

Unveiling the multifaceted outcomes of principal evaluations

International
Journal of
Educational
Management

1

Ahmed M. Alkaabi

United Arab Emirates University, Al Ain, United Arab Emirates

Asma Abdullah

Sharjah Education Academy, Sharjah, United Arab Emirates

Shaikah Al-Taneiji

United Arab Emirates University, Al Ain, United Arab Emirates, and

Suhair Almaamari

UAE Ministry of Education, Al Ain, United Arab Emirates

Received 4 February 2025
Revised 7 July 2025
18 September 2025
Accepted 19 September 2025

Abstract

Purpose – This study examines how public-school principals in the United Arab Emirates (UAE) perceive—and act upon—the outcomes of their mandatory end-of-year evaluations. The research addresses the intersection of three elements that follow principal evaluation outcomes—incentives, improvement plans, and punitive measures—by asking: What are the experiences of school principals concerning the outcomes of their final evaluations conducted throughout the academic school year? Stemming from this main question, the following sub-questions will be addressed: What types of incentives do school principals receive? How does the current system capture both good and poor performance? What improvement plans, if any, are provided when school principals receive low scores? What punitive measures are taken against repeatedly low-performing principals?

Design/methodology/approach – This qualitative study drew on multiple data sources, including semi-structured interviews with six principals and three principal supervisors, along with the concurrent collection of relevant documents. Thematic analysis was employed to synthesize codes, trace similarities and differences across participants, and map overarching patterns of meaning.

Findings – The analysis revealed four interrelated themes. First, promised financial and symbolic rewards often failed to materialize, limiting their motivational effect. Second, high evaluation ratings did not translate into clear or attainable promotion opportunities, constraining professional growth. Third, improvement plans—particularly those developed collaboratively—provided the most actionable and relevant guidance for refining leadership practices. Finally, the absence of strong consequences for persistent underperformance undermined accountability and system credibility. Taken together, the findings suggest that the current model privileges procedural compliance over developmental impact, thereby constraining its potential to enhance school leadership and student learning.

Originality/value – This study centers the perspectives of school leaders in a Gulf-state context and contributes to global scholarship on principal evaluation. It offers insight into how evaluation mechanisms operate in non-Western systems. Additionally, the findings yield a transferable framework for understanding how incentives, developmental feedback, and accountability structures interact. Practical recommendations—transparent rewards, expanded advancement pathways, and coaching linked to improvement plans—are also provided to help transform principal evaluation into a more constructive driver of educational improvement.

Keywords Principal evaluation, Principal supervisor, Incentives, Improvement plans, Punitive measures

Paper type Research article

Introduction

Principal evaluation is a cornerstone of educational administration because it shapes the quality of school leadership. Amid heightened accountability, principals are expected to

© Ahmed M. Alkaabi, Asma Abdullah, Shaikah Al-Taneiji and Suhair Almaamari. Published by Emerald Publishing Limited. This article is published under the Creative Commons Attribution (CC BY 4.0) licence. Anyone may reproduce, distribute, translate and create derivative works of this article (for both commercial and non-commercial purposes), subject to full attribution to the original publication and authors. The full terms of this licence may be seen at [Link to the terms of the CC BY 4.0 licence](#).

Conflicts of interest: The authors declare no potential conflicts of interest with respect to the research, authorship, or publication of this article.



International Journal of Educational
Management
Vol. 40 No. 8, 2026
pp. 1-22
Emerald Publishing Limited
e-ISSN: 1758-6518
p-ISSN: 0951-354X
DOI 10.1108/IJEM-02-2025-0089

serve simultaneously as visionaries, instructional and curricular leaders, critical supervisors, facility managers, and respectful enforcers of policy mandates and initiatives (Zepeda *et al.*, 2020). Their role has shifted from building manager to instructional leader; they now occupy pivotal positions to guide, influence, and produce positive change in school performance (Al-Hamdan and Al-Yacoub, 2005; Gümüş and Bellibaş, 2020). Consequently, principals require a rigorous evaluation system that satisfies accountability demands while fostering continuous professional growth (Brauckmann and Pashiardis, 2010).

Renewed scholarly attention to the principalship reflects its critical influence on school outcomes. They argued that a robust, reliable evaluation system enables principals to diagnose areas for improvement and make informed decisions about their professional development, thereby narrowing the gap between current practice and desired performance (Davis *et al.*, 2011). Fitzpatrick *et al.* (2011) defined evaluation as “the identification, clarification, and application of defensible criteria to determine an evaluation object’s value (worth or merit) in relation to those criteria” (p. 7). Applied to school leadership, principal evaluation assesses a leader’s merit against established standards (Fuller *et al.*, 2015). The prominence of principal evaluation has been amplified by federal initiatives—Elementary and Secondary Education Act waivers, Race to the Top, and the reauthorized Every Student Succeeds Act—which emphasize instructional leadership and professional capacity (Donaldson *et al.*, 2021a, b; Grissom *et al.*, 2021a, b).

Evaluation, however, is only the first step. Once scores are assigned, actions follow: incentives that reinforce exemplary performance, improvement plans that promote growth, and punitive measures that address persistent underperformance (Anderson and Turnbull, 2016; DeMatthews *et al.*, 2020). These post-evaluation outcomes determine the process’s ultimate impact on individual leaders and the schools they guide. Post-evaluation actions are essential to effective leadership (Zepeda *et al.*, 2014a, b), school improvement (Elfers and Plecki, 2017), and, ultimately, enhanced student outcomes (Anderson and Turnbull, 2016; Grissom *et al.*, 2021a, b). Although principal evaluation has gained prominence during the past 2 decades, few studies investigate its consequences, particularly in the United Arab Emirates (UAE; Alkaabi, 2025; Alkaabi and Almaamari, 2020). This study addresses that gap by examining post-evaluation outcomes from the perspectives of principals and their supervisors.

The central research question asks how school leaders experience the outcomes of their annual evaluations. Four subsidiary questions guide the inquiry: (a) What incentives do principals receive? (b) How does the current system capture both good and poor performance? (c) Which improvement plans, if any, accompany low ratings? and (d) What punitive measures are taken against repeatedly low-performing principals? The study traces how evaluation scores translate into incentives, targeted development, and punitive responses, thereby offering practical insights for designing systems that are equitable, transparent, and conducive to leadership growth. The findings are expected to illuminate principals’ lived realities and encourage decision makers to adopt responsive measures. Given the limited evidence from non-Western contexts, this research contributes a culturally grounded perspective to global discussions of educational leadership and accountability. The subsequent literature review synthesizes research on common post-evaluation outcomes, including incentives, promotions, professional learning, and sanctions.

The anticipated findings will illuminate principals’ lived realities, reveal systemic blind spots, and encourage senior decision makers to adopt evidence-responsive measures. Because extant scholarship draws heavily on Western experiences, this research contributes a culturally grounded perspective to global conversations about educational leadership and accountability. The forthcoming literature review therefore elaborates on common post-evaluation outcomes—such as incentives, improvement plan, and punitive measures—and positions the UAE study within that broader evidence base.

Literature review

Conceptual framework

This study is anchored in the summative principal-evaluation model, which assesses school leaders' performance at the close of a designated cycle to support consequential decisions. In the comprehensive framework articulated by Clifford *et al.* (2014), summative evaluation serves three interrelated purposes: ensuring accountability, guiding professional growth, and shaping administrative action. The model rests on three intertwined dimensions—performance-based incentives, punitive measures for persistent underperformance, and improvement planning. These components are not merely procedural; they function as essential drivers of school improvement. Goff *et al.* (2016) contended that linking evaluation outcomes to differentiated support and rewards enhances both principal motivation and efficacy. Mitani (2025) reported that incentive-based systems, while yielding modest gains overall, generate particularly positive effects in high-need contexts, thereby demonstrating the value of context-specific incentives. Clifford *et al.* (2011) showed that evaluations combining formative and summative measures cultivate leadership development and inform decisions concerning retention, remediation, and recognition. Turnbull *et al.* (2016) asserted that improvement plans derived from evaluation data strengthen instructional leadership. Recent findings by Alkaabi (2025) indicate that principal supervision cannot achieve sustainable leadership enhancement unless improvement plans are fully embedded in evaluation practices. Positioned at the intersection of these elements, the present study investigates how incentives, improvement plans, and punitive measures are implemented and perceived by school principals. The following section explores each domain in greater depth.

Incentives

Incentives constitute a central mechanism for motivating employees to perform at high levels. Robust incentive systems enhance individual capability and align personal goals with organizational objectives (Fulmer and Li, 2022). Conversely, the absence of such systems can erode job satisfaction and constrain effort. Within education, inadequate performance-based incentives diminish principals' productivity (Assiry *et al.*, 2022). School leaders therefore receive a mix of monetary rewards—performance bonuses, salary increases, and merit pay—and non-monetary rewards such as leadership opportunities, public recognition, professional development, mentoring, and coaching (Schuermann *et al.*, 2009). Several districts acknowledged exemplary leadership primarily through non-monetary incentives, including improved working conditions and paid leave. As most principals are former teachers who regard their role as a form of stewardship, these non-monetary rewards can be particularly motivating (Goff *et al.*, 2016, p. 132). At times, even a single piece of guidance or a simple gesture can significantly influence individuals' behaviors and attitudes (Qablan and Al-Qaderi, 2009; El Zaatari and Ibrahim, 2021).

Federal policy has also shaped incentive structures: when the Elementary and Secondary Education Act was reauthorized in 2015, the Teacher Incentive Fund became the Teacher and School Leaders Incentive Program. Two Institute of Education Sciences studies subsequently evaluated the effectiveness of systems that combined performance feedback with monetary bonuses (Garet *et al.*, 2017; Wayne *et al.*, 2018). The first study provided principals with ratings and verbal feedback across key leadership domains, including the establishment of high expectations for student learning (Garet *et al.*, 2017). Although the ratings were not formally linked to employment decisions, the feedback was intended to guide professional improvement. The second study, extending Chiang *et al.* (2015), revealed that most principals continued to qualify for bonuses, implying that the earning criteria may have been insufficiently rigorous (Wayne *et al.*, 2018).

Policymakers still struggle to isolate principals' precise contributions to student achievement and overall school effectiveness. Many evaluation systems rely on value-added models to estimate educators' impact on student progress (Amrein-Beardsley, 2023;

Amrein-Beardsley and Geiger, 2019; Chiang *et al.*, 2016a, b; Goff *et al.*, 2016; Grissom *et al.*, 2015). Yet research underscores the complexity of attributing student outcomes solely to principals, whose influence is intertwined with multiple in-school and external factors (Fuller *et al.*, 2015; Shen *et al.*, 2016). Chiang *et al.* (2016a, b) compared four performance metrics to estimate principals' effects on student achievement and determined that same-year student test scores are poor predictors of principals' influence in subsequent years. Value-added and adjusted value-added models yielded more precise estimates, yet only about one-third of the variance in value-added ratings reflected principals' contributions. The authors therefore cautioned against relying exclusively on test-based measures to determine performance compensation.

Using longitudinal data from Pennsylvania elementary and middle schools, Chiang *et al.* (2012) examined the portion of the overall "school effect" attributable to leadership. By tracking schools that experienced leadership transitions over three years, they computed principal effects within grades and projected those effects to a different grade in the fourth year. Although principals did influence achievement, their contributions explained only about 15% of the total school effect. Mitani (2025) further complicated the use of performance-based rewards by analyzing statewide administrative data from Tennessee. Performance-Based Compensation Systems (PBCS) had no consistent effect on principals' job performance overall, although modest positive outcomes emerged in high-need schools. These results underscore how incentive systems can yield differential effects based on context and mirror longstanding concerns about the limitations of test-based metrics.

Design considerations are therefore paramount. When evaluations generate summative scores for accountability, high validity and reliability are essential. If the primary aim is formative feedback for growth, measurement precision may be less pivotal (Burkhauser *et al.*, 2013). In either case, principals' perceptions of fairness, accuracy, and transparency remain critical (Fitzpatrick *et al.*, 2011; Grissom *et al.*, 2015). Absent such perceptions, principals may disengage, disregard feedback, or manipulate data to avoid sanctions or secure rewards (Kane and Staiger, 2002). Process quality also shapes engagement. Nelson *et al.* (2021) found that principals judge evaluation systems more by procedural integrity and relationships with evaluators than by the ratings themselves. Perceived fairness, developmental orientation, and interpersonal trust thus prove integral to meaningful participation. Collectively, the literature signals that sole dependence on value-added metrics is neither valid nor reliable for awarding incentives. Multiple contextual factors—many beyond a principal's control—also influence school effectiveness, necessitating complementary indicators. Reviews consistently document limited reliability, inconsistent implementation, and weak alignment with leadership standards. Furthermore, many frameworks lack empirical evidence of their impact on practice, blur accountability and growth purposes, and employ diverse indicators with tenuous links to student outcomes (Clifford *et al.*, 2011; Davis *et al.*, 2011; Portin *et al.*, 2006).

Improvement plan

Principal evaluation systems should cultivate robust adult-learning cultures similar to those found in high-performing businesses, positioning human development as their central objective (Micheaux and Parvin, 2018). Kegan and Lahey (2016) introduced the idea of deliberately developmental organizations, characterized by radical transparency and continuous cycles of coaching, feedback, and candid communication. A 50-state review conducted by Fuller *et al.* (2015) indicated that most states identify professional growth as a primary goal of principal evaluation, and more than half link evaluation results to student achievement. Yet empirical evidence on how evaluations actually foster professional learning remains limited (Hvidston *et al.*, 2018). Whereas some studies offer detailed roadmaps for improvement (Hvidston *et al.*, 2015; Parylo *et al.*, 2012; Zepeda *et al.*, 2014a, b), others provide only broad guidance that leaves principals uncertain about next steps (Davis *et al.*, 2011; Fuller *et al.*, 2015; Goldring *et al.*, 2015). Across this literature, the importance of

supervisors who deliver high-quality feedback is repeatedly emphasized (Alkaabi and Almaamari, 2020 Ibrahim, 2012; Sun *et al.*, 2012).

Turnbull *et al.* (2016) evaluated the Principal Pipeline Initiative through interviews with district administrators, focus groups with novice principals, surveys, and document analysis. Among principals advised to strengthen instructional leadership, 86% reported receiving targeted support, although that assistance depended on additional supervisor training. Participants generally rated supervisors as knowledgeable and responsive, but mentors and coaches received slightly higher ratings. In a survey of 255 Rocky Mountain principals, Hvidston *et al.* (2015) likewise concluded that effective evaluation hinges on qualified, competent, and reliable supervisors. Despite these supports, researchers still identify gaps in the improvement plans intended to translate evaluation findings into enhanced leadership practice (Grissom *et al.*, 2015). Parylo *et al.* (2012) contended that plans must be individualized, feedback-driven, and anchored in sustained reflection. Implemented well, such plans clarify expectations, promote continuous growth, and elevate school performance by aligning principals' professional learning with institutional needs. Recognition of their value has increased, and recent work positions goal-oriented improvement plans as a core component of principal evaluation (Alkaabi, 2025).

Goal setting for principals should be informed by prior evaluation data; however, well-crafted goals and learning agendas do not automatically translate into substantive professional learning (Jones *et al.*, 2022). In a study of 95 principals, Sinnema and Robinson (2012) found that participants viewed goal setting primarily as a mandate to "doing important things" rather than "learning important things" (p. 157) and reported limited success in meeting those goals. Such findings amplify concerns about the consequences for principals who repeatedly receive low ratings. DeMatthews *et al.* (2020) examined how experienced principals engaged with the Texas Principal Evaluation and Support System (T-PESS), a model grounded in continuous improvement. Principals identified self-assessments, focused goal setting, and sustained evaluator coaching as the most powerful supports for building leadership capacity. Importantly, concentrating on a single professional goal—rather than several—was perceived as more manageable and impactful amid the complex, evolving demands of school leadership.

Trusting relationships with evaluators further enhanced the developmental value of the process by facilitating reflective dialogue and collaborative problem-solving. Alkaabi (2025) likewise underscored the centrality of improvement plans, arguing that such plans must be monitored during both formative and summative phases to remain dynamic instruments for strengthening instructional leadership. Ongoing monitoring ensures that plans address specific growth areas and promote sustained professional development rather than serving as mere formalities. In short, principal evaluations can foster professional growth when anchored in trust, clearly articulated goals, and continuous feedback. Their success, however, hinges on thoughtful implementation and sensitivity to local context, not on standardized procedures alone.

Punitive measures

Evaluations continue to inform school- and district-level personnel decisions, despite persistent concerns about the metrics employed—particularly value-added measures (VAMs). Roughly two-thirds of U.S. states report that principal evaluations can influence high-stakes determinations, and one-third either allow, recommend, or require districts to use these evaluations when making personnel decisions; nearly one-quarter further indicate that termination may be based on evaluation outcomes (Fuller *et al.*, 2015). In a Wallace Foundation-funded study of six districts, Anderson and Turnbull (2016) examined four initiatives designed to cultivate a competent pipeline of novice principals, including structured, on-the-job evaluations. Although the authors did not specify the exact number of principals dismissed because of poor ratings, administrators in every district acknowledged

that dismissal was a possible consequence. For example, New York legislation permits termination after two consecutive “ineffective” ratings. District leaders reported that principals receiving such ratings were frequently reassigned or not offered contracts for the subsequent year. Nonetheless, the study found that few novices were formally dismissed at the conclusion of their initial two- or four-year contracts. Instead, performance-management discussions often spurred voluntary attrition, suggesting that official dismissal statistics likely understate the total impact of evaluation-driven personnel actions.

Researchers caution policymakers and administrators against relying on principal evaluations primarily for short-term staffing decisions (Clifford *et al.*, 2014; DeMatthews *et al.*, 2020; Fuller *et al.*, 2015). Despite acknowledged limitations in both traditional and VAM-based models, most states still employ evaluation results for high-stakes purposes (Fuller *et al.*, 2015). A growing literature questions the validity of student-achievement VAMs as indicators of principal effectiveness, noting their sensitivity to exogenous factors such as student demographics (Chiang *et al.*, 2016a, b; Grissom *et al.*, 2015; Henry and Viano, 2016; Herrmann and Ross, 2016). Using longitudinal data on Grade 4–8 mathematics and reading outcomes across Pennsylvania, Chiang *et al.* (2016a, b) found that school-level VAM scores poorly predicted a principal’s sustained effectiveness. Likewise, in New Jersey, Herrmann and Ross (2016) reported that principals evaluated via median student-growth percentiles were less likely to receive “highly effective” ratings than peers assessed through alternative indicators such as progress toward professional goals. Collectively, these findings call into question the advisability of heavily weighting student-growth data when differentiating principal performance.

Achievement-growth metrics consistently correlate with students’ socioeconomic status, raising equity concerns about their use in principal-evaluation systems. Because these measures may systematically disadvantage leaders in schools enrolling large proportions of low-income students, their validity as indicators of principal effectiveness is contested. In an analysis of Tennessee’s statewide evaluation system, Grissom *et al.* (2018) found that principals serving high-poverty schools received lower performance ratings than peers in more advantaged settings, implying that demographic factors, rather than leadership quality, influenced the scores. Similarly, during Pennsylvania’s pilot of the Framework for Leadership (FFL), McCullough *et al.* (2016) observed only modest positive associations between FFL ratings and value-added scores—and solely for middle-school mathematics, not reading or writing. Collectively, these findings underscore the limitations of relying exclusively on achievement-growth data and highlight the need to incorporate contextual variables when assessing principal performance. Consequently, Clifford *et al.* (2014) argued that, when evaluations carry high-stakes consequences, they must meet rigorous psychometric standards to remain both technically sound and legally defensible.

The Principal Evaluation process in the United Arab Emirates

The Abu Dhabi Education Council (ADEC)—the authority overseeing education in the Emirate of Abu Dhabi, including Al Ain and Al Gharbia—has established a comprehensive framework of professional standards for school principals. Developed through iterative consultation and implemented across all public districts, these standards underpin both principal evaluation and leadership development. They specify what leaders must know, do, and demonstrate to fulfill their roles effectively and are grounded in empirical evidence linking leadership behaviors to improved student learning. Accordingly, principals are expected to: (a) lead strategically by articulating a clear vision, mission, and goals; (b) ensure effective teaching and learning by fostering continuous professional growth among staff; (c) make evidence-based decisions that advance organizational improvement; (d) support personnel through systematic professional development; and (e) cultivate productive relationships with stakeholders whose engagement is essential to achieving desired outcomes (ADEC, 2012).

ADEC's evaluation process unfolds through cyclical, interrelated phases designed to promote accountability and professional growth. Principal supervisors conduct multiple visits during the academic year, collect qualitative and quantitative evidence, and provide formative feedback. These visits culminate in an end-of-year summative evaluation guided by a rubric aligned with national professional standards. As Alkaabi (2025) explains, the rubric articulates the knowledge, dispositions, and practices essential for effective leadership and sustained school improvement. The process concludes with a formal defense meeting in which principals present evidence portfolios and engage in professional dialogue with their supervisors to justify performance ratings (Alkaabi, 2025). Although the UAE system aims to strengthen leadership and instructional quality, prior research has concentrated mainly on evaluative procedures and supervisory roles (Alkaabi, 2025; Alkaabi and Almaamari, 2020). Far less attention has been devoted to the consequences of evaluation outcomes—specifically, the implementation of incentives, improvement plans, and sanctions. To address this gap, the present study investigates the post-evaluation phase, examining how these elements are enacted, perceived, and experienced by principals, thereby extending understanding of the evaluation's impact beyond the summative judgment.

Methods

This qualitative study developed an in-depth, interpretive understanding of participants' social worlds and the meanings they assign to their experiences, perspectives, and histories (Patton, 2015). The design enabled the researcher to examine contemporary contexts in which limited empirical evidence exists regarding principal-evaluation outcomes and the roles of principals and supervisors. The investigation took place in the Al-Ain District—the Emirate of Abu Dhabi's second-largest district—which operates 103 schools and serves more than 50,000 students. The project extends research begun in 2021 on principal evaluation in the United Arab Emirates. The university's Institutional Review Board approved the study (STUDY00005027). Data are not publicly available because they contain sensitive information and are protected by participant-confidentiality agreements.

Nine individuals participated: six school principals and three supervisors. As the study is qualitative, the aim was to reach data saturation and capture diverse yet manageable perspectives rather than to achieve statistical generalizability. Including both principals and supervisors yielded a richer view of how evaluation outcomes are perceived, enacted, and experienced from supervisory and leadership standpoints. Purposeful sampling, guided by criteria delineated by Patton (2015), informed participant selection. Inclusion criteria required individuals to (a) have engaged in the formal principal-evaluation process, (b) have received evaluation outcomes for at least three consecutive years, and (c) be willing to reflect on their professional experiences. Potential participants were excluded if they had limited or no engagement with evaluation procedures or fewer than three years of participation. Table 1 presents participants' demographic characteristics. All ethical approvals have been obtained, and guidelines have been followed for the research.

Data sources

Interviews. To examine what occurred following the receipt of evaluation outcomes—specifically, the implementation of incentives, improvement plans, and punitive measures as experienced by school principals and their supervisors—semi-structured interviews served as the primary data source for this study. Seidman (2012) emphasized that the purpose of interviews was not to test hypotheses but to explore and interpret participants' lived experiences. Similarly, Kvale (1997) described qualitative interviews as a means of uncovering the meanings embedded in the key themes shaping each participant's world. Both perspectives underscored the importance of researchers engaging interviewees with genuine and respectful interest, recognizing their voices as central to constructing meaningful

Table 1. Demographic profile of research participants

Participants	Gender	Job position	Years of experience	School level
Majed	Male	Principal	21	Cycle I
Ali	Male	Principal	29	Cycle II
Saeed	Male	Principal	17	Cycle III
Sara	Female	Principal	21	Cycle I
Nouf	Female	Principal	26	Cycle II
Shaikah	Female	Principal	20	Cycle III
Julia	Female	Supervisor	5	Cycle I
Mais	Female	Supervisor	6	Cycle II
Ben	Male	Supervisor	8	Cycle III

Source(s): Table created by authors

understanding. Each participant engaged in a face-to-face, in-depth interview lasting approximately two hours, designed to elicit nuanced insights into their individual perspectives. The semi-structured format supported a conversational flow while maintaining alignment with the study's overarching purpose. This format also allowed the researcher to introduce new, context-relevant questions as they emerged—potentially supplementing or replacing predetermined prompts—to explore specific domains of interest more thoroughly (Glesne, 1999). Table 2 presents a sample of interview questions posed to school principals regarding practices implemented following evaluation outcomes.

Before each interview, the researchers coordinated the date, time, and location with each principal. At the start of the interview, the researchers provided a brief self-introduction and outlined the purpose and objectives of the interview. All interviews were recorded electronically using an IC Recorder and were later transcribed verbatim. Participants were encouraged to speak in the language that best allowed them to articulate their perspectives. To maintain confidentiality, the researchers removed any identifying information from the transcripts and replaced names with pseudonyms.

Documents. In this qualitative study, documentary evidence corroborated and enriched findings derived from other methods. Beyond supporting triangulation, documents offered inherent objectivity because they did not alter the research setting as an investigator's presence might. Merriam (1998) observed that documents are not subject to human manipulation during data collection and therefore constitute a reliable source of evidence. Yin (2014) added that documents possess four advantages: stability (they can be reviewed repeatedly), unobtrusiveness (they are not produced for the study), specificity (they record exact details such as names and dates), and broad coverage (they provide contextual insight across events and settings). With participants' consent, the researcher collected e-mail correspondence, official communications, evaluation agendas, and principal-evaluation rubrics and forms.

Data analysis. Thematic analysis was conducted to generate meaningful insights from an integrated data set. Bryman and Burgess (1994) emphasized that thematic analysis involves identifying, categorizing, defining, theorizing, explaining, and mapping patterns of meaning. Guided by this view, analysis remained central throughout the project—from initial data collection through manuscript preparation. Thematic analysis allowed the team to examine interview transcripts and detect recurring patterns and themes that reflected participants' core experiences and perspectives. Braun and Clarke (2006) stated that thematic analysis organizes qualitative data and distills complex themes, thereby clarifying a study's objectives and research questions.

Before beginning formal coding, the researchers read each transcript several times to achieve deep familiarity with the data. They then generated codes that highlighted similarities and differences in participants' experiences. Corbin and Strauss (2008) defined coding as the extraction of concepts from raw data and the development of those concepts in terms of their properties and dimensions (p. 159). The initial analytic phase relied on open coding and

Table 2. Sample of interview questions

Sample of interview questions for principals and supervisors

Questions related to incentives following evaluation outcomes

- What types of incentives, if any, have you received as a result of your annual evaluation?
- Are these incentives linked to specific evaluation scores or performance levels? If so, how is this connection communicated?
- Do you believe the current incentive system equitably rewards effective leadership? Please explain
- Can you describe an instance in which you or a colleague were recognized or rewarded based on an evaluation outcome?
- In your opinion, which forms of incentives (monetary or non-monetary) are most meaningful or motivating for school leaders?
- To what extent do evaluation-based incentives influence your motivation, leadership goals, or school improvement efforts?
- What challenges or concerns, if any, exist regarding the distribution or perception of incentives among principals?
- How could the incentive component of the evaluation system be improved?

Questions Related to Punitive Measures Following Evaluation Outcomes

- What types of punitive measures, if any, are implemented when a principal repeatedly receives low evaluation scores?
- Are there clearly defined guidelines or thresholds for when punitive measures are applied?
- How effectively are the consequences for underperformance communicated?
- Do you believe punitive measures are applied consistently across schools and leadership roles?
- Have you observed or heard of instances where principals were demoted, reassigned, or dismissed due to evaluation results? If so, please describe the situation
- In your view, do punitive consequences support leadership improvement or foster a climate of fear and compliance?

Questions Related to Improvement Plans Following Evaluation Outcomes

- Following your evaluation, were you provided with an improvement plan? If so, what did it entail?
- Was the plan tailored to your specific needs, or did it appear standardized?
- Who was involved in designing or guiding the improvement plan? Did you contribute to the planning process?
- What forms of support—such as coaching, mentoring, or professional development—were included in the improvement plan?
- How frequently was your progress monitored or reviewed during the improvement phase?
- In your opinion, did the improvement plan support your growth as a school leader? Please explain
- What recommendations would you offer to improve the usefulness and impact of improvement plans for principals?

Source(s): Table created by authors

line-by-line coding. Open coding involved systematically dividing the data into discrete units to identify emerging concepts and patterns. Line-by-line coding required meticulous segmentation of transcripts and documents so that each unit could be labeled and grouped into broader conceptual categories (Charmaz, 2006). Together, these techniques captured both explicit statements and nuanced, implicit meanings, thereby enriching the conceptual depth of the emerging themes.

After completing open and line-by-line coding, the researchers used thematic analysis to synthesize codes, trace similarities and differences across participants, and map overarching patterns of meaning. This iterative approach connected data sources and revealed themes throughout data collection and analysis. Table 3 presents a sample revised theme and its associated codes.

To ensure the quality of the data and establish trustworthiness in the findings, several measures were implemented, focusing on credibility, dependability, transferability, and confirmability (Lincoln and Guba, 1985). First, credibility was achieved through triangulation and member-checking. This involved comparing and triangulating data collected from interviews and documents to ensure the consistency of findings generated by different data-collection methods and sources over time. Member checks were also used to ensure that

Table 3. A selected sample of theme one

Open coding	Axial coding	Participants' words	Theme
No reward or acknowledgment for outstanding performance Equal treatment regardless of principal performance Recognition deficits reduce motivation Absence of a formal incentive system Desire for performance-based appreciation	The absence of differentiated incentives diminishes motivation and undermines the evaluation process by failing to recognize and reward high-performing principals	"It feels unfair to see hardworking principals being treated the same as low-performing ones . . . How would you feel if you worked hard and got nothing in return?" (Sara) "My efforts felt worthless." (Shaikah) "At least a thank-you letter or some acknowledgment would help." (Ali) "If we base incentives on this subjective evaluation, then principals will rightly complain that the supervisors are biased." (Ali) "Not having a reward system can be bad for principals . . . their dedication and ambition could suffer as a result." (Ben) "There's no structured way to appreciate good performance—nothing to push us forward." (Julia)	Unveiling the Core Importance of Incentives

Source(s): Table created by authors

participant responses were accurately represented and to avoid any data distortions or researcher biases (Creswell and Poth, 2017). Second, solid documentation was included in a dependability audit to ensure the accuracy of data collection and analysis methodologies, similar to Yin's recommendations for creating chains of evidence (2014). Third, to facilitate transferability, the researcher thoroughly described the context of the research project and its underlying assumptions. This allows other researchers to apply the study's findings to their own projects, promoting progress and unity in the field. Fourth, confirmability, which pertains to the neutrality of research conclusions, was ensured by documenting the research process via an audit trail. This trail informed each phase of the research and certified the non-biased nature of the findings (Patton, 2015). Peer debriefings were also conducted to assess any potential research bias that may have inadvertently affected the study.

Limitations

This study has several limitations. The small sample—six principals and three supervisors—limits statistical generalizability; however, the qualitative design prioritized depth over breadth to capture context-specific insights. The research was also limited to a single district within the Emirate of Abu Dhabi, which may affect the transferability of findings to other regions or systems. In addition, the reliance on self-reported interview data introduces potential for social desirability bias, despite the use of triangulation and member checking to enhance credibility.

Results

The study identified four primary themes that collectively address the central research question: (1) unveiling the core importance of incentives; (2) limited scope of promotion

opportunities; (3) enhancing leadership performance through yearly improvement plans; (4) lack of high-stakes consequences.

Theme one: unveiling the core importance of incentives

All participants expressed concern about the absence of incentives within the evaluation process, emphasizing that such recognition was essential to motivating sustained leadership performance. Principal Sara shared her experience:

It feels unfair to see hardworking principals being treated the same as low-performing ones. They don't receive any rewards, incentives, or even a simple thank you for their exceptional efforts in improving the school or maintaining its excellence. How would you feel if you worked hard and got nothing in return, while others who didn't do anything also received the same treatment?

In addition, Principal Sara shared her evaluation scores, which remained consistent over a three-year period without leading to any changes in status or recognition. Figure 1 presents her evaluation results across those years.

In addition to Sara's case, Principal Shaikha shared her evaluation data, which provided a detailed breakdown of her annual performance report. This level of transparency offered further insight into the areas emphasized within the evaluation framework, including both quantitative scores and qualitative performance areas. As shown in Figure 2, Shaikha's performance ratings reinforce the recurring concern among principals that consistent, high-level achievement often lacks corresponding recognition or advancement.

Participants noted that all principals were treated uniformly, without differentiating between high and low performing principals. Shaikah felt that her dedicated efforts were perceived as "worthless," while Ali wished for a simple acknowledgment, such as a "thank-you letter," for high performance. Such gestures, they believed, would foster a sense of self-worth among principals. Principals Nouf and Majed suggested that an effective evaluation system could benefit high-performing principals but disadvantage those who are less effective. They argued that a robust incentive system could influence behavior and effort levels. However, Ali cautioned that implementing such a system would require addressing existing flaws in the evaluation process. He stated, "If we base incentive on this subjective evaluation, then principals will rightly complain that the principal supervisors are biased in one way or another." He later acknowledged that even if student achievement were a metric, the current system could not definitively attribute this to principal leadership, potentially leading to unintended consequences.

Principal supervisors unanimously noted the lack of a structured framework for incentives, awards, or acknowledgments in the principal evaluation process. They believed that an integrated incentive system could enhance motivation and sustain commitment among principals. Ben, Mais, and Julia advocated for the inclusion of extrinsic rewards for principals who achieve an "accomplished" level or higher to maintain productivity and high-quality performance. Ben elaborated, "Not having a reward system can be bad for principals. It might make them less motivated, committed, and energetic. If they don't get what makes them feel motivated, their dedication and ambition could suffer as a result." Furthermore, Ben

Principal [REDACTED]

Academic Year	Annual Performance Summary	Performance Level
2022–2023	Annual Performance Summary	Exceeds Expectations
2023–2024	Annual Performance Summary	Exceeds Expectations
2024–2025	Annual Performance Summary	Exceeds Expectations

Figure 1. Evaluation ratings of a high-performing principal (2022–2025). Source: Data provided by participant Sara

Section No.	Performance Area	Final Performance Score
Section 1	Professional Commitment	15
Section 2	Professional Practices	48.61
Section 3	Feedback for Staff Performance Development	4
Section 4	Quality of Life for Stakeholders	5.00
Section 5	Impact on Students' Academic Progress	4.17
Section 6	Use of Information and Communication Technology	4.00
Section 7	Impact of Principal and Vice Principal on the Educational Community	4
Section 8	Attendance	5
—	Personal Development Plan	14
—	Distinguished Employee	—
—	Innovative Employee	3
—	Total Annual Performance Score	108.78
—	Annual Performance Level	Exceeds Expectations

Figure 2. Principal Shaikha's annual evaluation report by performance area. Source: Data provided by participant Shaikha

specifically urged to motivate principals based on their progress in yearly summative evaluations.

Theme two: limited scope of promotion opportunities

Participating principals expressed concerns about the restricted promotional opportunities stemming from the principal evaluation process, especially when they receive high evaluation scores. Shaikah noted that while a high score in Irtiqa's school inspection was essential for promotion to "executive principal," the irony lay in low performance evaluations. She stated:

I just knew . . . some principals . . . got selected to be the executive principal based on their school performance. There is not much information about this kind of job because it is new, but from what I heard and understood, it is similar to the job of cluster manager [principal supervisor]. The only difference is that those selected principals are more likely to remain working on their schools besides supervising other principals they are assigned to.

Similarly, Ali mentioned that principals whose schools received high scores in bi-annual inspections were considered for promotions. Majed added:

Principals whose school performance receives a high score in Irtiqa inspection are more likely to be nominated for a newly established position, which is similar in many ways to the cluster manager position. Principals have already been selected for interviewing and training. Basically, they will have to do the work of cluster manager in addition to leading their own schools.

Sara, Saeed, and other participating principals contended that promotions to the role of executive principal were not directly tied to their individual evaluations. Rather, these promotions were primarily associated with the biennial school evaluations conducted through the Irtiqa program. All principals expressed a desire for a promotion system that considered their individual evaluations, as they believed this would contribute significantly to their professional growth.

Among principal supervisors, there was a consensus regarding promotion opportunities. Julia explained that promotions to executive principal relied on the investigation reports generated every two years. These reports were distinct from principal evaluations. Julia noted, "This is a separate evaluation for schools and different from the principal evaluation," and suggested, "We should attach at least a partial incentive program directly linked to principal evaluation to foster higher motivation and boost performance." Ben further clarified the prerequisites for such promotions:

Such promotion is not directly related on their individual evaluation. Instead, it is based on two key factors: Firstly, Irtiqa’s report that carefully assesses and evaluates the quality and effectiveness of the school’s overall performance. Secondly, an individual interview is conducted to thoroughly gauge the principal’s abilities, skills, and knowledge to ensure they are the strong candidate to effectively take on this position. There should be promotion opportunities linked to the principal evaluation.

Theme three: enhancing leadership performance through yearly improvement plans

All participating principals indicated that those who performed poorly in their mandatory annual evaluations were given a yearly improvement plan to address their weaknesses in school leadership skills. Majed elaborated:

Principals who receive low scores in their evaluation are provided with a treatment plan to support them in their leadership role. The plan includes intensive training through workshops, PDs [professional development sessions], seminars, and consultations ... all aimed at helping the principals improve their performance.

Even principals who received high evaluation scores identified areas for continued professional growth. For example, Principal Nouf, whose ratings consistently reflected strong performance, shared a performance table aligned with specific indicators that revealed areas requiring further development (see Figure 3). The table presents a sample standard related to educational leadership and includes both the principal’s self-evaluation—denoted by an educator icon—and the supervisor’s evaluation—represented by a thumbs-up symbol. In this example, the supervisor disagreed with the principal’s self-assessment on a particular element, which indicated a need for improvement in the criterion concerning the promotion of educational research and its application to school improvement.

Nouf noted, “I requested a meeting with my supervisor to clarify these areas and to assist in setting targeted goals for development.” Similarly, Shaikha described her supervisor as highly cooperative and responsive to her request, promptly arranging a meeting to engage in a detailed discussion of her evaluation and to develop an improvement plan tailored to her

Standards	Performance Indicators	Needs Improvement	Partially Meets Expectations	Meets Expectations
Educational Research	3.5 Encourages educational research and applies it to improve student learning	<ul style="list-style-type: none"> • Has basic knowledge of modern educational and pedagogical research, its relevance to the school context, and its potential applications. • Understands the need for research and ongoing professional learning, and discusses this research, but there is no clear evidence of its application across all aspects of the school’s work. 	<ul style="list-style-type: none"> • Demonstrates comprehensive and up-to-date knowledge of educational and pedagogical research, its relevance to the school context, and its potential applications. • Reinforces and promotes a professional ethical commitment that emphasizes the importance of research and continuous professional learning. 👍 	<ul style="list-style-type: none"> • Demonstrates deep and up-to-date knowledge of a wide range of educational and pedagogical research, its relevance to the school context, and its potential applications. • Oversees research, inquiry, and continuous professional learning processes among the majority of staff. • Fosters a culture of peer learning among staff. • Takes personal responsibility for his/her own learning. • Regularly reviews data related to the impact of professional learning in the school. 🙅

Figure 3. Principal Nouf’s self- and supervisor ratings on the educational research performance indicator. Source: Data provided by participant Nouf

professional growth. She also shared the corresponding email response from her supervisor, which read: “No worries . . . just let me know the day, time, and place—your school, I suppose—and we can discuss whatever you need.”

Similarly, Sara, Nouf, and Ali emphasized the importance of offering appropriate support to principals who faced challenges in effectively managing their schools. Notably, they received support from a committee of high-performing principals to provide additional assistance when initial support proves insufficient. Saeed explained:

We were selected to give support on different issues that were problematic for the school principal who was not able to deal with his school . . . as in formulating internal policies and regulations . . . leading, teaching, learning, [and handling] uncontrolled student misbehaviors. This principal was new to the position and needed help. It goes without saying that his evaluation was down because of the chaotic situation his school faced. But now, it is getting much better.

Principal supervisors also shared similar experiences of providing intensive training and detailed plans to low-performing principals. Ben stated:

In simple terms, when principals struggle with managing their schools, it can be tough for them to oversee everything. The treatment plan is there to offer the help and guidance they need to overcome these difficulties and improve their leadership skills. The plan includes one-on-one weekly supervision, monthly professional development workshops, connecting with experienced principals in a network, and having ongoing support whenever they face any challenges.

Mais highlighted, “Without a structured plan and training, principals may not receive the necessary support to improve their performance.” Ben highlighted that the lack of intensive training could impede the essential development of leadership skills among principals. This could hinder their ability to effectively tackle challenges and adapt to evolving educational trends and demands. Consequently, he stated, “low-performing principals may find themselves trapped in repetitive actions” and making the same mistakes, thereby “jeopardizing the school’s overall success” and leading it towards “potential failure.”

Theme four: lack of high-stakes consequences

Both principals and principal supervisors identified a significant gap: the absence of sanctions tied to the summative evaluation, which could motivate principals to improve, especially those showing no growth over consecutive years. Sara noted, “Principals are unlikely to take the evaluation system as seriously as they should because they know they cannot be dismissed.” Shaikah concurred, stating that no consequences were directly linked to the summative evaluation. Nouf elaborated:

No instances of principals facing consequences like dismissal or discipline due to poor performance . . . you know. having a system with clear consequences would make principals take their job more seriously and work harder to meet expectations. The current system lacks the ability to accurately decide if termination is necessary or if excellent performance should be rewarded. This highlights the pressing need for revamp the principal evaluation system.

Ali, however, mentioned that some low-performing principals were transferred to other schools as the only punitive measure. He questioned the effectiveness of this approach, stating:

I’ve seen some principals getting transferred to another school as punishment when they couldn’t improve their previous school. But the question remains: if a low-performing principal is transferred, they might not do well in the new school either. I don’t think transferring is the best solution.

Majed expressed concerns about the current evaluation system, criticizing its subjective nature. He argued that incorporating high-stakes consequences could lead to increased attrition rates, affect the morale of principals, and discourage experienced teachers from pursuing administrative roles. He emphasized the need for “objective mechanisms” to ensure fairness when attaching serious consequences to low-performing principals. Julie observed, “It is very

difficult to raise a principal for termination, which is very easy in the US, and you would have a lot of authority.” She elaborated on the differences between the roles of principal supervisors in the US and the UAE:

If I was a cluster manger [principal supervisor] in the US, I would be given an equal role like a superintendent. But here, if I had a low principal, I would definitely document, document, document! And there is a process. When they officials ask for names of principals that you would like to raise for termination, you have to have a lot of documentation and it is a very severe process.

Ben noted the complexities of employment decisions in the UAE, particularly regarding the termination of Emirati teachers and leaders who received poor evaluations. He stated, “The final decisions on high-stakes employment matters are made by higher-level officials; we, as cluster managers, merely provide them with the evaluation results.” Ben also mentioned that he had never witnessed the termination of any principals during his tenure as a principal supervisor. Both participating principals and supervisors openly admitted that the evaluation outcomes were often subjective and prone to bias, due to the absence of objective performance data and assessment metrics. This lack of an objective framework raised questions about the reliability of making high-stakes decisions based solely on these evaluations. In such a context, requiring comprehensive documentation was the minimum that could be done.

Discussion

Summative evaluation, conceptualized by Clifford *et al.* (2014) as a multifaceted tool for accountability, professional growth, and administrative decision-making, did not fully meet those aims in the present study. Although the system incorporated post-evaluation improvement plans, it offered neither performance-based incentives nor meaningful consequences. Principals consistently reported an absence of differentiated recognition for excellence and a lack of clear sanctions for underperformance. These results indicate a need to recalibrate the evaluation framework so that it supports growth and fosters motivation through equitable acknowledgment and accountability. The discussion below situates each emergent theme within the extant literature and highlights both gaps and confirmations.

Theme one: unveiling the core importance of incentives

All participants noted that the current evaluation process provides no mechanism for rewarding exemplary performance. Whether a principal was rated “excellent” or “underperforming,” treatment and outcomes were identical, leaving many leaders feeling undervalued. This perception mirrors long-standing scholarly concerns that principal-evaluation systems rarely include structured, performance-linked rewards (Andrews, 1990; Garet *et al.*, 2017; Reeves, 2013; Wayne *et al.*, 2018). The resulting uniformity weakens motivation and diminishes the system’s capacity to inspire excellence. Conversely, research has shown that well-designed merit incentives can motivate school leaders and positively affect student outcomes (Pham *et al.*, 2021).

Participants also expressed a desire for even modest forms of recognition. Simple gestures—such as a certificate of appreciation or a written acknowledgment—could, in their view, bolster morale. This finding aligns with studies showing that both monetary and non-monetary incentives can enhance principal motivation (Schuermann *et al.*, 2009). Goff *et al.* (2016) observed that districts often reward outstanding teachers with improved working conditions, additional leave, or expanded roles; because most principals are former teachers, similar non-monetary incentives could prove equally effective for leaders. Collectively, these results support the development of tiered evaluation frameworks that both recognize high performance and provide clear pathways for continued growth.

Theme two: limited scope of promotion opportunities

Participants reported that advancement to “executive principal” is primarily determined by a school-inspection report evaluating overall school performance and by an interview assessing the candidate’s suitability, rather than by individual evaluation scores. They advocated for recognition and incentives directly tied to principal evaluations to strengthen motivation and performance. This view echoes research indicating that principals are more satisfied and more likely to remain in their positions when clear pathways for career advancement exist (Yan, 2020).

Theme three: enhancing leadership performance through yearly improvement plans

Participants agreed that principals who receive low evaluation ratings should be placed on well-designed improvement plans. Such plans provide targeted guidance, resources, and support to address specific growth areas. Prior studies have shown that intensive, individualized professional development can accelerate principals’ growth (Casserly *et al.*, 2013; Corcoran *et al.*, 2013). Without focused intervention, low-performing leaders risk repeating mistakes that jeopardize school success. Fuller *et al.* (2015) argued that evaluation should serve foremost as a mechanism for professional learning. Similarly, Alkaabi (2025) and Parylo *et al.* (2012) contended that effective instructional leadership depends on targeted feedback and continuous self-reflection.

Theme four: lack of high-stakes consequences

All principals noted that the lack of meaningful consequences undermines the evaluation system’s credibility; the belief that dismissal is unlikely reduces the seriousness with which some leaders approach their reviews. Although transfers were occasionally used, participants viewed them as insufficient to remedy persistent underperformance. They also criticized the process as subjective and devoid of robust, objective metrics. While many systems incorporate student-achievement data, scholars have questioned the validity of value-added growth scores because they are heavily influenced by contextual factors such as student demographics (Chiang *et al.*, 2016a, b; Grissom *et al.*, 2015; Henry and Viano, 2016; Herrmann and Ross, 2016). McCullough *et al.* (2016) therefore advised caution when interpreting value-added data in isolation. Overall, these findings underscore an urgent need to overhaul the principal-evaluation framework and to embed objective, fair measures capable of supporting high-stakes decisions (Fuller *et al.*, 2015).

Conclusion

This qualitative study examined principals’ experiences with the outcomes of final performance evaluations. Thematic analysis revealed four central findings: (1) unveiling the core importance of incentives; (2) limited scope of promotion opportunities; (3) enhancing leadership performance through yearly improvement plans; (4) lack of high-stakes consequences. These findings have direct implications for both policy and practice related to principal evaluation. As outlined by Clifford *et al.* (2014), comprehensive evaluation systems should explicitly connect evaluation outcomes to incentives, targeted improvement plans, and calibrated consequences for underperformance. However, the present findings indicate that integrating these elements into practice remains complex. To introduce meaningful incentives, evaluation frameworks must incorporate clearly defined performance tiers and transparent criteria—both of which were lacking, according to participants. This study strongly recommends that relevant school districts revise their evaluation processes to include structured incentive systems that differentiate performance levels and recognize excellence. Incentives may take various forms, including financial rewards, professional development opportunities, and formal public acknowledgment. Even in the absence of monetary resources, non-monetary recognition—such as a formal note of appreciation—was

cited by participants as highly motivating. For fairness and transparency, policy guidelines should clearly define the criteria governing these incentives. Promotion criteria should also be reevaluated to better align with individual principal evaluations. Strengthening the connection between performance and career advancement would enhance motivation and ensure that promotion decisions are informed by both school-level outcomes and individual leadership effectiveness.

Improvement planning emerged as the most consistently implemented component of post-evaluation practice. Participants reported that improvement plans, when applied effectively, offer a structured and constructive approach to professional development. Compared to incentives or punitive measures, this component was viewed as more feasible, less subjective, and more likely to produce actionable growth. These findings suggest that improvement planning should be institutionalized within principal evaluation systems, not only for low-performing principals but also to support ongoing development among high-performing leaders. Policy frameworks should define the structure of improvement plans, specifying professional development activities such as workshops, coaching, mentoring, and targeted training. Experienced principals should be engaged in supporting their peers, creating opportunities for collaborative professional learning. Supervisors should work jointly with both high- and lower-performing principals to design individualized improvement plans.

Ongoing assessments should be used to monitor progress and refine development goals as needed. Moreover, in the absence of objective measures, a balanced approach to consequences within the principal evaluation system is essential. The policy should outline a clear framework for both recognizing excellence and addressing chronic underperformance. A progressive set of consequences—ranging from intensive support to reassignment or further professional development—should be specified, depending on the severity and persistence of underperformance. Before any high-stakes decisions such as termination are made, it is critical to ensure that the evaluation process is objective, reliable, and accurate. Establishing fair and defensible punitive measures requires rigorous methods that ensure validity and reliability—an area where many evaluation systems, including the one examined in this study, continue to struggle. Although Clifford *et al.* (2014) provided a strong conceptual framework for integrating incentives, improvement plans, and consequences, educational leaders must overcome practical, political, and technical barriers to realize these benefits. However, this is not to suggest that high-stakes decisions should be avoided altogether; some principals may remain stagnant or continue to underperform over extended periods. School districts may consider the use of unbiased, independent evaluators in conjunction with principal supervisors to render fair judgments about performance. Drawing from multiple data sources would offer a more comprehensive view of a principal's effectiveness and inform decision-making more accurately than reliance on a single measure.

In terms of future research, this study sheds light on principals' experiences with evaluation outcomes. However, quantitative or mixed-methods research involving a larger sample could provide a more comprehensive understanding of the topic. Future studies might explore the effectiveness of different incentive structures, the complexities of promotion opportunities, strategies for delivering meaningful feedback, and the implementation of structured professional development plans. There is also a clear opportunity to develop and validate reliable evaluation metrics that more accurately measure principal performance, given the continuing emphasis on objectivity. Ultimately, evaluation outcomes have the potential to significantly influence the career trajectories of school principals. These outcomes affect not only motivation, commitment, and professional growth, but also career advancement within the educational system. Positive evaluations can serve as powerful motivators, foster a culture of excellence, and encourage principals to continue strengthening their leadership abilities. Conversely, negative outcomes may generate a sense of urgency and prompt underperforming principals to seek targeted support and engage in professional development. In both cases, evaluation outcomes play a critical role in shaping principals' sense of purpose, their commitment to student success, and the broader development of effective school leadership.

Acknowledgments

The authors express their sincere gratitude to the principals and supervisors who participated in this study. Their contributions provided essential insights into key aspects of principal evaluation.

References

- Abu Dhabi Education Council (2012), *Abu Dhabi Education Reform: the Road to 2030*, [PDF], Abu Dhabi Education Council, available at: https://centres.insead.edu/innovation-policy/events/policy-breakfasts/documents/ad_edu_ref_pres-2012april.pdf
- Al-Hamdan, J.M. and Al-Yacoub, A.M. (2005), "Evaluation of section heads' performance at Kuwait secondary schools", *International Journal of Educational Management*, Vol. 19 No. 3, pp. 233-245, doi: [10.1108/09513540510591011](https://doi.org/10.1108/09513540510591011).
- Alkaabi, A.M. (2025), "A qualitative multi-case study of supervision in the principal evaluation process in the United Arab Emirates", *International Journal of Leadership in Education*, Vol. 28 No. 1, pp. 53-80, doi: [10.1080/13603124.2021.2000032](https://doi.org/10.1080/13603124.2021.2000032).
- Alkaabi, A.M. and Almaamari, S.A. (2020), "Supervisory feedback in the principal evaluation process", *International Journal of Evaluation and Research in Education*, Vol. 9 No. 3, pp. 503-509, doi: [10.11591/ijere.v9i3.20504](https://doi.org/10.11591/ijere.v9i3.20504).
- Amrein-Beardsley, A. (2023), "A Florida teacher wrongfully terminated? Alleged (mis)uses of value-added model (VAM) estimates for high-stakes teacher evaluation decisions", *Educational Assessment, Evaluation and Accountability*, Vol. 35 No. 3, pp. 313-352, doi: [10.1007/s11092-023-09411-8](https://doi.org/10.1007/s11092-023-09411-8).
- Amrein-Beardsley, A. and Geiger, T.J. (2019), "Potential sources of invalidity when using teacher value-added and principal observational estimates: artificial inflation, deflation, and conflation", *Educational Assessment, Evaluation and Accountability*, Vol. 31 No. 4, pp. 465-493, doi: [10.1007/s11092-019-09311-w](https://doi.org/10.1007/s11092-019-09311-w).
- Anderson, L.M. and Turnbull, B.J. (2016), "Evaluating and supporting principals: building a stronger principalship", *Policy Studies Associates*, Vol. 4, available at: <https://eric.ed.gov/?id=ED570471>
- Andrews, C. (1990), "Evaluating principals", *Research Roundup*, Vol. 6 No. 2, pp. 1-6, ERIC Clearinghouse on Educational Management, available at: <https://files.eric.ed.gov/fulltext/ED318131.pdf>
- Assiry, A.A., Alnemari, A., Adil, A.H., Karobari, M.I., Sayed, F.R., Marya, A., Basheer, S.N. and Marya, C.M. (2022), "Extensive evaluation of the overall workplace experience and job satisfaction of dentists in Saudi Arabia", *BioMed Research International*, Vol. 2022 No. 1, 4968489, doi: [10.1155/2022/4968489](https://doi.org/10.1155/2022/4968489).
- Brauckmann, S. and Pashiardis, P. (2010), "The clash of evaluations: in search of the missing link between school accountability and school improvement—experiences from Cyprus", *International Journal of Educational Management*, Vol. 24 No. 4, pp. 330-350, doi: [10.1108/09513541011045263](https://doi.org/10.1108/09513541011045263).
- Braun, V. and Clarke, V. (2006), "Using thematic analysis in psychology", *Qualitative Research in Psychology*, Vol. 3 No. 2, pp. 77-101, doi: [10.1191/1478088706qp063oa](https://doi.org/10.1191/1478088706qp063oa).
- Bryman, A. and Burgess, B. (Eds) (1994), *Analyzing Qualitative Data*, 1st ed., Routledge, London, doi: [10.4324/9780203413081](https://doi.org/10.4324/9780203413081).
- Burkhauser, S., Gates, S.M., Hamilton, L.S., Li, J.J. and Pierson, A. (2013), *Laying the Foundation for Successful School Leadership*, (RR-419-RC), RAND Corporation, available at: https://www.rand.org/pubs/research_reports/RR419.html
- Casserly, M., Lewis, S., Simon, C., Uzzell, R. and Palacios, M. (2013), *Principal Evaluations and the Principal Supervisor: Survey Results from the Great City Schools*, Council of the Great City Schools, available at: <https://www.cgcs.org/cms/lib/DC00001581/Centricity/Domain/87/Principal%20evaluation%20report.pdf>
- Charmaz, K. (2006), *Constructing Grounded Theory: a Practical Guide Through Qualitative Analysis*, Sage.

- Chiang, H., Lipscomb, S. and Gill, B. (2012), "Assessing the feasibility of using value-added models for principal evaluations", (Mathematica Policy Research Working Paper), Mathematica Policy Research.
- Chiang, H., Wellington, A., Hallgren, K., Speroni, C., Herrmann, M., Glazerman, S. and Constantine, J. (2015), *Evaluation of the Teacher Incentive Fund: Implementation and Impacts of pay-for-performance After Two Years*, (NCEE 2015-4020). National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education, available at: <https://ies.ed.gov/ncee/pubs/20154020/pdf/20154020.pdf>
- Chiang, H., Lipscomb, S. and Gill, B. (2016a), "Is school value added indicative of principal quality?", *Education Finance and Policy*, Vol. 11 No. 3, pp. 283-309, doi: [10.1162/EDFP_a_00184](https://doi.org/10.1162/EDFP_a_00184).
- Chiang, H., McCullough, M., Lipscomb, S. and Gill, B. (2016b), *Can Student Test Scores Provide Useful Measures of School Principals' Performance?*, (NCEE 2016-002). National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education, available at: <https://ies.ed.gov/ncee/pubs/2016002/pdf/2016002.pdf>
- Clifford, M., Hansen, U., Lemke, M., Wraight, S., Menon, R. and Brown-Sims, M. (2011), *A Practical Guide to Designing Comprehensive Principal Evaluation Systems*, National Comprehensive Center for Teacher Quality, available at: <https://eric.ed.gov/?id=ED532778>
- Clifford, M., Hansen, U.J. and Wraight, S. (2014), *Practical Guide to Designing Comprehensive Principal Evaluation Systems*, Center on Great Teachers and Leaders, available at: <https://files.eric.ed.gov/fulltext/ED555668.pdf>
- Corbin, J.M. and Strauss, A.L. (2008), *Basics of Qualitative Research: Techniques and Procedures for Developing Grounded Theory*, 3rd ed., Sage.
- Corcoran, A., Casserly, M., Price-Baugh, R., Walston, D., Hall, R. and Simon, C. (2013), *Rethinking Leadership: the Changing Role of Principal Supervisors*, The Wallace Foundation, available at: <http://www.wallacefoundation.org/knowledge-center/Documents/Rethinking-Leadership-The-Changing-Role-of-Principal-Supervisors.pdf>
- Creswell, J.W. and Poth, C.N. (2017), *Qualitative Inquiry and Research Design: Choosing Among Five Approaches*, 4th ed., Sage.
- Davis, S.H., Kearney, K., Sanders, N.M., Thomas, C. and Leon, R. (2011), *The Policies and Practices of Principal Evaluation: a Review of the Literature*, WestEd, available at: https://www2.wested.org/www-static/online_pubs/resource1104.pdf
- DeMatthews, D.E., Scheffer, M. and Kotok, S. (2020), "Useful or useless? Principal perceptions of the Texas principal evaluation and support system", *Journal of Research on Leadership Education*, Vol. 16 No. 4, pp. 279-304, doi: [10.1177/1942775120933920](https://doi.org/10.1177/1942775120933920).
- Donaldson, M., Mavrogordato, M., Dougherty, S.M., Al Ghanem, R. and Youngs, P. (2021a), "Principal evaluation under the elementary and secondary every student succeeds act: a comprehensive policy review", *Education Finance and Policy*, Vol. 16 No. 2, pp. 347-361, doi: [10.1162/edfp_a_00332](https://doi.org/10.1162/edfp_a_00332).
- Donaldson, M., Mavrogordato, M., Youngs, P., Dougherty, S.M. and Al Ghanem, R. (2021b), "Doing the 'real' work: how superintendents' sensemaking shapes principal evaluation policies and practices in school districts", *AERA Open*, Vol. 7, pp. 1-16, doi: [10.1177/2332858420986177](https://doi.org/10.1177/2332858420986177).
- El Zaatari, W. and Ibrahim, A. (2021), "What promotes adolescents' sense of school belonging? Students and teachers' convergent and divergent views", *Cogent Education*, Vol. 8 No. 1, 1984628, doi: [10.1080/2331186X.2021.1984628](https://doi.org/10.1080/2331186X.2021.1984628).
- Elfers, A.M. and Plecki, M.L. (2017), *Washington's Teacher and Principal Evaluation System: Examining the Implementation of a Complex Policy*, Center for the Study of Teaching and Policy, University of Washington, available at: <https://ospi.k12.wa.us/sites/default/files/2023-10/tpimplementationuwreport2017.pdf>
- Fitzpatrick, J.L., Sanders, J.R. and Worthen, B.R. (2011), *Program Evaluation: Alternative Approaches and Practical Guidelines*, 4th ed., Pearson.

- Fuller, E.J., Hollingworth, L. and Liu, J. (2015), "Evaluating state principal evaluation plans across the United States", *Journal of Research on Leadership Education*, Vol. 10 No. 3, pp. 164-192, doi: [10.1177/1942775115614291](https://doi.org/10.1177/1942775115614291).
- Fulmer, I.S. and Li, J. (2022), "Compensation, benefits, and total rewards: a bird's-eye (re)view", *Annual Review of Organizational Psychology and Organizational Behavior*, Vol. 9 No. 1, pp. 147-169, doi: [10.1146/annurev-orgpsych-012420-055903](https://doi.org/10.1146/annurev-orgpsych-012420-055903).
- Garet, M.S., Wayne, A.J., Brown, S., Rickles, J., Song, M. and Manzeske, D. (2017), *The Impact of Providing Performance Feedback to Teachers and Principals*, (NCEE 2018-4001). National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education, available at: <https://files.eric.ed.gov/fulltext/ED578873.pdf>
- Glesne, C. (1999), *Becoming Qualitative Researchers: an Introduction*, 2nd ed., Longman, available at: https://search.worldcat.org/title/Becoming-qualitative-researchers-%3A-an-introduction/oclc/40135345?utm_source=chatgpt.com
- Goff, P., Goldring, E. and Canney, M. (2016), "The best laid plans: pay-for-performance incentive programs for school leaders", *Journal of Education Finance*, Vol. 42 No. 2, pp. 127-152, available at: <https://eric.ed.gov/?id=EJ1169815>
- Goldring, E.B., Mavrogordato, M. and Haynes, K.T. (2015), "Multisource principal evaluation data: principals' orientations and reactions to teacher feedback regarding their leadership effectiveness", *Educational Administration Quarterly*, Vol. 51 No. 4, pp. 572-599, doi: [10.1177/0013161X14556152](https://doi.org/10.1177/0013161X14556152).
- Grissom, J.A., Kalogrides, D. and Loeb, S. (2015), "Using student test scores to measure principal performance", *Educational Evaluation and Policy Analysis*, Vol. 37 No. 1, pp. 3-28, doi: [10.3102/0162373714523831](https://doi.org/10.3102/0162373714523831).
- Grissom, J.A., Blissett, R.S.L. and Mitani, H. (2018), "Evaluating school principals: supervisor ratings of principal practice and principal job performance", *Educational Evaluation and Policy Analysis*, Vol. 40 No. 3, pp. 446-472, doi: [10.3102/0162373718783883](https://doi.org/10.3102/0162373718783883).
- Grissom, J.A., Egalite, A.J. and Lindsay, C.A. (2021a), *How Principals Affect Students and Schools*, The Wallace Foundation, available at: <https://wallacefoundation.org/sites/default/files/2023-09/How-Principals-Affect-Students-and-Schools.pdf>
- Grissom, J.A., Egalite, A.J. and Lindsay, C.A. (2021b), "How principals affect students and schools: a systematic synthesis of two decades of research", *The Wallace Foundation*, available at: <https://www.wallacefoundation.org/knowledge-center/Documents/How-Principals-Affect-Students-and-Schools.pdf>
- Gümüş, S. and Bellibaş, M.Ş. (2020), "The relationship between professional development and school principals' leadership practices: the mediating role of self-efficacy", *International Journal of Educational Management*, Vol. 34 No. 7, pp. 1155-1170, doi: [10.1108/IJEM-10-2019-0380](https://doi.org/10.1108/IJEM-10-2019-0380).
- Henry, G.T. and Viano, S.L. (2016), "An evaluation of the North Carolina educator evaluation system for school administrators: 2010-11 through 2013-14", *Consortium of Educational Research and Evaluation—North Carolina*, available at: <https://fi.ncsu.edu/resource-library/an-evaluation-of-the-north-carolina-educator-evaluation-system-for-school-administrators-2010-11-through-2013-14/>
- Herrmann, M. and Ross, C. (2016), *Measuring Principals' Effectiveness: Results from New Jersey's First Year of Statewide Principal Evaluation*, (REL 2016-156). Regional Educational Laboratory Mid-Atlantic, National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education, available at: <https://files.eric.ed.gov/fulltext/ED568434.pdf>
- Hvidston, D.J., Range, B.G. and McKim, C.A. (2015), "Principals' perceptions regarding their supervision and evaluation", *AASA Journal of Scholarship and Practice*, Vol. 12 No. 2, pp. 20-34, available at: <https://link.gale.com/apps/doc/A421771216/AONE?u=anon~7b17bc36&sid=googleScholar&xid=5a7137c3>
- Hvidston, D.J., McKim, C.A. and Holmes, W.T. (2018), "What are principals' perceptions? Recommendations for improving the supervision and evaluation of principals", *NASSP Bulletin*, Vol. 102 No. 3, pp. 214-227, doi: [10.1177/0192636518802033](https://doi.org/10.1177/0192636518802033).

- Ibrahim, A.S. (2012), "Induction and mentoring of novice teachers: a scheme for the United Arab Emirates", *Teacher Development*, Vol. 16 No. 2, pp. 235-253, doi: [10.1080/13664530.2012.688676](https://doi.org/10.1080/13664530.2012.688676).
- Jones, E., Bergin, C. and Murphy, B. (2022), "Principals may inflate teacher evaluation scores to achieve important goals", *Educational Assessment, Evaluation and Accountability*, Vol. 34 No. 1, pp. 57-88, doi: [10.1007/s11092-021-09366-8](https://doi.org/10.1007/s11092-021-09366-8).
- Kane, T. and Staiger, D. (2002), "The promise and pitfalls of using imprecise school accountability measures", *The Journal of Economic Perspectives*, Vol. 16 No. 4, pp. 91-114, doi: [10.1257/089533002320950993](https://doi.org/10.1257/089533002320950993).
- Kegan, R. and Lahey, L. (2016), "An everyone culture: becoming a deliberately developmental organization", *Stanford Social Innovation Review*, doi: [10.48558/X1XV-3Z72](https://doi.org/10.48558/X1XV-3Z72).
- Kvale, S. (1997), *The Qualitative Research Interview*, Sage Publications, London.
- Lincoln, Y.S. and Guba, E. (1985), *Naturalistic Inquiry*, Sage, available at: https://www.researchgate.net/profile/Jan-Armstrong/publication/256294652_Naturalistic_Inquiry/links/02e7e5222605c57278000000/Naturalistic-Inquiry.pdf
- McCullough, M., Lipscomb, S., Chiang, H., Gill, B. and Cheban, I. (2016), *Measuring School Leaders' Effectiveness: Final Report from a Multiyear Pilot of Pennsylvania's Framework for Leadership*, (REL 2016-106). Regional Educational Laboratory Mid-Atlantic, National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education, available at: <https://files.eric.ed.gov/fulltext/ED563446.pdf>
- Merriam, S.B. (1998), *Qualitative Research and Case Study Applications in Education*, 2nd ed., Jossey-Bass.
- Micheaux, D.J. and Parvin, J.L. (2018), "Principal evaluation as a tool for growth", *The Learning Professional*, Vol. 39 No. 2, pp. 52-55, available at: <https://learningforward.org/wp-content/uploads/2018/05/principal-evaluation-as-a-tool-for-growth.pdf>
- Mitani, H. (2025), "Performance-based compensation systems and principal job performance", *American Educational Research Journal*, Vol. 62 No. 2, pp. 414-455, doi: [10.3102/00028312241303760](https://doi.org/10.3102/00028312241303760).
- Nelson, J.L., Grissom, J.A. and Cameron, M.L. (2021), "Performance, process, and interpersonal relationships: explaining principals' perceptions of principal evaluation", *Educational Administration Quarterly*, Vol. 57 No. 4, pp. 641-678, doi: [10.1177/0013161X211009295](https://doi.org/10.1177/0013161X211009295).
- Parylo, O., Zepeda, S.J. and Bengtson, E. (2012), "Principals' experiences of being evaluated: a phenomenological study", *Educational Assessment, Evaluation and Accountability*, Vol. 24 No. 3, pp. 215-238, doi: [10.1007/s11092-012-9150-x](https://doi.org/10.1007/s11092-012-9150-x).
- Patton, M.Q. (2015), *Qualitative Research and Evaluation Methods*, 4th ed., Sage.
- Pham, L.D., Nguyen, T.D. and Springer, M.G. (2021), "Teacher merit pay: a meta-analysis", *American Educational Research Journal*, Vol. 58 No. 3, pp. 527-566, doi: [10.3102/0002831220905580](https://doi.org/10.3102/0002831220905580).
- Portin, B., Feldman, S. and Knapp, M.S. (2006), *Purposes, Uses, and Practices of Leadership Assessment in Education*, Center for the Study of Teaching and Policy, University of Washington, available at: <https://eric.ed.gov/?id=ED494235>
- Qablan, A. and Al-Qaderi, S. (2009), "How to change university faculty members' attitudes and behavior in the context of education for sustainable development", *Applied Environmental Education and Communication*, Vol. 8 No. 3, pp. 184-194, doi: [10.1080/15330150903269407](https://doi.org/10.1080/15330150903269407).
- Reeves, D.B. (2013), *Assessing Educational Leaders: Evaluating Performance for Improved Individual and Organizational Effectiveness*, 2nd ed., Corwin Press.
- Schuermann, P.J., Guthrie, J.W., Prince, C.D. and Witham, P.J. (2009), *Principal Compensation and Performance Incentives*, Center for Educator Compensation Reform, U.S. Department of Education.

- Seidman, I. (2012), *Interviewing as Qualitative Research: a Guide for Researchers in Education and the Social Sciences*, 4th ed., Teachers College Press.
- Shen, Z., Simon, C.E. and Kelcey, B. (2016), "The potential consequence of using value-added models to evaluate teachers", *JEP: Journal of Education Policy*, pp. 1-12.
- Sinnema, C.E.L. and Robinson, V.M.J. (2012), "Goal setting in principal evaluation: goal quality and predictors of achievement", *Leadership and Policy in Schools*, Vol. 11 No. 2, pp. 135-167, doi: [10.1080/15700763.2011.629767](https://doi.org/10.1080/15700763.2011.629767).
- Sun, M., Youngs, P., Yang, H., Chu, H. and Zhao, Q. (2012), "Association of district principal evaluation with learning-centered leadership practice: evidence from Michigan and Beijing", *Educational Assessment, Evaluation and Accountability*, Vol. 24 No. 3, pp. 189-213, doi: [10.1007/s11092-012-9145-7](https://doi.org/10.1007/s11092-012-9145-7).
- Turnbull, B.J., Anderson, L.M., Riley, D.L., MacFarlane, J.R. and Aladjem, D.K. (2016), "The principal pipeline initiative in action: building a stronger principalship", *Policy Studies Associates*, Vol. 5, available at: <https://www.wallacefoundation.org/knowledge-center/Documents/Building-a-Stronger-Principalship-Vol-5-The-Principal-Pipeline-Initiative-in-Action.pdf>
- Wayne, A., Garet, M., Wellington, A. and Chiang, H. (2018), *Promoting Educator Effectiveness: the Effects of Two Key Strategies*, (NCEE 2018-4009). National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education, available at: <https://ies.ed.gov/ncee/pubs/20184009/pdf/20184009.pdf>
- Yan, R. (2020), "The influence of working conditions on principal turnover in K–12 public schools", *Educational Administration Quarterly*, Vol. 56 No. 1, pp. 89-122, doi: [10.1177/0013161X19840391](https://doi.org/10.1177/0013161X19840391).
- Yin, R.K. (2014), *Case Study Research: Design and Methods*, 5th ed., Sage.
- Zepeda, S.J., Lanoue, P.D., Price, N.F. and Jimenez, A.M. (2014a), "Principal evaluation—Linking individual and building-level progress: making the connections and embracing the tensions", *School Leadership and Management*, Vol. 34 No. 4, pp. 324-351, doi: [10.1080/13632434.2014.928681](https://doi.org/10.1080/13632434.2014.928681).
- Zepeda, S.J., Parylo, O. and Bengtson, E. (2014b), "Analyzing principal professional development practices through the lens of adult learning theory", *Professional Development in Education*, Vol. 40 No. 2, pp. 295-315, doi: [10.1080/19415257.2013.821667](https://doi.org/10.1080/19415257.2013.821667).
- Zepeda, S.J., Alkaabi, A. and Tavernier, M. (2020), "Leadership and supervision", *Oxford Research Encyclopedia of Education*, Oxford University Press, Oxford, doi: [10.1093/acrefore/9780190264093.013.617](https://doi.org/10.1093/acrefore/9780190264093.013.617).

Further reading

Stake, R.E. (1995), *The Art of Case Study Research: Perspectives on Practice*, Sage.

Corresponding author

Ahmed M. Alkaabi can be contacted at: ahmed.alkaabi@uaeu.ac.ae