intelligence and Employee Performance: The Mediating Role of Knowledge Sharing, Knowledge Retrieval, and Meaning Systems*. Allameh Tabataba'i University Press, 2025.

[11] R. Abbasi and M. Esmaeili, "Artificial Intelligence and Digital Human Resource Processes: Applications and Challenges," *Human Resource Studies,* vol. 14, no. 1, pp. 116-140, 2024.

[12] A. Mirzaei, M. Naghavi, R. Vaezi, and H. Aslipour, "Identifying the Dimensions of Smart Human Resource Governance in Iran's Public Sector: A Case Study of the National Iranian Oil Company," *Human Resource Studies,* vol. 15, no. 1, pp. 584-606, 2025.

[13] J. Salehi Sedghiani, H. Hamzavi, and F. Asadi, "Identifying and Ranking the Drivers of Improving the Human Capital Performance Evaluation System in Iranian Government Organizations," *New Research in Performance Evaluation,* vol. 4, no. 2, pp. 156-176, 2025.

[14] R. Sotudeh, "Explaining the Model of Performance Evaluation and Accountability of Iran's Audit Organization from the Experts' Point of View," (in English), *Audit Knowledge,* vol. 25, no. 98, pp. 348-375, 2025. [Online]. Available: http://danesh.dmk.ir/article-1-3159-fa.html.

[15] H. Mokhtarpour Asl and M. J. Kameli, "A Sustainable Human Resource Performance Evaluation Model," *Development Management Process,* vol. 35, no. 1, pp. 85-119, 2022.

[16] M. Daneshvar, F. Karimi Jafari, and Z. Moradi Golriz, "Presenting a Performance Evaluation Model for Organizational Units Using a Combined DEMATEL-ELECTRE-Intuitionistic Fuzzy Approach," *Decision Making and Operations Research,* vol. 9, no. 3, pp. 584-606, 2024.