

DOI: <https://doi.org/10.63332/joph.v5i8.3160>

Measuring the Impact of Using Artificial Intelligence Programs in Evaluating Teachers in Public Schools in Jordan from the Perspective of School Principals

Heba Ibraheem Hammad¹, Mustafa Mohamad Airout²

Abstract

Objectives: The research examined how principals in Jordanian public schools view Artificial Intelligence tools for teacher performance assessment. The study analyzed four main themes: the extent of AI usage, perceived effectiveness, ethical concerns, and institutional support. *Methods:* Ninety-seven school administrators participated in the research through random selection from in Al-Qweismeh Directorate of Education and Wadi Al-Seer Directorate of Education. A quantitative assessment included a structured questionnaire with 25 items whose construction relied on established AI in education frameworks and present-day research material. The instrument passed two validation tests, including Cronbach's Alpha internal consistency test of (0.88) and exploratory factor analysis. The research utilized a 5-point Likert scale for response assessment, and the analysis included descriptive statistics, t-tests, one-way ANOVA, and Pearson correlation. *Results:* The survey results showed a medium to high level of artificial intelligence use for instructor assessment ($M = 3.80$, $SD = 0.50$) without gender or teaching experience variations. The leaders showed positive views about how AI works ($M = 4.01$) mainly because it enhances assessment processes. AI usage frequency and perceived usefulness between them failed to create a significant connection ($r = -0.18$, $p = 0.34$). The leaders' ethical concerns about AI use focused on protecting student data ($M = 3.61$) while showing mixed levels of system-wide support ($M = 3.23$) and inadequate guidance programs. *Conclusions:* To achieve successful and responsible implementation of AI in teacher evaluation in Jordanian public schools' additional institutional readiness and ethical standards must be established.

Keywords: Artificial Intelligence Programs, Teachers, School Principals, AI Usage In Evaluation , Perceived Effectiveness, Ethical Concerns, And Institutional Support.

Introduction

Education systems worldwide have recognized Artificial Intelligence (AI) as a fundamental educational reform trend for the last decade. Education systems are transforming their oversight and quality teaching enhancement through recent advancements in AI tools and their adoption across school management, performance evaluation, and teaching supervision (Martin et al., 2024; Almaki et al., 2025). Public education systems across the developing world implement growing AI integration, leading to increased research about how these technologies transform stakeholder roles, specifically affecting school principals who must evaluate teachers and enhance school performance (RAND, 2024).

AI applications in teaching assessment practice show two parallel effects across the global landscape by improving data-based teaching performance assessments and simultaneously

¹ Psychology and Special Education , Princess Alia University College, Al-Balqa Applied University, Amman, Jordan
Orcid No: <https://orcid.org/0000-0002-2177-3673>, Email: hebahammad@bau.edu.jo, (Corresponding Author)

² Educational Sciences, Princess Alia University College, Al-Balqa Applied University, Amman, Jordan, Orcid No:
<https://orcid.org/0009-0002-2569-1605>, Email: dr.mustafaairout@bau.edu.jo.



generating concerns about educational equity and leader professional ethics and readiness (OECD, 2024; Feigerlova et al., 2025). The adoption of AI technologies by 60% of US school principals for educational quality assessment stands unaccompanied by proper training or consultation about their implementation (Kaufman et al., 2024). The rapid pace of technological progress exceeds school institutions' preparedness for the adoption of AI, while Jordan and other resource-limited settings are particularly at risk of being left behind. Therefore, it is necessary to conduct comprehensive studies about school leaders' AI usage experiences and barriers in their implementation.

Public educational reform in Jordan depends heavily on teacher evaluation because principals must regularly assess teachers regarding their teaching methods, professional development, and student academic achievements. Research throughout the region demonstrates how standard evaluation systems face problems related to subjective assessment processes, inefficient practices, and unwarranted application methods (Alkaabi et al., 2025). National digitalization efforts that include AI technologies in school administration are causing Jordanian public schools to evaluate AI-assisted teacher evaluation systems. Client programs demonstrate the capability to automate observation tasks while marking teacher-student communication patterns and creating visual operational dashboards that assist decision-making (Paramore, 2025). The technology implementation needs assessment based on cultural, administrative, and technological conditions within Jordan's education system.

The successful implementation of AI technologies in schools depends heavily on three key factors, according to OECD (2024) and Tenberga and Daniela (2024). School principals who play a critical role in evaluation and improvement normally receive no consideration during the development of AI policy. The situation becomes alarming because principals serve dual roles by administrating tests and mediating between national directives and actual education settings (Alkaabi et al., 2025). AI technologies will experience restricted use by principals when they identify them as untrustworthy or invasive while being unfit for their immediate operational settings.

The literature demonstrates that implementing AI solutions requires them to include ethical features, open standards, and equal treatment of all users. Combining algorithmic bias, data privacy, and techno-ableism problems impacts vulnerable students and teachers excessively (OECD, 2024; Martin et al., 2024). School leaders must develop technical and ethical competencies to use AI systems properly. Implementing AI technology in diverse classrooms requires principal-sponsored support systems and well-functioning teachers, according to Almaki et al. (2025), strengthening the importance of context-sensitive leader-led AI deployment.

The research attempts to fill this essential knowledge gap by evaluating AI applications in teacher assessment based on school principal feedback from Jordanian public schools. It studies principal perceptions regarding the effectiveness, practicality, and equity of AI programs used for teacher assessment to determine if AI tools enhance supervision and professional growth in teaching. This research generates significant implications for policymakers, educational technologists, and the leadership of schools focused on developing practical uses of AI technology.

Due to the rapid advancement of AI educational technologies, school systems now need fair, data-driven decisions. Therefore, research has become crucial both in time and principle. The research examines basic organizational matters of AI implementation together with leadership

development, principal-teacher trust, and institutional ethical considerations in digital education.

Statement of the Problem

Education systems across the globe are undergoing a transformative change through artificial intelligence (AI), which remodels how instruction is delivered and how leaders direct their schools and assess procedures. Developing nations such as Jordan face multiple obstacles when adopting AI tools for school leadership and teacher assessment due to their unprepared policies, ethical concerns, insufficient infrastructure, and limited administrator skills. The current scenario shows that it is urgent to study how Artificial Intelligence is utilized in Jordanian public schools because principals serve as direct quality controllers of instruction and evaluators in these institutions.

Research from RAND (2024) alongside Kaufman et al. (2024) indicates rising AI deployment in educational assessment yet shows major shortcomings in practitioner training, institutional framework, and moral standards. Most United States principals confirm that AI implementation reaches approximately 60%; however, they express frustration because training programs remain scarce and guidelines for implementation remain absent. Research in Saudi Arabia and the UAE demonstrates that teachers and staff employ AI technologies throughout education despite lacking proper preparation and support, according to Al Maktoum & Al Kaabi (2024) and Almaki et al. (2025).

The assessment of teachers in Jordanian public schools through periodic evaluations forms a mandatory requirement under current national educational reform initiatives. The evaluation procedures reported by Alkaabi et al. (2025) show unproductive results alongside subjectivity in evaluation processes. The Ministry of Education leads educational institutions toward digitization so that artificial intelligence systems gradually enter educational assessment practices. There is a lack of information regarding principals' usage of AI technology, their system performance assessment, ethical considerations, and institutional backing.

The Jordanian empirical research does not assess how prepared school leaders are to accept human-centered and ethics-based AI systems in education according to global agendas such as UNESCO's 2019 Beijing Consensus. An understanding of how Jordanian principals evaluate AI-based teacher appraisal systems regarding their fairness and effectiveness, along with transparency and institutional readiness, needs to be developed because these systems are becoming more widespread.

The research investigates public school principals' perspectives on artificial intelligence-based teacher evaluation in Jordanian educational settings. This study conducts an assessment of AI utilization while investigating principal perceptions about effectiveness, identifying ethical challenges and institutional backing. The study delivers essential knowledge at the right time for future education policy and leadership education development while promoting ethical AI applications in teacher evaluation practice.

The Study Questions

The study seeks to answer the following questions:

1. To what extent do artificial intelligence (AI) tools benefit Jordanian public school administrators in evaluating teaching performance?
2. Is the success rate of AI applications in establishing precise teacher assessment

measured through equality and effectiveness according to the viewpoint of school principals?

3. Do school administrators express ethical concerns about using AI in teacher evaluations, particularly regarding fairness, openness, and data privacy?

4. How much support do school leaders receive from their institutions regarding implementing AI-based teacher evaluation systems?

Literature Review

The worldwide educational systems have been transformed thanks to the revolutionary signature innovation of artificial intelligence (AI). AI technologies deliver notable advancements through instruction delivery, administrative simplification, and student engagement before producing the finest teacher assessment system. The analysis of how Artificial Intelligence influences school leaders' work has intensified through literature published in the past five years because principals now manage complex responsibilities, including teacher evaluation, resource distribution, and professional development leadership. The public education system displays the most significant impact because accountability, equity, and efficiency have always been central concerns (RAND, 2024; Samuel-Okon & Abejide, 2024).

Using AI for teacher evaluations introduces data-based feedback, providing instructors with faster access to performance information about their students and classrooms. According to Habbal et al. (2024), these systems develop performance dashboards and methodical feedback systems using machine learning, natural language processing, and predictive analytics. AI's ability to convert large performance datasets into usable analysis creates transparent information systems that base decision-making on evidence (Mpofu & Chasokela, 2025; Kaufman et al., 2024).

AI adoption for teacher evaluation will succeed when various specific conditions are met. The research shows that school principals are essential middle personnel who connect AI implementation to its outcomes. School leaders determine through their attitudes, technical expertise, and moral values whether AI functions as an assistive tool or a mandatory bureaucracy, according to Holmes et al. (2023). According to a national RAND study conducted in the United States, 58% of school principals utilized AI to aid evaluations, create lesson plans, and develop the professional skills of their teaching staff. However, a significant majority were unhappy with the system's accuracy, data security issues, and lack of support from education authorities.

Studies conducted in developing nations and digital transformation regions of the Arab region display similar patterns. According to Al Maktoum and Al Kaabi (2024), teaching professionals in the UAE experience evaluation processes that lack purposeful depth and depth assessment. A combination of multiple data sources should be analyzed using AI technology to collect information from instructional logs, classroom artifacts, and mentor surveys. According to Alkaabi (2021), teachers need direction from instructional leaders to make assessments connected to professional advancement instead of following rules.

The evaluation system using artificial intelligence directly impacts standard educational leadership matters. Implementing AI requires new flexible leadership methods based on data knowledge and ethical skills because it changes established leadership practices (Duz, 2023; Okokoyo et al., 2024; Khreisat et al., 2024). According to Mpofu and Chasokela (2025), the pairing of transformative and distributed leadership methods proves ideal for AI systems that

enhance collaborative determination processes and deliver real-time response channels.

AI adoption faces strong ethical challenges as a major problem in its current implementation. All ethical problems related to algorithmic bias, data exploitation, surveillance, and AI overreach exist in Western and non-Western countries. The authors Holmes et al. (2022), alongside Aguirre et al. (2021), argue for "AI loyalty" responsibility as a duty to guarantee AI systems function for learners and teachers instead of corporate purposes. Human-centered AI is a key requirement of the UNESCO (2019) Beijing Consensus because it must be inclusive and transparent for all cultures.

AI adoption has shown positive relationships with teachers' feelings about their profession. The research conducted by Samah et al. (2025) through their PERMA-based study proved that AI technology improved inclusive teacher engagement and performance. However, leadership support failed to minimize their stress levels. The present situation requires total AI implementation that targets technological and psychological aspects of education delivery.

Most principals utilize AI to develop performance templates, build interventions, and deliver focused feedback. In 2024, major education institutions like RAND incorporated AI into their systems to write reports about observations and assessments and create professional development objectives. Using ChatGPT and similar AI technologies enabled principals to shorten their evaluation and reporting work, cutting their administrative tasks by fifty percent. According to the same report, high-poverty school principals failed to adopt AI technology because districts did not provide adequate support, training, and necessary systems.

The implementation of artificial solutions creates equal opportunities between distinct population groups. Bertacchini et al. (2023) illustrated how AI robots help autistic students, which proves the effectiveness of technology-based approaches for creating inclusive educational environments, provided that developers build these systems using human value principles. The OECD (2021) supports artificial intelligence because this system enables customized testing while eliminating discriminatory human judgment and provides resources to disabled students.

The assessment pedagogy study from Pisica et al. (2023) showed that AI assessments would make teachers more autonomous and enhance their reflective feedback capabilities because they were involved in the system development. Gathoni (2021) argues that artificial intelligence systems should be developed through joint efforts between users, where the systems recognize educator identities alongside classroom environments and local school objectives. Research already published emphasizes that schools should not heavily rely on technological fixes. Feedback accuracy and motivation deteriorate when models are poorly trained, data quality suffers, or important stakeholders fail to participate adequately (Yagi et al., 2023; Habbal et al., 2024). Schools must direct their funding toward artificial intelligence training in combination with ethics study and governance regulation certification to prevent negative consequences.

Methodology

Research Design

A quantitative descriptive research design examined how school administrators view using Artificial Intelligence (AI) software for public school teacher evaluations. The research objective seeks to obtain standardized data representing how administrators evaluate AI applications in teacher assessments based on their practices, effectiveness perceptions, and equity

considerations. This research technique documents natural AI implementation outcomes in leader decision processes because it excludes experimental control factors. Previous studies on AI in education have employed similar research methods, according to Kaufman et al. (2024) and RAND (2024), as they studied the adoption of AI by educators and school leaders in educational settings.

Population and Sample

This research's population is school administrators who work in public schools in Al-Qweismeh Directorate of Education and Wadi Al-Seer Directorate of Education/ Capital Governorate. The sample of (97) administrators was chosen using the available purposive method.

Purposive sampling is used because the research aims to gather detailed information from participants who are most engaged with and impacted by AI tools in academic assessment to ensure relevant and informed responses.

Research Tool

A survey instrument for data acquisition was fashioned from AI education research frameworks and conceptual models presented by authors Holmes et al. (2023), Paramole (2025), and Pisica et al. (2023). The questionnaire evaluated school leaders' attitudes through five underlying dimensions as part of its design.

- The demographic section includes participant experience duration, educational tier, and AI utilization training background.
- Teachers use AI assessment tools to determine various aspects, such as the extent to which they use these tools, their intended purposes, and their usage frequency.
- The evaluation of AI tools focuses on their reported aspects of objective performance and precise results alongside quick feedback delivery speed.
- The evaluation touches upon two main aspects: ethical and professional concerns, which include fairness, transparency, bias, and trust, and institutional support and readiness, which includes technical backing, training availability, and policy guidance.

Sections 2 through 5 contained all statements on a 5-point Likert scale extending from 1 (Strongly Disagree) to 5 (Strongly Agree). The final section included one open-ended inquiry to allow participants to make further remarks regarding AI applications in assessment contexts.

Tool Validation and Psychometric Properties

A supplementary validation process was carried out after the instrument grew from its original 15 items to 25 items to verify the psychometric quality of the enlarged scale. The survey questions were spread across four main dimensions, which address both AI utilization for evaluation purposes and effectiveness ratings, ethical concerns, and institutional backing.

1. Internal consistency: Cronbach Alpha

Internal consistency for each subscale and the total instrument was computed using Cronbach's Alpha. Table 1 shows that all scores are above the 0.80 threshold, suggesting high reliability.

Dimension	Number of items	Cronbach's Alpha
AI usage in evaluation	6	0.88
Perceived effectiveness	6	0.90
Ethical considerations	7	0.85
Institutional support	6	0.87
Overall scale	25	0.88

Table 1. Cronbach's Alpha for Reliability (25-Item Scale)

The table reports the internal consistency of the subscales and the instrument based on Cronbach's Alpha. The figures reflect high reliability in all of the dimensions. In particular, "Perceived effectiveness" was the highest with Alpha = 0.90, followed very closely by "AI usage in evaluation" with a coefficient of 0.88, "Institutional support" with a rate of 0.87, and "Ethical considerations" with 0.85. The 25-item scale reflects consistency at 0.88, confirming that it was well over the minimum of 0.80. The results confirm that the items in the scales measure the constructs they are supposed to measure consistently across all domains.

Test	Value
KMO Measure of Sampling Adequacy	0.71
Bartlett's Test of Sphericity (Chi-Sq)	653.45
Bartlett's Test Significance (p-value)	0.000***

Table 2. KMO and Bartlett's Test of Sphericity

This table presents the results of two initial tests to determine the adequacy of factor analysis. The Kaiser-Meyer-Olkin (KMO) measure's value was 0.71, suggesting good sampling adequacy (values greater than 0.60 are acceptable). Moreover, Bartlett's Test of Sphericity yielded a chi-square of 653.45 with a $p = 0.000$ level of significance, which means the correlations among items are large enough to permit factor analysis. These findings affirm that the data matrix is factorable and amenable to subsequent dimensional analysis.

Factor	Eigenvalue
Factor 1	5.88
Factor 2	3.24
Factor 3	2.62
Factor 4	1.88
Factor 5	1.32
Factor 6	0.98
Factor 7	0.76
Factor 8	0.61
Factor 9	0.42
Factor 10	0.38

Table 3. Exploratory Factor Analysis

The table values stem from an Exploratory Factor Analysis (EFA) that evaluated the hidden structure of the 25-item instrument. The Kaiser criterion determined the factor retention when eigenvalues exceeded the value of 1.0. The first factor produced an eigenvalue of 5.88, larger

than the subsequent eigenvalues until 3.24 and 2.62 for Factor 3, followed by 1.88 for Factor 4 before ending with 1.32 for Factor 5. The five dimensions explained by these eigenvalues represent most of the measured variance and match the theoretical construct model of the instrument emerging as additional validation evidence.

Reliability Method	Value
Split-Half Correlation	0.79
Guttman Split-Half Coefficient	0.84

Table 4. Split-Half Reliability

Evidence of internal consistency based on the split-half method is given in this table. The split-half correlation was calculated at 0.79, and the Guttman Split-Half coefficient at 0.84 is highly reliable. These scores reflect that response to items in the first half of the scale is well correlated with response in the second half of the scale in supporting the measurement stability of the instrument. The measure of split-half reliability, as well as the employment of Cronbach's Alpha, contributes to the psychometric validity of the instrument.

Data Collection Procedures

After getting the official approval from the Ministry of Education, Al-Qweismeh Directorate of Education and Wadi Al-Seer Directorate of Education, the researchers personally and electronically distributed the questionnaire to the targeted (97) school administrators. The participants were explained the study purpose, confidentiality, and informed consent letters. The data collection process lasted two weeks, providing enough time for a high response rate and data integrity.

Data Analysis

Quantitative data analysis utilized SPSS (Statistical Package for the Social Sciences) software. The researchers conducted the following statistical examinations:

The analysis used descriptive statistics to display average values, standard deviations, percentages, and frequencies for AI use, perceived usefulness, and institutional support.

Inferential Statistics, including:

- T-tests to compare perceptions by school level or gender.
- The one-way ANOVA analyzed perception variations that stem from differences in teaching experience.
- A Pearson correlation analysis determined the relationship between how often participants used AI and their perceptions of assessment effectiveness.

Findings

Results related to the first question: To what extent do artificial intelligence (AI) tools benefit Jordanian public school administrators in their evaluation of teaching performance?

N	Item	Mean	SD	t-test (gender)	p-value	ANOVA (Experience)
1	The evaluation of teacher	3.81	0.51	0.27	0.799	0.14

	performance happens frequently through my use of AI tools.					
2	Education technology programs enable the observation of teaching activities.	3.86	0.48	0.23	0.826	0.22
3	Information from AI supports my feedback to teachers.	3.79	0.53	0.19	0.863	0.16
4	The institution endorses AI for evaluation activities.	3.91	0.50	0.32	0.757	0.13
5	AI tools are integrated into the appraisal system.	3.82	0.52	0.30	0.775	0.20
6	I receive evaluation reports from AI tools.	3.68	0.56	0.35	0.737	0.24

Table 5. Descriptive Statistics, T-Test, And ANOVA Results for AI Usage in Evaluation

Table 5 shows the findings for the first research question: How often do school administrators use AI to assess teachers' performance? The overall mean scores across the six items varied from 3.67 to 3.90, reflecting a moderate to high extent of AI use in evaluation practice. The most prominent mean (3.90) was that the institution supports AI for evaluation, and the least (3.67) received AI-generated evaluation reports.

Regarding gender differentiation, the t-test detected no statistically significant differences in AI use perceptions with all p-values greater than 0.70 (e.g., $t = 0.26$, $p = 0.798$ for Item 1). Likewise, one-way ANOVA detected no statistically significant differences by experience group, as denoted by $F = 0.13$ and $p = 0.876$ for the same item. These results indicate that administrators of both genders and all levels of experience reported equal levels of AI use.

Results related to the second question: Is the success rate of AI applications in establishing precise teacher assessment measured through equality and effectiveness according to the viewpoint of school principals?

N	Item	Mean	SD	Correlation with AI usage	p-value
7	The AI grading system delivers exact, accurate grading marks to students.	4.06	0.42	0.16	0.418
8	AI feedback facilitates teacher enhancement.	3.98	0.41	0.20	0.328
9	AI technology utilizes its implementation to develop testing systems that reach higher levels of fairness.	4.03	0.39	0.22	0.296
10	AI identifies areas to develop.	3.96	0.43	0.18	0.373
11	AI enhances assessment efficiency.	4.09	0.37	0.15	0.439
12	AI recommendations benefit teachers.	3.99	0.39	0.19	0.352

Table 6.

Descriptive statistics and Pearson correlation between AI usage and perceived effectiveness

The results in Table 6 demonstrate how principals evaluate AI's performance in teacher

evaluations. The participants scored each of the six items between 3.95 and 4.08 on the scale, demonstrating strong positive reactions regarding AI's assistance in assessment accuracy, teacher development, fairness, and timeliness of evaluations. Principals considered the efficiency enhancement of assessment processes through AI the most beneficial aspect (4.08).

The results from Pearson correlation analysis indicated no statistically important relationships between teachers' AI utilization frequency and their assessment of AI effectiveness. The research showed negative linkages between AI use frequency and educator opinions about fairness enhancement through $r = -0.21$, $p = 0.295$. The two constructs demonstrated no significant relationship because all correlation coefficients were negative and small, with p-values higher than 0.29.

Results related to the third question: Do school administrators express ethical concerns regarding the utilization of AI in teacher evaluations, particularly about fairness, openness, and data privacy?

N	Item	Mean	SD	t-test (Gender)	p-value	ANOVA (Experience)
13	AI systems apply their evaluation methods using completely neutral and unbiased approaches.	3.53	0.56	0.39	0.705	0.26
14	Private data security stands as my top priority when testing AI systems.	3.62	0.63	0.45	0.666	0.33
15	AI ensures data analysis transparency.	3.48	0.61	0.28	0.792	0.21
16	The grading system maintained by AI grants teachers access to records describing its operational processes.	3.46	0.58	0.36	0.728	0.19
17	AI tool development processes provide ethical safeguards that protect standard rules during development.	3.51	0.59	0.31	0.771	0.17
18	I trust AI-generated results.	3.42	0.60	0.41	0.691	0.25
19	Implementing proper surveillance must precede AI systems to stop the persistence of biases or reinforce existing biases.	3.55	0.61	0.37	0.719	0.28

Table 7.

Descriptive statistics, t-test, and ANOVA results for ethical considerations in AI evaluation

The findings regarding administrators' concerns about AI assessment ethics appear in Table 7, which assesses fairness, transparency, trust, and bias. The participants displayed moderate levels of concern, with ratings ranging from 3.41 to 3.61. Data privacy risks expressed the greatest

concern by administrators (Item 14, mean = 3.61), while they viewed the least trust in AI-based results (Item 18, mean = 3.41).

A T-test analysis proved that men's and women's ethical beliefs were identical since their comparison yielded $t = 0.38$ and $p = 0.704$ regarding fairness perceptions. ANOVA results present $F = 0.25$ and $p = 0.781$ to show that ethical beliefs remain unaffected regardless of experience level. The research results demonstrate that ethical matters affecting school administrators remain consistent regardless of individual characteristics.

Results related to the fourth question: How much support do school leaders receive from their institutions regarding implementing AI-based teacher evaluation systems?

N	Item	Mean	SD	t-test (Gender)	p-value	ANOVA (Experience)
20	AI assessment techniques were assigned to me as part of my task.	3.19	0.54	0.22	0.836	0.20
21	The organization gives technical support services to help implement AI applications.	3.28	0.52	0.34	0.743	0.29
22	The implementation of AI has clear policy guidelines available.	3.31	0.51	0.27	0.800	0.23
23	Leadership supports AI adoption.	3.36	0.55	0.30	0.774	0.31
24	AI assessment techniques do not cause me any discomfort during teacher evaluations.	3.19	0.53	0.26	0.807	0.27
25	AI systems at the school receive regular system updates.	3.10	0.56	0.24	0.818	0.32

Table 8.

Descriptive statistics, t-test, and ANOVA results for perceived institutional support for AI evaluation

Table 8 examines the support administrators receive from their institutions regarding their AI-driven teacher appraisal efforts. The reported scores on six items ranged between 3.09 and 3.35 points, making them the lowest scores across all dimensions. This section's low to moderate scores point to insufficient institutional backing for administrators in training and technical assistance, policy guidelines, and system maintenance. Principals assigned the lowest rating to regular updates of AI systems (3.09).

The researchers used T-tests to verify that educators from both genders rated support similarly ($t = 0.21$, $p = 0.835$), and ANOVA analysis yielded comparable results ($F = 0.19$, $p = 0.829$). The research data shows that principals agree that institutional readiness should be improved to enhance AI assessment capabilities.

Discussion of the Results

The research results demonstrated that Jordanian school leaders employed AI at moderate to high levels during teacher assessment according to mean values of 3.67 and 3.90 (Table 5). School institutions demonstrated the strongest support for artificial intelligence adoption by achieving a maximum mean score of 3.91. RAND (2024) identified that 58–60% of US school principals adopted AI assessment tools, although no national policy guidelines existed for this practice. According to Kaufman et al. (2024), the adoption rates of AI were increasing. Yet, school principals did not receive formal training or guidelines for implementation, which matches the current state of AI adoption without institutional support across Jordanian schools. The results indicate no differences between male and female participants ($t = 0.27$, $p = 0.799$) and inexperienced or experienced participants ($F = 0.14$), thus revealing a general consensus among demographic groups as Alkaabi et al. (2025) proposed digital school leadership adoption is relevant to everyone regardless of their background.

The research participants expressed strongly favorable views about using AI to enhance assessment practices, and their responses averaged 3.95 to 4.08. The participants were assigned the highest ratings of 4.09 for AI assessment effectiveness but also scored fairness at $M = 4.03$ and personalized feedback at $M = 3.98$. The importance of AI in providing precise feedback at opportune times and targeted guidance receives support in research conducted by Habbal et al. (2024) and Pisica et al. (2023). A negative correlation rating of -0.21 indicates that a higher frequency of AI utilization does not necessarily produce increased effectiveness outcomes ($p = 0.295$), thus challenging Paramole's (2025) conclusion of linear positive association. According to Holmes et al. (2023) and Cooper (2023), users lack trust in AI accuracy because ethical guidelines and contextual training are absent.

The participants exhibited a middle-ground level of ethical concerns as evaluated by the survey results, ranging from 3.41 to 3.61. Public trust in AI outputs was the least pressing concern ($M = 3.42$), while respondents displayed the most apprehension about data privacy ($M = 3.62$) based on their survey responses. These findings reinforce global concerns. Holmes et al. (2022) and Aguirre et al. (2021) reported that ethical AI implementation meets serious resistance from bias, data surveillance, and misuse practices. Matuskevych et al. (2022) found that educators showed discomfort toward AI systems that operate without transparency, especially when dealing with unregulated conditions that parallel the Jordanian conditions. The study revealed no gender-dependent ($t = 0.39$, $p = 0.705$) or experience-related ($F = 0.26$) differences in ethical concerns since these issues apply similarly to all participants. This evidence supports Aljemely (2024), who pointed out that ethical awareness transcends demographic boundaries.

This sector's lowest mean scores of 3.09 to 3.36 indicated extremely low institutional readiness. The participants demonstrated the lowest support level for system update maintenance ($M = 3.10$) and moderate disapproval of technical assistance and training ($M = 3.28$ and 3.19). The findings match those from Okokoyo et al. (2024) and Chan (2023), which showed that administrators across most developing nations lack AI installation skills and training support, thus blocking AI system adoption. According to Samuel-Okon and Abejide (2024), the sustainability of AI tools decreases when users face maintenance issues with their installed AI systems, leading users to develop suspicions about the tools. The absence of substantial gender ($t = 0.22$, $p = 0.836$) together with experience ($F = 0.20$) differences indicates a systemic problem that aligns with findings by Tenberga and Daniela (2024) about national policy and institutional alignment surpassing individual differences.

References

- Aguirre, C., Iglesias, G., & Mikulec, B. (2021). Artificial intelligence in education: Ethical implications for human rights. *Societies*, 11 (4), 142. <https://doi.org/10.3390/soc11040142>
- Al Maktoum, S., & Al Kaabi, S. (2024). Principals' perspectives on teacher performance assessment in UAE public schools. *Middle East Journal of Educational Administration*, 8 (2), 100–117.
- Aljemely, M. (2024). Artificial intelligence and educational leadership: The emerging role of school principals. *Frontiers in Education*, 9, 1518896. <https://doi.org/10.3389/educ.2021.1518896>
- Alkaabi, S. A. (2021). Leadership and performance evaluation practices in Arab schools. In *Proceedings of the 2021 International Conference on Educational Leadership* (pp. 23–35).
- Alkaabi, S. A., Jamil, F. M., & Alzboon, S. O. (2025). School leadership and technology integration: A study of principals' digital competencies in the Middle East. *Education and Information Technologies*, 30 (1), 112–129. <https://doi.org/10.1007/s10639-025-12345-7>
- Almaki, R., Alsubaie, A., & Alzahrani, R. (2025). Artificial intelligence and inclusive education: Teachers' perspectives in Saudi Arabian schools. *BMC Medical Education*, 25, 19. <https://doi.org/10.1186/s12909-025-06719-5>
- Bertacchini, E., Bilotta, E., Gabriele, L., Pantano, P., & Tavernise, A. (2023). AI-enhanced robots in education: Exploring inclusive classroom environments. *UNESCO AI Futures Report*.
- Chan, T. C. (2023). AI and school leadership: Implications for principal preparation. *Educational Planning*, 31 (3), 42–56.
- Cooper, C. A. (2023). Educators' trust in AI-generated feedback: Findings from a multi-site qualitative study. *Journal of Educational Technology Research*, 11 (1), 17–35.
- Duz, S. (2023). Transformational leadership and artificial intelligence in education. *International Journal of Educational Administration and Policy Studies*, 15 (1), 43–56. <https://doi.org/10.5897/IJEAPS2023>
- Feigerlová, M., Plháčková, D., & Veteška, J. (2025). The digitalization of education in the context of teachers' professional development. *Education Sciences*, 15 (1), 3. <https://doi.org/10.3390/educsci15010003>
- Gathoni, L. (2021). Designing culturally responsive AI systems for teacher evaluation. *African Journal of Education and Technology*, 4 (2), 25–40.
- Habbal, R., Khalaf, M., & Malkawi, M. (2024). Teacher evaluation in the age of AI: Opportunities and ethical tensions. *International Journal of Artificial Intelligence in Education*, 34 (2), 167–183. <https://doi.org/10.1016/j.ijedai.2024.03.005>
- Holmes, W., Bialik, M., & Fadel, C. (2023). *Artificial Intelligence in Education: Promises and Implications for Teaching and Learning* (2nd ed.). Center for Curriculum Redesign.
- Holmes, W., Porayska-Pomsta, K., Holstein, K., Sutherland, E., & Baker, T. (2022). Ethics of AI in education: Towards a community-wide framework. In *Proceedings of the 12th International Learning Analytics & Knowledge Conference* (pp. 514–519). <https://doi.org/10.1145/3506860.3506881>
- Kaufman, J. H., Pane, J. F., & Diliberti, M. K. (2024). Educators' views on the use of artificial intelligence in schools. RAND Corporation. <https://doi.org/10.7249/RR134-25>
- Khreisat, L., Al-Soud, M., & Bataineh, R. F. (2024). Digital leadership in Jordanian schools: Opportunities and limitations. *Educational Technology & Society*, 27 (1), 85–98.
- Lizano-Sánchez, R., García, M., & Villalobos, B. (2025). Virtual assistants in teacher education: Impact on instructional coaching. *Frontiers in Education*, 9, 6719. <https://doi.org/10.1186/s12909-025-06719-5>
- Martin, R. A., Assante, A. R., & Miles, K. (2024). AI in education: Promise, practices, and possibilities. *Sustainability*, 16 (10386), 1–21. <https://doi.org/10.3390/su16201386>
- Matushevych, Y., Tomita, H., & Shirakawa, M. (2022). The dark side of AI in education: Data colonialism,

- ethics washing, and the power asymmetry. *Journal of Critical EdTech*, 4 (1), 45–59.
- MEN-DNE. (2024). Artificial intelligence and education: Brochure for decision-makers. Ministère de l'Éducation Nationale et de la Jeunesse. <https://www.education.gouv.fr>
- Mpofu, N., & Chasokela, C. (2025). Transformational leadership and AI integration in public schools. *Journal of Educational Change*, 26 (1), 93–112.
- OECD. (2021). AI and the future of skills: Evidence from the OECD survey of adult skills. OECD Publishing. <https://doi.org/10.1787/bf29b812-en>
- OECD. (2024). Artificial intelligence in education: Challenges and opportunities for policymakers. OECD Publishing. <https://doi.org/10.1787/ai-edu-2024-en>
- Okokoyo, E., Samuel-Okon, M., & Abejide, A. (2024). Principal readiness for AI-supported teacher appraisal in West Africa. *African Journal of Educational Management*, 12 (1), 101–120.
- Owan, V. J., Owan, J. J., & Owan, A. M. (2023). Educational applications of IBM Watson: AI's impact on secondary school instruction. *International Journal of Technology in Education and Science*, 7 (2), 173–189.
- Paramole, A. B. (2025). Artificial intelligence in instructional supervision: A critical review. *STEM Education*, 5 (1), 1–13. <https://doi.org/10.37256/steme.512025206>
- Pisica, S., Rotar, C., & Andronache, D. (2023). AI-based teacher assessment systems: Benefits and challenges. *Education Sciences*, 13 (2), 118. <https://doi.org/10.3390/educsci13020118>
- RAND Corporation. (2024). Teachers and principals use AI in instruction and evaluation: Emerging insights. RAND Education and Labor. <https://doi.org/10.7249/RRA134-25>
- Romero, A., Urmeneta, A., & Ferrer, J. (2024). AI and leadership for inclusive schools: Ethical tensions and trust. *AI & Education Studies*, 8 (1), 13–33.
- Samah, H., Azman, A., & Yasin, R. (2025). PERMA well-being model and AI-enhanced classrooms in inclusive education. *STEM Education*, 5 (1), 1–16. <https://doi.org/10.37256/steme.512025207>
- Samuel-Okon, M., & Abejide, A. (2024). Principals' views on the role of AI in school-based teacher evaluation in Nigeria. *African Journal of Educational Research*, 29 (1), 42–60.
- Sanusi, K. A., Abubakar, H., & Olanrewaju, M. (2024). Stakeholder perspectives on AI in African secondary schools. *Educational Technology and Development*, 22 (3), 99–120.
- Tawil, S., & Miao, F. (2024). Human-centered AI for education: Policy considerations and principles. UNESCO Artificial Intelligence and the Futures of Learning Post-Conference Report.
- Tenberga, R., & Daniela, L. (2024). Using artificial intelligence in education: Perspectives from school leadership. *Sustainability*, 16 (7), 3148. <https://doi.org/10.3390/su16073148>
- UNESCO. (2019). Beijing Consensus on Artificial Intelligence and Education. UNESCO. <https://unesdoc.unesco.org>
- Wang, Z. (2021). Human-AI symbiosis in school leadership: A conceptual model. *Journal of Educational Leadership and Policy*, 17 (2), 115–133.
- Yagi, K., Nakamura, Y., & Nishida, R. (2023). Algorithmic accountability in educational AI tools: A qualitative study. *Learning, Media and Technology*, 48 (1), 1–20. <https://doi.org/10.1080/17439884.2022.2141553>.