

1-22-2024

## Exploring teachers' experiences within the teacher evaluation process: A qualitative multi-case study

Sana Butti Al Maktoum  
*Zayed University*

Ahmed M. Al Kaabi  
*United Arab Emirates University*

Follow this and additional works at: <https://zuscholars.zu.ac.ae/works>



Part of the [Education Commons](#)

---

### Recommended Citation

Maktoum, Sana Butti Al and Kaabi, Ahmed M. Al, "Exploring teachers' experiences within the teacher evaluation process: A qualitative multi-case study" (2024). *All Works*. 6317.  
<https://zuscholars.zu.ac.ae/works/6317>

This Article is brought to you for free and open access by ZU Scholars. It has been accepted for inclusion in All Works by an authorized administrator of ZU Scholars. For more information, please contact [scholars@zu.ac.ae](mailto:scholars@zu.ac.ae).

# Exploring teachers' experiences within the teacher evaluation process: A qualitative multi-case study

Sana Butti Al Maktoum & Ahmed M. Al Kaabi

**To cite this article:** Sana Butti Al Maktoum & Ahmed M. Al Kaabi (2024) Exploring teachers' experiences within the teacher evaluation process: A qualitative multi-case study, Cogent Education, 11:1, 2287931, DOI: [10.1080/2331186X.2023.2287931](https://doi.org/10.1080/2331186X.2023.2287931)

**To link to this article:** <https://doi.org/10.1080/2331186X.2023.2287931>



© 2024 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.



[View supplementary material](#)



Published online: 22 Jan 2024.



[Submit your article to this journal](#)



Article views: 96



[View related articles](#)



[View Crossmark data](#)



Received: 01 June 2023  
Accepted: 21 November 2023

\*Corresponding author: Sana Butti Al Maktoum, College of Education, Zayed University, Dubai, United Arab Emirates  
E-mail: [sana.almaktoum@zu.ac.ae](mailto:sana.almaktoum@zu.ac.ae)

Reviewing editor:  
Sammy King Fai Hui, Curriculum & Instruction, The Education University of Hong Kong, HONG KONG

Additional information is available at the end of the article

## EDUCATIONAL ASSESSMENT & EVALUATION | RESEARCH ARTICLE

# Exploring teachers' experiences within the teacher evaluation process: A qualitative multi-case study

Sana Butti Al Maktoum<sup>1\*</sup> and Ahmed M. Al Kaabi<sup>2</sup>

**Abstract:** This multi-case study explores teachers' experiences of the teacher evaluation process implemented in schools across the UAE. Data were collected using interviews and document analysis and covered the seven emirates using the same evaluation process; seventeen teachers—15 female and 2 male teachers—participated in online and face-to-face semi-structured interviews. Our objective was to examine teachers' experiences in the yearly evaluation cycle enacted in public schools during the 2021–2022 academic year, from the formative evaluation process to the summative evaluation review. It uncovers the overall quality of the evaluation cycle, role of administrators, how the formative evaluation process promotes professional growth, and challenges and outcomes of summative evaluation. An analysis of the collected findings reveals four themes related to teachers' experiences as recipients of the evaluation process: (1) unreliable indicators to judge teacher quality, (2) lack of motive to provide evidence of performance, (3) episodic superficial feedback, and (4) compliance versus the satisficing mindset of teachers and evaluators. These findings have implications for practice and further



Sana Butti Al Maktoum

### ABOUT THE AUTHOR

Alkaabi is an assistant professor in the Foundations of Education Department—College of Education at United Arab Emirates University. He is currently serving as the Director of the Emirates Institute for Learning Outcomes Assessment at UAEU as well as Coordinator of the Master of the Educational Innovation Program. His educational qualifications include a Ph.D. in Educational Administration and Policy with an emphasis on supervision from the University of Georgia–Athens, USA (2019). Dr. Alkaabi was the recipient of two distinguished academic awards: the Ray Bruce Award in 2017 for his dedicated work and projects in the field of instructional supervision, and the Faculty Award in 2019 for his accomplishments in the Educational Administration and Policy Doctoral Program at the University of Georgia. His research interests reflect his expertise in school leadership, specifically in the areas of supervision, evaluation, induction, and professional development.

### PUBLIC INTEREST STATEMENT

This study was conducted in the United Arab Emirates to investigate teachers' experience in the evaluation process in public schools. Evaluation is commonly used as an accountability measure and to help improve teaching and learning. Seventeen teachers from the seven Emirates articulated their thoughts on the evaluation process as an unfair image of their actual practices in class. Teachers are held accountable through various evaluation variables such as students' attainment scores, omitting other contextual factors within the school, the learner, and the outside community. Therefore, the UAE governmental and educational stakeholders, part of the evaluation system, need to consider the effects of the current evaluation tool voiced by teachers before it manifests and negatively affects teaching and learning in schools.

research to inform stakeholders of teachers' raw experiences within the evaluation process and promote positive communication channels with teachers to improve the cyclical education process.

**Subjects: Education Policy & Politics; School Leadership, Management & Administration; Education - Social Sciences**

**Keywords: UAE schools; professional growth; formative evaluation; summative evaluation; cyclical education process; documentation; feedback; incentives**

## 1. Introduction

Over the past decade, there have been drastic changes in educational systems worldwide, particularly in terms of evaluation reforms. Generally, an evaluation is meant to elevate performance, and its primary goal is to improve teaching and learning (Darling-Hammond, 2015; Donaldson, 2016). However, it also extends to other variables such as ensuring that students attain high test scores and achievement (Kim & Youngs, 2016), especially since countries worldwide have measured students on international assessments such as PISA. In the UAE, the results of international assessments are inducted as part of educational reform (Badri et al., 2013; Ibrahim et al., 2020; Matsumoto, 2019). However, results from the PISA of 2018 show lower average scores in reading literacy among learners (OECD, 2021). Disparities thereof are evident as immigrants score higher than nationals, opposing the realities in other countries (Marquez et al., 2022). Nationals are commonly equated with the higher-scoring majority in public schools, where the type of evaluation analyzed in this study occurs. In the past, the UAE followed subject supervisors who visited classes to evaluate teachers (Gaad et al., 2006). Indeed, the former initiatives have been criticized based on biases that may occur, as principals have the freedom to select supervisors, which encourages superficial evaluation (Farah & Ridge, 2009). Gaad et al. (2006) further clarified that evaluation was based on supervisors' ratings across instructional variables such as planning, teaching skills, knowledge of the subject, and innovation in teaching. Other variables include teachers' professional development and personality. However, current evaluations have been replaced with a more systematic form that encompasses multiple variables to weigh both teacher and student performance, such as the number of classes, student performance in international assessment, and professional training. The official manual from the evaluation platform shows four main skills (development, presentation, soft, and technical skills) and an equal percentage of 25% to calculate the overall performance for a calendar year (see S1).

Research on evaluation in the context of the UAE has been largely confined to higher education (Mercer, 2007), with some conducted in schools, particularly among novice teachers (Ibrahim, 2012) and others reporting on principals and supervisors in schools (Alkaabi, 2021; Alkaabi & Almaamari, 2020). To the best of our knowledge, no study has been conducted on school teachers and their experience with the evaluation. Therefore, it is crucial to investigate teachers' routine evaluations and everyday practices in the seven emirates of the UAE to fill this gap in the literature (Hammad & Hallinger, 2017; Oplatka & Arar, 2017). To this end, this qualitative multi-case study seeks to answer the following questions: How are school teachers supervised and evaluated? How is feedback delivered, and how effective is it in improving teachers' professional practices? What types of support do school administrators provide during the evaluation process? What challenges do teachers face during evaluation? What are the incentives and consequences for those with high and low scores?

## 2. Literature review

This section will explore critical aspects of evaluation, encompassing evaluation variables and contextual factors, motivation, performance, incentives, the role of instructional leaders, and the specific nature of evaluation in the UAE.

### **2.1. Evaluation variables and contextual factors**

Realistic reflection of teachers' evaluation and performance in their overall grades has been increasingly discussed in the literature (e.g., Darling-Hammond, 2014; Moran, 2017; Tobiason, 2019). The main challenge is to define an effective teacher and address the discrepancies among stakeholders, particularly regarding accountability. The latter is more challenging in education than in other fields (Doğan & Ayduğ, 2023). However, one important understanding is that teachers are not the sole variable of accountability, and other factors such as school leaders, the school community, students, and parents play a role (Cohen & Goldhaber, 2016; Darling-Hammond, 2000; Levinson, 2011; Lingenfelter, 2003).

One significant challenge faced by the evaluation system is measuring students' performance in relation to teachers' performance and measuring teachers' performance using a number of variables, including the correlation with students' test results from state-mandated tests. As state-mandated tests measure students' reading proficiency, Lenhoff et al. (2017) examined the association between teachers' evaluation scores and students' reading results. The study found a correlation between teachers rated as "effective" and "highly effective" and higher reading achievement of their students. The present study measured correlation statistically, but doing so only raised questions regarding its validity, particularly as it excluded other considerations such as teachers' teaching methods, pupils' readiness, and social or emotional background. Contextual factors, such as the latter (Cohen & Goldhaber, 2016; Darling-Hammond & McLaughlin, 2011), and practical factors, such as curriculum and policy (Darling-Hammond, 2004; Ibrahim et al., 2020), can affect teacher performance and are normally not measured. Tucker and Stronge (2005), for instance, expressed doubts about correlating student performance with teacher evaluations and having fixed standards for all. They emphasized the role of student growth as a more effective measure than achievement scores. Tying student grades to teacher evaluation results can create "demoralization" in the teaching profession as this approach responds more to grades than progress. Other evaluative measures include low-inference classroom observations based on standardized rubrics, student learning objectives (SLOs), student surveys, and portfolios. Using multiple measures, such as multiple observations and student surveys, is deemed beneficial (Close & Amrein-Beardsley, 2018; NCTQ, 2019). SLOs use classroom-level measures to determine student growth and teacher effectiveness (Crouse et al., 2016; Gill et al., 2013).

Proof of documentation as evidence of teaching and learning in the classroom can be used as an accountability measure (Seitz, 2008), but, more importantly, it can be used to document student progress at different intervals (Joseph et al., 2014).

In a study conducted within Urban Teacher Residency programs that integrated multiple essential sources, an endeavor was undertaken to help educators understand and assess teachers' teaching quality and performance. Kawasaki et al. (2020) collected data from seven distinct sources: (1) observation rubrics, (2) teaching artifacts, (3) instructional logs, (4) Value-added measures, (5) Assessment based pedagogical content knowledge, (6) surveys of teachers and mentors, and (7) teacher portfolios. Despite the considerable effort and challenges faced during data collection, these measures captured various aspects of teaching quality with varying depths and details. Consequently, employing these measures collectively could address the intricate nature of teaching quality in ways that are both theoretically and empirically grounded, and can be applied for program enhancement.

Kawasaki et al. (2020) determined that there is no optimal, entirely scientific, or objective method for weighing or combining multiple measures of teacher evaluations. Any framework discussed inherently contains some degree of subjectivity; the real concern is not the presence of subjective, nonscientific considerations, but their location, nature, and magnitude. Clearly articulating the assumptions and judgments that shape the design of a teacher evaluation system and its goals, components, and procedures could empower educators to oversee the system's functioning more effectively, implement necessary modifications, and ultimately provide evidence

affirming the validity of the conclusions about teacher efficacy and the system's utility in enhancing teaching practices.

Darling-Hammond (2004) emphasized the need to consider schools holistically when discussing accountability and standards; as such, she asserts that teacher training can elevate students' scores (Darling-Hammond, 2004). Teacher training is another controversial topic that has been narrated by teachers as robust and mechanical, targeted at raising teachers' scores while disregarding student growth (Warren & Ward, 2019).

Similarly, Martell (2016) asserted that the top-down education system affects professional development, as certain goals are addressed as student outcomes on standardized tests that abort teachers' views and needs. Thus, teacher training is the sum of its effectiveness (Elmore & Burney, 1997). Furthermore, teacher training and professional development can be beneficial if accompanied by incentives for teachers.

### **2.2. Motivation, performance, and incentives**

Incentives have been a topic along the lines of evaluation and can be accompanied by multi-faceted evaluation, including classroom observation, student surveys, student scores (Springer, 2009), or career-ladder programs and benefits (Dee & Keys, 2004). Incentives are used as a motivating variable but can also be utilized as tenure for low performance. Internationally, merit pay is used in the less disadvantaged, economically or academically failing schools to retain teachers in the profession (Blackmore et al., 2023; Springer et al., 2016). On the other hand, performance pay has been targeted explicitly for raising students' achievement, such as the Colorado and Denver trials (Chiang et al., 2015). Other studies declare that to do so, students' scores need to be tracked for three rounds to declare any valid association between evaluation and student scores (Hanushek & Rivkin, 2010). Singapore, which is like the UAE context in two variables: the aim to score in international tests and the diversity within the population, uses performance-based evaluation (Steiner, 2010). In the UAE, the performance management system is a newly enacted system (Waxin & Bateman, 2016), evidently in minimal small-scale studies (Al Bustami, 2014) and the effect on the educational sector, whereas a study in the health sector reports that monetary rewards are the most favored amongst health professionals (Younies et al., 2008). Even though a lack of incentives can hamper motivation and has been shown to lower morale amongst teachers in public and private schools in the UAE (McKnight et al., 2016), it is unrealistic to give incentives to all (Gagnon, 2016). Thus, there is a need to consider non-monetary rewards such as community support (Nyamubi, 2017) and teacher collegiality, which can be more beneficial than incentives (Blackmore et al., 2023). Indeed, other variables like organizational climate have more substantial outcomes in comparison to incentives (Destler, 2016), one important variable to climate is the role of leaders in schools (Arif et al., 2019).

### **2.3. Instructional leadership role (supervisory role and feedback)**

Principals in schools commonly monitor evaluation and supervision processes; thus, they play an essential role in this regard. Another significant role of the leadership is providing feedback. Feedback is the essence of supervisory measures and is the element that sets it apart from the evaluation. The former provides support in the form of classroom visits and feedback, while the latter is used as a compliance measure (Zepeda & Ponticell, 1998).

Before delving into further details regarding the relationship between supervision and evaluation, it is crucial to highlight how feedback practices may deteriorate due to the additional responsibilities and duties placed upon school administrators, not to mention the demands of accountability. Research indicates that many school districts implemented new teacher evaluation systems in ways that may undermine administrators' capacity to offer frequent and effective feedback. In most states, the additional responsibility of conducting time-intensive teacher observations has been added to administrators' existing duties without the provision of extra support or training (Kraft & Gilmour, 2016; Neumerski et al., 2018). Consequently, administrators might resort

to “satisficing” behaviors, such as conducting brief observations and delivering generic positive feedback (Halverson et al., 2004). In the context of teacher evaluation, “satisficing” refers to a potential issue where administrators might opt for a quick, superficial evaluation process that meets minimum requirements or standards, rather than conducting a more comprehensive and in-depth assessment of a teacher’s performance.

For example, in Chicago, researchers discovered that administrators dominated post-observation discussions and seldom posed open-ended, higher-order questions that encourage teachers to reflect on their teaching practices (Sartain et al., 2011). Principals who primarily viewed the evaluation process as an accountability tool tend to invest minimal time in providing constructive feedback (Kraft & Gilmour, 2016). Therefore, the quantity and quality of feedback that teachers receive through the evaluation process significantly depend on the skills, capacity, and objectives of school leaders (Donaldson & Woulfin, 2018).

In the literature, supervision, and evaluation, the two are often conflated, which can compromise the quality of both processes. Various metaphors highlight the fragile relationship between teachers, who are recipients of supervision, and leaders, who sometimes resort to unproductive tactics. These tactics can lead to an environment in which both parties are distrustful; instructional practices remain private; and teacher agency, efficacy, and collaboration are overshadowed. This is because supervision, evaluation, and professional development are imposed on them (Alkaabi, 2021, 2023; Zepeda et al., 2020). Support is provided when real solutions to classroom issues are discussed after observation. Indeed, feedback targeting instructional strategies is a common theme in teachers’ statements (Mireles-Rios & Becchio, 2018). Instructional strategies, in turn, need to move from superficial recommendations, wherein teachers are measured by whether they remember to write the date on the board (Gallagher, 2019) to essential recommendations. Further, at the preservice teacher training stage, emphasis on constructive feedback is highly needed (Mohebi & ElSayary, 2022).

Educators have observed a challenging relationship between the formative and summative aspects of supervision and evaluation. Based on the literature, Alkaabi (2023) contended that, when evaluators offer formative feedback, and then transition to providing summative judgments on performance, potential issues such as role conflict, eroding trust, and delivering inconsistent messages can arise. Popham (2013) recommended that supervisors actively participate in both formative and summative evaluations but should do so distinctly. By incorporating formative goals into daily practice and using them as steps toward the summative phase of an evaluation, leaders can identify those who may be underperforming. Moreover, they can devise professional development plans to ensure that decisions and actions foster growth (Alkaabi, 2021).

#### **2.4. Nature of evaluation and leadership roles in the UAE**

Similar to international behavior, the UAE also faces the engagement of other stakeholders in educational change and policy (Bruns et al., 2019), often keeping the teacher the last to be engaged in any decision-making process. In fact, in the UAE, there are no teacher unions, and globally, there is tension between teacher unions’ influence in the face of high-stakes accountability (Ministry of Education, 2022). Further, the nature of educational leadership in the UAE is characterized by a top-down method evidently present in the several entities supervising schools, and the larger authority of educational policy is the Ministry of Education (Badri, 2019). In fact, the numerous educational entities and a highly centralized system subtract leaders in schools from the decision-making (Al-Taneiji & McLeod, 2008; Stephenson et al., 2012) and, in the case of evaluation, transfer it to another entity outside the school and the Ministry of Education. The Human Resources Management Information System (HRMIS) enacts the evaluation for teachers in schools and other governmental employees and is responsible for incentives and promotions. Even though there are other local evaluative measures, such as Irtiq’a for the emirate of Abu Dhabi, to monitor public and private schools, focused on raising the ranking of international student assessments (Morgan & Ibrahim, 2020). Overall, the UAE aims to reach the top 20 countries in the PISA

results and the top 15 countries in the TIMSS (UAE Vision 2021, 2018). Thus, the blueprint for evaluative measures is guided by the federal vision (MoCa, 2019) and international benchmarks (Morgan & Ibrahim, 2020; Pataki, n.d.). Thus, international trends reflect the nature of evaluation in the UAE; the PISA and TIMSS results judge schools' overall quality (Morgan & Ibrahim, 2020) and serve as a benchmarking method worldwide (Ibrahim et al., 2020). In addition to enacting evaluative measures for teachers, stakeholders recruit teachers. Badri points out the qualification of teachers has a direct effect on students being assessed internationally; the latter is deemed the job of stakeholders and not schools, which in most cases have direct observation and assessment of their current teaching staff (Badri, 2019). Regionally, the Middle East has been documented as having issues in governance and accountability and in the education domain to be able to address the rapid changes encountered in schools and higher education (Kamel, 2014). Governance and accountability can be challenging and confusing since many entities work together to serve the educational community. However, the nation has recently been on the mission of centralizing and merging various regulatory entities (Gallagher, 2019).

### 3. Materials and methods

This study adopted a qualitative research approach to align with its objectives of investigating the implementation of evaluation and supervision procedures in public schools. Qualitative research, as defined by Rubin and Rubin (2016), involves gathering detailed accounts of human behavior and beliefs within specific contexts.

Denzin and Lincoln (2003) emphasized that qualitative research employs an interpretive and naturalistic perspective, focusing on understanding phenomena in their natural settings and interpreting them based on individuals' attributed meanings. The primary objective of this study is to delve into aspects such as factors impacting evaluation, the delivery and effectiveness of feedback in enhancing teaching practices, the support offered by school administrators, challenges teachers face during the evaluation process, and the incentives and consequences associated with varying evaluation scores.

To comprehensively address these objectives, a multi-case study design will be employed. Creswell and Clark (2017) highlight the case study's popularity due to its capacity to provide an in-depth understanding of specific individuals, identified problems, or unique situations through thorough investigation. Furthermore, the case study design offers the advantage of data triangulation, which involves collecting data through interviews, observations, and document analysis. This triangulation enhances credibility by examining data from multiple perspectives, aiming to identify consistencies across various data sources.

Data for this study were collected through semi-structured interviews with 17 teachers from seven Emirates exposed to the public school evaluation process. Analysis of the qualitative data revealed how formative evaluation was enacted during the 2021–22 academic school year, including the manner in which school administrators supervised and assessed teachers, how feedback was delivered to teachers, what kind of support school administrators provided, the challenges that teachers faced in the evaluation process, and the incentives that were applied to receive high evaluation scores. Information from each of these areas offers a better understanding of teachers' experiences.

#### 3.1. Context of research site and participants

The UAE government continues to raise the bar in order to provide excellent education and exceptional standards of living. In 2018, 11 million dirhams of the total federal budget of 51.4 million were allocated to the Ministry of Education (UAE Cabinet, 2023). In addition, structural changes and continuous educational reforms have occurred in the UAE (Gallagher, 2019). Recently, the Ministry of Education in the UAE has shifted from being the primary monitor of schools to being a structural entity. At the same time, the Emirates School Establishment (ESE) acts as



a groundwork supplier, monitoring program strategies and national and international exams (Emirates School Establishment, 2022).

Regarding performance management, the HRMIS was launched in January 2013 for all federal government employees (FAHR, 2023). Schoolteachers are considered federal government employees and are part of the management system for all services, including evaluation tools. Teachers are a critical part of the educational sector, and the total number of teachers serving the seven emirates was 211,153 in 2019 (Ministry of Education, 2022). Given their status and importance, there is a vested interest in studying the components of evaluation tools and determining the extent to which they can improve education from teachers' perspectives. Teacher participants (see Table 1 below) were chosen using representative sampling, with special attention paid to including at least one participant from each Emirate. Further, the participants varied in the subject area, cycle "school level," and experience. As for the evaluation score, all participants received accomplished, and only two received above expectations.

### 3.2. Data collection methods

Interviews were used as the primary data collection method in this case study to gather more in-depth information about the evaluation processes and supervisory practices. As Seidman (2012) stated, the purpose of interviews is not to test hypotheses, but rather to understand the experiences of others and make sense of them. Each teacher participant was interviewed in person or online for one hour about their experiences with the evaluation process. The semi-structured interviews allowed for flexibility while paying attention to the overarching purposes (see Glesne, 1999). Before each interview, the researchers and participating teachers notified the Emirates School establishment by letter. The researchers contacted all participants via email to send them consent forms and let them know that the interviews would be recorded, while allowing the option of activating the camera for online interviews to adhere to the country's cultural norms.

Additional data sources, including various documents, were added to this multi-case study to triangulate the data with other sources. Documents were utilized, in part because they could not be influenced by the interviewers (Merriam, 1998). The data sources consisted of formal documents from the evaluation platform shared by school personnel, (1) the system of management tool—phase one- *January to February 2018*, (2) self-assessment—an evidence-based systematic tool, (3) self- service human resources user guide, (4) performance management tool—Interim review, the last and fifth document is a lesson observation shared by a teacher. Bowen (2009) indicates that the presence of at least two documents can aid in document analysis in corroborating or divergence of the results (Bowen, 2009).

Overall, the ethical checklist and procedures proposed by Patton (2014) were used to ensure that ethical considerations and guidelines were properly followed. The research purpose and methods were illuminated, the benefits and reciprocity for participants were explained, any potential risks were introduced clearly, assurance of confidentiality and informed consent were fulfilled, data access and ownership were clarified, and the research data were examined through strict confines established by the Research Ethics Committees (Ref. ERS\_2021\_8414).

### 3.3. Data analysis

The researchers used the constant comparative method when analyzing the data to enable them to distinguish between experiences within the teachers' accounts. The role of the two researchers was essential to eliminate any misunderstandings, and cross-checking was implemented post-translationally (from Arabic to English) for all interviews. Prior to cross-checking, the researchers combined the data to create coding using Atlas-ti, a software platform used to divide large amounts of data into codes and quotes. Subsequently, small segments were matched to overarching themes. A constant revision between the researchers was essential to determine the reliability and accuracy of the theme-code combinations.

**Table 1. Overview of teacher participants—all names are pseudonyms**

Participant	Gender	School Level/Subject	Nationality	Emirate	Years of Experience	Evaluation Score
Sameera	Female	Elementary/Arabic	Emirati	Dubai	25	Accomplished
Basma	Female	Elementary/Arabic and Islamic	Emirati	Dubai	15	Accomplished
Fatema	Female	Elementary/Science	Emirati	Dubai	9	Above Expectations
Laila	Female	Elementary/Math	Jordanian	Dubai	17	Accomplished
Sami	Male	Elementary/English	Egyptian	Abu Dhabi	7	Accomplished
Huda	Female	Elementary	Emirati	Abu Dhabi	6	Accomplished
Rami	Male	Elementary/Science	Egyptian	Abu Dhabi	10	Accomplished
Mai	Female	Secondary/Math	Palestinian	Sharjah	25	Accomplished
Manal	Female	Elementary/Science	Emirati	Sharjah	23	Accomplished
Mariam	Female	Elementary/English	Emirati	Fujairah	8	Accomplished
Isra	Female	Elementary/English	Emirati	Fujairah	8	Above Expectations
Nada	Female	Secondary/Biology and Science	Emirati	Ras Al Khaima	18	Accomplished
Omniah	Female	Kindergarten/English	Emirati	Ras Al Khaima	8	Accomplished
Sara	Female	Elementary/Design and Technology	Emirati	Um Al Quwain	14	Accomplished
Nora	Female	Kindergarten/English	Emirati	Um Al Quwain	13	Accomplished
Haya	Female	Elementary/Social studies	Emirati	Ajman	17	Accomplished
Wedad	Female	Elementary/Islamic Social Studies and Arabic	Emirati	Ajman	14	Accomplished

Individual interview transcripts were analyzed as part of the qualitative framework of this study using a thematic analysis approach to discover themes across the data from the interviews with each participant (Maguire & Delahunt, 2017). In the first phase of the analysis, the researchers conducted a preliminary content check to verify the translation accuracy. In the second phase, researchers open-coded the transcripts to identify phrases or individual words that conveyed the most fundamental meanings behind the emotions, activities, and thought processes contained therein. Themes were formed in phase three of the analysis, each derived from similarly grouped elements or quotes found across the participants' responses (Patton, 2002). The researchers double-checked how accurately these themes represented the groups of participants' statements in the fourth and final phases.

### **3.4. Trustworthiness of the study**

To ensure that the data and findings were of high quality and trustworthiness, attention was paid to the study's credibility, confirmability, dependability, and transferability (Lincoln & Guba, 1985). The first factor, credibility, is often regarded as the most important in determining the trustworthiness of a study, as it focuses on proving the truth of the findings (Weiss, 1998). In this study, credibility was attained through triangulation and member-checking. Seale (1999) stated that triangulation increases the credibility of research by considering several data sources instead of a single source to provide multiple viewpoints.

In terms of confirmability, or the extent to which the study results were unbiased, the researchers monitored the progress of the investigation using an audit trail as evidence (see Patton, 2014). The justification for each action and choice made during the study was viewed as a result of the audit trail. Additionally, peer consultations were held with researchers of similar studies in the field of education to identify any potential biases that might have still been present (Lincoln & Guba, 1985). The researchers contributed different views on the research process and content of the results.

Dependability, used interchangeably with reliability, refers to the consistency or repeatability of research and data collection processes that yield similar outcomes (Lincoln & Guba, 1985). The main objective of dependability is to "minimize errors and biases in a study" (Yin, 2014, p.49). By carefully examining the integrity of the present study, the researchers were able to maintain their dependability and ensure that other researchers would reach the same conclusions if the same research processes were used.

Regarding transferability, the researchers included a detailed account of the research environment and the assumptions made during the research process to underscore the extent to which the findings may be applicable or transferable outside the scope of the study. Although the results of this study are not generalizable because of factors such as the small sample size, they might be transferable to school districts with traits similar to those found in this study (Lincoln & Guba, 1985).

## **4. Results**

The primary objective of this study is to investigate the implementation of evaluation and supervision procedures in public schools. This investigation takes into account various factors influencing the evaluation process, the delivery and effectiveness of feedback in improving teaching practices, the support provided by school administrators, the challenges faced by teachers during evaluations, and the incentives and consequences associated with high and low evaluation scores. The analysis of the collected data revealed four key themes that directly address the research questions: (1) unreliable indicators to judge teacher quality, (2) lack of motive to provide evidence of performance, (3) episodic superficial feedback, and (4) compliance versus the satisficing mindset of teacher and evaluators.

#### **4.1. Theme one: Unreliable indicators to judge teacher quality**

In this theme, five indicators, presented as subthemes, appeared in the evaluation system that directly affected the evaluation and final grades teachers received on their evaluations, including evaluation indicators, student grades versus student progress, physical attendance, professional development, and teaching loads. The evaluation indicators, type, and weight were subjects of controversy among teachers. The second subtheme outlined the correlation between student and teacher grades, where teacher evaluation results replaced student progress with student grades. Student attendance highlights how students' physical attendance affects teachers' evaluation scores. Sub-theme four further discusses professional development, along with supporting quotes from the participants. The last indicator, and perhaps the most important, is teaching load and its effect on yearly evaluation. The collective indicators were derived from the first research question on how school teachers are supervised and evaluated. The following sections discuss each of the subthemes with additional examples and elaborations.

##### *4.1.1. Evaluation indicators: Type and weight*

Teacher participants revealed a few issues with the indicators in the evaluation system, namely addressing non-subject-related indicators accounting for 50 percent of the overall criteria. Sami explained, "Fifty percent are related to the profession and community, which I do not think are subject-related." Wedad elaborated more on the criteria: "They do not measure our performance. If you look at the indicators, many of them have nothing to do with what happens inside the classroom." Some of the indicators, as explained by teachers, are general in nature, such as the use of technology, infusion of global topics and national identity within the subject matter, and conveying happiness and positivity to learners within the school community. Sameera added that the latter is an easy "high" score given by the principal to all of her staff. The former topics were aligned when examining the lesson observation document and the principal's feedback, as it showed the emphasis on relating lesson topics to real-life examples and national identity (see S2).

Basma elaborated on the matter: "Do not be fooled here, this standard equates for only five percent." She compared other indicators that accounted for 10 percent of the score: "those you have no control over, such as the class quantity, and also for the professional development."

The latter dilemma of percentages was further investigated by exploring the documents; none of the documents clearly indicate how percentages are distributed. The formal documents from the evaluation platform loosely specify 25% for the following, development, presentation, soft, and technical skills, as mentioned earlier (see S1). Another document, obtained from a science teacher, had the goals, such as positive learning experience and incorporating high expectations of learners, but empty percentages awaiting the principal and headquarters' approval (see S3).

Some indicators proved unachievable, as Sami mentioned in an example he gave: "All students must participate ... I have 30 students in [the] classroom, which makes [it] hard for me to let every student participate in one classroom." Isra concurred, stating, "Too many things they need, and they want to see it in the classroom, but no time." Although some of the participants were able to articulate the evaluation criteria, others were unsure about the language used. Rami explained, "Teachers should demonstrate excellence in all practices. What is the definition of 'excellence'?" Another evaluation indicator that led to controversy was the manner in which the indicators were calculated. Sara expressed her view: "The grade is so ambiguous, and we have no clue what it is." Huda added, "How things are calculated across all evaluations is unknown!" Wedad further explained that the percentage was shown and the grades were added, but "we still do not know how the overall grade came about."

Language teachers declared that basic reading and writing skills were not evident in the criteria, expressing their sentiments, "I am ... an Arabic teacher. My goal is to teach reading and writing." One indicator flagged by Fatema was the use of innovation and creativity in the class. She confirmed, "This is more specific to the subject matter of mine, which is science." Sami, another

participant, explained, “These tools are designed as a generic tool for evaluating the teachers in all [Ministry of Education] schools with no considerations to the specific contexts of specific schools, subject areas, geographical areas, or any other aspect.” Laila concurred and added The evaluation is for all the subjects. They do not customize it for subjects ... How can you have goals for all the teachers? How can you join scientific and literary subjects? They are not equal! [For example] arts and physical education.

Basma, who had experienced conflict with the use of the one-size-fits-all evaluation approach, objected to a specific criterion placed in the rubric that suggested that national tests were only subjected to a certain school cycle. She expressed her thoughts, “How is it reasonable to have this K-12 test to be used in my case, as I teach grade one.”

#### 4.1.2. *Student grade versus student progress*

Teachers expressed their struggles with their students’ grades as an indicator of their teaching skills. Indeed, the nature of tying student grades to teacher grades in an evaluation system often drives credibility concerns. Mai explained, “For example, you would hear someone complaining about a student level as low, and then give her a grade of 90. How is this possible?” In other instances, supporting the “failing student is encouraged by the administration,” Manal asserted. Rami explored the juxtaposition of grades and progress, contributing his thoughts to this mix:

The principal would not think or take into account if this teacher has high-achieving students or low-achieving students. The growth made does not count or [is not] considered. He would evaluate you without putting other overwhelming factors that may largely influence your evaluation.

Huda elaborated on the same idea by adding:

Talking about low-achieving students, the considerate evaluator will not take for granted the end result of their achievement! At least this evaluator will see the growth made, or difference made, between terms and then give more sensible judgment about my performance.

In subjects who adhered to accumulated knowledge, such as math, other factors played a role in the grade versus progress balance. Laila explained, “A student in grade four in math depends on his metacognition of the former years such as grade three and two. Some other subjects do not require this. So, they basically measure what he has studied this particular year.” Similarly, comprehending language depends in large measure on reading and writing. For instance, Basma, an Arabic teacher, was given “a very, very weak class,” and she had to follow “a special literacy program.” She expressed her frustrations, “Imagine they know the class is weak and yet the students’ grades are merged with mine? I was affected. That is not fair.”

#### 4.1.3. *Physical attendance*

Student attendance also had a direct correlation with teachers’ grades in the evaluation system, which is often detrimental to teachers. Mariam explained that if a student is regularly absent, this affects the evaluation. Therefore, “the evaluation was based on their attendance.” Wedad added, “Attendance of the student has nothing to do with the teacher. It is the responsibility of the parents or the student himself, but it is part of the evaluation!” Omniah interpreted the process as “... certainly the responsibility of the student and the parent.” She continued, “plus I cannot force him to come to school.”

#### 4.1.4. *Professional development*

The participants’ perspectives on professional development, overall grades, and connections with each other diverged significantly. One participant, Mai, stated. “We attend weekly training online or in person. In all evaluations, the teacher must do it all and receive two, which meets expectations.”

Another participant, Huda, added her view: “PD is not linked to our evaluation results.” Several teachers saw workshops as necessary task in order to receive good evaluation scores at the end of the year. However, other teachers were unsure about this issue. As S3 shows, one teacher even outlined the specific goals that they were to meet by the end of the year to receive high evaluation scores. The diverging viewpoints of the participants revealed no small degree of unpredictability among the evaluators and school leaders.

The second issue within the professional development domain was the content of the workshops provided to teachers. As the school teachers recounted, the content was decided by the training department of the Ministry of Education, which followed generic goals. However, the teachers agreed that conversations about everyday tasks would be more valuable. Manal elaborated, “We first try to get over the scientific part of our lesson that comes from the training department that we have to cover—and mostly useless. We try to finish it hastily to discuss real matters such as exams.”

#### 4.1.5. *Teaching load*

Another indicator of the evaluation system was the teaching load, specifically, the number of hours taught in a week—which was vital for increasing or decreasing teachers’ overall grades at the end of the year. Teachers collectively declared that their teaching load was not within their power and was normally assigned based on ministry procedures, administrative decisions, and section availability. The correlation between workload and the teachers’ overall grades was considered illogical. Basma, for instance, shared her interpretation: “Say you teach 24 classes per week, you get ‘meets expectations.’” Sameera added, “So, it is quantity versus quality.” Mariam also asserted her view, stating, “So the more classes, the better, without looking at the ways of teaching.” Mai explained that the difference in one class accounted for a variation between earning a “meets expectations” and “above expectations.” Omniah experienced similar circumstances as a teacher in two different cycles and was evaluated accordingly. She explained:

In high school, you can easily teach 24 classes per week, but in kindergarten, I can teach 12 classes. But this is not a true measure of my performance . . . In kindergarten, I deploy more effort, as young people require more attention.

#### 4.2. **Theme two: Lack of motive to provide evidence of performance**

Another important follow-up to the criteria in the evaluation system is proof of documentation, which in turn addresses two research questions: one, on the challenges teachers face during evaluation, and two, documentation consequently affects incentives and consequences. As explained by the teachers, the online evaluation system had the option of uploading as many documents as possible for each criterion to address accomplishments and credibility. The latter proved controversial when it came to the purpose and mechanics of providing documentation, such as what was accepted and what was not, and the reasoning behind it. Huda explained that it was not clear-cut and “there was no unified expectation regarding what to accept and reject in terms of evidence.” Sami held a similar view in his interpretations.

I really do not have any problem collecting evidence but really hate when I am told, “This is not [what] we are looking for.” Then, why did you not specifically tell me what kind of evidence to collect in the first place?

Teachers were divided according to the purpose of providing evidence. Some participants indicated that it was compulsory at some point. Other participants stated that documentation was only required upon request for promotion or if there was an error or discrepancy in the calculation.

In one of the documents obtained from the evaluation manual, there was an indication of the evidence needed to support practice, “avoid using memory to prove accomplishments done in a year’s time.” The latter statement supports the need for proof in line with any achievements done by teachers.

#### **4.3. Theme three: Episodic superficial feedback**

One of the core concepts of evaluation is feedback, and the research question of this study aimed to understand how feedback is delivered and how much support is given to teachers. The findings showed that several teachers believed that the feedback they received was superficial and episodic in nature, simply given to check off their managerial duties as administrative personnel on the evaluation system and the Ministry of Education. “It is a kind of routine, but I feel it is done in [a] perfunctory way to comply with the Ministry of Education.” Sami added that the post-evaluation was “not accompanied by a discussion or [a] plan to make me improve.” In one case, Sami’s school principal pointed out his limited “lesson objectives” and asked him to “make more.” He elaborated, “We did not have a conversation to talk more about the reasons for having [just] one.” Sami asserted that his justification for having one lesson objective—that the content is challenging—was received with dismay. Huda had a similar experience. She explained that she had received feedback and comments but not a discussion. Rami added that if you are “lucky,” you will get a discussion, but it is mostly “bullet-point statements.” The latter was evident as a common practice when exploring the documents, specifically lesson observation and feedback (see S4), which shows bullet points and general statements such as in need to emphasize reading and writing and high-order thinking in class.

Further, Laila described the feedback as “not specific and very general” due to the principal’s lack of content knowledge. A similar dilemma was shared by Isra, who felt that administrators were limited when observing her lessons because she taught English by speaking English as a mode of instruction.

The participants agreed that the feedback they received from the administration was episodic and almost all participants identified the evaluation as a one-time event when asked whether it was an ongoing process. When asked about the duration of their evaluations, most participants described the process as summative. However, these two participants conflicted. Sami explained: “When the evaluator visits me and throws [out] a comment with a score, I feel it’s summative.” In like manner, Isra defined the evaluation in the following way: “Let’s say, if they are coming four times throughout this term, the first three would be formative and the last one would be summative.” Rami, conversely, stated, “You are watched all the time! This gives the feeling of being summative despite of hav[ing] three observations because you are being evaluated three times!”

The documents obtained from the evaluation platform outline broad evaluation goals without specific details (such as international assessment and student attendance, refer to S3), which leaves room for varying practices that can be shaped and influenced by school administrators in different schools. Indeed, the second document, lesson observation goals, obtained from schools and provided by the Ministry of Education, shows classroom-specific goals (such as quality of teaching and learning, refer to S5).

#### **4.4. Theme four: Compliance versus the satisficing mindset of teachers and evaluators**

A salient and logical path toward improving evaluation scores is the implementation of an incentive, typically a financial incentive or promotion. Thus, this topic has been crafted as a research question to understand if the robust evaluation tool includes incentives and consequences. Most participants shared that they had not received any incentives, including those who disclosed that they were given “high” evaluation scores in the last evaluation round. A few participants disclosed that they had previously received certificates. Teachers collectively echoed the sentiments of Sara, who said, “In the field of teaching, the certificate has no value. It is a paper placed in a folder.” Isra added her view: “Financial rewards from time to time. Let us be honest here, not everyone cares about the [a] certificate.” Regarding low-achieving teachers, several explained that this was not announced. Some even shared that they would typically be transferred

to another school and undergo compulsory remedial programs. Wedad explained that in some cases, “the ones that go for training are the ones with 20 years’ experience.”

The updated evaluation documents clearly state the following, “the direct manager should provide feedback on the employees’ performance and allow sufficient time for discussion before the system closes for revision.” Also, the direct manager should avoid indicators such as “always” and “never” and present vague information that is not supported by raw data or realistic examples.

## 5. Discussion

One of the primary aims of this study was to discover how school teachers are supervised and evaluated, revealing that evaluation is vague, unfair, and challenging. Teachers’ evaluation perspectives are similar to those in other parts of the world, such as Korea and the United States (Danielson, 2011; Park, 2006; Popham, 2013). As the interviews clarified, the evaluation variables were universal and quantifiable measures were used. Scholars agree on the need for qualitative measures in line with quantitative evaluation (Chong et al., 2014) and an evaluation that can yield an understanding of teaching and learning in class (Harris et al., 2014). In consensus, teacher evaluation has two main purposes: as an accountability measure and as a tool of improvement (Stronge, 1995, 2005); the former has been depicted through student standardized tests and the latter in professional training, and both are problematic (Darling-Hammond, 2004; Warren & Ward, 2019) if not used effectively. In the era of student growth measures and increased teacher accountability, teachers have not always seen evaluations as a method to improve their skills (Clipa, 2011; Robertson-Kraft & Zhang, 2018; Tucker & Stronge, 2005). Indeed, the inclusion of standardized tests causes teachers to leave teaching altogether, especially because they correlate with their evaluations (Darling-Hammond & Wise, 1985; Downey, 2016; González & Darling-Hammond, 1997). Similar attitudes were evident among the participants in this study, as they explained that evaluations based on student grades could corrupt the essence of teaching, such as transparency in grading. The latter is an international dilemma wherein teachers are held accountable for their students’ progress, disregarding any contextual factors that may come into play (Buckner, 2018; Close et al., 2019; Darling-Hammond, 2020; Morgan & Ibrahim, 2020). Another problematic variable is using standardized tests as a measure of teacher performance to reward merits, which can be inaccurate due to its single-subjects and engaging in non-authentic assessment (Kim & Mikiewicz, 2021). Discovering similar findings as authentic assessments are encouraged, such as student portfolios, and teachers were aware that those authentic assessments can elevate their overall score of the evaluation. However, as narrated by teachers, it is time-consuming, and the platform can be restricted and locked at other times.

Teachers are on a quest to improve and hone their skills to improve or sustain a job (Popham, 2013). Still, little is known about the effect of the support they are given, whether it is professional development or feedback from supervisors—in most cases, principals or vice principals. For instance, principals influence schools, and their academic supervision and feedback can reap benefits (Hardono et al., 2017), but engaging principals with two roles to address formative and summative teacher evaluations can also be challenging and time-consuming (Kraft & Gilmour, 2016; Popham, 2013) especially since there is a lack of autonomy in schools in the UAE evident in the current study. In fact, the literature asserts the importance of autonomy for schools to explore reward systems in line with the data set of students in the school

(Schultz, 2017). The evaluation platform in the UAE is led by policymakers who may not have the background in education to be able to form a solid understanding of evaluative measures and teachers’ performance. The findings of this study mirror international practice wherein policymakers design evaluations, but school leaders are the ones who rate and assign percentages for teachers (Lipsky, 2010) and should also be held accountable in the evaluation process (Thompson & Samuels-Lee, 2020). Thus, in reality, administrators engage in satisficing behaviors, including conducting short observations and providing generic positive feedback (Halverson & Clifford, 2006). latter was proven in this study and documented in the UAE context, where supervision occurs hastily with minimal feedback (Al Bustami, 2014; Alkaabi & Almaamari, 2020). Similarly, the participants in this study expressed feedback as



episodic and superficial; they described the support they received as artificial and resembled a checklist for administrative purposes that provided little benefit to real classroom issues. In a similar study, teachers expressed the vagueness of career progression and the overall assessment of teaching (Goe et al., 2020). Superficial feedback may be due to principals' organizational duties and the absence of real issues in the classroom (Cohen & Goldhaber, 2016). The latter was evident in teachers' narratives of the feedback given in the classroom. Increasing the length of classroom observations to gain insight into teaching (Firestone & Donaldson, 2019) or lesson studies and lesson observations as a tool for improving teaching and learning (Abdallah et al., 2023; Fernandez, 2002; Fernandez & Yoshida, 2004) can be more beneficial and sustainable. Achenushure et al. (2020) address the lack of mentoring programs in the UAE, which may account for teacher attrition. Indeed, the absence of mentoring in line with the heavy evaluation scheme followed in schools may be troublesome, as reported in this study. Furthermore, the absence of an induction program can increase the burden on novice teachers and even lead them to consider leaving the teaching profession (Abdallah & Alkaabi, 2023).

Another important variable to evaluate teachers is the presence of incentives. Kraft and Gilmour (2016) state that merit pay is limited to top performers due to budget restrictions, and on the other hand, veteran teachers are particularly not motivated by the pay-for-performance method (Chiang et al., 2015). The latter has been voiced by veteran teachers in this study when asked about incentives who instead aimed for promotion to head of department or administrative position. The latter is an important measure to induce motivation in teachers and emphasize their career advancement, which feeds into motivational theories (Shikalepo, 2020) and is a common practice in other parts of the world to apply vertical promotion in line with evaluation (Isoré, 2009).

The findings of this study resonate with and extend prior literature in several critical ways. Firstly, our research aligns with the existing body of knowledge that highlights the universal challenges associated with teacher evaluation, as observed in various countries such as Korea and the United States (Danielson, 2011; Park, 2006; Popham, 2013). It underscores the need for a balanced approach that combines both qualitative and quantitative measures (Chong et al., 2014) and emphasizes evaluation as a tool for both accountability and improvement (Stronge, 1995, 2005). Moreover, our study echoes the concerns raised in the literature regarding the use of standardized tests as a sole measure of teacher performance, which can lead to unintended consequences (Darling-Hammond & Wise, 1985; Kim & Mikiewicz, 2021). Similarly, the preference for authentic assessments such as student portfolios, while acknowledging their time-consuming nature, reflects a growing discourse on alternative evaluation methods (Close et al., 2019). Furthermore, our findings shed light on the critical role of school leaders, who often play a pivotal role in teacher evaluation (Kraft & Gilmour, 2016), and the need for administrators to provide meaningful, constructive feedback rather than superficial evaluations (Halverson & Clifford, 2006). Overall, as seen from the study findings and literature showcases insights from the UAE context while reinforcing the broader challenges and considerations surrounding teacher evaluation in education systems worldwide.

## 6. Conclusion

This study illuminated the teacher evaluation process, particularly focusing on the process, quality and delivery of feedback, inherent challenges in the evaluation process, and subsequent incentives and consequences. The thematic analysis identified four themes related to feedback quality: (1) (Unreliable indicators to judge teacher quality), (2) lack of motive to provide evidence of performance, (3) episodic superficial feedback, and (4) compliance versus the satisficing mindset of teacher and evaluators. These results have implications for future research, emphasizing the need for robust evaluation processes to enhance teaching quality and support teachers. This study had several practical implications. The consensus among teachers for practice teachers suggests that quantitatively assessing their practices is "unfair," and that genuine teaching practices are best observed in the classroom. Evaluators should consider teachers' backgrounds, including previous evaluation feedback, years of experience, and the subjects taught. Analyzing past evaluations can equip evaluators better. It is also crucial to consider factors such as school environment, teaching load, available resources, students' prior academic achievements, and other broader school

context elements (Baker et al., 2010; Kraft & Gilmour, 2016). When used appropriately, standardized tests can significantly improve teaching and learning by guiding professional development and addressing students' specific needs (Darling-Hammond & Rustique-Forrester, 2005)

Teacher evaluations should incorporate multiple sources to ensure a holistic view of teacher performance. Evaluators must also be aware of potential conflicts between evaluation and supervision. Transitioning from formative assessments to summative judgments can introduce challenges, such as role conflict, trust erosion, and mixed messages. To address these issues, supervisors should engage in both formative and summative evaluations, but only distinctly. This method converts formative goals into daily activities, leading to a summative evaluation phase. This strategy not only identifies underperformers, but also fosters professional growth and ensures quality decisions that promote improvement.

Feedback during both evaluation types should be constructive, supportive, and actionable, highlighting teachers' strengths and areas of growth. Effective feedback fosters professional development and enhances teaching and learning qualities. Finally, an effective incentive system for teacher evaluation is vital to acknowledge their dedication and hard work. Such a system can boost motivation, promote excellence, and improve school performance.

In future studies, several areas have emerged as potential avenues for understanding and improving evaluation components. First, there is a need to understand better how criteria with unified definitions and interpretations can benefit the teacher evaluation process and help reduce anxiety and feelings of discrimination, as reported by teachers. Subsequent research can further explore the diverse facets of teacher evaluation. This includes scrutinizing the tools employed during observation visits, investigating if supervision should be seamlessly integrated with evaluation or kept distinct to foster learning from errors and development, and undertaking comprehensive studies on the merits of integrating student achievement into teacher assessments, among other topics. The current study can be extended to a quantitative study to identify how common some complaints are and determine which elements of the evaluation system are more problematic and in need of revision. Thus, there is a need to explore further the nuanced relationship between supervision and evaluation and clinical supervision in the UAE. Hence, teacher evaluations are an indispensable aspect of our education system, as they play a crucial role in ensuring that the quality of education provided to students is of the highest standard. Through teacher evaluations, teachers' experiences and skills are carefully assessed and scrutinized, ensuring that they are fully equipped to impart the knowledge and skills necessary to promote student learning and academic growth.

#### Acknowledgments

The authors thank all the participating teachers. Their engagement and input provided valuable insights into the UAE's evaluation process in the United Arab Emirates.

#### Author details

Sana Butti Al Maktoum<sup>1</sup>  
E-mail: [sana.almaktoum@zu.ac.ae](mailto:sana.almaktoum@zu.ac.ae)  
ORCID ID: <http://orcid.org/0000-0003-1808-9034>  
Ahmed M. Al Kaabi<sup>2</sup>  
ORCID ID: <http://orcid.org/0000-0001-7220-8087>

<sup>1</sup> Zayed University, Dubai, United Arab Emirates.

<sup>2</sup> Foundations of Education Department, United Arab Emirates University, Al Ain, United Arab Emirates.

#### Disclosure statement

No potential conflict of interest was reported by the author(s).

#### Data availability statement

The data supporting the findings of this study are available on request from the corresponding author, [SA, and AA]. The data are not publicly available due to [restrictions such as containing information that could compromise the privacy of research participants].

#### Supplementary material

Supplemental data for this article can be accessed online at <https://doi.org/10.1080/2331186X.2023.2287931>

#### Citation information

Cite this article as: Exploring teachers' experiences within the teacher evaluation process: A qualitative multi-case study, Sana Butti Al Maktoum & Ahmed M. Al Kaabi, *Cogent Education* (2024), 11: 2287931.

#### References

- Abdallah, A. K., & Alkaabi, A. M. (2023). Induction programs' effectiveness in Boosting New teachers' instruction and student achievement: A critical Review. *International Journal of Learning, Teaching and Educational Research*, 22(5), 493–517. <https://doi.org/10.26803/ijlter.22.5.25>
- Abdallah, R. K., Al Maktoum, S. B., & Al Mansoori, M. K. (2023). The road to lesson observation as a tool to school improvement: Accountability vs. Perfunctory. In A. K. Abdallah & A. M. Alkaabi (Eds.), *Advances in Educational Marketing, administration, and leadership* (pp. 222–252). IGI Global. <https://doi.org/10.4018/978-1-6684-7818-9.ch012>

- Achenushure, M., Nereo, R., & Pius, A. (2020). Mentoring and the difference it makes in teachers' work: A literature review. *European Journal of Education Studies*, 7(6).
- Al Bustami, G. (2014). Improving the teacher's evaluation methods and tools in Abu Dhabi schools—case study. *Athens Journal of Social Sciences*, 1(4), 261–274. <https://doi.org/10.30958/ajss.1-4-3>
- Alkaabi, A. M. (2021). A qualitative multi-case study of supervision in the principal evaluation process in the United Arab Emirates. *International Journal of Leadership in Education*, 1–28. <https://doi.org/10.1080/13603124.2021.2000032>
- Alkaabi, A. M. (2023). Revitalizing supervisory models in Education: Integrating adult learning Theories and stage Theories for enhanced teaching and learning outcomes. In A. K. Abdallah & A. M. Alkaabi (Eds.), *Advances in Educational Marketing, administration, and leadership* (pp. 253–277). IGI Global. <https://doi.org/10.4018/978-1-6684-7818-9.ch013>
- Alkaabi, A., & Almaamari, S. (2020). Supervisory feedback in the principal evaluation process. *International Journal of Evaluation and Research in Education (IJERE)*, 9(3), 503–509. <https://doi.org/10.11591/ijere.v9i3.20504>
- Al-Taneiji, S., & McLeod, L. (2008). Towards decentralized management in United Arab Emirate (UAE) schools. *School Effectiveness and School Improvement*, 19(3), 275–291. <https://doi.org/10.1080/09243450802246384>
- Arif, S., Zainudin, Z., & Hamid, A. (2019). Influence of Leadership, organizational culture, work motivation, and job satisfaction of performance principles of senior high school in Medan city. *Budapest International Research and Critics Institute (BIRCI-Journal): Humanities and Social Sciences*, 2(4), 239–254. <https://doi.org/10.33258/birci.v2i4.619>
- Badri, M. (2019). School emphasis on academic success and TIMSS Science/Math achievements. *International Journal of Research in Education & Science*, 5(1), 176–189.
- Badri, M. A., Mohaidat, J., Ferrandino, V., & El Mourad, T. (2013). The social cognitive model of job satisfaction among teachers: Testing and validation. *International Journal of Educational Research*, 57, 12–24. <https://doi.org/10.1016/j.ijer.2012.10.007>
- Baker, E. L., Barton, P. E., Darling-Hammond, L., Haertel, E., Ladd, H. F., Linn, R. L., Ravitch, D., & Rothstein, R. (2010). *Problems with the Use of Student Test Scores to Evaluate Teachers*. <https://www.epi.org/publication/bp278/>
- Blackmore, J., Hobbs, L., & Rowlands, J. (2023). Aspiring teachers, financial incentives, and principals' recruitment practices in hard-to-staff schools. *Journal of Education Policy*, 1–20. <https://doi.org/10.1080/02680939.2023.2193170>
- Bowen, G. A. (2009). Document analysis as a qualitative research method. *Qualitative Research Journal*, 9(2), 27–40. <https://doi.org/10.3316/QRJ0902027>
- Bruns, B., Macdonald, I. H. & Schneider, B. R. (2019). The politics of quality reforms and the challenges for SDGs in education. *World Development*, 118, 27–38.
- Buckner, E. (2018). *The Other Gap: Examining Low-Income Emiratis' Educational Achievement*. <https://publications.alqasimifoundation.com/en/the-other-gap-examining-low-income-emiratis-educational-achievement>
- Chiang, H., Wellington, A., Hallgren, K., Speroni, C., Herrmann, M., Glazerman, S., & 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. <https://eric.ed.gov/?id=ED559723>
- Chong, S., Chatterji, M., Kevin, G., & Welner, P. (2014). Academic quality management in teacher education: A Singapore perspective. *Quality Assurance in Education*, 22(1), 53–64. <https://doi.org/10.1108/QAE-05-2012-0023>
- Clipa, O. (2011). Teacher perceptions on teacher evaluation: The purpose and the assessors within the assessment process. *Procedia - Social & Behavioral Sciences*, 29, 158–163. <https://doi.org/10.1016/j.sbspro.2011.11.220>
- Close, K., & Amrein-Beardsley, A. (2018). Learning from what doesn't work in teacher evaluation. *Phi Delta Kappan*, 100(1), 15–19. <https://doi.org/10.1177/0031721718797115>
- Close, K., Amrein-Beardsley, A., & Collins, C. (2019). Mapping America's teacher evaluation plans under ESSA. *Phi Delta Kappan*, 101(2), 22–26. <https://doi.org/10.1177/0031721719879150>
- Cohen, J., & Goldhaber, D. (2016). Building a more complete understanding of teacher evaluation using classroom observations. *Educational Researcher*, 45(6), 378–387. <https://doi.org/10.3102/0013189X16659442>
- Creswell, J. W., & Clark, V. L. P. (2017). *Designing and conducting mixed methods research*. Sage publications.
- Crouse, K., Gitomer, D. H., & Joyce, J. (2016). An Analysis of the Meaning and Use of Student Learning Objectives. In K. Kappler Hewitt & A. Amrein-Beardsley (Eds.), *Student Growth Measures in Policy and Practice*. Palgrave Macmillan. [https://doi.org/10.1057/978-1-137-53901-4\\_11](https://doi.org/10.1057/978-1-137-53901-4_11)
- Danielson, C. (2011). Evaluations that help teachers learn. *Educational Leadership*, 68(4), 35–39. <https://doi.org/10/vol68/num04/abstract.aspx>
- Darling-Hammond, L. (2000). Teacher quality and student achievement. *Education Policy Analysis Archives*, 8, 1–1. <https://doi.org/10.14507/epaa.v8n1.2000>
- Darling-Hammond, L. (2004). Standards, accountability, and school reform. *Teachers College Record*, 106(6), 1047–1085. <https://doi.org/10.1177/016146810410600602>
- Darling-Hammond, L. (2014). One piece of the whole: Teacher evaluation as part of a comprehensive system for teaching and learning. *American Educator*, 38(1), 4.
- Darling-Hammond, L. (2015). Can Value Added Add Value to Teacher Evaluation? *Educational Researcher*, 44(2), 132–137. <https://doi.org/10.3102/0013189X15575346>
- Darling-Hammond, L. (2020). Accountability in teacher education. *Action in Teacher Education*, 42(1), 60–71. <https://doi.org/10.1080/01626620.2019.1704464>
- Darling-Hammond, L., & McLaughlin, M. W. (2011). Policies that support professional development in an era of reform. *Phi Delta Kappan*, 92(6), 81–92. <https://doi.org/10.1177/003172171109200622>
- Darling-Hammond, L., & Rustique-Forrester, E. (2005). The consequences of student testing for teaching and teacher quality. *Teachers College Record: The Voice of Scholarship in Education*, 107(14), 289–319. <https://doi.org/10.1177/016146810510701411>
- Darling-Hammond, L., & Wise, A. E. (1985). Beyond standardization: State standards and school improvement. *The Elementary School Journal*, 85(3), 315–336. <https://doi.org/10.1086/461408>
- Dee, T. S., & Keys, B. J. (2004). Does merit pay reward good teachers? Evidence from a randomized experiment. *Journal of Policy Analysis and*

- Management*, 23(3), 471–488. <https://doi.org/10.1002/pam.20022>
- Denzin, N. K., & Lincoln, Y. S. (2003). *Strategies of qualitative inquiry* (2nd ed.). SAGE Publications.
- Destler, K. N. (2016). Creating a performance culture: Incentives, climate, and organizational change. *The American Review of Public Administration*, 46(2), 201–225. <https://doi.org/10.1177/0275074014545381>
- Doğan, Ö. & Ayduğ, D. (2023). Accountability and Organizational Effectiveness in Education. In A. Abdallah & A. Alkaabi (Eds.), *Restructuring Leadership for School Improvement and Reform* (pp. 339–357). IGI Global. <https://doi.org/10.4018/978-1-6684-7818-9.ch017>
- Donaldson, M. L. (2016). Teacher Evaluation Reform: Focus, Feedback, and Fear. *Educational Leadership*, 73(8), 72–76.
- Donaldson, M. L., & Woulfin, S. (2018). From tinkering to going “Rogue”: How principals use agency when enacting new teacher evaluation systems. *Educational Evaluation and Policy Analysis*, 40(4), 531–556. <https://doi.org/10.3102/0162373718784205>
- Downey, M. (2016). Survey of Georgia teachers reveals ‘a workforce that feels devalued and constantly under pressure. *Atlanta Journal Constitution*. <http://getschoolblog.myajc.com/2016/01/06/survey-of-georgia-teachers-reveals-a-workforce-that-feels-devalued-and-constantly-under-pressure>
- Elmore, R. F., & Burney, D. (1997). *Investing in teacher learning: Staff development and instructional improvement in Community School District*. National Commission on Teaching & America’s Future, Teachers College. <https://eric.ed.gov/?id=ED416203>
- Emirates School Establishment. (2022). *About Us | Emirates Schools Establishment*. <https://www.es.gov.ae/about>
- FAHR. (2023). *About Bayanati | Bayanati—HR Management System for Federal Government*. <https://www.fahr.gov.ae/bayanati/portal/en/introduction.aspx>
- Farah, S., & Ridge, N. (2009). *Challenges to curriculum development in the UAE*. Dubai School of Government.
- Fernandez, C. (2002). Learning from Japanese approaches to professional development: The case of lesson study. *Journal of Teacher Education*, 53(5), 393–405. <https://doi.org/10.1177/002248702237394>
- Fernandez, C., & Yoshida, M. (2004). *Lesson study: A Japanese approach to improving mathematics teaching and learning*. Taylor & Francis Group. <http://ebookcentral.proquest.com/lib/uaeu-ebooks/detail.action?docID=234239>
- Firestone, W. A., & Donaldson, M. L. (2019). Teacher evaluation as data use: What recent research suggests. *Educational Assessment, Evaluation and Accountability*, 31(3), 289–314. <https://doi.org/10.1007/s11092-019-09300-z>
- Gaad, E., Arif, M., & Scott, F. (2006). Systems analysis of the UAE education system. *International Journal of Educational Management*, 20(4), 291–303. <https://doi.org/10.1108/09513540610665405>
- Gagnon, D. (2016). ESSA and rural teachers: New roads ahead? *Phi Delta Kappan*, 97(8), 47–49. <https://doi.org/10.1177/0031721716647019>
- Gallagher, K. (2019). *Education in the United Arab Emirates*. Springer.
- Gill, B., Bruch, J., & Booker, K. (2013). Using alternative student growth measures for Evaluating teacher performance: What the literature says. REL 2013-002. In *Regional Educational Laboratory Mid-Atlantic*. Regional Educational Laboratory Mid-Atlantic. <https://eric.ed.gov/?id=ED544205>
- Glesne, C. (1999). *Becoming qualitative researchers: An introduction*. Longman.
- Goe, L., Alkaabi, A. K., & Tannenbaum, R. J. (2020). Listening to and supporting teachers in the United Arab Emirates: Promoting Educational success for the nation. *ETS Research Report Series*, 2020(1), 1–18. <https://doi.org/10.1002/ets2.12289>
- González, J. M., & Darling-Hammond, L. (1997). *New concepts for new challenges: Professional development for teachers of immigrant youth*. Center for Applied Linguistics; Delta Systems Co.
- Halverson, R. R., & Clifford, M. A. (2006). Evaluation in the wild: A distributed cognition perspective on teacher assessment. *Educational Administration Quarterly*, 42(4), 578–619. <https://doi.org/10.1177/0013161X05285986>
- Halverson, R., Kelley, C., & Kimball, S. (2004). Implementing teacher evaluation systems: How principals make sense of complex artifacts to shape local instructional practice. *Educational Administration, Policy, and Reform: Research and Measurement*, 153–188.
- Hammad, W., & Hallinger, P. (2017). A systematic review of conceptual models and methods used in research on educational leadership and management in Arab societies. *School Leadership & Management*, 37(5), 434–456. <https://doi.org/10.1080/13632434.2017.1366441>
- Hanushek, E. A., & Rivkin, S. G. (2010). Generalizations about using value-added measures of teacher quality. *American Economic Review*, 100(2), 267–271. <https://doi.org/10.1257/aer.100.2.267>
- Hardono, H., Haryono, H., & Yusuf, A. (2017). Principal Leadership, academic supervision, and work motivation in improving teacher performance. *Educational Management*, 6(1), 11–19.
- Harris, D. N., Ingle, W. K., & Rutledge, S. A. (2014). How teacher evaluation methods matter for accountability: A comparative analysis of teacher effectiveness ratings by principals and teacher value-added measures. *American Educational Research Journal*, 51(1), 73–112. <https://doi.org/10.3102/0002831213517130>
- Ibrahim, A. S. (2012). Induction and mentoring of novice teachers: A scheme for the United Arab Emirates. *Teacher Development*, 16(2), 235–253. <https://doi.org/10.1080/13664530.2012.688676>
- Ibrahim, A., Alhosani, N., & Vaughan, T. (2020). Impact of language and curriculum on student international exam performances in the United Arab Emirates. *Cogent Education*, 7(1), 1808284. <https://doi.org/10.1080/2331186X.2020.1808284>
- Isoré, M. (2009). *Teacher Evaluation: Current Practices in OECD Countries and a Literature Review*. <http://repositorio.minedu.gob.pe/handle/20.500.12799/2541>
- Joseph, L. M., Kastein, L. A., Konrad, M., Chan, P. E., Peters, M. T., & Ressa, V. A. (2014). Collecting and documenting evidence: Methods for helping teachers improve instruction and promote academic success. *Intervention in School and Clinic*, 50(2), 86–95. <https://doi.org/10.1177/1053451214536043>
- Kamel, S. (2014). Education in the Middle East: Challenges and opportunities. In N. Azoury (Eds.), *Business and education in the Middle East* (pp. 99–130). Palgrave Macmillan UK. [https://doi.org/10.1057/9781137396969\\_9](https://doi.org/10.1057/9781137396969_9)
- Kawasaki, J., Quartz, K. H., & Martinez, J. F. (2020). Using multiple measures of teaching quality to strengthen

- teacher preparation. *Education Policy Analysis Archives*, 28, 128–128. <https://doi.org/10.14507/epaa.28.5011>
- Kim, S. J., & Mikiewicz, P. (2021). Merit pay, case-by-case: Variables affecting student achievement, teacher retention, and the problem of standardized tests. *Cogent Education*, 8(1), 1920560. <https://doi.org/10.1080/2331186X.2021.1920560>
- Kim, J., & Youngs, P. (2016). Promoting instructional improvement or resistance? A comparative study of teachers' perceptions of teacher evaluation policy in Korea and the USA. *Compare: A Journal of Comparative and International Education*, 46(5), 723–744. <https://doi.org/10.1080/03057925.2015.1057478>
- Kraft, M. A., & Gilmour, A. F. (2016). Can principals promote teacher development as evaluators? A case study of principals' views and experiences. *Educational Administration Quarterly*, 52(5), 711–753. <https://doi.org/10.1177/0013161X16653445>
- Lenhoff, S. W., Pogodzinski, B., Mayrowetz, D., Superfine, B. M., & Umpstead, R. R. (2017). District stressors and teacher evaluation ratings. *Journal of Educational Administration*, 56(2). <https://doi.org/10.1108/JEA-06-2017-0065>
- Levinson, M. (2011). Democracy, accountability, and education. *Theory & Research in Education*, 9(2), 125–144. <https://doi.org/10.1177/1477878511409622>
- Lincoln, Y. S., & Guba, E. G. (1985). *Naturalistic inquiry*. sage.
- Lingenfelter, P. E. (2003). Educational accountability: Setting standards, improving performance. *Change: The Magazine of Higher Learning*, 35(2), 18–23. <https://doi.org/10.1080/00091380309604089>
- Lipsky, M. (2010). *Street-level bureaucracy: Dilemmas of the individual in public service*. Russell Sage Foundation.
- Maguire, M., & Delahunt, B. (2017). Doing a thematic analysis: A practical, step-by-step guide for learning and teaching scholars. *All Ireland Journal of Higher Education*, 9(3), 3351–33514.
- Marquez, J., Lambert, L., Ridge, N. Y., & Walker, S. (2022). The PISA performance gap between national and expatriate students in the United Arab Emirates. *Journal of Research in International Education*, 24(1), 22–45. <https://doi.org/10.1177/14752409221090440>
- Martell, C. C. (2016). Teaching emerging teacher-researchers: Examining a district-based professional development course. *Teaching Education*, 27(1), 88–102. <https://doi.org/10.1080/10476210.2015.1042855>
- Matsumoto, A. (2019). Literature Review on Education reform in the UAE. *International Journal of Educational Reform*, 28(1), 4–23. <https://doi.org/10.1177/1056787918824188>
- McKnight, K., Yarbro, J., Graybeal, L., & Graybeal, J. (2016). *United Arab Emirates: What makes an effective teacher?*. Pearson.
- Mercer, J. (2007). Challenging appraisal orthodoxies: Teacher evaluation and professional development in the United Arab Emirates. *Journal of Personnel Evaluation in Education*, 18(4), 273. <https://doi.org/10.1007/s11092-007-9024-9>
- Merriam, S. B. (1998). *Qualitative research and case study applications in Education revised and expanded from case study research in Education* (2nd edition ed.). Josey-Bass.
- Ministry of Education. (2022). *Open Data: Teacher Distribution by Ring and Educational District Public Schools*. <https://www.moe.gov.ae/En/OpenData/pages/home.aspx>
- Mireles-Rios, R., & Becchio, J. A. (2018). The evaluation process, Administrator feedback, and teacher self-efficacy. *Journal of School Leadership*, 28(4), 462–488. <https://doi.org/10.1177/105268461802800402>
- MoCa, U. A. E. (2019). *Government Performance*. <https://www.moca.gov.ae/en/area-of-focus/government-performance>
- Mohebi, L., & ElSayary, A. (2022). Evaluating preservice teachers' performance in a blended field experience course during the quarantine of COVID-19. *Journal of Educators Online*. 19(3). <https://zuscholars.zu.ac.ae/works/5434/>
- Moran, R. M. (2017). The impact of a high stakes teacher evaluation system: Educator perspectives on accountability. *Educational Studies*, 53(2), 178–193. <https://doi.org/10.1080/00131946.2017.1283319>
- Morgan, C., & Ibrahim, A. (2020). Configuring the low performing user: PISA, TIMSS and the United Arab Emirates. *Journal of Education Policy*, 35(6), 812–835. <https://doi.org/10.1080/02680939.2019.1635273>
- NCTQ. (2019). *NCTQ state of the States 2019: Teacher and principal evaluation Policy*. National Council on Teacher Quality. <https://www.nctq.org/pages/State-of-the-States-2019-Teacher-and-Principal-Evaluation-Policy>
- Neumerski, C. M., Grissom, J. A., Goldring, E., Drake, T. A., Rubin, M., Cannata, M., & Schuermann, P. (2018). Restructuring instructional leadership: How multiple-measure teacher evaluation systems are redefining the role of the school principal. *The Elementary School Journal*, 119(2), 270–297. <https://doi.org/10.1086/700597>
- Nyamubi, G. J. (2017). Determinants of Secondary School Teachers' Job Satisfaction in Tanzania. *Education Research International*, 2017, 1–7. <https://doi.org/10.1155/2017/7282614>
- OECD. (2021). PISA 2018 assessment and analytical framework. <https://doi.org/10.1787/b25efab8-en>
- Oplatka, I., & Arar, K. (2017). The research on educational leadership and management in the Arab world since the 1990s: A systematic review. *Review of Education*, 5(3), 267–307. <https://doi.org/10.1002/rev3.3095>
- Park, S. (2006). Teachers' perceptions on the teacher evaluation in the organizational context. *Journal of Elementary Education*, 19(1), 261–291.
- Pataki, G. (n.d.). Working paper a Review of advanced teacher professional development models in public schools. *UNESCO and Regional Center for Educational Planning*. <https://rcepunesco.ae/en/KnowledgeCorner/ReportsandStudies/ReportsandStudies/A%20Review%20of%20Advanced%20Teacher%20Professional%20Development%20Models%20in%20Public%20Schools.pdf>
- Patton, M. Q. (2002). *Qualitative research & evaluation methods*. SAGE.
- Patton, M. Q. (2014). *Qualitative research & evaluation methods: Integrating theory and practice*. Sage publications.
- Popham, W. J. (2013). On serving two Masters: Formative and summative teacher evaluation. *Principal Leadership*, 13(7), 18–22.
- Robertson-Kraft, C., & Zhang, R. S. (2018). Keeping great teachers: A case study on the Impact and implementation of a Pilot teacher evaluation system. *Educational Policy*, 32(3), 363–394. <https://doi.org/10.1177/0895904816637685>
- Rubin, H. J., & Rubin, I. S. (2016). *Qualitative interviewing: The art of hearing data*. sage. [https://books.google.com/books?hl=en&lr=&id=bgekGK\\_xpYsC&oi=fnd](https://books.google.com/books?hl=en&lr=&id=bgekGK_xpYsC&oi=fnd)

- fnd&pg=PP1&dq=Rubin,+H.+J.,+%26+Rubin,+I.+S.+ (2016).+Qualitative+interviewing:+The+art+of+hearing+data.+Sage.+&ots=tJcCkLr5Sf&sig=jIbTe3GFJLutWJXp6PktQD9dA
- Sartain, L., Stoelinga, S. R., & Brown, E. R. (2011). *Rethinking teacher evaluation in Chicago: Lessons learned from classroom observations, Principal-Teacher Conferences, and District Implementation. Research Report.* ERIC. <https://eric.ed.gov/?id=ED527619>
- Schultz, B. D. (2017). *Teaching in the cracks: Openings and opportunities for student-centered, action-Focused Curriculum.* Teachers College Press.
- Seale, C. (1999). *The quality of qualitative research. Introducing qualitative methods.* Sage.
- Seidman, I. (2012). *Interviewing as qualitative research: A guide for researchers in education and the social sciences.* Teachers College Press.
- Seitz, H. (2008). The power of documentation in the early childhood classroom. *YC Young Children*, 63(2), 88–93. <https://doi.org/10.1007/s10643-008-0242-7>
- Shikalepo, E. E. (2020). The role of motivational theories in shaping teacher motivation and performance: A Review of related literature. *International Journal of Research and Innovation in Social Science*, 4. [https://www.researchgate.net/profile/Elock-Shikalepo/publication/340923164\\_The\\_Role\\_of\\_Motivational\\_Theories\\_in\\_Shaping\\_Teacher\\_Motivation\\_and\\_Performance\\_A\\_Review\\_of\\_Related\\_Literature/links/5ea447a0299bf112560ca429/The-Role-of-Motivational-Theories-in-Shaping-Teacher-Motivation-and-Performance-A-Review-of-Related-Literature.pdf](https://www.researchgate.net/profile/Elock-Shikalepo/publication/340923164_The_Role_of_Motivational_Theories_in_Shaping_Teacher_Motivation_and_Performance_A_Review_of_Related_Literature/links/5ea447a0299bf112560ca429/The-Role-of-Motivational-Theories-in-Shaping-Teacher-Motivation-and-Performance-A-Review-of-Related-Literature.pdf)
- Springer, M. G. (2009). *Rethinking teacher compensation policies: Why now, why again?* Brookings Institution Press.
- Springer, M. G., Swain, W. A., & Rodriguez, L. A. (2016). Effective teacher retention bonuses: Evidence from Tennessee. *Educational Evaluation and Policy Analysis*, 38(2), 199–221. <https://doi.org/10.3102/0162373715609687>
- Steiner, L. (2010). Using competency-based evaluation to drive teacher excellence: Lessons from Singapore. Building an opportunity culture for America's teachers. In *Public impact.* Public Impact. <https://eric.ed.gov/?id=ED539994>
- Stephenson, L., Dada, R., & Harold, B. (2012). Challenging the traditional idea of leadership in UAE schools. *On the Horizon*, 20(1), 54–63. <https://doi.org/10.1108/10748121211202071>
- Stronge, J. H. (1995). Balancing individual and institutional goals in Educational personnel evaluation: A conceptual framework. *Studies in Educational Evaluation*, 21(2), 131–151. [https://doi.org/10.1016/0191-491X\(95\)00010-R](https://doi.org/10.1016/0191-491X(95)00010-R)
- Stronge, J. H. (2005). *Evaluating teaching: A guide to current thinking and best practice.* Corwin Press.
- Thompson, C. S., & Samuels-Lee, L. (2020). Jamaican teachers' perspectives on the desirability of performance-based payment: Lessons for education policy makers and school administrators. 27(2), 63–83.
- Tobiason, G. (2019). Countering expert uncertainty: Rhetorical strategies from the case of value-added modeling in teacher evaluation. *Minerva*, 57(1), 109–126. <https://doi.org/10.1007/s11024-018-9359-z>
- Tucker, P. D., & Stronge, J. H. (2005). *Linking teacher evaluation and student learning.* ASCD.
- UAE Cabinet. (2023). <https://uaecabinet.ae/en/details/news/the-uae-cabinet-approves-aed514-billion-federal-budget-for-2018>
- UAE Vision 2021. (2018). *First-rate Education System.* Default. <https://www.vision2021.ae/en/national-agenda-2021/list/first-rate-circle>
- Warren, A. N., & Ward, N. A. (2019). “It didn't make me a better teacher”: Inservice teacher constructions of dilemmas in high-stakes teacher evaluation. *School Effectiveness and School Improvement*, 30(4), 531–548. <https://doi.org/10.1080/09243453.2019.1619185>
- Waxin, M.-F., & Bateman, R. (2016). *Human Resource Management in the United Arab Emirates.* <https://dspace.aus.edu/xmlui/handle/11073/16345>
- Weiss, C. H. (1998). *Evaluation: Methods for studying programs and policies.* Prentice Hall.
- Yin, R. K. (2014). *Case study research: Design and methods* (5th ed.). SAGE Publications.
- Younies, H., Barhem, B., & Younis, M. Z. (2008). Ranking of priorities in employees' reward and recognition schemes: From the perspective of UAE health care employees. *The International Journal of Health Planning and Management*, 23(4), 357–371. <https://doi.org/10.1002/hpm.912>
- Zepeda, S. J., Alkaabi, A. M., & Tavernier, M. D. (2020). Leadership and supervision. *Oxford Research Encyclopedia of Education.* <https://doi.org/10.1093/acrefore/9780190264093.013.617>
- Zepeda, S. J., & Ponticell, J. A. (1998). At cross-purposes: What do teachers need, want, and get from supervision? *Journal of Curriculum and Supervision*, 14(1), 68.