Received: 19 Oct 2025 | **Revised:** 10 Dec. 2025 | **Accepted:** 20 Jan. 2026 | **e-ISSN:** 3085-7813

Alumni.in

ARTICLE

INTEGRATING ARTIFICIAL INTELLIGENCE INTO TEACHER PERFORMANCE EVALUATION: EVIDENCE FROM UNDERGRADUATE INSTITUTIONS IN ANHUI PROVINCE

[1]**Song Jiao**. Faculty of Educational Sciences, Universitas Pendidikan Indonesia. **ORCID:** https://orcid.org/0009-0005-2562-3735

[2*]**Sururi Sururi**. Faculty of Educational Sciences, Universitas Pendidikan Indonesia **ORCID:** https://orcid.org/0009-0007-8290-4044

[3]**Aan Komariah**. Faculty of Educational Sciences, Universitas Pendidikan Indonesia **ORCID:** https://orcid.org/0000-0003-2269-3592

[4]**Taufani Chusnul Kurniatun**. Faculty of Educational Sciences, Universitas Pendidikan Indonesia **ORCID:** https://orcid.org/0000-0003-1736-7258

Corresponding Author:
Sururi Sururi
E-mail: sururi@upi.edu

Editor in chief
Altieres De Oliveira Silva
Alumni.In Editors

How to cite this article:

songjiao, Sururi, S., Aan Komariah, & Kurniatun, T. C. (2026). Integrating Artificial Intelligence into Teacher Performance Evaluation: Evidence from Undergraduate Institutions in Anhui Province. ***Journal of Sustainable Competitive Intelligence*** *16*, e0606. https://doi.org/10.37497/eagleSustainable.v16i.606

ABSTRACT

Purpose: China's recent educational evaluation reforms and the rapid advancement of artificial intelligence (AI) have exposed significant limitations in traditional undergraduate teacher performance evaluation systems, including overly one-dimensional criteria, excessive emphasis on research output, and insufficient monitoring and feedback mechanisms. Building on the author's prior research on teacher evaluation reforms in Anhui Province, this study investigates how AI can be integrated into teacher performance evaluation to enhance strategic alignment, evaluative fairness, and developmental orientation within undergraduate institutions.

Methodology/approach: The study adopts a mixed-methods approach grounded in strategic human resource management theory. First, a systematic review of domestic and international literature on teacher performance evaluation, AI in educational management, and strategic HRM is conducted. Second, an AI-based performance evaluation model is developed by adapting a strategy-oriented evaluation framework for undergraduate faculty in Anhui Province, incorporating data analytics, intelligent evaluation mechanisms, and multi-source assessment systems. The proposed model is empirically tested using questionnaire data collected from faculty members in selected undergraduate institutions in Anhui Province, employing statistical techniques such as reliability and validity analyses and structural equation modeling.

Originality/Relevance: By applying intelligent data analysis and multi-dimensional evaluation tools, the proposed model reduces excessive reliance on research metrics and promotes a more balanced recognition of teaching, research, and service contributions. The study addresses national policy requirements to move beyond the "five-only" evaluation framework and responds to faculty dissatisfaction with current evaluation practices.

Key findings: The findings demonstrate that AI-supported evaluation can improve objectivity, accuracy, strategic alignment, and feedback mechanisms, fostering continuous faculty development.

Theoretical/methodological contributions: The integration of AI with strategic human resource management and performance evaluation theory offers a novel framework for higher education management and provides an empirically validated model for transformative teacher evaluation reform, particularly in Anhui Province.

Keywords: AI, Performance evaluation of teachers, Strategic human resource management, Higher education institutions, Anhui Province

DOI: https://doi.org/10.37497/eagleSustainable.v16i.606

Open access

RESUMO

Objetivo: As recentes reformas na avaliação educacional na China e o rápido avanço da inteligência artificial (IA) evidenciaram limitações significativas nos sistemas tradicionais de avaliação do desempenho docente no nível de graduação, incluindo critérios excessivamente unidimensionais, ênfase desproporcional na produção científica e mecanismos insuficientes de monitoramento e feedback. Com base em pesquisas anteriores do autor sobre reformas na avaliação docente na Província de Anhui, este estudo investiga como a IA pode ser integrada à avaliação do desempenho docente para aprimorar o alinhamento estratégico, a equidade avaliativa e a orientação ao desenvolvimento nas instituições de ensino superior de graduação.

Metodologia/abordagem: O estudo adota uma abordagem de métodos mistos fundamentada na teoria da gestão estratégica de recursos humanos. Inicialmente, realiza-se uma revisão sistemática da literatura nacional e internacional sobre avaliação de desempenho docente, IA na gestão educacional e gestão estratégica de recursos humanos. Em seguida, é desenvolvido um modelo de avaliação de desempenho baseado em IA, adaptado de um framework estratégico de avaliação para docentes de graduação na Província de Anhui, incorporando análise de dados, mecanismos inteligentes de avaliação e sistemas de avaliação multifuente. O modelo proposto é testado empiricamente por meio de dados coletados via questionários aplicados a docentes de instituições de graduação selecionadas na Província de Anhui, utilizando técnicas estatísticas como análises de confiabilidade e validade e modelagem de equações estruturais.

Originalidade/Relevância: Ao aplicar análise inteligente de dados e ferramentas avaliativas multidimensionais, o modelo proposto reduz a dependência excessiva de métricas de produção científica e promove um reconhecimento mais equilibrado das contribuições em ensino, pesquisa e extensão. O estudo responde às exigências das políticas nacionais de superação do modelo avaliativo “five-only” e às insatisfações docentes com os sistemas atuais de avaliação.

Principais resultados: Os resultados indicam que a avaliação apoiada por IA pode melhorar a objetividade, a precisão, o alinhamento estratégico e os mecanismos de feedback, promovendo o desenvolvimento contínuo do corpo docente.

Contribuições teóricas/metodológicas: A integração da IA com a teoria da gestão estratégica de recursos humanos e da avaliação de desempenho oferece um novo referencial para a gestão do ensino superior e apresenta um modelo empiricamente validado para a reforma transformadora da avaliação do desempenho docente, especialmente nas instituições da Província de Anhui.

Palavras-chave: Inteligência artificial, Avaliação de desempenho docente, Gestão estratégica de recursos humanos, Instituições de ensino superior, Província de Anhui

1 INTRODUCTION

The most important human resources in higher education institutions are the teachers. They determine the quality of teaching, research, talent cultivation, and social service the university is able to provide. Therefore, to provide an effective governance mechanism, and to drive the allocation of resources, professional development, structure the organization, and achieve goals for the system, teacher performance evaluations become vital in the higher education systems across the globe (OECD, 2013; Armstrong, 2021). A system of evaluation that is scientific, objective, and fair, one that measures the results of work of the individual teacher, is one that has a large impact in the guiding and motivating of professional work and the behavior of teachers, and synchronizing the efforts of the individual with the institution's mission.

Research that has gained notoriety over the past twenty years has questioned the traditional models of teacher evaluation that are unidirectional. The models that are heavily dependent on summative assessments, a narrow range of quantifiable measures, and administrative foresight. Evaluation systems that are driven on metric systems closed on themselves that are result-oriented, tend to not consider the effort on teaching, the quality of teaching, the role of collaborative teaching and professional development, and in many cases become academic killjoys, burn the people out, and create work for the system (Darling-Hammond, Amrein-Beardsley, Haertel, & Rothstein, 2012). Consequently, many countries have started to try comprehensive evaluation systems that are developmental in nature and consider multiple dimensions, inclusive of feedback from various stakeholders over intervals (Alwaely et al., 2023).

University teacher evaluation reform in China has become a central focus in the modernization of higher education governance. Universities in China became subject to the “Double First-Class” initiative, which has placed two conflicting sets of expectations on the institutions: the first is an increase in the system’s academic competitiveness, and the second is the traditional role of the universities as institutions of purposeful and systematic human resource development. However, the evaluation of the teaching staff at Chinese universities, at least until recently, has been conceptually framed within a reductive and negative “five only” syndrome: only papers, only titles, only degrees, only awards, only projects. This syndrome has led to the instrumentalization of teaching, and the primary function of teaching has been neglected (Ministry of Education of the People’s Republic of China, 2020). The Chinese government, in this regard, has been streamlining policies to propose the fracturing of this evaluative logic in favor of the construction of an evaluative system to articulate quality, classification, and development (Zhu, 2018).

While there has been a lot of policy support, there are still inconsistencies when it comes to evaluations and the reforms that are being put in place at an institutional level. A number of studies highlight the continuing existence of the majority of undergraduate institutions using rudimental appraisal systems that are overly focused on a few research outputs, lack differentiation among faculty roles, and have little room for student and peer evaluations, and feedback and improvement processes (Qu, 2024; Zhang & Jia, 2024). These issues are very common in teaching and application-oriented undergraduate institutions, given their strategic mission of emphasizing undergraduate teaching effectiveness and servicing the local economy and not the higher academic output (Kasman & Lubis, 2022).

Anhui Province offers a suitable and informative research context to investigate these

issues. As a key province in central China, Anhui has been developing applied undergraduate institutions to aid in the region's industrial and sociospatial development. However, empirical studies show that the teacher performance evaluation practices in many of the Anhui undergraduate universities still rely on traditional academic-oriented models. As such, there is a lack of alignment between evaluation criteria and institutional strategic positioning (Zhu, 2018; Qu, 2024). Faculty members who are primarily responsible for teaching reform, collaborative work with industry, or student mentorship often feel that their efforts are inadequately recognized, leading to a loss of motivation and a diminished sense of professional identity.

From a conceptual perspective, systemic issues can be analyzed via Strategic Human Resource Management (SHRM). Adopting positions such as Delery and Doty (1996) and Chadwick and Dabu (2008) argue from the integration of human resource strategy and management to attain sustained performance gain. Therefore, as a core practice of HRM, performance evaluation serves to exert control and act as a mechanism to convert organizational goals into individual goal expectations, assisting in the attainment of the desired organizational goal through capability development and reinforcing desired behaviors (Armstrong, 2021).

In the context of higher education, applying principles of SHRM entails the alignment of teacher performance evaluation systems to the strategic objectives of the university, taking into consideration the differentiated roles of the various categories of teachers and fostering sustained professional growth (OECD, 2013; Delery & Doty, 1996). However, traditional systems of evaluation fail to address, in a timely manner, the comprehensive multidimensional data of performance, and in this regard, the practical application of many principles of SHRM in university performance management has been largely constrained, particularly in large, complex, and diverse academic units (Chadwick & Dabu, 2008).

Changes in artificial intelligence (AI), big data, and digital technologies are affecting human resource management in every field. Brougham and Haar (2018) and Jarrahi (2018) note that AI systems can capture and evaluate behavior and performance data, assist in multi-source assessments, and provide feedback, assessments, and predictions in real time. AI systems have been used in corporate settings for performance evaluation, talent management, and support decision-making while enhancing objectivity, efficiency, and decreasing administrative workload (Kellogg, Valentine, & Christin, 2020).

In education, AI has been used in learning analytics, intelligent tutoring systems, and academic management (Wang et al., 2024; Pham & Sampson, 2022). AI can provide more accurate assessments of learning outcomes and analyze teaching and learning behavior and assist in decision-making in education (Luckin & Holmes, 2016). AI is recognized as a means of improving quality in education and management systems when effectively and ethically designed (Zawacki-Richter et al., 2019; Nguyen et al., 2023). Existing research regarding AI in education is primarily concerned with student learning processes rather than the management and performance evaluation of teachers. In higher education, the application of AI to the evaluation of teachers' performance is particularly unexplored due to the intricacies of the profession, the roles and responsibilities of the teachers, the evaluation autonomy, and the variety of evaluation purposes. In the case of China, this research gap is particularly noticeable because, in China, the reform of teacher evaluation and digital transformation is occurring at the same time but often in isolation (Qu, 2024).

Evaluation of teacher performance using AI can be tailored to meet some of the challenges of

the systems in use today (Pei & Lu, 2023; Zhang & Jia, 2024). For instance, AI systems for teacher performance evaluation can collect and assess diverse documentation from teaching evaluation, research, service, and peer evaluation to construct a more rounded profile of performance. This can reduce the reliance on a single measure and aid in the differentiated evaluative classification and role assignments of teachers (Yildiz Durak & Onan, 2025). AI can improve evaluation measurement by using analytics. Instead of using evaluations that are, in a way, dead end and passive, evaluations can be reformative and supportive of evaluative growth (Bauer et al., 2025). AI can analyze evaluation data to better delineate data, criteria, and sources which improves the perception of evaluation in terms of fairness and transparency (Nguyen et al., 2023).

That said, the potential use of AI in assessing teacher performance raises several challenges regarding the coherence of strategy, teacher buy-in, and governance. Without a defined positive strategic outcome, AI might capture and exacerbate bias in evaluation and increase the evaluative pressure on the teachers. From an SHRM perspective, AI should be an enabler, facilitating the positive alignment of strategy, purposeful capability and engagement, and is not supposed to be an imposition of a purely technical or managerial surveillance system (Brougham & Haar, 2018; Kamalov et al., 2023).

In the undergraduate institutions of Anhui Province, where evaluation reform, digital transformation, and a strategic shift are happening concurrently, assessing the possible impact of AI on teacher performance evaluation is both opportune and essential (Qu, 2024). Given the context, to assist in the evaluation reform process, it is critical to know how to design AI-enhanced evaluative systems to be strategically aligned, capture and recognize varied teacher contributions, and to be equitable and constructively supportive (Alwaely et al., 2023; Pei & Lu, 2023).

Considering this context, this study investigates the use of artificial intelligence in the evaluation of teacher performance through the lens of strategic human resource management, focusing on undergraduate institutions located in Anhui Province. This study attempts to bridge the gap between SHRM theory and AI performance management by establishing and validating a teacher evaluation framework that is strategy-oriented and empirically driven (Delery & Doty, 1996; Wang et al., 2024). More specifically, the study attempts to explain the following: the extent to which AI improves the alignment of institutional strategy to teacher performance evaluation; the extent to which evaluative intelligence systems affect teachers' perceptions of equity and motivation; and the extent to which AI evaluative systems engender continuous professional development (Yildiz Durak & Onan, 2025; Bauer et al., 2025).

By means of theoretical speculation and empirical analysis, this study seeks to contribute to the scholarship of higher education management, specifically the intersection of SHRM theory, the use of artificial intelligence, and teacher performance evaluation (Pham & Sampson, 2022). From a practical standpoint, this study seeks to provide undergraduate institutions evidence-based recommendations and policy options to support higher education in China with its digital transformation by providing a more sophisticated means of developing and evaluating faculty (Ministry of Education of the People's Republic of China, 2020; Kasman & Lubis, 2022).

Research Objectives

1. To research on the role of Artificial Intelligence in improving the congruence between institutional strategic objectives and teacher performance assessment in Anhui Province undergraduate institutions.
2. In order to determine how AI-aided evaluation systems can influence teacher perceptions of equity, inspiration and professional growth.
3. To create and empirically test the strategy-oriented, AI-enabled teacher performance evaluation model based on the Strategic Human Resource Management theory.

Research Questions

1. What is the effect of Artificial Intelligence integration on the strategic alignment of institutional goals with teacher performance assessment?
2. How do the use of AI-based evaluative tools affect the perception of fairness, motivation, and professional growth opportunities by teachers?
3. Is there a way to make a strategy-focused, AI-based evaluation model more effective and developmental regarding the overall quality of teacher performance evaluation in undergraduates?

2. THEORETICAL FRAMEWORK

The evaluation of teacher performance in higher education is a multidimensional process that relies on various evaluation criteria (Alwaely et al., 2023; OECD, 2013). To achieve construct scientific credibility, fairness and the strategic relevance of the evaluation is needed (Darling-Hammond et al., 2012). In undergraduate institutions, most notably those with an application-oriented mission, the evaluation of teacher performance must achieve a balance among teaching quality improvement, research, and social service, and the professional development of the faculty (Kasman & Lubis, 2022; Zhu, 2018). The traditional evaluation systems that have been employed in the past focus on administrative control and outcomes that are quantitative in nature. It is clear that the diverse goals have not been achieved. Hence, there is a need for a more comprehensive theoretical framework to determine how evaluation systems in higher education have been articulated with institutional goals, teacher advancement, and professional development (Qu, 2024; Armstrong, 2021).

The study draws a framework from the intersection of strategic human resource management, performance management, and the most recent artificial intelligence (AI) based evaluation systems (Delery and Doty, 1996; Chadwick and Dabu, 2008; Pham and Sampson, 2022; Wang et al., 2024). The theories suggest that evaluation systems can substantially shift their focus from more of a results-oriented approach to goals that are more developmental and positive control strategic systems (Brougham and Haar, 2018; Kamalov et al., 2023). The framework presents a model of AI-based performance evaluation for higher educational institutions in the Anhui Province, China (Pei & Lu, 2023; Zhang & Jia, 2024).

2.1 Theories of Strategic Human Resource Management

The Strategic Human Resource Management (SHRM) theory studies the relationship between the actions in the human resource function and the goals of the firm (Delery \& Doty,

1996; Chadwick \& Dabu, 2008). While traditional personnel management deals with administrative functions and processes, SHRM considers human resources as a pivotal driver of responding to the organization's performance and sustainable growth (Armstrong, 2021). Core SHRM practices include performance review, training, and development, and these must be aligned with the strategies of the institution as a whole, including the institution's incentive systems and career advancement processes (Brougham & Haar, 2018).

In the case of higher education SHRM theory, there is an implication that the administrative routine of evaluating the performance of teachers is in some ways a strategic activity that transforms the development goals of the university into performance expectations for the individual (OECD, 2013; Alwaely et al., 2023). This is particularly important for universities that are primarily undergraduate level, especially for those with an emphasis on application, as there is a strong regional development focus (Qu, 2024). Quality of undergraduate teaching, the development of practical talents, and teaching in service to regional development are common strategic goals (Zhu, 2018). Teacher performance evaluation systems that are absent of these strategic goals risk administrative misdirection of teachers and compromise the institution's mission.

Utilizing SHRM theory in the case of the teacher performance evaluation, the author suggests that the evaluation must align with the institution's positioning and developmental goals (Delery & Doty, 1996). This means that the evaluation criteria, the processes, and the resultant uses must be aligned with the strategic goals so that each teacher's performance supports specific goals of the institution. This kind of alignment strengthens the evaluative performance role in guiding and motivating and contributing to the greater faculty developmental structures (Chadwick & Dabu, 2008; Armstrong, 2021).

2.2 Performance Management Theories and Multiple Evaluations

Performance management theories assist in defining an evaluative framework that supports accountability and developmental processes (Armstrong, 2021; Delery & Doty, 1996). Contemporary performance management focuses on the iterative processes such as goal setting, performance monitoring, feedback, and improvement as opposed to summative evaluation (Podsakoff et al., 2003). Evaluation within this context is an ongoing and integral part of management as opposed to a one-off measurement exercise.

In the specific case of evaluating university teachers, performance management principles underscore the use of a multi-faceted evaluation system (OECD, 2013; Alwaely et al., 2023). Different, yet, interdependent facets of the work of the teacher include teaching, research, community service, and professional development (Kasman & Lubis, 2022). When a single aspect is given disproportionate focus such as research output, it creates negative teacher behaviors and this ultimately affects the quality of teaching (Darling-Hammond et al., 2012). Therefore, an evaluative system of balance and equity is necessary to achieve comprehensive recognition of performance (Kamalov et al., 2023).

This research takes the position that, in the context of performance of evaluations, the criteria should be multidimensional with respect to both the outcome and the process of the evaluations themselves (Pei & Lu, 2023; Wang et al., 2024). Such a situation would allow for an evaluation that is predicated on the performance of the teachers to foster professional growth in a positive manner and avoid the detrimental outcomes of overly simplistic performance measures (Zawacki-Richter et al., 2019).

2.3 Developmental Evaluation with Respect to Feedback and Incentives

Developmental evaluation, in this case, looks to emphasize the process, in this case, the evaluation of the performance, even beyond the potential of rank ordering and distribution of rewards (Darling-Hammond et al., 2012). In this case, evaluations are supposed to bring about learning, reflection, and the succession of improvement (Pham & Sampson, 2022). In this case, feedback is the most critical mechanism to help evaluators identify their strengths and weaknesses and adjust behavior accordingly (Yildiz Durak & Onan, 2025).

Most, if not all, academic institutions at the undergraduate level do not appear to have mechanisms of feedback with respect to performance evaluations (Kasman & Lubis, 2022). Performance is assessed, and results are applied to pay increases and promotions with little or no communication between evaluators and all teaching staff. This diminishes the evaluative, in this case, the developmental component of the evaluation and frustrates the teaching staff (Nguyen et al., 2023). From a theoretical perspective, feedback, if is to be effective, has to be provided in a timely manner, be specific, and be tied to developmental activities (Luckin & Holmes, 2016; Bauer et al., 2025).

Incentive systems further shape the impact of evaluation systems (Armstrong, 2021). When incentives are misaligned with evaluation outcomes, evaluations tend to lose their energizing effect. Thus, development-focused incentive systems should include immediate incentives and prospects for future career advancement in order to foster ongoing professional commitment (Darling-Hammond et al., 2012; Kamalov et al., 2023).

2.4 Artificial Intelligence–Supported Performance Evaluation

The use of artificial intelligence for the evaluation of teacher performance can also be improved (Wang et al., 2024; Pei & Lu, 2023). With the help of AI, large-scale evaluation data can be collected and analyzed, and the evaluation can be drawn from multiple sources and can be ongoing and include feedback (Zhang & Jia, 2024; Pham & Sampson, 2022). The integration of artificial intelligence in teacher evaluation systems can help improve SHRM and performance management systems and the evaluation process in its accuracy, transparency, and efficiency (Chiu et al., 2023; Bauer et al., 2025).

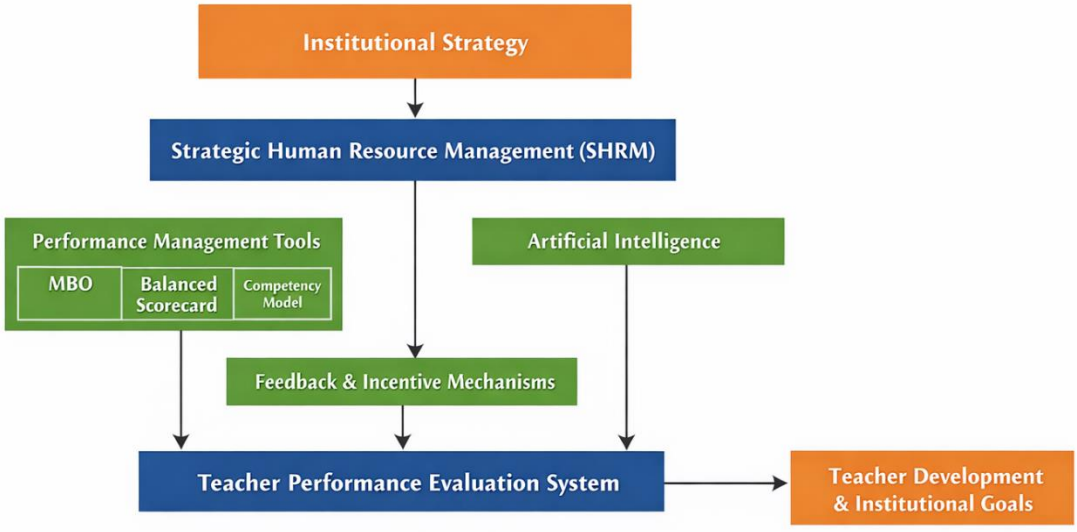
AI can analyze performance evaluations, faculty evaluations, service evaluations, and peer evaluations to generate all-encompassing evaluations (Zhang et al., 2026; Wang, Wang, & Solhi, 2026). AI may ultimately be less evaluative subjective, providing a greater case for the reliability of the evaluation and reducing the evaluative risks and consequences related to the subjectivity of the evaluation (Nguyen et al., 2023). Additionally, the evaluative process can go from being event-based to being continuous because intelligent analytics can capture evaluative data in real time; this also contributes to the evaluative process's focus on improvement (Yildiz Durak & Onan, 2025).

The integration of potentially groundbreaking AI technologies into teacher performance evaluations may be beneficial, but, unregulated AI systems evaluating teacher performance can reinforce biases that need to be considered. (Kellogg et al., 2020; Nguyen et al., 2023). It is concerning that the more sophisticated the systems of AI are in evaluating teacher performance, the more SHRM and developmental evaluation theory will be needed to ensure the systems remain highly and solely focused on the strategic and human-centered objectives (Delery & Doty, 1996; Kamalov et al., 2023; Alwaely et al., 2023).

2.5 Conceptual Framework of the Study

The process of building a conceptual framework involves the integration of strategic human resource management, assessment, artificial intelligence, and performance management systems (Delery and Doty, 1996; Chadwick and Dabu, 2008; Armstrong, 2021). Within the framework, variables can be identified. The institutional strategy is linked to the purpose of evaluation; the performance management system is linked to the evaluative purpose and the indicators; the feedback and reward system is linked to the cycle of improvement; and the artificial intelligence system is linked to the evaluation sophistication and the systems integration of the disparate data (Alwaely et al., 2023; Pei and Lu, 2023; Zhang and Jia, 2024).

As illustrated in Figure 1, this framework underpins the analysis of the teacher performance evaluation systems of undergraduate institutions in Anhui Province and the first of its kind, AI-driven, strategy-based evaluation framework (Qu, 2024; Zhu, 2018). It seeks to transform the teacher performance evaluation systems, shifting them from control systems to development systems that promote the growth of teachers and the institutional sustainability (Darling-Hammond et al., 2012; OECD, 2013; Kamalov et al., 2023).



3 METHOD

The evaluator focuses on the use of AI in teacher performance evaluation systems aligned with strategic human resource management from a more quantitative and theoretical modeling perspective. The approach aimed at scientific rigor, empirical correctness, and alignment with the relevant theory in the preceding section. The evaluator applied survey research and statistical analysis to the interrelated constructs of strategic alignment, evaluation and feedback, and teacher performance in the undergraduate institutions of Anhui province.

The methodological framework consists of four main parts: research design, sampling and data collection, development of measurement instruments, and a set of methods for data analysis. This framework helps the evaluator to achieve a greater degree of alignment between research objectives and theoretical and empirical constructs.

3.1 Research Design

This study employs a cross-sectional survey in an effort to address the limited scholarly work on the relationships between latent variables in an organizational and educational management perspective (Creswell, 1994). A quantitative methodology is chosen to test the theoretical propositions and to reinforce the objectivity and the extent to which the findings can be generalized. The explanatory research design focuses on defining the dimensions of AI-enabled teacher performance evaluation systems and teacher performance, motivation, and development outcomes.

A conceptual research model is developed based on the integrated theoretical framework, which positions strategic human resource management orientation as the foundational construct, the performance management tools (management by objectives, balanced scorecard, and competency-based evaluation) as the core evaluative mechanisms, and artificial intelligence as the technological resource, with the feedback and incentive mechanisms as the mediators of teacher performance outcomes. Given its ability to assess complicated causal relationships and numerous latent variables, structural equation modeling is chosen as the principal analytical method (Leguina, 2015).

3.2 Sample Selection and Data Collection

The target population comprises of full-time faculty of tertiary educational institutions in Anhui Province, China. Since provinces in the People's Republic of China (PRC) are the primary administrative subdivisions of the country, they are of educational interest, so provinces in the PRC are used as the primary unit of analysis in the study. Anhui was chosen as an illustrative province as it is a typical province in the PRC that is transitioning to application-oriented undergraduate education, resulting in a population of teachers whose performance are evaluated using a bottom-up approach, as a result justified by policies guided from above.

To stratify sampling, in order to enhance the quality of sampling, all universities included in the study were grouped by institutional type, that is, public undergraduate universities, and application-oriented undergraduate universities. After grouping, teachers were selected at random from each of the faculties and different academic rungs of each of the educational institutions. This approach minimizes sampling bias and provides balanced representation of different categories of teachers, i.e., teaching vs. research.

The primary method used to collect data was the administration of a structured questionnaire, which was anonymous to the respondents, and was a combination of online and paper formats. In order to make the questionnaire reliable, a preliminary survey was conducted amongst a select few educators in order to identify the clarity of each construct operationalized within the questionnaire. After the preliminary survey, a few modifications were implemented to enhance the clarity of some items as well as the accuracy of the constructs that were being measured. After the modifications were instituted, the final survey was conducted after an interval of three months, motivationally based on the fact that respondents were guided to be honest by the fact that the participants were allowed to respond constructively (Podsakoff, MacKenzie, Lee, & Podsakoff, 2003).

3.3 Measurement Instrument

The initial version of the questionnaire was adapted from existing measurement scales in the areas of strategic human resource management, performance management, and educational evaluation research as contextualized for the higher education sector. Each of the measurement items was captured on a five-point Likert (1932) scale commonly deployed in organizational research to measure attitudes and perceptions, where 1 means "strongly disagree" and 5 means "strongly agree".

The questionnaire was divided into five sections. The first section is demographic, to capture information on the respondent's gender, age, academic rank, years of teaching, and institution type. The second section focuses on the alignment of performance evaluation with institutional strategy and job responsibilities, assessing evaluative criteria for strategic evaluative function. The third section focuses on the performance management spectrum describing the presence of clear goals, balanced indicators, and relevance to the competencies of the evaluators. The fourth section assesses the role of artificial intelligence in the performance evaluation process in terms of the intelligence's ability to integrate data, transparently evaluative, and timely provide evaluative feedback. The fifth section assesses the interrelation of feedback and incentive mechanisms and how these mechanisms affect the motivation and professional development of teachers.

Questionnaire items were sent for content validation to experts in management of higher education and human resources for content validation. Construct validity and reliability were assessed and complemented with statistical approaches: confirmatory factor analysis, Cronbach's Alpha, and composite reliability (Leguina, 2015).

3.4. Data Analysis Techniques

Several steps were conducted concerning data analysis using relevant statistical packages. As a first step, descriptive statistical analysis was conducted to provide a summary of the demographic characteristics of the sample and the overall response distribution. As a second step, reliability analysis (using Cronbach's Alpha) was conducted and, for each of the constructs, assessed with values above 0.70 as the threshold for acceptance (Kline, 1999).

As a third step, validity analysis, through confirmatory factor analysis (CFA) assessed the convergent and discriminant validity of the measurement model. The Kaiser-Meyer-Olkin (KMO) and Bartlett's sphericity test for factor analysis (Kaiser, 1974) were included to test the adequacy of the data for factor analysis. As a fourth step, the hypothesized relationships among the latent variables were tested using structural equation modeling (SEM). As per the recommendations (Hu & Bentler, 1999) multiple measures were used to assess model fit, including the comparative fit index, Tucker-Lewis index, root mean square error of approximation, and standardized root mean square residual.

The last effects of feedback and incentive mechanisms are assessed concerning AI-enabled evaluation systems and teacher performance outcomes. Such an analytical technique is effective in evaluating both the direct and indirect impact of the model in question.

4. RESULTS

This part describes and examines the empirical findings of the study. In line with the

research objectives, the results are structured to depict the research setting, assess the measurement model's reliability and validity, evaluate the structural relationships among the variables, and discern the role of artificial intelligence in the evaluation of teachers' performance systems from strategic human resource management. Each subsection presents the empirical findings related to the research objectives, including descriptive statistics, measurement model evaluation, and structural model results.

4.1 Characterization of the Research Content and Sample

The empirical research study targeted undergraduate colleges within Anhui Province, China. Out of 420 questionnaires distributed, 368 responses were deemed valid, which leads to an effective response rate of 87.6%.

Demographic details of the participants are provided in Table 1. Out of the total sample, there were 182 male (49.5%) and 186 female (50.5%) participants. In terms of academic position, the largest proportions were from the lecturer (38.6%) and associate professor (26.6%) categories. The distribution of teaching experience was relatively even, with 42.4% of participants indicating they had over ten years of experience. The participant sample aligns with the standard distribution of faculty across undergraduate colleges in Anhui Province.

Table 1. Demographic Characteristics of Respondents

Variable	Category	Frequency	Percentage (%)
Gender	Male	182	49.5
	Female	186	50.5
Academic Rank	Assistant Lecturer	96	26.1
	Lecturer	142	38.6
	Associate Professor	98	26.6
	Professor	32	8.7
Teaching Experience	≤5 years	88	23.9
	6–10 years	124	33.7
	>10 years	156	42.4

4.2 Reliability and Validity Analysis of the Measurement Model

Before hypothesis testing, the evaluation of construct validity and internal consistency reliability was conducted. All the examined indicators display Cronbach's alpha value over 0.70, demonstrating sufficient internal consistency.

Table 2 presents that composite reliability for all the constructs is between 0.908 and 0.935, and all constructs also have AVE values above 0.50. The KMO value is 0.872, and Bartlett's test of sphericity is significant at $\chi^2 = 3264.18$, $p < 0.001$, which supports the appropriateness of the data for factor analysis.

Table 2. Reliability and Validity Statistics

Construct	Cronbach's Alpha	Composite Reliability	AVE
Strategic Alignment	0.892	0.914	0.681
Performance Management Tools	0.905	0.927	0.704
Artificial Intelligence Application	0.918	0.935	0.721
Feedback & Incentives	0.887	0.908	0.664
Teacher Performance Outcomes	0.901	0.923	0.698

KMO value = 0.872; Bartlett's test of sphericity: $\chi^2 = 3264.18$, $p < 0.001$

4.3 Structural Model Results and Hypothesis Testing

Using Structural Equation Modeling (SEM), the proposed relationships between different latent variables were tested. Using the model fit indices, the following values were achieved which indicate good fit to the data (CFI = 0.943; TLI = 0.936; RMSEA = 0.048; SRMR = 0.042).

Standardized path coefficients can be found in Table 4. Performance management tools ($\beta = 0.287$, $p < 0.001$) and strategic alignment ($\beta = 0.214$, $p < 0.001$) both had a positive effect on teacher performance. The application of artificial intelligence had a positive direct effect ($\beta = 0.198$, $p < 0.001$) on teacher performance and also had a positive effect on feedback and incentive mechanisms ($\beta = 0.352$, $p < 0.001$). Additionally, feedback and incentive mechanisms had a positive effect on teacher performance ($\beta = 0.301$, $p < 0.001$).

Table 3. Structural Path Coefficients

Hypothesized Path	Standardized Coefficient (β)	t-value	p-value
Strategic Alignment → Teacher Performance	0.214	4.32	<0.001
Performance Tools → Teacher Performance	0.287	5.41	<0.001
AI Application → Feedback & Incentives	0.352	6.18	<0.001
Feedback & Incentives → Teacher Performance	0.301	5.89	<0.001
AI Application → Teacher Performance	0.198	3.96	<0.001

Model fit indices: CFI = 0.943; TLI = 0.936; RMSEA = 0.048; SRMR = 0.042

4.4 Mediating Effect of Feedback and Incentive Mechanisms

A bootstrapping method was used to analyze the mediating impact of feedback and incentive mechanisms. Table 4 indicates that the application of artificial intelligence impacts teacher performance through feedback and incentives, with an indirect effect of 0.106 (95% CI [0.067, 0.158]). The direct effect continued to be significant, point out the evidence of partial mediation.

Table 4. Mediation Effect Analysis

Effect Type	Path	Effect Size	95% CI
Direct Effect	AI → Performance	0.198	[0.121, 0.273]
Indirect Effect	AI → Feedback → Performance	0.106	[0.067, 0.158]
Total Effect	AI → Performance	0.304	[0.219, 0.381]

5. DISCUSSION

The empirical outcomes firmly back the theoretical framework of the study (Figure 2). First, the positive effects of performance management tools and strategic alignment show that systems for evaluating teacher performance need to be integrated with the institutional strategy to work (Delery & Doty, 1996; Chadwick & Dabu, 2008; Armstrong, 2021). This is especially pertinent for undergraduate colleges in Anhui Province, where there has been a systemic lack of alignment between the evaluation systems and the institutional mission (Qu, 2024; Zhu, 2018).

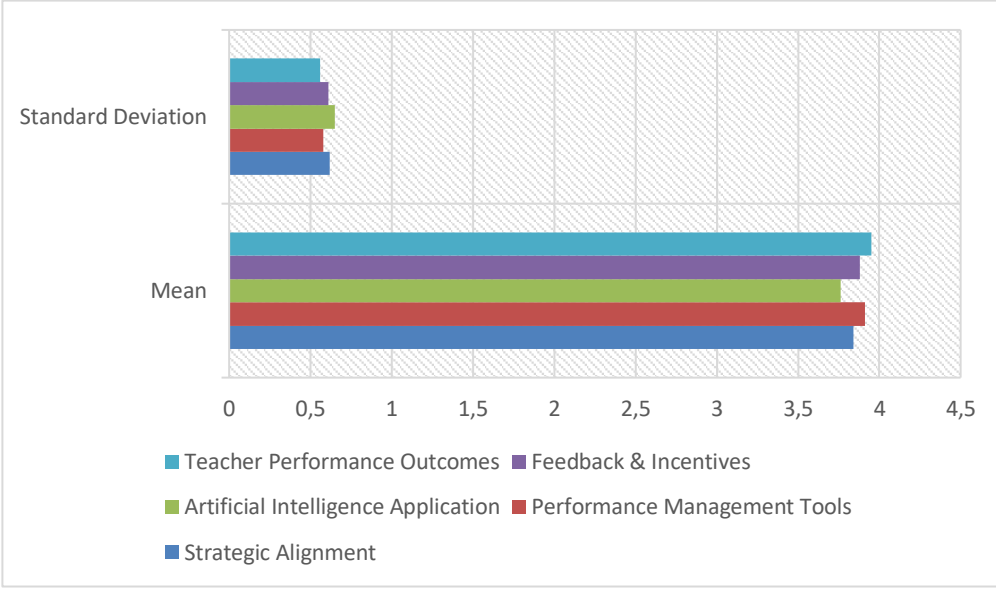


Figure 2. Mean Scores of Key Constructs

Secondly, the significant impact of artificial intelligence provides proof that evaluation systems, enhanced by digital technologies, can greatly increase accuracy, transparency, and responsiveness (Pei & Lu, 2023; Pham & Sampson, 2022; Wang et al., 2024). However, the analysis of mediation means that, in isolation, technology is inadequate. AI has the most impact when combined with good systems of feedback and positive reinforcement (Zhang et al., 2026; Bauer et al., 2025; Yildiz Durak & Onan, 2025). This supports the notion that evaluation systems should not be "judgmental" but should be "developmental" (Darling-Hammond et al., 2012; Alwaely et al., 2023).

Lastly, the findings indicate that SHRM-driven, AI-supported evaluation systems increase teacher motivation, and also help in fairness and professional growth

(Kamalov et al., 2023; Brougham & Haar, 2018). These are central to the enhancement of teaching and the sustainable development of undergraduate colleges during the higher education transition (OECD, 2013; Ministry of Education of the People's Republic of China, 2020).

6. FINAL CONSIDERATIONS

This paper has discussed how the implementation of the Artificial Intelligence (AI) in the performance evaluation system of teachers can revolutionize management of higher education institutions within the framework of undergraduate university in Anhui Province, China (Zhu, 2018; Qu, 2024). Relying on the substantive assumptions of Strategic Human Resource Management (SHRM), performance management, and developmental assessment (Delery & Doty, 1996; Chadwick & Dabu, 2008; Armstrong, 2021), the research shows that AI should not be seen as a technical facilitator but a strategic facilitator that can make institutional goals correspond to the individual academic achievement, facilitate evaluative quality, and encourage the ongoing professional growth (Alwaely et al., 2023; OECD, 2013).

Empirical data prove that AI-based teacher assessment systems help to make the correlations between institutional strategy and teaching outcomes considerably more robust (Pei & Lu, 2023; Zhang & Jia, 2024). According to the structural model, the implementation of AI has a positive impact on the performance of teachers both directly, by more precise, data-driven assessment, and indirectly, by enhancing feedback and incentive systems (Zhang et al., 2026; Wang et al., 2024; Bauer et al., 2025). These conclusions support the fact that technological capability should be combined with people-oriented management so that performance evaluations are used in the developmental but not the punitive way (Darling-Hammond et al., 2012; Kamalov et al., 2023).

Theoretically, the research will contribute to currently developing area of higher education management in that it will introduce and prove a structure of evaluative AI-driven, based on SHRM (Pham & Sampson, 2022; Chiu et al., 2023). This framework is a reformulation of teacher evaluation as a dead and summative process to a dynamic and goal-focused system with a lot of feedback (Luckin & Holmes, 2016; Yildiz Durak & Onan, 2025). The synthesis of the theory therefore transcends the technology vs. governance dichotomy and proves that both can be used correspondingly (Nguyen et al., 2023; Jarrahi, 2018).

In practice, the paper can offer policymakers and administrators of higher education practical information on the construction of AI-sustained evaluation systems, which are equitable, clear, and adaptive (Ministry of Education of the People's Republic of China, 2020). These systems are capable of minimizing levels of bias, identifying diverse forms of contributions like teaching, research, and service, and providing prompt advice to support the professional development (Kellogg et al., 2020; Wang, Wang, & Solhi, 2026). The results are particularly helpful to the institutions within Anhui Province but are also relevant to other areas where they are going through the assessment reforms and digitization (Kasman & Lubis, 2022).

However, there are multiple restrictions that should be mentioned. Although the quantitative design is highly empirically valid (Hu & Bentler, 1999; Kaiser, 1974; Kline, 1999), longitudinal or qualitative research design could be utilized in the future and help researchers observe behavioral changes across time and get a better understanding of the experiences of teachers who live in AI-enabled evaluative settings (Creswell, 1994). More so, the ethical governance, algorithmic transparency, and data privacy questions should be monitored

continuously to ensure that the AI systems do not uphold new types of inequity or surveillance (Nguyen et al., 2023).

To sum everything up, the implementation of artificial intelligence into teacher performance assessment is one of the key steps in reforming the sphere of higher education governance (Zawacki-Richter et al., 2019; Wang et al., 2024). AI-driven evaluative systems will be able to create a culture of equity, motivation, and academic excellence in a sustainable manner when they are oriented towards the strategic human resource management principles (Brougham & Haar, 2018; Delery & Doty, 1996).

REFERENCES

Alwaely, S. A., El-Zeiny, M. E., Alqudah, H., Mohammad Alamarnih, E. F., Ibrahim Salman, O. K., Halim, M., & Saleem Khasawneh, M. A. (2023). The impact of teacher evaluation on professional development and student achievement. *Environmental & Social Management Journal / Revista de Gestão Social e Ambiental*, *17*(7). https://doi.org/10.24857/rgsa.v17n7-022

Arcinas, M. M., Meenakshi, M., Bahalkar, P. S., Bhaturkar, D., Lalar, S., Rane, K. P., & Raghuvanshi, A. (2025). An efficient course recommendation system for higher education students using machine learning techniques. *Bulletin of Electrical Engineering and Informatics*, *14*(2), 1468–1475. https://doi.org/10.11591/eei.v14i2.7711

Armstrong, M. (2021). *Armstrong's handbook of human resource management practice* (15th ed.). Kogan Page.

Baek, C., & Doleck, T. (2023). Educational data mining versus learning analytics: A review of publications from 2015 to 2019. Interactive Learning Environments, *31*(6), 3828–3850. https://doi.org/10.1080/10494820.2021.1943689

Bauer, E., Sailer, M., Niklas, F., Greiff, S., Sarbu-Rothsching, S., Zottmann, J. M., & Fischer, F. (2025). AI-based adaptive feedback in simulations for teacher education: An experimental replication in the field. *Journal of Computer Assisted Learning*, *41*(1), e13123. https://doi.org/10.1111/jcal.13123

Brougham, D., & Haar, J. (2018). Smart technology, artificial intelligence, robotics, and algorithms (STARA): Employees' perceptions of our future workplace. *Journal of Management & Organization*, *24*(2), 239–257. https://doi.org/10.1017/jmo.2016.55

Chadwick, C., & Dabu, A. (2008). Human resources, human resource management, and the competitive advantage of firms: Toward a more comprehensive model of causal linkages. *Organization Science*, *20*(1), 253–272. https://doi.org/10.1287/orsc.1080.0375

Chandrakant, S. M. (2025). AI-powered teaching assistants: Enhancing educator efficiency with NLP-based automated feedback systems. *International Journal of Science and Research Archive*, *14*(3), 9–18.

Chen, W., & Min, L. (2025). The evaluation model of physical education teaching performance based on deep learning algorithm. Scalable Computing*: Practice and Experience*, *26*(1), 458–466. https://doi.org/10.12694/scpe.v26i1.3858

Chiu, T. K., Xia, Q., Zhou, X., Chai, C. S., & Cheng, M. (2023). Systematic literature review on opportunities, challenges, and future research recommendations of artificial intelligence in education. *Computers and Education: Artificial Intelligence*, *4*, 100118. https://doi.org/10.1016/j.caeai.2023.100118

Creswell, J. W. (1994). Research design: Qualitative, quantitative, and mixed methods approaches. *Sage*.

Darling-Hammond, L., Amrein-Beardsley, A., Haertel, E., & Rothstein, J. (2012). *Evaluating teacher evaluation*. *Phi Delta Kappan*, *93*(6), 8–15. https://doi.org/10.1177/003172171209300603

Delery, J. E., & Doty, D. H. (1996). Modes of theorizing in strategic human resource management: Tests of universalistic, contingency, and configurational performance predictions. *Academy of Management Journal*, *39*(4), 802–835. https://doi.org/10.2307/256713

Hu, L. T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis. Structural Equation Modeling, *6*(1), 1–55. https://doi.org/10.1080/10705519909540118

Jarrahi, M. H. (2018). Artificial intelligence and the future of work: Human–AI symbiosis in organizational decision making. *Business Horizons*, *61*(4), 577–586. https://doi.org/10.1016/j.bushor.2018.03.007

Kaiser, H. F. (1974). An index of factorial simplicity. Psychometrika, *39*(1), 31–36. https://doi.org/10.1007/BF02291575

Kamalov, F., Santandreu Calonge, D., & Gurrib, I. (2023). New era of artificial intelligence in education: Towards a sustainable multifaceted revolution. Sustainability*, 15*(16), 12451. https://doi.org/10.3390/su151612451

Kasman, K., & Lubis, S. K. (2022). Teachers' performance evaluation instrument designs in the implementation of the Merdeka curriculum. *Jurnal Kependidikan*, *8*(3), 760–775. https://doi.org/10.33394/jk.v8i3.5674

Kellogg, K. C., Valentine, M. A., & Christin, A. (2020). Algorithms at work: The new contested terrain of control. *Academy of Management Annals*, *14*(1), 366–410. https://doi.org/10.5465/annals.2018.0174

Klimova, B., & Pikhart, M. (2026). Use of artificial intelligence in foreign language education in higher education settings. *In* Social robots and artificial intelligence in education (pp. 327–351). Springer.

Kline, R. B. (1999). Book review: Psychometric theory (3rd ed.). *Journal of Psychoeducational Assessment*, *17*(3), 275–280. https://doi.org/10.1177/073428299901700307

Lampropoulos, G., & Evangelidis, G. (2025). Learning analytics and educational data mining in extended reality environments: A systematic review. Applied Sciences, 15*(2), 971.* https://doi.org/10.3390/app15020971

Leguina, A. (2015). A primer on partial least squares structural equation modeling (PLS-SEM). *International Journal of Research & Method in Education*, *38*(2), 220–221. https://doi.org/10.1080/1743727X.2015.1005806

Likert, R. (1932). A technique for the measurement of attitudes. Archives of Psychology*, 140*, 1–55.

Luckin, R., & Holmes, W. (2016). Intelligence unleashed: *An argument for AI in education*. Open University.

Ministry of Education of the People's Republic of China. (2020). Overall plan for deepening education evaluation reform in the new era. http://www.moe.gov.cn

Nguyen, A., Ngo, H. N., Hong, Y., Dang, B., & Nguyen, B. P. T. (2023). *Ethical principles for artificial intelligence in education*. *Education and Information Technologies*, *28*(4), 4221–4241.

OECD. (2013). Teachers for the 21st century: Using evaluation to improve teaching. OECD Publishing. https://doi.org/10.1787/9789264193864-en

Pei, Y., & Lu, G. (2023). Design of an intelligent educational evaluation system using deep learning. IEEE Access, *11*, 29790–29799. https://doi.org/10.1109/ACCESS.2023.3260979

Pham, S. T., & Sampson, P. M. (2022). The development of artificial intelligence in education: A review in context. *Journal of Computer Assisted Learning*, *38*(5), 1408–1421. https://doi.org/10.1111/jcal.12687

Podsakoff, P. M., MacKenzie, S. B., Lee, J. Y., & Podsakoff, N. P. (2003). Common method biases in behavioral research. *Journal of Applied Psychology*, *88*(5), 879–903. https://doi.org/10.1037/0021-9010.88.5.879

Qu, H. M. (2024). Campus education administration, teacher motivation, and performance management in China. *International Journal of Research Studies in Management*, *12*(16). https://doi.org/10.5861/ijrsm.2024.1319

Wang, S., Wang, F., Zhu, Z., Wang, J., Tran, T., & Du, Z. (2024). Artificial intelligence in education: A systematic literature review. Expert Systems with Applications, *252*, 124167.

Wang, S., Wang, Y., & Solhi, M. (2026). Teacher–student interpersonal behaviours and GenAI feedback on students' achievement emotions in EFL writing. *European Journal of Education*, *61*(1), e70395.

Yildiz Durak, H., & Onan, A. (2025). A systematic review of AI-based feedback in educational settings. *Journal of Computational Social Science*, *8*(4), 96.

Zand, J., Suvakov, M., Overgaard, S. M., Kozikowski, C. G., Jimison, R. C., Nelson, K. B., & Garovic, V. D. (2026). Building artificial intelligence and data literacy across the health care research workforce. Mayo Clinic Proceedings, *101*(1), 17–25.

Zawacki-Richter, O., Marín, V. I., Bond, M., & Gouverneur, F. (2019). Systematic review of research on artificial intelligence applications in higher education. *International Journal of Educational Technology in Higher Education*, *16*(1), 39. https://doi.org/10.1186/s41239-019-0171-0

Zhang, C., Hu, M., Wu, W., Chen, Y., Wang, K., Gao, T., & Fu, Q. (2026). AI versus teacher feedback in developing pre-service teachers' teaching self-efficacy. *Teaching and Teacher Education*, *172*, 105355.

Zhang, T., & Jia, S. (2024). Construction and implementation of application-oriented undergraduate teaching evaluation systems. *International Journal of Web-Based Learning and Teaching Technologies*, *19*(1), 1–19. https://doi.org/10.4018/IJWLTT.338386

Zhu, Y. (2018). Research on the reform of Chinese university teachers' evaluation system. *In* Proceedings of the International Conference on Education Science and Social Development. https://doi.org/10.2991/essd-18.2018.50