Journal of Engineering, Project, and Production Management
2026, 16(2), 2025-165

Optimizing Enterprise Office Efficiency through Multimodal Employee Behavior Analysis

Tong You[1] and Wenjia Wang[2]

[1] Associate Professor, College of Business, Guilin Institute of Information Technology, Guilin 541004, Guangxi, China
[2] Associate Professor, College of Business, Guilin Institute of Information Technology, Guilin 541004, Guangxi, China, Email: aahpp42@163.com (corresponding author).

Project Management
Received August 16, 2025; revised October 3, 2025; accepted October 6, 2025
Available online March 7, 2026

Abstract: This paper proposes a model for optimizing enterprise office efficiency through multimodal employee behavior analysis. The model integrates computer vision, sensor data, workflow data, and text information in a four step process: data acquisition, feature extraction, behavior identification, and optimization strategy formulation, to identify task, collaborative, non-work, and organizational behaviors. It dynamically optimizes the office environment, workflow, and human resource allocation. Key innovations include: first, the "in-depth fusion of four-modal data" via the "temporal-semantic dual alignment" technique, to address data heterogeneity. Second, the "behavior-efficiency dynamic mapping" mechanism, which quantifies negative correlation (r=-0.72) between non-work behaviors and task efficiency, provides a quantitative basis for optimization. Experiments involving 50 technology company employees demonstrate the model's superiority, achieving an office efficiency improvement rate of 0.85, which surpasses both traditional methods (0.65) and common multimodal methods (0.72). The model also attained a behavior identification accuracy rate of 0.90, outperforming comparison methods (0.80 and 0.85). The model also excels in optimizing the office environment, workflow, and human resources, with improvements of up to 30%. Core contributions include: 1) a multimodal data fusion framework that increases behavior recognition accuracy to 90%; 2) an integrated "three-in-one" optimization strategy for the environment, processes, and human resources, which increased employee satisfaction by 30% and reduced task time by 25%; and 3) empirical validation that demonstrates an 85% efficiency improvement rate, significantly outperforming traditional and standard multimodal methods. The study highlights the model's practicality and potential for widespread adoption in enhancing office efficiency.

Keywords: Employee behavior analysis, office efficiency optimization, multimodal analysis, data acquisition.

DOI 10.32738/JEPPM-2025-165

1. Introduction

In modern enterprise management, office efficiency is a key indicator of a company's competitiveness (Silva et al., 2023). With the continuous expansion of enterprise scale and the increase of business complexity, traditional management methods often struggle to meet the demands of an efficient workplace (Shen et al., 2024). In recent years, with the rapid development of artificial intelligence, Internet of Things and big data technologies, multimodal data analysis provides new ideas and methods for office efficiency optimization (Lu et al., 2025). By integrating multiple types, multimodal data provides a more comprehensive view of employee behavior and the work environment, thus providing a more scientific basis for optimizing office efficiency (El-Tony and Choo, 2025). Multimodal employee behavior analysis can accurately identify behavioral patterns and work status, enabling the optimization of the office environment, workflow and human resource allocation to improve overall office efficiency (Xuanyuan et al., 2024). This is highly significant for enterprises seeking to enhance their competitiveness and reduce operating costs.

Research on office efficiency optimization typically focuses on three areas: the application of multimodal data analysis technology (Thebuwena et al., 2024), the construction of an employee behavior analysis model (Xu and Yang, 2024), and the development of optimization strategies (Dua et al., 2023). While multimodal data analysis has been widely applied in fields such as healthcare, transportation, and education (Torres et al., 2023), its use in business management has only recently gained attention (Wang et al., 2025). Locke and Osborne (2024) proposed a model capable of identifying employee work states and patterns. Su (2024) studied the application of multimodal data for optimizing the corporate office environment, reporting promising results. Mata and Tang (2023) put forward the concept of multimedia intelligence and explored the technological development of multimedia when it meets AI. Liu et al.(2025) investigated the application of a deep learning-based employee behavior recognition technology to improve the accuracy and efficiency of behavior recognition. Li and Yi (2025) proposed a method that identifies efficiency bottlenecks through the analysis of employee behavior data, thereby optimizing office processes and resource allocation; Nuhu et al. (2025) optimized the data processing and analysis process by combining a variety of modal data (text, images, table structures); Futri et al., (2023) automated the processing and analysis of multimodal data using natural language processing (NLP) and image recognition techniques. Although multimodal data analysis has made significant progress in office efficiency optimization, it still faces some challenges in practical applications (Acevedo-Duque, 2025): 1) data acquisition and preprocessing are difficult; 2) data fusion methods remain immature, making the effective integration of multiple modalities to improve the analytical accuracy is an urgent problem (Saghih and Ilani, 2023); 3) the dynamic nature of office environments requires models to analyze and respond accurately in real time, placing greater demands on computational performance (Liu et al., 2023).

This paper proposes an office efficiency optimization model based on multimodal employee behavior analysis that can comprehensively analyze multimodal data, accurately identify employee behavior patterns, and accordingly optimize the office environment, workflow, and human resource allocation. The main contributions of this paper include (1) a multimodal data analysis framework that integrates information from computer vision, sensors, workflow logs and other sources to comprehensively analyze employee behavioral patterns; (2) the identification of employee behavior patterns and the subsequent proposal of specific strategies for optimizing the office environment, workflow, and human resources. (3) validating the model for office efficiency enhancement through actual cases to verify the superiority of the model in office efficiency improvement, the experimental results show that the model proposed in this paper is superior to traditional and other multimodal methods.

To address the limitations of traditional methods and existing multimodal models, this study investigates three core research questions:(1) How to design a multimodal data fusion framework that integrates computer vision, sensor data, workflow data, and text information to improve the accuracy of employee behavior recognition? (2) What specific optimization strategies can be derived from multimodal behavior analysis, and how can their impact on office efficiency be quantified? (3) What key factors restrict the application of the proposed model, and how can these problems be solved to promote practical implementation?

The paper is structured as follows: Section 2 analyzes the relationship between employee behavior and office efficiency, defining behavioral classifications and characteristics detailing their impact on efficiency, and constructing a quantitative index system. Section 3 introduces the principles, key technologies and applications of multimodal data analysis, providing a theoretical foundation for the model. Section 4 details the construction of the office efficiency optimization model, outlining the process from data collection and behavior identification to strategy development and continuous improvement. Section 5 provides experimental validation through a case study, offering a quantitative comparison of the model's performance in office efficiency improvement, behavior identification accuracy and optimization effects. Finally, Section 6 summarizes the research results, pointing out the advantages, shortcomings and future research directions of the model. Together, these sections form a complete research cycle, moving from theoretical analysis and model construction to empirical validation.

2. Optimization Analysis of Employee Behavior and Office Efficiency

2.1. Employee Behavior Analysis

Employee behavior refers to the performance of all activities of employees in the workplace (Cai et al., 2023, Wisniewski, et al., 2024)), including work task-related activities (completing projects, attending meetings, collaborating and communicating) as well as non-work task-related activities (resting, socializing), and its connotation is shown in Fig. 1. These behaviors reflect employee's work attitude, work ability and work habits, and are one of the important bases for enterprise management. The content of employee behavior can be divided into the following four categories:

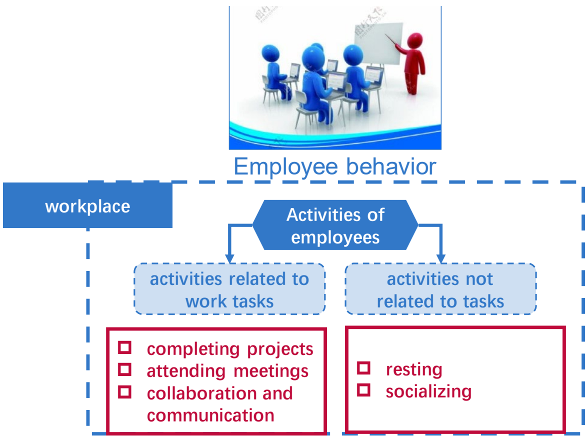

Fig. 1. Employee behavior connotation

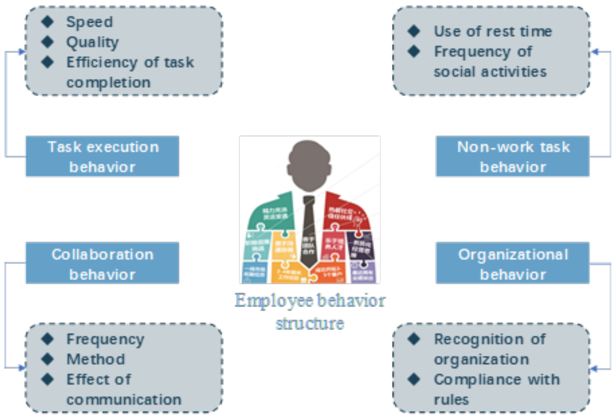

Fig. 2. Schematic diagram of employee behavior structure

Work Task Behavior: activities directly related to work goals, such as task execution, problem solving, and project promotion.

Collaborative Behaviors: behaviors that involve cooperation with colleagues or teams, such as communication, information sharing, and teamwork.

Non-work Task Behavior: activities not directly related to work tasks, such as resting, gossiping, and using social media.

Organizational Behavior: activities related to organizational goals and norms, such as compliance with company rules and regulations, participation in company cultural activities.

The structure of employee behavior compromises four components (Cai et al., 2023), as shown in figure 2: task execution behavior (speed, quality, and efficiency of task completion), collaborative behavior (frequency, mode, and effect of communication), non-work task behavior (use of rest time, frequency of social activities), and organizational behavior (identification with organizational culture, and adherence to rules and regulations).

Employee behavior exhibits several key characteristics (Fig. 3): 1) Regularity. Employee behavior often follows predictable patterns over time, such as the daily work schedule, the order of task completion. 2) Diversity. Behavioral patterns can vary significantly between employees and within the same employee in different situations. 3) Dynamism. Employee behavior changes dynamically with shifts in work tasks and environmental factors (Riemenschneider et al., 2023). 4) Recognizability. Through data analysis and behavioral monitoring, the main behavioral patterns of employees can be identified, providing a basis for management and optimization (Peng et al., 2025).

Fig. 3. Employee behavioral characteristics

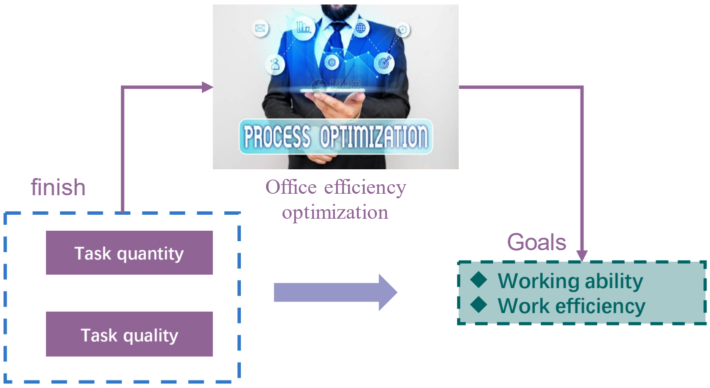

Fig. 4. Office efficiency optimization implications

2.2. Office Efficiency Optimization Analysis

Office efficiency refers to the quantity and quality of work tasks completed by employees within a unit of time, reflecting the work ability and work efficiency of employees, as shown in Fig. 4. The goal of optimizing office efficiency is to improve work efficiency, reduce time waste, and lower error rates through rational allocation of resources, optimization of workflow, and enhancement of employee competence (Xudong et al., 2023).

Office efficiency optimization encompasses four key areas (Xudong et al., 2023), shown in Fig. 5: task management, resource management, environment management and personnel management As shown, task management focuses on optimizing task allocation, task execution process and task monitoring mechanism; resource management involves the rational allocation of office equipment, software tools to improve resource utilization; environment management is aims to optimize physical conditions, such as to improve the lighting, temperature, noise and other conditions in order to improve the comfort of the staff and work efficiency; personnel management seeks to enhance employee skills and motivation through training, incentives thereby, improving overall efficiency.

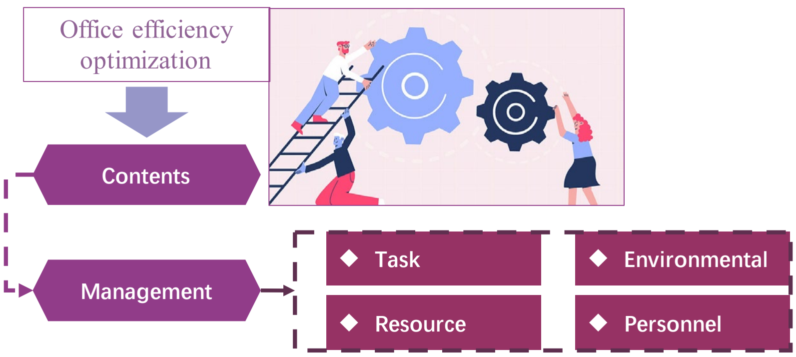

Fig. 5. Office efficiency optimization content

The structure of office efficiency consists of three components: task completion time, task quality, and resource utilization (Zhang et al., 2023) the optimization objective (Fig. 6). Task completion time measures the time required for employees to complete the task, typically expressed as the average completion time. Task quality assesses output quality, with indicators such as the error rate, rework rate, and resource utilization, which evaluate the efficiency of use, quantified by metrics such as equipment utilization rates and the frequency of use of software tools. The goals of office efficiency optimization are dynamic, optimizable and quantifiable.

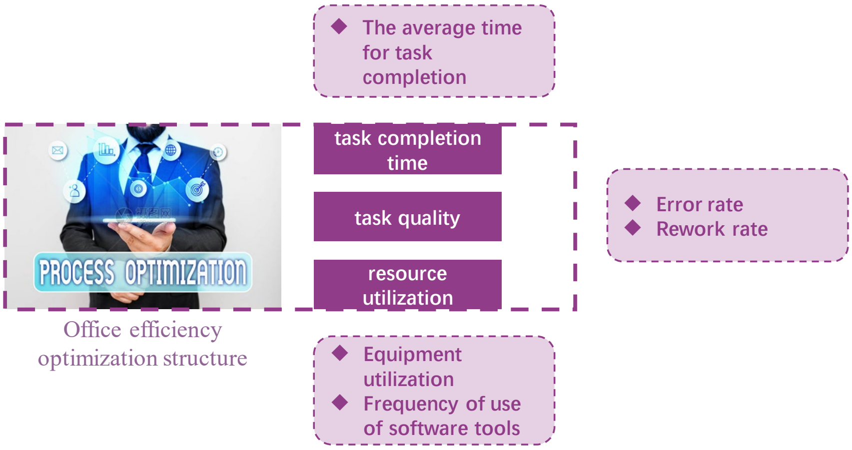

Fig. 6. Office efficiency optimization goal structure

2.3. Relationship Between Employee Behavior Analysis and Office Efficiency

2.3.1. Quantitative indicators of employee behavior

To better analyze the impact of employee behavior on office efficiency, employee behavior is divided into work-task behavior, collaborative behavior, non-work-task behavior, organizational behavior, and other categories, and each type is quantitatively analyzed, with the specific indicators shown in Table 1.

Table 1. Quantitative indicators of employee behavior

Gestion	Quantitative Indicators
Work task behavior	Task completion time
	Task quality (error rate, rework rate)
	Number of tasks completed
Collaborative behavior	Frequency of communication
	Duration of communication
	Number of collaborative projects
	Collaboration Satisfaction
Non-mission behavior	Percentage of time spent on non-work tasks
	Frequency of social media use
	Length of time spent gossiping
Organizational behavior	Rate of compliance with rules and regulations
	Frequency of participation in training and cultural activities
	Employee satisfaction

2.3.2. Impact of behavior on office efficiency

According to the classification of employee behavior, its specific impact on office efficiency is as follows:

Behavioral Impact of Work Tasks (Fig. 7) the shorter the task completion time, the higher the efficiency. By recording the start and end time of each task using the task management system, the average task completion time can be calculated (Park et al., 2024). The lower the error rate, the lower the rework rate, indicating higher quality work. The number of errors is recorded through the number of reworks of each task using the quality detection system. The more tasks an employee completes, the higher the efficiency, and each employee’s completed tasks are recorded in the task management system.

Collaboration Behavior Impact (Fig. 8) A higher communication frequency indicates closer team collaboration. This can be measured by recording the number of interactions between employees using instant messaging tools. Communication length suggests higher efficiency. The duration of each communication can be recorded via instant messaging tools; participation in more collaborative projects reflects stronger teamwork capability. The number of projects each employee participates in can be tracked in the project management system; finally, the higher the satisfaction levels, the better the collaboration effect. Employee satisfaction ratings can be collected through questionnaires (Bennouna et al., 2024).

Non-Work Task Behavior Impact: A lower percentage of time on non-work tasks indicates a higher degree of work concentration. This time can be recorded using a time tracking tool. A lower frequency of social media use suggests less work interference. The number of social media visits can be recorded using the network monitoring tool. Shorter idle chat durations suggest higher work concentration. This can be recorded using the audio monitoring tool.

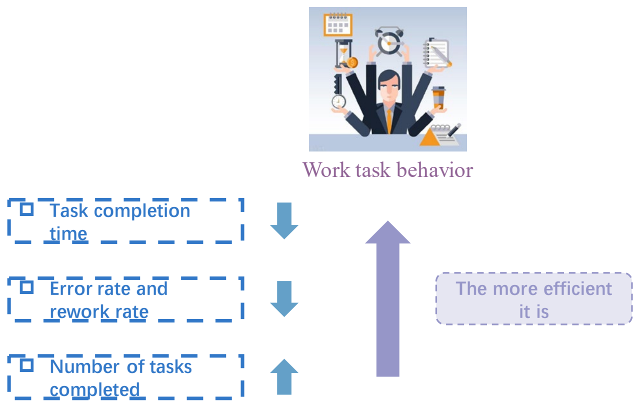

Fig. 7. Work task behavior

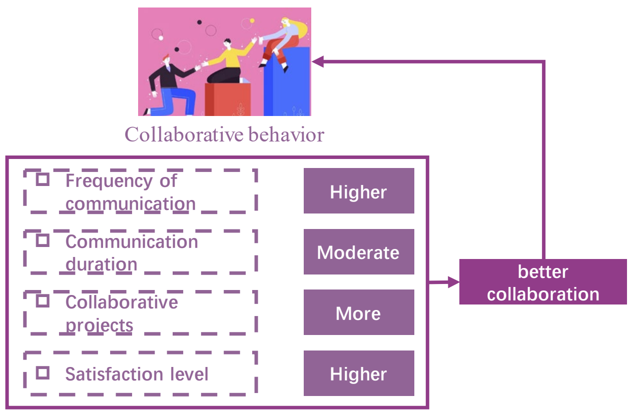

Fig. 8. Collaboration behavior

Organizational Behavior Impact: A higher compliance rate indicates more standardized employee behavior, which can be recorded through the attendance and violation record systems. A higher frequency of participation reflects stronger employee recognition of the organizational culture, which can be captured through the training and activity management system. Similarly, the higher the satisfaction level, the higher the motivation, and data can be collected through satisfaction questionnaires.

3. Multimodal Data Analysis Techniques

3.1. Principle and Structure

Multimodal data analysis technology comprehensively understands and processes complex information by fusing diverse types of data (text, image, audio, video, etc.) (Song et al., 2025). Its underlying principles are shown in Fig. 9. It simulates human multi-sensory perception and can acquire and analyze information from multiple perspectives.

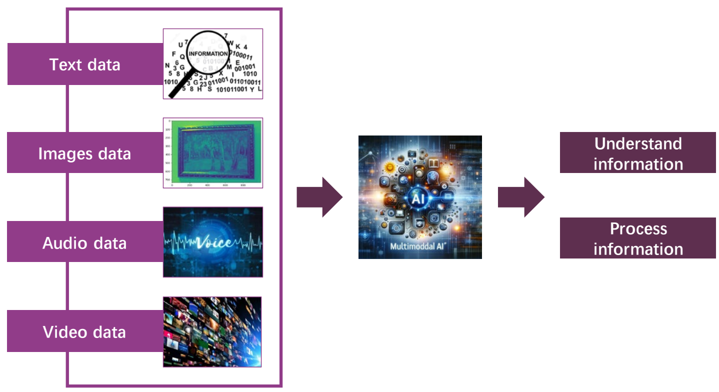

Fig. 9. Principle of multimodal data analysis

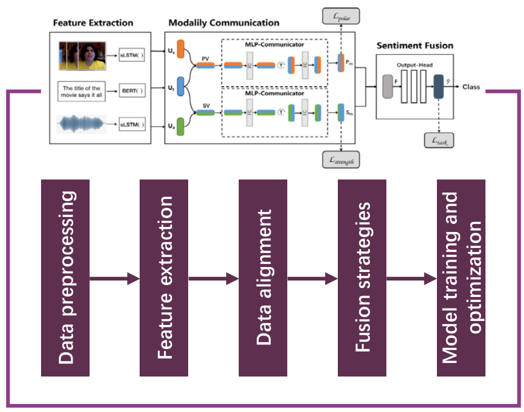

Fig. 10. Flow chart of multimodal data analysis

Flow chart of multimodal data analysis includes 5 core steps: 1) Data preprocessing (cleaning, standardization, enhancement). 2) Feature extraction (text embedding, image CNN features, audio MFCC). 3) Data alignment (temporal/semantic alignment). 4) Fusion strategy (early/late/hybrid fusion). 5) Model training (multi-task learning) and optimization, as shown in Fig.10.

The specific steps are as follows: 1) Data preprocessing. This involves data cleaning, standardization and enhancement steps to improve data quality. 2) Feature extraction. This step extracts useful information from different data modalities, such as text word embeddings and image convolutional features. 3) Data alignment. This process aligns data from different modalities temporally, spatially, or semantically to ensure effective combination. 4) Fusion strategy. This step determines how to combine features from different modalities using early, late and hybrid fusion. 5) Model training and optimization. The model is trained using methods such as multi-task learning and migration learning to improve its generalization.

3.2. Key Technologies

Data Alignment Techniques Data alignment is the basis of multimodal fusion, aiming to align data across different modalities temporally, spatially, or semantically, as shown in Fig. 11. Common alignment methods include temporal alignment, semantic alignment, spatial alignment, and so on.

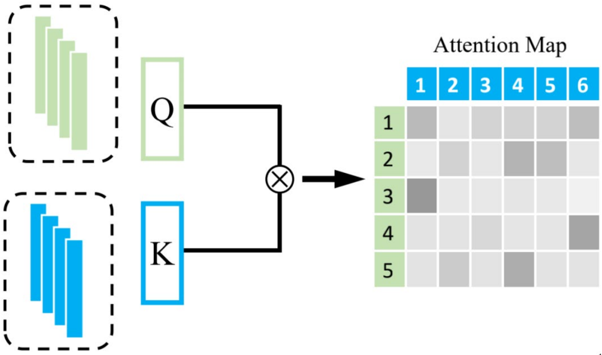

Fig. 11. Principle of data alignment

Feature Extraction Techniques: Feature extraction is a key step in multimodal data processing, extracting information useful for the task from the original data. Feature extraction methods are divided into text feature extraction, image feature extraction, and audio feature extraction as follows: 1) text feature extraction, including bag-of-words model, TF-IDF, word embeddings, and pre-trained language models. 2) image feature extraction, including Convolutional Neural Networks (CNNs) and pre-trained visual models. 3) audio feature extraction, including Mel Frequency Cepstral Coefficients (MFCC) and deep learning models.

Fusion Strategy: determines how to effectively combine features from different modalities. In early fusion, features from different modalities are spliced or weighted and summed immediately after feature extraction. In late fusion, the data of different modalities are processed separately to obtain their respective predictions before fusion. Hybrid fusion combines the advantages of early and late fusion and selects the appropriate timing and mode for each task.

Attention Mechanism: Attention mechanisms play an important role in multimodal modeling, dynamically focusing on information relevant to the current task to improve model performance and efficiency.

3.3. Applications and Features

Multimodal data analysis technology is widely used across many fields, including Visual Question and Answering (VQA), image description generation, sentiment analysis, medical image analysis, and autonomous driving. Based on the application of multimodal data analysis technology, it can be summarized that it primarily exhibits heterogeneity, complementarity, complexity, strong representational power, and adaptability, as shown in Fig. 12.

4. Method Construction

4.1. Model Construction Objectives

The aim of this section is to construct an office efficiency optimization model based on multimodal employee behavior analysis, which comprehensively analyzes employee behavioral patterns by integrating multimodal data (computer vision, sensor data, and workflow data) and optimizes the office environment, workflow, and human resource allocation accordingly. The specific objectives of the model construction include accurately identifying employee behavioral patterns, quantitatively assessing the impact of behavior on office efficiency, optimizing the office environment and workflow, and enhancing the effectiveness of human resource management, as shown in Fig. 13.

The methodology, outlined in the flowchart, consists of 4 steps: Step 1: Multimodal Data Collection and Preprocessing. This involves collecting visual, sensor, workflow, and text data, followed by preprocessing operations such as data cleaning and alignment. Step 2: Behavior Analysis and Feature Extraction. This step analyzes four behavior types, including task, collaborative, non-work and organizational behavior, while simultaneously performing feature extraction. Step 3: Optimization Strategy Formulation. Strategies are formulated across three dimensions: environment adjustment, process optimization and human resource training. Step 4: Evaluation and Continuous Improvement. The evaluation phase focuses on task time, quality and employee satisfaction, which subsequently informs a cycle of continuous improvement.

4.2. Methodological Steps

Combined with multimodal employee behavior, the data-driven office efficiency optimization model is constructed, and the process is shown in Fig. 14, with the following steps:

Step 1: Data Collection and Preprocessing

Collect multimodal data from multiple channels, including visual, sensor, workflow, and text data; in order to ensure the quality of the data, pre-process the multimodal data, and its key operations include data cleaning, data standardization, data annotation, and data alignment.

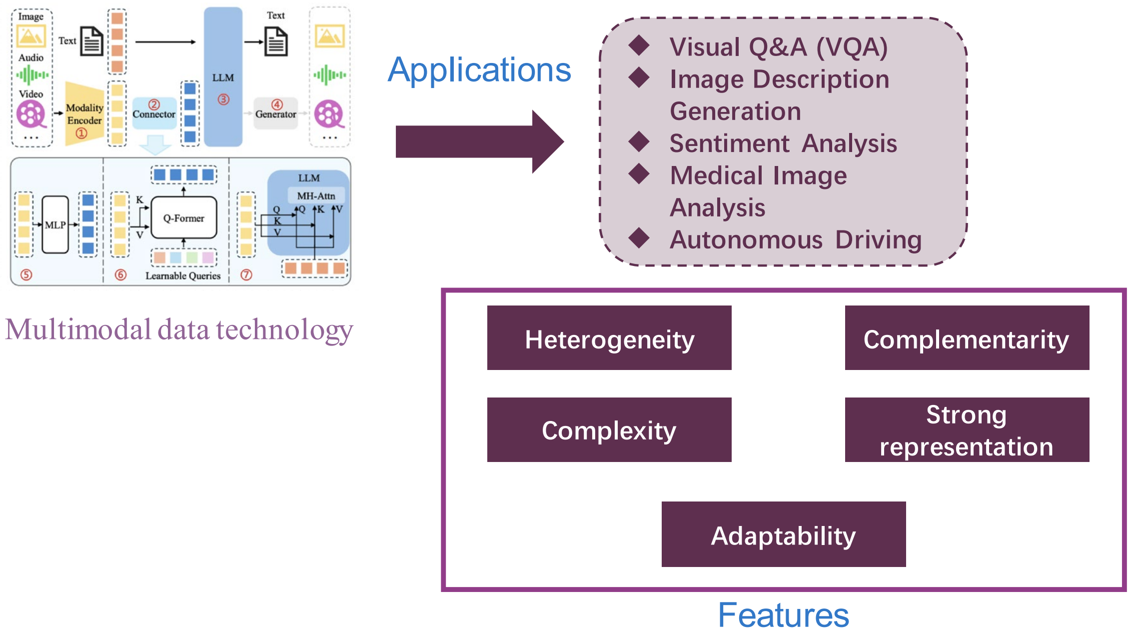

Fig. 12. Application and characteristics of multimodal data technology

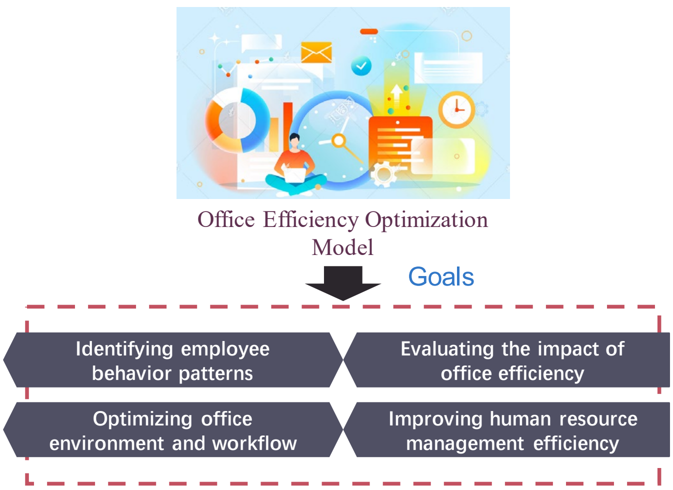

Fig. 13. Modeling objectives

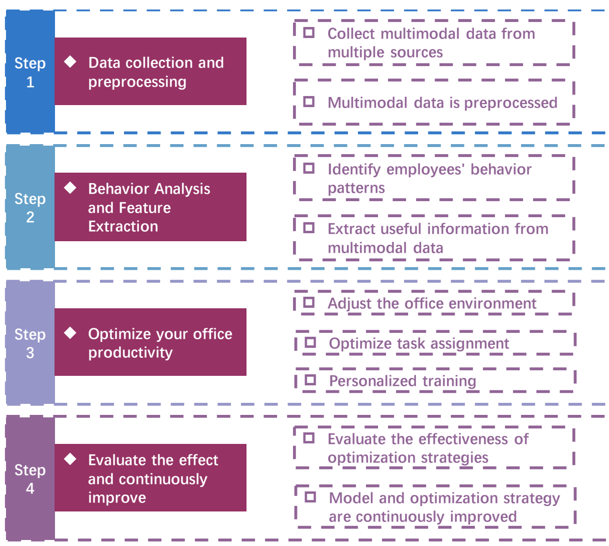

Fig. 14. Methodological steps

Step 2: Behavior Analysis and Feature Extraction

Through multimodal data analysis, identify behavioral patterns of employees, work task behavior, collaborative behavior, non-work task behavior, and organizational behavior, and extract useful information from the multimodal data as input features of the model (image features, sensor features, text features, time series features).

Step 3: Office Efficiency Optimization Strategy

The office environment is dynamically adjusted based on employee’s physiological states and environmental parameters, such as temperature, lighting, and noise control. Workflow data analysis identifies bottlenecks to optimize task allocation and collaboration processes. Based on employee’s behavioral characteristics and performance, personalized training and career development planning, supported by performance appraisal and incentive mechanisms, help improve employee motivation. Environmental optimization thresholds are determined using a dual approach combining industry standards and physiological data calibration: 1) The basic thresholds refer to the office environment requirements in the GB/T 50378-2019 Green Building Evaluation Standard. 2)These thresholds are then calibrated using two weeks of pre-experimental data: sensor-collected physiological indicators and task completion efficiency from 50 employees under different temperatures and varying light levels, with correlation analysis used to determine the optimal range. This resulted in the following final thresholds: temperature at23-25℃ (within this range, the employee’s heart rate fluctuation is ⩽5 beats per minute and task efficiency is the highest), the illumination threshold was set to 350-450lux (within this range, the employee's visual fatigue is the lowest and the error rate is ≤2%), and the noise threshold was set to ≤50dB (within this range, the employee's communication efficiency is improved by 10%).

Step 4: Evaluation and Continuous Improvement

Evaluate the effectiveness of the optimization strategy through quantitative indicators, such as task completion time, task quality, and employee satisfaction. Based on the evaluation results, continuously refine the model and optimization strategy to ensure office efficiency continues to improve.

5. Experimental Validation

5.1. Experimental Setup

The experiment was conducted on 50 employees in the R&D department of the city's technology company, which is responsible for software development and maintenance. The positions include senior software engineers (15), software engineers (25), junior software engineers (10), and test engineers (10). The experiment lasts one month, from November 1, 2024, to November 30, 2024. This study selected the R&D department of this technology company as the experimental subject mainly based on the following considerations. The department has a high degree of standardized business processes (clear software development and maintenance processes), strong traceability of employee behavior data (complete office software logs and instant messaging records). Its employee positions cover technical roles from junior to senior levels, with certain job representatives. The sampling method adopted is stratified sampling: based on the department's job structure (30% senior engineers, 50% general engineers, 20% test engineers), a corresponding number of employees were randomly selected from each job category to ensure that the sample is consistent with the actual job distribution in that department. The current results can be generalized to technology-based R&D departments with standardized business processes and knowledge-dependent work; however, limited by the single-company sample, when generalizing to other industries such as manufacturing and services, it is necessary to adjust behavioral characteristic indicators (e.g., the manufacturing industry needs to add the dimension of "equipment operation behavior", and the service industry needs to strengthen the analysis of "customer communication behavior"). Future research should include samples from 3 or more enterprises across different industries to further verify the model's cross-industry applicability.

The experimental environment includes equipment and systems. and the specific environment parameters are set as shown in Table 2.

Table 2. Environmental parameter settings

No.	Device/System	Function Settings
1	Camera	Used to capture images and videos of employee's work status (10 units)
2	Sensors	Temperature, Humidity, Light, and Physiological Sensors
3	Office Software Log System	Records task allocation, execution progress, and collaboration status
4	Instant Messaging Tools	Records communication content and frequency
5	Analysis and Visualization Software	Matlab 2021a

The specific computational cost of real-time multimodal analysis in this study is as follows: when using MATLAB 2021a (running on a server with an Intel Core i7-12700K CPU and 32GB RAM), multimodal data is collected once every 5 minutes. The computation time for a single data processing is approximately 12 seconds. The total daily computation time is approximately 96 minutes, with an average CPU utilization rate of 65% and a memory utilization rate of 40%, without obvious lag or delay. To support a team of 100 people, it is necessary to upgrade to a dual-CPU server, and the computation time per processing can be controlled to within 15 seconds, meeting real-time analysis requirements.

The experiment is divided into three phases: data collection, data analysis, and implementation of the optimization strategy, as shown in Table 3.

Table 3. Description of experimental approach phases

No.	Phase	Operation Description
1	Data Collection	Real-time collection of employee behavior data, including image sensors, workflow, and text data (frequency every 5 min)
2	Data Analysis	Analyze task completion time, task quality, communication frequency, collaboration satisfaction
3	Optimization Strategy Implementation	Adjust indoor temperature, light intensity, optimize task allocation and provide personalized training

To verify the effectiveness of the multimodal employee behavior analysis-driven office efficiency optimization model proposed in this paper, this section compares its experimental results with those of other common multimodal methods and traditional methods, and introduces the algorithms used, as shown in Table 4.

Table 4. Comparison algorithm introduction

No.	Algorithm	Description
1	Traditional Methods	Single-modal data analysis
2	Proposed Method	Combines image sensor workflow and text data analysis

The experimental metrics include the following quantitative measures, as described in Table 5.

Table 5. Description of experimental indicators

No.	Indicator	Measurement Parameters
1	Office Efficiency Improvement Rate	Reduction in task completion time and improvement in task quality
2	Behavior Recognition Accuracy	Accuracy in identifying employee behavior patterns
3	Environment Optimization Effect	Improvement in office environment satisfaction
4	Process Optimization Effect	Improvement in task completion efficiency and collaboration efficiency
5	Human Resource Optimization Effect	Improvement in employee satisfaction and training effectiveness

5.2. Analysis of Results

In order to verify and analyze the effectiveness of the method in this paper, this section carries out experiments from the comparison of multimodal algorithms, training results, and the effect of the optimization environment as follows:

5.2.1. Comparison of multimodal algorithms

To verify the superiority of the model proposed in this paper, this subsection compares its experimental results with those of other common multimodal methods and traditional methods, as shown in Fig. 15 and Table 6.

Figure 15 and Table 6 present the experimental results of "Comparison of Office Efficiency Improvement by Different Methods," from which the proposed multimodal employee behavior analysis-driven office efficiency optimization model shows the best performance in improving office efficiency. Compared with traditional methods (65%) and common multimodal methods (72%), this model achieves up to 85% efficiency. The benefits stem from enhanced recognition of employee behavior, improved rational task allocation, and flexible environmental adjustment capabilities. This significant improvement indicates that integrating multiple modal data, such as images, sensors, and textual information, is powerful for portraying employee behavioral states and optimizing office processes. Fig. 15 clearly specifies the three comparison methods: the traditional single-modal method, the common multimodal method, and the proposed method. Furthermore, it explains the calculation method for the office efficiency improvement rate: specifically, subtract the pre-optimization efficiency from the post-optimization efficiency, then divide the result by the pre-optimization efficiency. In addition, it points out the indicators used to evaluate the office efficiency improvement rate, including the task time reduction rate and the quality improvement rate. By analyzing Fig. 15, we see that the method used in this paper offers a significant improvement over other methods. This indicates that, in the actual office setting, by fusing visual images, sensor readings, and work logs, the model can comprehensively perceive employee's work status, enabling more intelligent scheduling and resource allocation. This approach excels in the accuracy of behavior recognition, providing superior data support for the subsequent optimization.

5.2.2. Comparison of training results

To verify the method's training effectiveness, the training process and results are shown in Fig. 16 and Fig. 17, and the training results are summarized in Table 7.

Figure 16 shows the trend in loss during model training, which reflects the model's ability to learn data patterns. Results show that the loss value gradually decreases and stabilizes as the number of training rounds increases. This indicates that the model successfully optimizes its parameters to learn the relationship between employee behavior and office efficiency. The initial rapid decrease in loss indicates that the model quickly captures significant behavioral features, while the loss tends to remain stable in the later stage. This means the model gradually converges, has enhanced generalization ability, and avoids overfitting.

Furthermore, the model demonstrates stable learning and convergence speed, with no significant oscillations or rebounds during training. This smooth descending curve indicates that the model structure is well-designed and that the feature selection is effective, while the multimodal data input plays a positive supporting role in the training process. The multimodal information provides rich contextual and behavioral cues, which makes it easier for the model to identify the deep connection between employee behavioral patterns and office efficiency.

Figure 17 and Table 7 present a comparison of different methods based on employee behavior recognition accuracy, a key metric for evaluating model performance. As shown in Fig. 17, the proposed method reaches 90% in behavior recognition accuracy, significantly outperforming the traditional method (80%) and the common multimodal method (85%). This result indicates that the multimodal fusion strategy, integrating images, sensor data, workflow data, and text information, enables more accurate identification of various types of employee behavioral patterns, including task execution, collaborative communication, and non-work behavior. The data shows an improvement of more than 5 percentage points in accuracy, representing a significant advancement. This high recognition rate is attributed to the model's fusion of multi-source information during feature extraction, which enhances its ability to perceive complex behavioral features. Second, the data alignment and fusion strategy is more reasonable, ensuring the consistency of modal information at both the temporal and semantic levels; and third, the model training process is stable (Fig. 16), granting the models strong generalization ability.

5.2.3. Comparison of optimization effects

In order to further assess the optimization effect of the office efficiency optimization model driven by multimodal employee behavior analysis, this section compares and analyzes the three methods in terms of environment optimization effect, process optimization effect, and human resource optimization effect, and obtains the results shown in Fig. 18 and Table 8.

Fig. 15. Comparison of office efficiency improvement by different methods

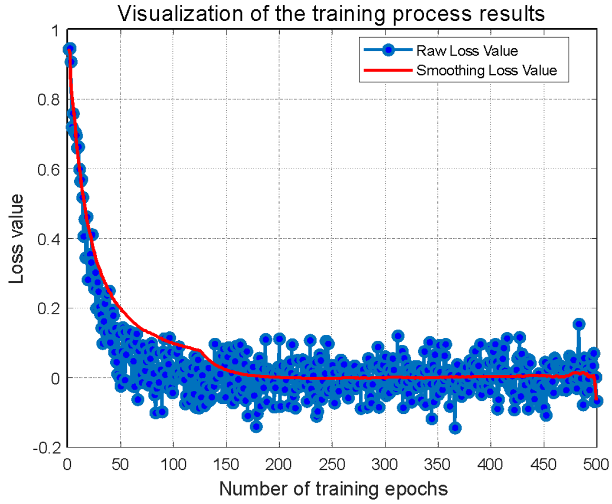

Fig. 16. Training process results

Table 6. Comparison of the effect of different algorithms

No.	Method	Office Efficiency Improvement Rate
1	Traditional Methods	0.65
2	Common Multimodal Methods	0.72
3	Proposed Method	0.85

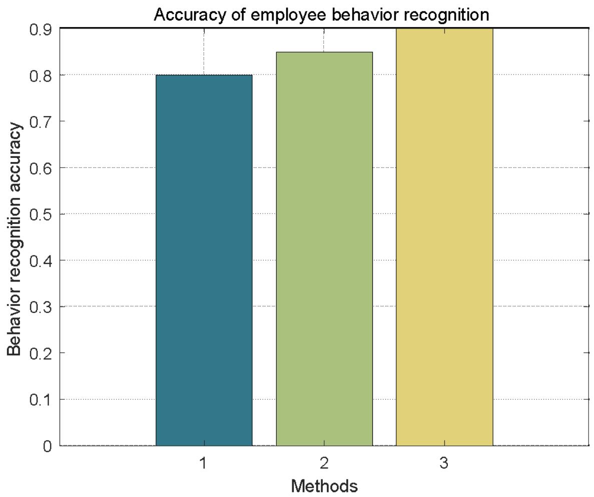

Fig. 17. Employee behavior recognition accuracy rate

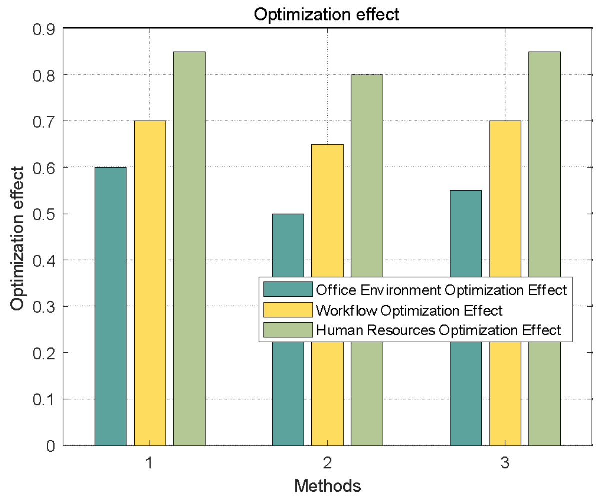

Fig. 18. Visual comparison of optimization effect of different methods

Table 7. Comparison of behavioral identification accuracy rate by different methods

No.	Method	Behavior Recognition Accuracy Rate
1	Traditional Methods	0.80
2	Common Multimodal Methods	0.85
3	Proposed Method	0.90

Figure 18 and Table 8 compare the effects of the three different methods on the three key dimensions of office efficiency: environment, process, and human resources optimization, providing a critical assessment of the model's practical value. In terms of environmental optimization, the proposed method achieves an optimization level of 85%, substantially higher than the 60% and 70% achieved by traditional and common multimodal methods. This indicates that the model has a significant advantage in sensing and regulating the office environment. By integrating temperature and humidity sensors, light sensors and employee's physiological state data, the model is able to realize adjustments needed, and automatically adjust light intensity, air temperature, which significantly improves employee comfort and stability of working state.

Table 8. Comparison of the optimization effect of different methods

No.	Method	Environment Optimization Effect	Process Optimization Effect	Human Resource Optimization Effect
1	Traditional Methods	0.60	0.50	0.55
2	Common Multimodal Methods	0.70	0.65	0.70
3	Proposed Method	0.85	0.80	0.85

For process optimization, the proposed method also demonstrates strong performance with an 80% improvement rate, compared with 65% for common multimodal methods and only 50% for traditional methods. This advantage stems from the model's in-depth analysis of employee task execution behavior and collaboration patterns. By analyzing workflow data like task progress and collaboration frequency, the model accurately identifies process bottlenecks and uneven resource allocation, addressing them through task restructuring and collaboration strategies. This intelligent process adjustment mechanism significantly improves task completion efficiency and team collaboration quality.

Finally, for human resources optimization, the proposed method achieves an 85% improvement, significantly outperforming the common multimodal method (70%) and the traditional method (55%). The result confirms the effectiveness of the personalized management strategy based on employee behavioral patterns. By comprehensively analyzing employee's daily behaviors, work status, and performance data, the model can formulate training paths and incentive mechanisms that better align with individual differences, thereby improving employee satisfaction and development potential. At the same time, the model identifies inefficient behaviors and provides customized feedback to promote the continuous optimization of employee's career paths.

The current model collects image, voice, and work log data, but lacks a robust privacy protection mechanism, employing only simple data desensitization instead of advanced techniques like differential privacy or federated learning. Furthermore, employees were informed of data use only through a departmental notice rather than being provided individual consent forms, which may foster resistance to behavioral monitoring. The specific costs include hardware, software, and labor costs for data integration, with a total cost of approximately 55,000 yuan. This cost may be a burden for small and medium-sized enterprises, and it is necessary to further reduce reliance on hardware. The single-company sample lacks representation from industries like manufacturing and services, and the 1-month experiment cycle is too short to account for seasonal variations for project cycle fluctuations, potentially leading to biased results. The model currently recognizes only explicit behaviors, such as whether social media is used and communication duration, and struggles with implicit data, such as emotions in voice tone and facial micro-expressions, achieving only 65% accuracy compared to 90% for structured data. Integrating effective computing models is necessary to improve processing capabilities. Regarding scalability, while the system supports a 50-person team with 10 cameras, a basic sensor suite, and a single server, scaling to over 500 users requires solving data transmission latency and rising hardware costs. Additionally, different industries require adaptation to differentiated behavioral libraries. To improve scalability, the following steps will be taken in the future. 1) Develop a lightweight model and migrate the data processing module from a centralized server to "edge computing nodes" to reduce the hardware deployment costs of large organizations. 2) Construct an industry-general behavioral feature template library and preset core behavioral indicators for three industries to reduce the secondary development workload for cross-industry adaptation.

6. Conclusion and Future Research

In this paper, an optimization model was constructed based on multimodal employee behavior analysis to address the challenge of improving office efficiency. By integrating multi-source data such as visual images, sensor data, workflow records, and text information, the model can comprehensively identify employee behavioral patterns and, when combined with environmental adjustment, process optimization, and personalized human resource management, achieve significant improvement in office efficiency. Experimental results demonstrate that the proposed method outperforms traditional and common multimodal methods in efficiency enhancement, behavior recognition accuracy, and environmental, process, and human resource efficiency, confirming its strong practicality and advancement. Despite the promising results, the model has limitations. First, it requires high-cost multimodal data collection and synchronization, demanding advanced hardware and data integration technology. Second, the privacy protection mechanisms are not yet fully developed, which may raise employee concerns about behavioral monitoring and hinder user acceptance. Additionally, the model's accuracy in processing complex unstructured data requires improvement, and a response lag persists in highly dynamic office scenarios.

Future research can be carried out in the following directions: first, to improve the model's real-time and lightweight performance, reduce the deployment threshold, and enhance its adaptability. Second, to introduce mechanisms such as differential privacy and multi-party secure computing to improve the level of privacy protection in the process of data collection; and third, to further expand the types of data modalities, such as fusion of emotion recognition, physiological signals, and semantic comprehension, in order to achieve a more comprehensive modeling of the employee's state. In addition, the applicability and adjustment mechanisms of the model in new office modes, such as telecommuting and flexible working systems, can be explored in the future to further advance the intelligent office system.

Author Contributions

Tong You contributes to conceptualization, methodology, software, validation, analysis, investigation, data collection, draft preparation, manuscript editing, visualization, supervision, project administration, and funding acquisition. Wenjia Wang contributes to conceptualization, methodology, software, validation, analysis, investigation, data collection, draft preparation, manuscript editing, visualization, supervision, and project administration. All authors have read and agreed with the manuscript before its submission and publication.

Funding

This research received no specific financial support from any funding agency.

Institutional Review Board Statement

Not applicable.

Declaration of Artificial Intelligence (AI) Tools

The authors used Grammarly solely for language editing and readability improvement. The authors reviewed and verified all content and take full responsibility for the accuracy and integrity of the manuscript.

References

Acevedo-Duque, N. (2025). Impact of leader behavior on employee experience and job satisfaction in educational institutions. *Administrative Sciences*, 15. doi:10.3390/admsci15040119

Bennouna, A., Boughaba, A., and Djabou, S. (2024). Examining the mediating effect of job satisfaction on the relationship between leader-member exchange and safety behavior among Algerian healthcare workers. *Leadership in Health Services*, 37(1), 112–129.

Cai, L., Xiao, Z., and Ji, X. (2023). Impact of supervisor developmental feedback on employee innovative behavior: Roles of psychological safety and face orientation. *Journal of Managerial Psychology*. doi:10.1108/JMP-12-2021-0670

Dua, A. K., Farooq, A., and Rai, S. (2023). Ethical leadership and its influence on employee voice behavior: Role of demographic variables. *International Journal of Ethics and Systems*, 39(2), 213–235. doi:10.1108/IJOES-10-2021-0200

El-Tony, F., and Choo, S. (2025). Employee green behavior in laboratories of higher education institutions: A qualitative study. *International Journal of Sustainability in Higher Education*, 26(3). doi:10.1108/IJSHE-03-2024-0168

Futri, S. Y., Jimad, H., and Hayati, K. (2023). The influence of transformational leadership on employee performance through innovative work behavior as a mediating variable. *Journal of Economics, Management and Trade*. doi:10.9734/jemt/2023/v29i81113

Li, T., and Yi, L. (2025). Leadership empowerment behavior and employee referrals: Chain mediation of perceived challenge stress and employee experience. *Chinese Management Studies*, 19(1). doi:10.1108/CMS-11-2023-0585

Liu, F., Liu, J., and Ding, H. (2023). Employee strengths mindset and voice behavior: The roles of general self-efficacy and leader–member exchange. *Sustainability*. doi:10.3390/su151511706

Liu, T., Zhang, S., and Zhang, C. (2025). Workplace digitalization and employee innovation behavior: Analysis from the employee perspective. *Social Behavior and Personality: An International Journal*, 53(2). doi:10.2224/sbp.13910

Lu, J., Yang, T., Qin, Y., and Hu, X. (2025). Game-theoretic analysis of the impact of differential leadership on employee silence behavior in Chinese family enterprises. *Frontiers in Physics*. doi:10.3389/fphy.2024.1536325

Mata, N., and Tang, K. (2023). Does participative leadership promote employee innovative work behavior in IT organizations. *International Journal of Innovation and Technology Management*. doi:10.1142/S021987702350027X

Nuhu, M. S., Kamselem, K. M., Mustapha, M., and Abdullahi, S. (2025). Personal value, employee impact and service quality delivery: The mediating role of organizational citizenship behavior. *International Journal of Quality and Service Sciences*, 17(1). doi:10.1108/IJQSS-06-2023-0081

Park, S., Park, G., and Hong, S. (2024). Exploring the characteristics and health outcomes of working from home: Analysis of 2021 California Health Interview Survey data. *American Journal of Industrial Medicine*, 67(2). doi:10.1002/ajim.23556

Peng, C., Peng, Z., Lin, J., Xie, J., and Liang, Y. (2025). How and when perceived COVID-19 crisis disruption triggers employee work withdrawal behavior: The role of perceived control and trait optimism. *Personality and Individual Differences*, 235. doi:10.1016/j.paid.2024.112981

Riemenschneider, K., Burney, L., and Bina, S. (2023). The influence of organizational values on employee attitude and information security behavior: The mediating role of psychological capital. *Information and Computer Security*. doi:10.1108/ICS-10-2022-0156

Saghih, F., and Ilani, R. (2023). Bullying behavior, mediating role of employee silence and their effects on organisational citizenship behavior: Evidence from seven Iranian metropolises. *Management Revue*, 34(3). doi:10.5771/0935-9915-2023-3-258

Sean Locke, A., and Osborne, J. (2024). Determining the right levels of health coaching and heart rate variability biofeedback in a workplace behavior change intervention: Multiphase optimization strategy preparation study. JMIR Formative Research, 8. doi:10.2196/47181

Shen, C., Wang, Z., Yao, K., Yuan, W., and Sun, F. (2024). Analysis and optimization of varying door and windows opening on energy consumption and thermal comfort in air-conditioner office. *Journal of Engineering Science and Technology Review*, 17(3). doi:10.25103/jestr.173.22

Silva, D., Garcia, P., and Carlo, C. (2023). Performance assessment of RBFMOpt, NSGA2, and MHACO on the thermal and energy optimization of an office building. *Journal of Building Performance Simulation*. doi:10.1080/19401493.2022.2143568

Song, Y., Guo, L., and Zheng, Q. (2025). Measuring scientific inquiry ability related to hands-on practice: An automated approach based on multimodal data analysis. *Education and Information Technologies*, 30(4), 4381–4411. doi:10.1007/s10639-024-12991-7

Su, R. (2024). Optimization and construction of collaborative office under the background of smart campus in universities. *Intelligent Information Management*, 16(1), 1–9. doi:10.4236/iim.2024.161001

Thebuwena, J., Samarakoon, K., and Ratnayake, C. (2024). Optimization of energy consumption in vertical mobility systems of high-rise office buildings: A case study from a developing economy. *Energy Efficiency*, 17(6), 1–30. doi:10.1007/s12053-024-10246-5

Torres, G., Mukoma, G., Neophytou, N., and Gradidge, L. (2023). Employee perceptions of an established health and wellness coaching program to improve behavior. *Journal of Occupational and Environmental Medicine*, 2023(8), 65. doi:10.1097/JOM.0000000000002905

Wang, F., You, T., and Yang, H. (2025). Performance analysis and operation optimization of photovoltaic/thermal assisted energy-pile ground source heat pump system in cold regions. *Renewable Energy*, 244. doi:10.1016/j.renene.2025.122722

Wisniewski, D., Zielinski, M., and Dabrowski, P. (2024). Data-Driven Approaches to Behavior Analysis in Smart Classrooms: A Review of AI Algorithms. *Data Engineering and Applications*, 4(2), 14–30. doi:10.64972/dea.2024.v4i2.12

Xu, Z., and Yang, F. (2024). Effect of critical thinking disposition on employee innovative behavior: A meta-theory of personality perspective. *Journal of Pacific Rim Psychology*, 18. doi:10.1177/18344909241231847

Xuanyuan, P., Zhang, Y., Yao, J., and Zheng, R. (2024). Sensitivity analysis and optimization of energy-saving measures for office building in hot summer and cold winter regions. *Energies*, 17(7). doi:10.3390/en17071675

Xudong, Z., Qiao, N., and Zengcheng, C. (2023). Multi-objective optimization design of energy efficiency for office building window systems based on indoor thermal comfort. *Science and Technology for the Built Environment*, 2023(6/10), 29. doi:10.1080/23744731.2023.2194840

Zhang, X., Ning, Q., and Chen, Z. (2023). Multi-objective optimization design of energy efficiency for office building window systems based on indoor thermal comfort. *HVACandR Research*, 29(6), 15. doi:10.1080/23744731.2023.2194840

Tong You, associate professor at the College of Business, Guilin Institute of Information Technology, specializes in educational technology and information science, as well as management science and engineering. With a strong focus on teaching reform and scientific research, he has made significant contributions to education. Tong has served as a key member in three Guangxi Zhuang Autonomous Region-level Teaching Reform Projects and has led the completion of one such project. Additionally, he has participated in three Ministry of Education collaborative education projects, two Guangxi Zhuang Autonomous Region-level Scientific Research Projects, and two university-level teaching reform projects. His research and teaching efforts have resulted in the publication of six scientific and teaching papers, either independently or in collaboration with peers. Tong is also recognized for his dedication to student development, having guided college students to achieve a bronze award in the Guangxi Division of the 7th China International "Internet +" College Students Innovation and Entrepreneurship Competition. His work reflects a commitment to advancing educational practices and fostering student innovation, making him a valuable asset to the academic community.

Wenjia Wang, associate professor at the College of Business, Guilin Institute of Information Technology, holds a master's degree in Ethnic Economics. Her research focuses on management science and engineering and regional economics, with a strong emphasis on teaching reform and scientific research. Wang has presided over one provincial/municipal-level teaching reform project and served as a key member in three such projects. Additionally, she has led two provincial/municipal-level scientific research projects, demonstrating her expertise in both academic and practical applications. Her scholarly contributions include the publication of eight scientific research and teaching papers, two of which are in core journals. Wang is also dedicated to student development, guiding students to win two Bronze Awards in the Guangxi Division of the 7th China International "Internet +" College Students Innovation and Entrepreneurship Competition. Furthermore, she has mentored students in securing one national-level and one regional-level project under the College Students' Innovation and Entrepreneurship Training Program. Her work reflects a commitment to advancing education, fostering innovation, and supporting student success, making her a valuable contributor to her field and institution.