

A Study to evaluate the teacher evaluation matrix of Al-Shola schools in Sharjah – UAE from the perspectives of school members and in relation to students' results

دراسة حول تقييم مصفوفة تقييم أداء المعلمين المطبقة في سلسلة مدارس الشعلة في الشارقة - دولة الإمارات العربية المتحدة وفقاً لوجهات أعضاء المجتمع المدرسي ومقارنةً بنتائج تحصيل الطلبة

By: Maha Ibrahim Barakeh

Student ID number 2013101040

A Dissertation submitted to: ***The British University in Dubai***

In partial fulfillment of the requirements for the degree of

Masters degree in education

(Educational management leadership and Policy)

Faculty of Education

Dissertation Supervisor:

Professor Abdulai Abukari

May-2016

This dissertation consists of: 17,960 words

Abstract

Teacher evaluation represents a great challenge encountered by school leadership. The complexity of such evaluation is derived from the complex nature of the teaching and learning process, in addition to the diversity of definitions introduced to describe effective teaching. These challenges require a high level of awareness by school leaderships about the need to evaluate the multi-dimensional teaching performance through compatible evaluation systems that are not limited to the scene inside the classroom.

This research was conducted as a case study to evaluate a multi-lateral teacher evaluation matrix that was designed and applied by a chain of private schools in Sharjah UAE. The researcher followed a triangulation to combine three study phases. Both, qualitative and quantitative approaches were used during these phases. The perspectives of school leadership and teachers were incorporated through an interview and a questionnaire; in an approach to evaluate the matrix against good-practice criteria extracted through literature review. Quantitative analysis of the relationship between teachers' performance evaluations and students' results was used as an additional approach to evaluate the matrix and to support the analysis of school members' perspectives about the targeted system. This analysis was also conducted in an attempt to fill the gap found in previous studies; where quantitative analysis to study the relationship between teachers' evaluations and students' results was lacking.

The study findings revealed a shared support of the teacher evaluation matrix among school leadership and teachers. The diversity of the measurements used in the matrix increased the validity and reliability of evaluation; moreover, it facilitated evaluation-based professional development. The quantitative analysis proved a strong positive relationship between the evaluation matrix's results and students' results, indicating the efficiency of the matrix as a valid teacher evaluation system that is related to teacher's influence on students.

ملخص البحث

يشكل تقييم أداء المعلمين أحد التحديات الكبيرة التي تواجه القيادة المدرسية، وإن التحديات المرهونة بهذا النوع من التقييم ناجمة عن التوسع والتنوع المرتبطان بعملية التعليم والتعلم، كما أنه ناجم عن تعدد وتنوع التعريفات المطروحة في وصف التعليم الفعال. إن هذه التحديات تحتم على القيادات المدرسية امتلاك نسبة كافية من الوعي حول الحاجة إلى تقييم الأداء المتعدد الأبعاد للمعلم باستخدام أنظمة مماثلة لا تقتصر على المشهد الصفّي كمصدر وحيد للتقييم.

من خلال هذا البحث تم إجراء دراسة حالة لتقييم مصفوفة تقييم أداء المعلم الشاملة التي تم إعدادها وتطبيقها من قبل سلسلة مدارس خاصة في الشارقة – الإمارات، حيث قامت الباحثة باتباع طريقة التثليث في دمج التحليلين النوعي والكمي من خلال ثلاثة مراحل تضمنها البحث. وقد تخلل البحث استطلاعاً لوجهات نظر مجتمع الدراسة متمثلاً في القيادة المدرسية والمعلمين حول مصفوفة تقييم أداء المعلم من خلال مقابلة شخصية واستبانة، وذلك سعياً لتقييم المصفوفة وفقاً لمعايير فعالية تقييم المعلمين المستخلصة من خلال الإطار النظري للبحث. كما تم استخدام التحليل الكمي لدراسة العلاقة بين تقييمات أداء المعلمين الناتجة عن المصفوفة ونتائج الطلبة، ويمثل هذا التحليل مصدراً إضافياً لتقييم المصفوفة ودعم الاستنتاجات التي أفرزها تحليل آراء مجتمع الدراسة. إضافةً إلى ذلك، فقد تم إجراء هذا التحليل سعياً لردم الفجوة التي لمستها الباحثة من خلال الإطار النظري، والمتمثلة في نقص الدراسات الكمية السابقة التي تقيس طبيعة العلاقة بين تقييمات أداء المعلمين ونتائج الطلبة.

أظهرت نتائج الدراسة وجود تأييد مشترك بين القيادة المدرسية والمعلمين لاستخدام مصفوفة تقييم أداء المعلمين في مدارس الشعلة الخاصة، حيث أدى تنوع وشمولية الأدوات المستخدمة في المصفوفة إلى زيادة نسبة المصادقية والموثوقية في تقييم الأداء، كما ساهم في التنمية المهنية للمعلمين المرتبطة بنتائج تقييم الأداء. هذا وقد أثبت التحليل الكمي للدراسة وجود علاقة طردية قوية بين نتائج مصفوفة تقييم أداء المعلم ونتائج الطلبة مما يشير إلى فعالية المصفوفة كنظام تقييم يتصف بالمصادقية ويرتبط بأثر المعلم على الطلبة.

Dedication

*To my father, **Ibrahim Barakeh**, the pioneer educational leader, who struggled against the harshest obstacles to achieve his noble dreams and leave a remarkable mark in this world, and who I had always and will always consider to be the most influential role model in my life. Thank you for inspiring me in education and in life, and for always supporting me in achieving the highest objectives anyone could dream of.*

*To my mother, **Samira El-Mosleh**, who surrounded me with unconditional love and care. Thank you for who you are. Thank you for encouraging me and believing in my abilities and determination, and for always succeeding in easing my worries and nourishing me with peace and optimism.*

*To my husband, my soul mate and my amazing partner, **Tariq Abu-Jazar**, who was a key contributor to my career and educational achievements. Thank you for your love and patience; for believing in me, and for your constant advice and support. Thank you for always being there when I needed confidence and motivation.*

Last but not least, to all my family members and friends. Thank you for always being there for me, and for being a great source of encouragement and inspiration during this stressful, yet rich journey.

Acknowledgment

First and above all, I praise God for blessing me with this opportunity, and granting me the strength and determination to complete this research.

*A sincere gratitude and acknowledgment to my dissertation supervisor, **Prof. Abdulai Abukari**, for his continuous support and enlightening feedback. Your care and advice was highly motivational and deeply appreciated.*

*Additionally, I would like to express my deep gratitude to all my professors, **Prof. Clifton Chadwick, Prof. Sufian Forawi, Prof. David Solomon and Prof. John McKinney**. I was blessed to benefit from your profound experiences, and gain knowledge that will guide me through career and life.*

*Finally, I would like to thank our kind dean, **Prof. Eman Gad** for her care and support throughout this journey.*

Table of Contents

Title	Page number
Chapter 1: Introduction.....	1
1.1 Overview.....	1
1.2 Statement of the problem.....	1
1.3 Research Objectives.....	2
1.4 Research Questions.....	3
1.5 Research Hypothesis.....	3
1.6 Background of the research.....	3
1.7 The significance of the research.....	8
1.8 The structure of this paper.....	8
Chapter 2: Literature Review.....	10
2.1 Overview.....	10
2.2 Teacher evaluation.....	10
2.3 Teacher evaluation systems: Lessons learned through the evolutionary process....	10
2.3.1 (1940's-1950's).....	11
2.3.2 (1960's-1970's).....	12
2.3.3 (1980's).....	12
2.3.4 (1990's).....	13
2.3.5 (2000-present).....	15
2.4 Standards of effective teacher evaluation system.....	16
2.5 Teacher evaluation tools.....	20
2.5.1 Classroom observations.....	20
2.5.2 Teacher self-assessment.....	21
2.5.3 Teaching artifacts and planning documents.....	21
2.5.4 Students and parents surveys.....	21
2.5.5 Teachers' portfolios.....	22
2.5.6 Value-added models.....	22
2.5.7 Other teacher evaluation tools.....	23
2.6 Teacher evaluation for professional development.....	23
2.6.1 Danielson's three-tracks approach to evaluation-based professional enhancement.....	23
2.6.2 The influence of teacher evaluation on teachers' preparation programs.....	24
2.6.3 The need for leadership-level policies to link evaluation to professional development.....	25
2.7 Summary.....	26
Chapter 3: The present study.....	27
3.1 Overview.....	27
3.2 Research Design.....	27
3.2.1 Case studies.....	27
3.2.2 Research methodology.....	28
3.2.3 The three phases of the study.....	29

(A) Phase I: A Qualitative interview- Perspectives of policymakers.....	29
▪ Purpose.....	29
▪ Methodology.....	29
▪ Sampling.....	29
▪ Instrument.....	29
▪ Data Collection.....	31
(B) Phase II: A Quantitative questionnaire- Perspectives of evaluated teachers.....	31
▪ Purpose.....	31
▪ Methodology.....	32
▪ Sampling.....	32
▪ Instrument.....	33
▪ Validity and reliability.....	34
▪ Piloting the questionnaire.....	35
▪ Data Collection.....	35
(C) Phase III: Quantitative analysis and correlation coefficient- The relationship between teachers' performance evaluation results and students' results.....	35
▪ Purpose.....	35
▪ Methodology.....	35
▪ Sampling and Data Collection.....	36
3.3 Ethical consideration.....	37
Chapter 4: Results and interpretations.....	39
4.1 Overview.....	39
4.2 Study results and interpretations.....	39
4.2.1 Phase I: Qualitative interview- Perspectives of policymakers.....	39
▪ Results.....	39
▪ Discussion and interpretations.....	44
4.2.2 Phase II: Quantitative questionnaire- Perspectives of evaluated teachers.....	46
▪ Results and data analysis.....	46
▪ Discussion and interpretations.....	53
4.2.3 Phase III: Quantitative analysis and correlation coefficient- The relationship between teachers' performance evaluation results and students' results.....	59
▪ Results and data analysis.....	59
▪ Discussion and interpretations.....	64
Chapter 5: Summary, Conclusions and recommendations.....	66
5.1 Summary.....	66
5.2 Conclusions.....	66
5.2.1 School members' perspectives.....	67
5.2.2 Teachers' evaluations - students' results relationship.....	70
5.3 Limitations of the study.....	71
5.4 Recommendations.....	71
References.....	74
Appendices.....	80

List of Tables	Page number
Table 3-1 Summary of questionnaire participants' biographical data.....	33
Table 3-2 Teachers' questionnaire scheme.....	34
Table 4-1 Teachers' questionnaire results– Area (1): Evaluation of the targeted teacher evaluation system according to effectiveness criteria (Proportional percentile).....	48
Table 4-2 Teachers' questionnaire results– Area (2): Evaluation of school's use of evaluation data in professional development of teachers (Proportional Percentile).....	49
Table 4-3 Teachers' questionnaire results– Area (3): Importance of each teacher evaluation tool used within the matrix according to teachers (Part I).....	51
Table 4-4 Teachers' questionnaire results– Area (3): Importance of each teacher evaluation tool used within the matrix according to teachers (Part II).....	52
Table 4-5 Teachers' questionnaire results– Area (4): Reliability of teacher evaluation tools from the perspectives of teachers.....	53
Table 4-6 A comparison sheet between teacher evaluations obtained through the targeted evaluation matrix and the end-results of students taught by the evaluated teachers at the academic year: 2014/2015.....	60
Table 4-7 Pearson correlation coefficient (r) between teacher evaluations obtained through the targeted evaluation matrix and the end-results of students taught by the evaluated teachers at the academic year: 2014 / 2015.....	62
List of Figures	Page number
Figure 4-1 Teachers' questionnaire results– Area (1): Evaluation of the targeted teacher evaluation system according to effectiveness criteria (Graphical analyses).....	49
Figure 4-2 Teachers' questionnaire results– Area (2): Evaluation of school's use of evaluation data in professional development of teachers (Graphical analyses).....	50
Figure 4-3 Graphical comparison between teacher evaluations obtained through the targeted matrix and the end-results of students taught by the evaluated teachers at the academic year: 2014 / 2015.....	61
Figure 4-4 Pearson correlation coefficient (r) between teacher evaluations obtained through the targeted evaluation matrix and the end-results of students taught by the evaluated teachers at the academic year: 2014 / 2015.....	64

Chapter 1: Introduction

1.1 Overview

“The most effective supervision and evaluation systems empower teachers to accurately assess their own practice and self-diagnose areas for growth.”

- Paul Mielke and Tony Frontier (2012)

The power of teacher evaluation is an evident truth to most members of the educational society. This power can only be earned if teacher evaluation policies are based on two substantial approaches: (1) Efficient and valid evaluation procedures, (2) Investment in the path of professional development. When applied efficiently, teacher evaluation does not only reflect teachers' performance, but the performance of the school as an entity. This assumption is derived from the fact that the core business for schools is education, and in order to provide high-quality education, schools are required to have high-quality teachers.

For decades, teacher evaluation had been extensively analyzed in literature through different approaches, including its purposes (Scriven, 1981; Papay, 2012), methodology of evaluation (Davey, 1991; Brandt, 1996; Al Bustami, 2014), the relationship between teachers' effectiveness and overall school effectiveness (Ellett & Teddlie, 2003; Iwanicki, 1989), evaluation-based professional development (Danielson & McGreal, 2000), and standards-based teacher evaluation (Stufflebeam, 1988; NCLB, 2002; Milanowski, Kimball & White, 2004). Literature also introduced numerous handbooks to guide the process of teacher evaluation (Shinkfield & Stufflebeam, 1995; Darling-Hammond, 2000; Peterson, 2004; Stronge, 2005; Hinchey, 2010; Goe, Holdheide & Miller, 2014).

1.2 Statement of the problem

In 2011, a chain of schools in Sharjah-UAE has designed and started applying a new multi-measurement teacher evaluation system. The system was named: **The multi-lateral matrix**, because it measures teachers' performance in different areas using different tools, thus; forming a multi-lateral matrix that reflects each

teacher's performance individually. According to the Oxford University online dictionary (2016), the term (multilateral) refers to an item that has many sides or has contributions from multiple parties, and the term (Matrix) can be defined as a structure designed by an organization to arrange responsibilities or communication between several individuals. Both definitions apply to the targeted system which explains its name.

An empirical study was applied to evaluate the Teachers' evaluation matrix which the researcher had personally designed along with other school policymakers. Therefore; the researcher has professional and personal interest in evaluating the policy in a form of self-evaluation especially that this policy is being applied in the schools for almost five years now and has never been particularly evaluated by school members or by external researchers.

The evaluation of the matrix would provide opportunities for improvement of the current policy, and is expected to add positively to the body of teacher evaluation research.

1.3 Research objectives

The main purpose of this research is to introduce and evaluate a new teacher evaluation system applied by a chain of schools in the local region. The newly developed system represents an area for analysis and critique; where this paper aims at:

1. Introducing a new teacher evaluation system.
2. Assessing the system's efficiency against good-practice criteria.
3. Evaluating the system's design, implementation and role in professional development from the perspectives of school members.
4. Validating the data obtained through the evaluation system by comparing it with a second source of information; students' results.

1.4 Research questions

The objectives of this research raise four main questions:

- 1 In terms of design and implementation, does the teacher evaluation matrix of Al-Shola schools possess characteristics defined by literature to describe effective teacher evaluation?
- 2 What are the perspectives of school leadership and teachers about the targeted teacher evaluation matrix and evaluation tools included within?
- 3 How does the school use the data obtained from the matrix?
- 4 Is there a relationship between teachers' evaluations resulting from the targeted system and achievement results of their students?

1.5 Research hypothesis

For the first three research questions no hypothesis was suggested, because these questions were set to explore unexpected perspectives and opinions of school members; however, for the fourth question, the following hypothesis was suggested:

- There is a proportional positive relationship between teacher evaluations and achievement results of students taught by those teachers.

According to Fraenkel, Wallen & Hyun (1993), a research hypothesis is when the researcher predicts a certain outcome or relationship between variables to be concluded through the study.

1.6 Background of the research

- **The teachers evaluation matrix of Al-Shola Private Schools:**

Al-Shola private schools are a chain of schools that were established on 1983 in Sharjah – UAE. The beginning was modest with merely 160 students in the mother school on the founding year, where it grew in quantity and quality until the opening of its new branch holding the same name in 2004. Adapting with the rising demand and conveying to its ambition; Al-Shola educational organization established its most recent branch on 2014 with a new name: Manarat AlSharjah School.

Today, Al-Shola private schools serve around six thousands students from different Arabic nationalities.

The quality of education has always been a major concern for Al-Shola schools. This was illustrated through several educational approaches including: full-time academic supervisors, periodic on-campus training for teachers, periodical teachers' evaluations, internal school committees that review and audit students' assignments periodically, in addition to multiple annual surveys to measure the satisfaction of students and parents. However, the key step in the schools' quality approach was the establishment of an internal quality control department in 2008. This proactive measurement came in coherence with the overall orientation of UAE's educational sector towards focusing on the quality of education. Around that time, governmental educational organizations started to release several policies to evaluate schools all over the country (e.g. MOE's school accreditation, KHDA and ADEC schools' inspection...).

Taking in consideration the key role of teachers in introducing high quality education, the quality control department in Al-Shola organization concentrated its efforts on evaluating and improving teachers' performance, where the final outcome was the design of a new system to evaluate each teacher's performance in a comprehensive manner. The system was initially suggested and designed by the organization's leader and the head of quality control department , and was later edited according to the feedback of schools' educational supervisors, to be approved and put into action in 2010.

The idea behind this new teacher evaluation system started with the question: What are the roles required from a teacher within the school community? The answer yielded a differentiation of three main roles:

- 1. An academic role.**
- 2. A professional role.**
- 3. A social role.**

Differentiation of these roles produced ten specific areas that represent teacher's performance:

- | | |
|---|---------------------------------------|
| 1. Influence on students. | 6. Personal traits. |
| 2. Subject knowledge and adequacy. | 7. Social participation. |
| 3. Planning efficiency. | 8. Professional commitment. |
| 4. Classroom management. | 9. Professional relationships. |
| 5. Employing resources. | 10. Professional development. |

The school then designed a set of tools to measure these areas. Each area could be measured using several tools, and vice versa; each tool covers several areas with different weights. To ensure objectivity; data collection tasks were distributed among several school members according to specialty.

The Evaluation Matrix includes **twelve evaluation tools**:

1. Students' Results (150 points)

To determine the teacher's grade for his/her influence on students' achievement results, the quality control department in the school calculates the average end-year scores of the sections taught by the teacher in his/her subject.

2. Lesson Planning Evaluation (75 points)

Schools' academic superintendants or subject coordinators use a designated numeric-scale form to evaluate each teacher's skills in lesson planning.

3. Classroom Visits (100 points)

After the implementation of the new teacher evaluation system, the schools' academic supervisors redesigned the previous classroom visit form according to the criteria included within the new evaluation system. Through the new form, superintendants evaluate teachers' classroom performance using a numeric scale in addition to descriptive follow-up notes and recommendations.

4. Students' assignments review reports (75 points)

Al-Shola schools apply an auditing system through which students' assignments are reviewed periodically by academic supervisors or senior teachers. For each teacher, the reviewer uses a designated form to evaluate and grade teacher's individual performance in supervising his/her students' assignments. The average of grades earned by each teacher in these reports is used in his/her evaluation matrix.

5. Professional development Report (50 points)

The professional development supervisor of the school is responsible for planning and conducting teachers' training. Additionally, this supervisor uses an annual form to evaluate each teacher's efforts in training and research; considering that all teachers are tasked by the school administration with conducting an annual educational research or project.

6. Academic superintendant's Report (75 points)

Academic superintendants or subjects' coordinators are annually required to evaluate each teacher's overall academic performance. The designated form follows a multi-criteria numeric grading scale.

7. Section Supervisor's Report (100 points)

According to Al-Shola schools' functional structure, teachers of each stage (KG, Elementary, Secondary...etc) follow a section supervisor administratively. At the end of the year, each supervisor evaluates his/her teachers' professional commitment and relationships through a multi-criteria numerical form.

8. Commitment to laws and regulations Report (50 points)

The school has a list of regulations for teachers, where each teacher receives a copy and training courses to explain its contents. Whenever a teacher violates regulations, it is usually documented at the HR department. At the end of the year, this data is collected, and certain grades are omitted for each teacher who violates regulations according to the significance of his/her violation.

9. Attendance Report (50 points)

The HR department provides monthly reports that document each teacher's attendance. Accordingly, each teacher earns a grade for attendance based on a numerical scale that omits points from the teacher's total grade for each unexcused absence or lateness.

10. Resources & Technology Report (50 points)

The Information Technology department collects data about teachers' incorporation of technology into teaching through records designed for this purpose. Each time the teacher uses technology inside classroom (e.g. data show, laptop, smart technology...) he/she documents it in the records, where the section supervisor approves the records' contents. For each ICT use, the teacher earns 1.5 points with a maximum of 50 points.

11. Students' Surveys (150 points)

Teacher performance evaluation surveys are distributed among a sixty students' sample for each teacher twice per academic year. Surveys results' analysis grants each teacher a grade that is influenced by the ratio of positive/negative students' responses. Surveys are only executed among students at the age of twelve and above.

12. Participation in extracurricular activities Report (75 points)

Teachers are asked to document their participation in extra-curricular activities (e.g. school trips, competitions, sports and cultural activities...) through an annual form. A numeric grading scale grants each teacher certain points for his/her efficiency in extra-curricular activities.

Within the matrix, each tool is assigned a certain weight according to what the school leadership considers to be more important regarding teachers' performance. Eventually, the matrix yields a grade out of 1000 for each teacher. This process is operated for each teacher annually, where the resultant are individual reports for teachers, and a results' summary prepared for the purpose of

school leadership development plans. The setup that combines all these elements takes a form of a multi-lateral matrix (**Appendix.1**).

1.7 The significance of this research

The facts that the targeted policy was directly designed by the researcher and other professionals in the same organization and has never been published before, and that it has only been applied in the targeted schools, assure the uniqueness of the research topic. Moreover, while most previous literature in the area of teacher evaluation suggested the use of multi-tools evaluation systems, no study was found to include as many and as comprehensive elements and setup as the targeted policy. Considering all of that, the study is expected to produce the following outputs for the targeted organization and the general educational society:

- Introduce a new effective teacher evaluation policy to the educational society.
- Evaluate the existing teacher evaluation system in Al-Shola Schools and diagnose any design, implementation or employment setbacks. Thus; provide useful feedback for improvement to the targeted organization.

1.8 The structure of this paper

In this chapter, the researcher had the chance to explain the research objectives, questions and hypothesis, and explain how this research is expected to add to the body of educational research. Moreover, the background of the study was provided by explaining the targeted policy, its concepts, and components.

Chapter 2 includes review of literature related to the field of teacher evaluation. It starts with terminology definition followed by a historical analysis to compare teacher evaluation approaches over the years, thus; learn important lessons that would guide this research. Following is a summary of criteria that defines effective teacher evaluation systems. Literature review also includes an analysis of commonly discussed teacher evaluation tools. The researcher chose to analyze this area individually because evaluation tools represent the main components of

the targeted evaluation matrix. The final section discusses previous literature suggestions on using teacher evaluation data for professional development.

Explanation of the present study is provided in Chapter 3. This includes methodology, research tools, sampling, and data collection. The study findings are discussed in Chapter 4 with interpretations of the results.

Finally, Chapter 5 offers an overall summary, conclusions, and recommendations based on connecting all the elements of the research including literature, findings and interpretations discussed in previous sections. In addition to the limitations associated with this study.

Chapter 2: Literature review

2.1 Overview

The quality of education depends on numerous factors; however, the key contributor to this quality remains to be the teacher. If school improvement reforms are planned, the need for qualified teachers grows, and success of reform recommendations requires teachers who are capable of successful implementation and practice (Clark, 1993).

Despite the shared agreement on the need for an effective, accurate, and performance-differentiating teacher evaluation; variation in evaluation approaches between past and recent trends is noticed.

In this chapter, early examples and recent transformations in teacher evaluation trends will be summarized. Furthermore, the best practices described by reviewed literature will be concluded, in addition to an analysis of how literature discussed teacher evaluation tools. As for the final section, evaluation-based professional development is discussed.

2.2 Teacher evaluation

Teacher evaluation was defined by Darling-Hammond *et al.* (1983) as a collection of data and the use of that data in forming judgments about teaching. Moreover, Collins English dictionary (2016) defined teacher evaluation as the process of testing teachers against maintaining teaching standards.

2.3 Teacher evaluation systems: Lessons learned through the evolutionary process

Differences in teacher evaluation approaches are caused by the complex nature of what is being measured. While in other fields such as industry, products can be precisely evaluated using simple standardized checklists, in education the product is more complicated, and the numerous definitions introduced to identify what constitutes good teaching added to the complexity of teacher evaluation.

Such challenges, in addition to the infallibility character related to teachers by society in the old days led to the delay of formal forms of evaluation (Shinkfield & Stufflebeam, 1995). Formalization approaches were intangible until the late 1920's; even then, the development wheel in this field remained relatively slow until the 1970's (Shinkfield & Stufflebeam, 1995; Danielson & McGreal, 2000).

At the beginning of the 19th century, teachers were informally evaluated by administrators, with no significant recorded data until the documented efforts of Bobbitt (1912), who is considered as one of the most influential researchers in this field at that time. His successful attempts to link theory with practice in education represented support for schools to enhance the validity of evaluation procedures, and his suggestions to base judgments about teachers on students' learning outcomes are still being followed to the present day.

In the 1920's and 1930's supervision and evaluation were described as being democratic by several educators of that time including Dewey (1916; 1929, in Scherrer, 2009) who focused on calling for democratic evaluation by involving teachers in the evaluation process rather than subjecting them to it. However, most documented cases revealed poor practice, and evaluation depended on personal affiliations rather than performance (Ellett & Teddlie, 2003).

2.3.1 1940's-1950's: Spontaneous, superficial, and inaccurate evaluation

Examining teacher evaluation in this era revealed great emphasize on teachers' traits as calibers to indicate performance and influence. Danielson & McGreal (2000) described these traits as weak variables, as there was no proved connection between teachers' traits and students' progress. Teacher evaluation systems during this era were centrally based on embracing/disgracing teachers traits (e.g., Charisma, acceptance) which lacked validity and professionalism.

2.3.2 1960's-1970's: Evaluation based on teacher effectiveness research, the beginning of reform enactment

Danielson & McGreal (2000) detected a significant burst in teaching research in the 1960's which had great impact on teacher evaluation. Supervision and evaluation skills of administrators became more advanced, and educators began extensive attempts to define effective teaching in order to design evaluation tools. This was later named, *research on teacher effects*.

Shinkfield & Stufflebeam (1995) discussed growth in framing evaluation in the 1970s, catalyzing the launch of many advanced teacher evaluation models later in 1990s. Despite lacking certainty in many evaluation areas, a noticeable shift towards the enactment of reform procedures at nation-level and districts-level was recorded in the 1970s, especially in the U.S.

2.3.3 1980's: Classroom observation dominance and dissection, The Hunter Model, setting benchmarks

In the 1980's, teacher evaluation was greatly relying on classroom visits as the dominant evaluation tool; however, the orientation toward dissecting classroom observations criteria was ascending. Classroom observation criteria suggested by Dunkleberger (1982) included proper planning, achievable goals, resources employment, feedback for students, students' motivation, classroom activities, and classroom management. However, Dunkleberger (1982) had several setbacks; no assessment methodology was provided, in addition to the narrow scope that limited evaluating teachers to observations implicated only during classroom visits (Clark, 1993).

Moreover, the 1980's witnessed a marked dominance of the Hunter (1983) model that linked teaching to the behavioristic learning theory by emphasizing the concepts of motivation, transmission, and learners' *tabula-rasa* state of mind. The Hunter model, which had its origins in the early 70's, was analyzed by Shinkfield & Stufflebeam in 1995. They acknowledged the model for being

formative and motivational in nature, seeking excellence in teaching, and focusing on teacher's performance rather than personality, and on guidance rather than admonishment. Despite Shinkfield's and Stufflebeam's (1995) acknowledgments, they criticized the high time and money costs related to this model. Furthermore, many practical cases lacked the necessary commitment by all school community members, which is a success key in the case of the hunter model since it requires extensive training and conferences.

The pursuit to increase the quality of teacher evaluation was translated into a significant distinction in 1980's, which later developed to become a universal benchmark. That was the release of the personnel evaluation standards by *The Joint Committee on Standards for Educational Evaluation*, published by committee chair Daniel Stufflebeam (1988). The committee intensified its efforts at the late 1980's to set standards for educational evaluation systems; such standards were differentiated into four categories: Propriety, utility, feasibility, and accuracy. Million (1987) also recommended standardized teacher evaluation through his multi-strategic approach; however, he failed to establish a wider scope at teacher's overall role inside and outside the classroom.

Despite the dominance of classroom observations, awareness about high dependence on such tool was raised even at that time. Scriven (1981) introduced six possible risk factors associated with classroom observations: Unnatural teachers' behavior, insufficient number of samples, possible observer bias, time cost, narrow raters' perspectives, and evaluators' favoring of certain teaching styles over the others.

2.3.4 1990's: The formal beginning of multi-dimensional evaluation

Continuation of the 1980's conventional limited methods used in measuring teacher competencies stimulated frustration among principals and teachers in the early 1990's (Brandt, 1996). Limiting the role of teachers to the minimal of lecturing and transmission was disappointing to a large number of teachers who

were already practicing beyond that minimal area. Furthermore, principals and administrators were also complaining about the 1980's limitations as they represented obstacles to providing more comprehensive judgments about teaching practices (Weiss & Gary, 1998).

Rebellion against traditional practices inherited from the past decade triggered the 1990's to become a significant reform era. New evaluation systems were developed to reflect teachers' performance regularly (Sclan, 1994). Researchers focused on providing access to comprehensive systems that were not limited to the scene inside the classroom, which came in coherent with the general objectives of education that targeted constructivist teachers who aim to generate long-term learners (Weiss & Gary, 1998).

One remarkable design was introduced by Davey (1991), who viewed educational performance assessment as a testing strategy which follows subsequent processes: (1) Defining teacher's roles, (2) Using multiple, systematic evaluation procedures, and (3) Employing multiple sufficiently-trained evaluators. Davey's approach appeared promising; however, districts that applied his approach recorded poor outcomes. Flaws in Davey's (1991) outline were self admitted by Davey himself; in that the complexity of teaching and learning and many elements that influence teachers' performance such as knowledge, traits and abilities, were not well-targeted by his approach.

The 1990's also marked the early efforts of Charlotte Danielson (1996), one of the most known pioneers in launching new trends in teacher evaluation. The Danielson's model (1996) suggested a formative purpose-oriented approach that defined several teaching functions, and enlisted measurable variables to each function, constituting a system that represents a significant reference until present time. According to Danielson (1991), each teacher was responsible for four main functions: Planning, classroom environment, instruction, and professional responsibilities.

2.3.5 2000-present: Quality standards, Focus on professional development and students' outcomes

The beginning of the 21st century marked remarkable progress in the field of teacher evaluation with the focus on quality standards under the influence of *The No Child Left Behind Act* development in 2002. The NCLB developed a set of standards to ensure high quality education in schools. This in turn obligated schools to pay extra focus on the quality of teachers they are hiring and maintaining. As a response to NCLB legislation, many districts and schools began to reconsider their teacher evaluation systems and apply benchmarks inspired by NCLB standards (Scherrer, 2009).

Focus on standardized teacher evaluation which began in 1980's gained increased preference following the year of 2000. Milanowski, Kimball & White (2004) discussed how standardized evaluation can have improving impacts on instruction and accountability, and is consistent with the new trends of standardized-learning. However, they argued about challenges attached to such evaluations, including data collection intensity and the need for performance evidence to enrich evaluation.

Danielson and McGreal (2000) defined two main purposes for teacher evaluation: Quality assurance and professional development. This contemporary shift in teacher evaluation purposes from focusing on teachers' appraisal and incentives decisions in the past into seeking excellence and high quality reveals major changes in educators' understanding. The efforts of Danielson and McGreal (2000) ascended from the 20th to the 21st century to emphasize the use of formative teacher evaluation (improvement), with the addition of summative characteristics (Accountability, competence).

Other recent studies (Tuytens & Devos, 2010) suggested the use of job descriptions in teacher evaluation. This was successfully applied by the Flemish government (Belgium) in 2007. Another approach presented by the Flemish

experience was extending the period of teacher evaluation up to four years, to enhance teachers' improvement.

Using evaluation in predicting teachers' impact on students' outcomes was another area of interest during the recent era (Rockoff, 2004; Rivkin *et al.*, 2005; Harris & Sass, 2006; Aaronson *et al.*, 2007; and Clotfelter *et al.*, 2007; Tuytens & Devos, 2010)

2.4 Standards of effective teacher evaluation systems

To obtain full understanding of good-practice criteria for teacher evaluation and be able to reflect these criteria on the present research, the researcher reviewed literature that discussed characteristics of effective teacher evaluation systems in terms of purpose, design, implementation, and practice.

Review and summary of related literature yielded the following criteria:

1. Clear definitions of evaluation purposes and objectives

Hinchey (2010) suggested that policymakers should have clear understanding about teachers' assessment purposes. A combination of both, formative and summative purposes must be considered to achieve better evaluation.

More importantly is the nature of these objectives; such objectives must be mutual between all members of the evaluation system, including teachers (Stronge, 2005). Scriven (1972) explained the significance of this focus: "*It is obvious that if the goals aren't worth achieving then it is uninteresting how well they are achieved*" (1972, pp. 126-127). If teacher evaluation goals are solely important for the institution, the targeted teachers are expected to show resistance to the system which will lead to mass failure.

In many cases, stakeholders might succeed at achieving consensus on evaluation goals; however, they tend to neglect the significance of relating those goals to all the decisions and processes involved with the system (Goe, Holdheide & Miller, 2014). Goe *et al.* (2014) suggested the following evaluation

objectives: Personnel decisions, professional development, teachers' accountability, improving students' test scores, and approaching reform efforts. Moreover, they insisted that these objectives must be considered by schools' policymakers throughout all the evaluation steps, including design and weighing of evaluation tools, and utilization of evaluation data.

2. Define effective teaching standards based on credible references

With the current educational climate focusing on standardized learning, an urging need for standardized teaching emerges. Typically, setting teaching standards starts with forming a proper definition for an effective teacher. This definition is then differentiated into detailed components of specific teaching competencies and expected outcomes (Goe, Holdheide & Miller, 2014).

In 2010, The Department of Accreditation of UAE's ministry of education published a manual to guide schools about accreditation standards including standards for quality teaching. These standards covered areas related to subject knowledge, proper objectives, interaction with students, teaching strategies and resources, promoting learning continuity, targeting students' needs, and monitoring students' progress. However, the accreditation manual did not cover other standards such as professional commitment, professional relationships and professional development. The provided guide enlisted these standards within other focus areas such as school's leadership and community, which indicates that the Ministry of education considers professional areas to represent the responsibility of the school's administration rather than the teachers'.

3. Emphasis on systematic communication

Evaluation communication includes public communication to educate all teachers about teaching standards, evaluation criteria and dimensions. In addition to private communication established through ongoing individual follow-up and feedback conferences (Stronge, 2005). According to Stronge (2005), many evaluation systems perceive communication only as an end-step to deliver

evaluation outcomes to teachers, which diminishes the fulfillment of evaluation desired objectives in achieving high quality teaching.

According to Stronge (2005), communication expands to cover the public community which has the right to explore teacher evaluation systems applied in schools to benefit from the experience and recognize schools' efforts in seeking effective teaching. However, precautions of maintaining confidentiality and personal information of teachers must be carefully deliberated.

4. Setting an appropriate climate for high-quality evaluation

A main challenge in any teacher evaluation is the resistance of evaluated teachers. For teachers, evaluations are usually perceived as being personal and emotional, which causes them to react against it, whether directly or indirectly (Stronge, 2005). This resistance can be minimized through schools' efforts in raising the trust level and positive relationships between evaluators and teachers (Stronge, 2005). Nevertheless, in certain cases of persistent weak teachers' performance and unresponsiveness to remedial guidance, a support-based evaluation climate can deliver misleading messages. In such cases, school administrators are required to establish a clear, firm and sometimes negative climate with weak teachers to achieve meaningful evaluation, and to preserve students' interest.

5. The use of multiple data sources

Stronge (2005) recommended the integration of multiple sources in teacher evaluation because it creates realistic judgments about teachers' performance; considering that the functions and responsibilities assigned to teachers are too diverse to be measured using a single instrument.

Multi-measurement teacher evaluation systems will not only facilitate teachers' differentiation according to effectiveness, but will also diagnose professional development areas and increase evaluation credibility; hence, any weakness in a single tool is compensated by the strengths of another (Weisberg *et al.*, 2009; Hinchey, 2010; Glazerman *et al.*, 2011; Dretzke, Sheldon & Lim, 2015).

In multi-measures evaluation systems, allocation and weighing of measures depend on various factors, including: measure's validity, measure's impact on students' achievement and progress, and the reliable and amount of information provided by the measure (Goe, Holdheide & Miller, 2014).

6. Proper selection and training of evaluators

Goe *et al.* (2014) pointed out that implementing teacher evaluation systems with fidelity and integrity requires selecting and training evaluators who can understand the purpose of evaluation and provide reliable evaluations.

Suggested criteria for selecting evaluators included: content knowledge, subject-experience, and evaluation-experience.

Every evaluator should receive professional training to a certain level, whether to introduce a new evaluation system to existing evaluators, or new evaluators to existing systems, or to ensure good practice continuity; however, the amount and nature of training depends on the types of evaluation measures. Observational measures require fundamental reliability training and training evaluators to invest data in improvement. Whereas value-added measures require training related to measurements techniques and data interpretation (Goe, Holdheide & Miller, 2014).

In the local region, Farah and Ridge (2009) criticized evaluators' selection in UAE. They reported that school principals tend to choose evaluators who produce positive evaluation reports rather than improving teaching standards. Moreover, Thorne (2015) considered that many evaluators in UAE reported questionable practice by depending on their own teaching conventional transmission approach, and lacking sufficient familiarity with updated pedagogical approaches.

7. Employing sufficient resources

Adequate resources allocation is a success key to produce high-quality evaluation systems (Danielson & McGreal, 2000; Hinchey, 2010). Such resources include: cost, time, patience and flexibility. Availability of financial and

human capital resources can influence the whole design of the system. Hence; every district/school should design and implement an evaluation system that is feasible with the available resources, where resources must be sufficient not only to cover the evaluation processes, but also the training and professional development plans afterwards (Goe, Holdheide & Miller, 2014).

8. Valid evaluation tools

Validity is a key consideration for any assessment, and it usually refers to supporting any judgment with evidence. In teacher evaluation, provided evidence must prove that the measurement tool is in fact measuring the area of teacher effectiveness it claims to target (Goe, Bell, & Little, 2008).

Measuring the validity of teacher evaluation tools can be difficult and depends on the area covered by the tool. For instance, if a tool is claimed to measure students' achievement, evaluators must define whether this tool measures students' learning or test scores. The difference between both is that test scores' validity might be intentionally affected by teachers through several ways that do not improve students' learning (e.g., teaching for the test, allow cheating). As noted, teachers' actions might influence evaluation validity (Papay, 2012).

2.5 Teachers evaluation tools

Screening available literature produced a set of examples on teacher evaluation tools which will be discussed to guide this research's present study.

2.5.1 Classroom observations

Classroom observations are considered a common tool to evaluate teaching due to its benefits in measuring instruction, classroom environment, interactions, and students' participation. However, researchers admit that they are only suitable for measuring direct teaching activities. In this context, observations are incapable of evaluating outside-classroom activities, nor can they measure authentic feelings and beliefs (Danielson & McGreal, 2000; Little, Goe & Bell, 2009). Duke & Stiggins (1986) suggested that evaluators should establish a balance between

managing limited time available for evaluating many teachers, and providing adequate time for profound observations to grasp a wide scope of teaching.

2.5.2 Teacher self-assessment

Danielson & McGreal (2000) invested large beliefs in teachers self assessment as an evaluation instrument. Under the condition of introducing standardized performance criteria, teachers can reflect their self performance and professional development adding actively to their evaluation and improvement. Little, Goe & Bell (2009) added to this argument by expressing self-assessment's role in measuring un-observable criteria and the uniqueness it adds to the overall evaluation scheme. However, they insist that it must not be relied on as the sole source of evaluation, due to its validity and reliability issues.

2.5.3 Teaching artifacts and planning documents

Teaching artifacts are documents prepared by teachers to organize students' activities. These could include assignments, projects and worksheets. According to Little, Goe & Bell (2009) planning documents are also considered teaching artifacts. Danielson & McGreal (2000) suggested analyzing planning documents independently as evaluation data sources and justified that by discussing how teacher's planning skills are too important to be evaluated among wide criteria; however, they shared the same conceptions with Little, Goe & Bell (2009) on the importance of teaching artifacts in evidencing classroom life, and reflecting teachers' abilities of planning and preparing assignments that satisfy students needs. Little, Goe & Bell (2009) raised cautions regarding this tool stating that it needs further reliability and validity research. Another issue is that research has no definite answer for the right amount of work sample to be analyzed.

2.5.4 Students and parents surveys

Danielson and McGreal (2000) called multi-perspectives' evaluation systems "360-degree systems". In schools, students and parents' surveys can be a great source of information about teacher's performance, but they require high caution in design

and application. Student's surveys must be suitable for their age, and must raise questions about teaching rather than the teacher as a person. As for parents' surveys, they need to be less detailed and ask for accessible information. Little, Goe & Bell (2009) stated that students' surveys could reflect the perspectives of students, who have the most experience with teachers, in addition to their incorporation into formative teachers' assessment. Follman (1992) expressed that while most studies support the use of students' ratings, researchers warn about using such ratings as a sole evaluative tool due to students' inability to evaluate teaching areas that require professional experience, such as curriculum, adequacy, and professional relationships.

2.5.5 Teachers' portfolios

Teachers' portfolios are distinct from teaching artifacts in that portfolios include materials created and collected by teachers for the purpose of evaluation, whereas teaching artifacts are samples of materials already being used in teaching and learning (Goe, Bell, & Little, 2008). Portfolios might include materials that are invisible through any other evaluative instrument (e.g. reflective writing, students/parents notes, teachers' awards..), and they provide an opportunity for teachers to demonstrate their performance and accomplishments. However, portfolios' reliability might be questionable. *"Portfolios can offer a very comprehensive and in-depth portrait of teaching activities; however, their complexity can raise concerns about the ability of scorers to evaluate them reliably"* Goe, Bell, & Little (2008, pp. 33). Danielson and McGreal (2000) briefly suggested the use of teachers portfolios in their argument to shift the passive role of teachers in traditional evaluation systems into an active role. In this context, the two authors emphasized the need to set standards for teachers' portfolios, and inform teachers of those standards.

2.5.6 Value-added models

The value-added model is one of the most recently suggested procedures introduced for evaluating teachers. Although it is increasingly gaining popularity

among researchers and educators, this model is highly controversial and has been criticized for its complexity and inconsistency. According to Darling-Hammond *et al.* (2012), the value-added model is mainly based on the concept of estimating the influence of individual teachers on students' achievement by applying statistical operations that use students' results of multiple years, while considering other factors that might influence students' achievement. These models were criticized by Darling-Hammond (2012) along with others for their inconsistency, being influenced by students' nature, and their incapability to cover all variables that affect students' results.

2.5.7 Other teacher evaluation tools.

Some studies suggested the use of other teacher evaluation tools including: samples of students' work, professional responsibilities and professional development documented activities (Danielson & McGreal, 2000).

2.6 Teachers evaluation for professional development

2.6.1 Danielson's Three-tracks' approach to evaluation-motivated professional enhancement

Danielson & McGreal (2000) classified three tracks to enhance professional practice through evaluation, depending on the category of targeted teachers:

Track 1: For beginner teachers: new teachers can enhance their practice by learning from the evaluation process itself, including good practice criteria, evaluation measures and steps, and the feedback they receive afterwards. This track is not only used to enhance practice, but also for personnel decision making about beginning teachers (e.g., retention, tenures...).

Track 2: Professional development for all tenured teachers: teachers learn through ongoing interactions with administrators during summative evaluations. Moreover, teachers improve their performance by designing professional growth plans as part of a formative evaluation approach. Normally, all tenured teachers

follow track (2), but whenever supervisors detect weaknesses in an individual teacher's performance, and the targeted teacher fails to respond to remedial feedback, he/she will be transferred to the teacher assistance program.

Track 3: The teacher assistance track. This involves tenured teachers who show declined performance in one or more effective teaching standards. The number of involved teachers may vary according to the rigidity of district's standards. Direct supervisors are responsible for involving teachers in this track according to teachers' satisfaction of four main areas: Planning, Classroom environment, Instruction, and Professional responsibilities. Unlike tracks (1) and (2), track (3) targets teachers individually, and this is usually done through three phases: **An awareness phase**, to identify, document and address concerns about teacher's performance. **An assistance phase**, to prepare growth plans and review progress. And, **a disciplinary phase**, to discuss weaknesses or violations with the teacher and indicate recommendations for specific remedial plans, training, contract nonrenewal, or tenure review.

2.6.2 The influence of teacher evaluation on teachers' preparation programs

Darling-Hammond (2010) discussed how teachers' performance assessment can help to improve teaching. She supported her argument by referencing the influence of Performance Assessment for California Teachers (PACT), launched in 2002 and has been used ever since for licensing California teachers, in improving teachers' preparation programs. The detailed data delivered to these programs through assessment provide explicit information about each candidate, and reveal performance patterns that are greatly beneficial for personnel professional development. Moreover, PACT provides state-level data which can be used in effective teacher licensure, discovering teachers' initiatives, and in teacher induction and in-service training plans.

2.6.3 The need for leadership-level policies to link evaluation to professional development

A qualitative study conducted by Al-Bustami (2014) in Abu Dhabi, UAE explored supervisors' and teachers' perspectives on the relationship between teacher evaluation and professional development. Almost all participants agreed that teacher evaluation; if done properly can be directly linked to enhancing professional practice; however, they raised suspicions about the possibility of such link in the lack of supporting policies in their schools. Under such circumstances, supervisors believed that they were limitedly encouraged to use evaluation in improving teaching and providing recommendations for teachers to enhance weak performance areas and amplify the strong ones.

Another study conducted by Chicago University on 2012 to examine a new teacher evaluation system in Illinois districts revealed several challenges related to this area (White *et al.*, 2012). Many teachers admitted the efficacy of the evaluation system in diagnosing performance weaknesses; however, they were less satisfactory at facilitating the improvement of these weaknesses. Evaluators also expressed the difficulty they faced in addressing teachers about performance setbacks and identifying the next improvement steps (White *et al.*, 2012). The Illinois study discussed three leadership-level policies implemented to promote evaluation-based professional growth: **(1) Building evaluators' capacity, (2) Relating evaluation to professional development.** Districts' administrators reportedly designed training plans based on weak areas, and build bases of evaluation-obtained data to move beyond tracking commitment toward improving teaching. **(3) Increasing the rate of classroom observations**, where more frequent observations increase the chance of detecting problems and addressing them early rather than rushing into judgments about teachers' efficiency.

2.7 Summary

Literature review provided scopes on the definition and progress of teacher evaluation. It became clear how teacher evaluation systems evolved through past decades, where it started with informal evaluation based on personality judgments, and developed to questioning and defining purposes for evaluation, with the introduction of benchmarks to frame evaluation. Basing evaluation on effective teaching standards also grew interest over the years, in addition to the transformation of systems that depended on a single tool (mainly classroom observations) to incorporating multiple tools and evaluators into the evaluation scheme. These scopes enlightened the researcher about the history and future of teacher evaluation.

The researcher then summarized eight main characteristics of effective teacher evaluation systems discussed by previous literature. These standards will later be used to evaluate the evaluation system targeted by this research. Additionally, teacher evaluation tools, which represent the main components of the targeted system, were discussed using examples from recent literature.

Representing a main area of this research, evaluation-based professional development was discussed through several approaches, including Danielson's three tracks' approach to evaluation-motivated professional enhancement, employment of evaluation in teacher preparation programs, and examples of leadership policies suggested to relate evaluation to professional development.

Through literature review, the researcher detected a certain gap in previous studies that needs to be addressed; where quantitative analysis to study the relationship between teachers' evaluations and students' results was lacking.

All these scopes and discussions are expected to facilitate the design and implementation of the present study in this research.

Chapter 3: The present study

3.1 Overview

In a pursuit to answer the research questions, an empirical study was conducted to evaluate the validity of the targeted teacher evaluation system by comparing its results to another source of evaluation; students' achievement. Moreover, the study explored the perspectives of different stakeholders, including schools' leadership and teachers about the targeted teacher evaluation matrix.

This chapter explains the phases of the present study including the case studies, the purpose of each phase, methodology, sampling, and research tools.

3.2 Research design

3.2.1 Case studies

Three private schools in the emirate of Sharjah - UAE were targeted through this research:

- (1) Al-Shola Private School for KG, Elementary and Girls (KG to grade 12): established in 1983, with around 3000 enrolled students and 142 teachers.
- (2) Al-Shola Private School for Boys (Grades 5 to 12): established in 2004, with around 2000 enrolled students, and 87 teachers. Separate from the first branch with independent licensure and property.
- (3) Manarat Al-Sharjah Private School (KG to grade 12): established in 2014, with around one thousand enrolled students, and sixty teachers.

All the three schools apply the national curriculum of UAE and are part of the same organization.

These schools were selected as case studies because they share the same teacher evaluation system, a system that was designed as an initiative by the school leadership team. Additionally, being part of this leadership team as the head of the quality control department that covers all the three schools, the researcher was highly-motivated to evaluate a school-level policy in these schools as a form of self-evaluation.

3.2.2 Research methodology

Methodology for this study was designed and applied to cover three parallel phases, where the methodology, sampling, research instruments, and data collection varies between the three phases according to their purposes. This type of triangulation between research methods is termed as 'between-method triangulation' and it usually involves examining the same phenomenon by a combination of several qualitative and quantitative approaches, which helps in achieving collective validity (Hussein, 2015). This methodology was selected because this study aims to evaluate a teacher evaluation system that is complex in nature and involves several stakeholders. Therefore, it requires multiple data sources and different perspectives.

Several advantages of concurrent mixing of quantitative and qualitative research approaches were discussed by research experts. One important advantage is the further explanation, augmentation; and in some cases, contradiction provided by qualitative data to results obtained quantitatively (Driscoll *et al.*, 2007).

Moreover, this research was formulated as a case study. In case studies, a current phenomenon is investigated in its context by employing multiple data sources (Yin, 1989). The intention of case studies is to focus on a certain policy rather than studying the entire organization which facilitates deep analysis of a particular problem/system by individual researchers (Yin, 1989; Bell, 2014). Additionally, case studies can differentiate between planned policies and actual practice (Anderson, 1998).

In December of 2015, the researcher submitted the research proposal to the British University in Dubai. Approval was obtained from the university at the same month, and a letter asking permission to conduct a research was submitted to the Owner/Manager of Al-Shola educational organization which runs the three targeted school. The letter explained the objectives of the study

and the data collection procedures. Permission to conduct the study was granted on January of 2016 which allowed the researcher to commence with the research phases described below.

3.2.3 The three phases of the study

The mixed methodology for this research consisted of three simultaneous phases:

(A) Phase I – A Qualitative interview: Perspectives of policymaker's about the teacher evaluation matrix

❑ Purpose

The main objective of this phase was to explore the perspectives of the key policymaker of the targeted teacher evaluation matrix. These perspectives will be used to evaluate the matrix through reflection over good practice criteria.

❑ Methodology

A qualitative interview was selected for data collection. The qualitative approach is usually promoted in educational research for its utility in describing personal experience, and its interpretative nature. Additionally, it provides an opportunity to identify the uniqueness of the case by collecting features and events that can only be known by the person close at hand (Stake, 2010).

❑ Sampling

In order to investigate all the aspects related to the targeted evaluation matrix, the researcher selected the key policymaker of the matrix. His position as the leader of the educational organization that runs all the targeted three schools, in addition to being the principal of one of those schools makes him the most suitable candidate for the interview.

❑ Instrument

Selection of the interview as a part of this research's instruments was based on the researcher's desire to obtain unique data and interpretations of the interviewee and to find out about unobservable areas related to the teacher

evaluation matrix. According to Stake (2010) these two purposes are among the main purposes for research interviews.

Bell (2014) describes research interviews as being adaptable, where interviewers can discover motives and feelings, and use follow-up questions to interpret answers. On the other hand, high subjectivity and an increased risk of bias can be associated with interviews. Therefore, Bell (2014) advises researchers to be cautious in selecting interview questions and interpreting the answers they receive.

Guidelines suggested by Bell (2014) for the wording of interview questions included: considering the targeted research areas, establishing a rational sequence for questions, and avoiding any questions that might be offensive or led by presumptions. Consistent with Bell's (2014) guidelines, criteria for the interview questions were defined through literature review that identified effective teacher evaluation.

Summary of such criteria produced eight main areas for the interview questions:

1. The purpose of the system.
2. Alignment of the system with effective teaching standards.
3. The nature of communication associated with the evaluation system.
4. The school climate established to facilitate teacher evaluation.
5. Data sources employed to reflect teachers' performance.
6. Qualifications and credibility of evaluators and evaluation tools.
7. Resources employed to support the teacher evaluation system.
8. Feedback and professional development opportunities offered through the teacher evaluation system.

Differentiation of the eight areas yielded fifteen questions; mostly were open-ended in addition to a few closed questions. While open questions are usually described to cover broader variables and possibilities, closed questions can

also be generated to provide some extent of explanation (Roulston, 2010). To achieve such extent, any closed (Yes/No) interview questions used in this research were followed by follow-up clarification suggestions.

The interview questions were approved by the dissertation coordinator after providing appropriate feedback. Moreover, questions were submitted to the interviewee prior to conducting the interview, and all questions were approved by the candidate.

The detailed fifteen questions used during the policymaker interview are demonstrated as **(Appendix.2)**.

□ Data collection

Two meetings were arranged with the manager of Al-Shola educational organization to cover all the generated questions. During these meetings, the interview was conducted in a semi-structured form. According to Roulston (2010), in a semi-structured interview a number of questions are prepared as a guide, which represents a starting point to initiate discussions with the interviewee. The semi-structured approach was selected to reflect the experience and unique perspectives of the targeted policymaker. Unlike structured interviews that use pre-set, multiple-choice answers; semi-structured interviews allow the interviewee to use his/her own words and perspectives.

(B) Phase II – A Quantitative questionnaire: Perspectives of evaluated teachers about the teacher evaluation matrix

□ Purpose

The key goal of this study was to evaluate the teacher evaluation system of Al-Shola schools; therefore, it was logical to include the perspectives of those whom the system was designed to evaluate – the teachers. The purpose of this phase was to recognize the perspectives of teachers on the targeted teacher evaluation matrix through a detailed questionnaire.

The preliminary step to initiating the design and application of the questionnaire was setting its objectives. The questionnaire seeks teachers' perspectives on: (1) Effectiveness of the matrix's design and implementation. (2) The importance and reliability of evaluation tools used in the matrix. And, (3) Feedback and professional development opportunities offered for teachers through evaluation.

❑ **Methodology**

The quantitative approach was followed to facilitate data collection and analysis from a large sample. Preliminary readings and the study plan guided the researcher in deciding what needs to be found through the questionnaire. Questionnaires are highly popular as a research tool due to its several merits; they allow the researcher to target a large number of participants which in turn increases the reliability of results. Moreover, they are associated with low time-consumption and a decreased bias risk when compared to research interviewing (Kothari, 2004; Stake, 2010).

❑ **Sampling**

The research questionnaire is directly related to the teacher evaluation matrix applied in the targeted schools, and most of its sections can only be perceived if the participant had been evaluated by the matrix. As a result, all the teachers who were evaluated by the matrix at the latest academic year and were still working in the school at the time of the study were selected as the sample. The sampling was based on the research purpose and the nature of the questionnaire; therefore, it is considered as a purposive/non-random sample (Parfitt, 1997).

A total number of 123 teachers who were evaluated by the end of the academic year 2014/2015 were selected. Biographic data of the sample is demonstrated in **Table 3-1**.

Table 3-1 Summary of questionnaire participants' biographical data

Field	Data
Total sample	123
Total respondents	95
Case.1 respondents (Al-Shola School-KG, Elementary, Girls)	39
Case.2 respondents (Al-Shola School-Boys)	36
Case.3 respondents (Manarat Al-Sharjah School)	20
Gender	45 Males - 50 Females
Age Range	22 – 55 years
Educational Background	Bachelors – Masters
(0 – 5) years of experience in the school	49
More than 5 years of experience in the school	46

❑ Instrument

A questionnaire was prepared to collect data about teachers' perspectives on the teacher evaluation matrix. Criteria that define effective teacher evaluation summarized through literature review were employed to guide the wording of statements/questions of the questionnaires. Such criteria were similar to those used in the policymaker interview; however, criteria that include information inaccessible to teachers were excluded.

Three main domains were targeted through the questionnaire, and three different responding approaches were employed depending on the nature of the investigated data. According to Bell (2014), answering options offered to survey participants can be in a verbal form, a list, a scale, ranking, category, or quantity. The scheme used in the teachers' questionnaire is demonstrated in **Table 3-2**.

Table 3-2 Teachers' questionnaire scheme

Domains	Sub-domains	No. questions	Responses' approach
Criteria for effective teacher evaluation systems	Purpose of evaluation	2	Likert Scale
	Teaching standards	3	
	Systematic communication	3	
	Climate for high-quality evaluation	2	
	Qualifications of evaluators	4	
Teachers' performance evaluation tools	Significance of evaluation tools	1	Ranking
	Reliability of evaluation tools	2	A multiple choices list
Feedback and professional development	Feedback	1	Likert Scale
	Professional development	2	

The detailed questionnaire is enclosed as **(Appendix.3)**

Considering that all the targeted teachers were of Arabic nationalities, the questionnaire was translated into Arabic language by the researcher under the guidance of the English Subject supervisor at the school **(Appendix.4)**.

□ Validity and reliability

The questions used in the teachers' survey were inspired by previous literature (Danielson & McGreal, 2000; Peterson, 2004; Stronge, 2005; Weisberg *et al.*, 2009; Hinchey, 2010; Glazerman *et al.*, 2011; Goe, Holdheide & Miller, 2014; Dretzke, Sheldon & Lim, 2015).

To provide further validation to the survey, all questions were approved by the dissertation supervisor after providing the required feedback. Validity and reliability of the questionnaire were also approached through *a pilot study* as described in the next section.

❑ Piloting the questionnaire

Bell (2014) emphasized research instruments' piloting to check the appropriateness of several factors, including the time required by participants to fill the instrument and the clearance of instructions and questions. Piloting questionnaires and other research tools provides an opportunity for editing and removal of ambiguous or unusable information.

For this research, the questionnaire was tested on a pilot group of twenty teachers as a trial. All of the pilot study participants stated that the questionnaire's purpose, instructions, and questions were clear to them, and that it took less than ten minutes to be answered; thus, the questionnaire was considered ready for the distribution stage.

❑ Data collection

After obtaining the permission of school officials, a total number of 123 questionnaires were distributed among the sampled teachers. Within a one week period, the researcher received a total number of 95 responses. The researcher was present during most of the questionnaires' answering to clarify any areas if required.

(C) Phase III – Quantitative analysis and correlation coefficient: The relationship between teachers' performance evaluation results and students' results

❑ Purpose

The main aim of this phase was to measure the validity of the teacher evaluation matrix applied in the targeted schools by comparing the data obtained through the matrix with the data of a second source of evaluation; students' results.

❑ Methodology

The researcher followed a quantitative approach to analyze the relationship between the two variables. According to Hopkins (2000), quantitative design

is the most appropriate approach when the researcher desires to quantify a relationship between two variables, especially if this relationship was to be expressed in a statistical manner. Moreover, the use of quantitative methodology in testing hypothesis was suggested by Creswell (2009).

The researcher quantitatively compared between the three targeted schools using two factors: averages of teachers' evaluations and averages of students' results for each subject.

Moreover, the statistical relationship between a sample of teachers' evaluations and their students' results was investigated. Considering the bivariate nature of the population; Karl Pearson's coefficient of correlation was selected as the statistical tool in this study (Kothari, 2004).

❑ Sampling and Data collection

The researcher contacted the targeted schools to obtain two types of information:

1. Teachers' performance results obtained through the latest teacher evaluation matrix of the academic year: 2014 – 2015, in the form of numeric results out of 1000.
2. Students' end results in the ministry of education's formal examinations of the same academic year.

Five subjects were selected as the sample of the research: Islamic studies, Arabic language, English language, Mathematics, and Science. The rationale for such sampling is that other subjects including social studies, arts, and physical education are not included in the ministry of education standardized examination system; therefore, reliable end-year students' results for such subjects cannot be obtained.

Moreover, the sample of the study was confined to data about teachers and students of grades six up to twelve. The rationale behind this selection was that these grades are covered by the ministry of education standardized

examination system, while younger grades are assessed at school level. Therefore, achievement results of higher grades represent an external and standardized source of evaluation; thus they are considered a more reliable and authentic reference for comparison with the targeted teacher evaluation system.

To validate the findings of this study phase, a random sample of (40) teachers was selected collectively from the three case studies. The researcher collected data about the individual evaluations of these teachers resulting from the matrix, and the average of end-year results of the sections taught by each teacher within the sample. The purpose of this procedure was to measure the relationship between the targeted matrix's results and students' results using a statistical approach, which would provide support to any judgments based on the averages' comparison explained above. To analyze the data, Pearson correlation coefficient was calculated to measure the nature and strength of the described relationship.

3.3 Ethical consideration

Ethics are an important issue and must be highly-valued in any research. For ethical consideration in this research, several procedures were followed. The researcher contacted the owner/manager of the targeted educational organization and the principals of the three targeted schools about the research proposal. After explaining study objectives, procedures, terms of confidentiality, and uses of study results, permission was granted by all of them to apply the research. In addition to that, the selected candidate for the research interview, the leader of Al-Shola educational organization, was invited to voluntarily participate in the interview. Prior to interview conduction, he received a copy of the questions planned for the interview, and signed an informed consent (**Appendix.5**).

An intellectual property issue was found to be related to this research. This issue emerged from the fact that the studied teacher evaluation system was designed internally by school members and had never been published. Respect to intellectual property was considered by discussing the issue with the organization leader and explaining that any data obtained through this study will exclusively be used for educational research. The organization leader granted permission with a request of using the name of the organization when discussing the teacher evaluation system.

All teachers selected for the questionnaire were invited for a volunteer participation in the study through informative circulations. The circulations included information about the purpose of the study, confidentiality terms, and the use of data collected during research. Informed consents were also signed by participating teachers before filling the questionnaire **(Appendix.6)**.

Chapter 4: Results and interpretations

4.1 Overview

In this chapter all the data collected through different research instruments will be analyzed and discussed. Data analysis and interpretation will focus on three areas: (1) Analyzing the school leadership's perspectives on the targeted teacher evaluation matrix. (2) Analyzing teachers' perspectives on the matrix and the evaluation tools included within. And, (3) Comparing between the three case studies using two factors; teachers' evaluations resulting from the targeted system and achievement results of students taught by those teachers. In addition to measuring the relationship between these two factors.

4.2 Study results and interpretations

4.2.1 Phase I: Perspectives of school leadership about the teacher evaluation matrix

▪ Results

To collect the most explicit data about the targeted evaluation matrix, the main policymaker who is also the leader of the organization that runs all the targeted schools was selected for a personal interview. During the interview he was asked about aspects that are inaccessible to teachers or other school members. The answers introduced during the fifteen-question interview were reported by the researcher, and then summarized and categorized into eight main areas:

Area (1): Teacher evaluation purpose

The first two questions were related to the purpose of evaluation defined by the school. Through the first question, the policymaker defined the school's main purpose to develop and implement the current evaluation matrix, which was to use a fair and comprehensive evaluation system. Moreover, he differentiated other secondary purposes including: (1) Using an evaluation tool that considers various activities and duties of the teacher instead of the traditional tools based on classroom performance. (2) Stressing on the fact that teachers should

practice social, cultural and ethical roles in their schools. (3) Using a tool that covers teacher's behaviour all over the year. (4) Covering certain areas of interest for teachers (e.g. training, career development...). And (5) Emphasizing the importance of students' and parents' role in evaluating teacher performance.

The second question was related to the nature of evaluation purposes, in response, the policymaker described the school's evaluation purposes as a combination of both formative and summative, where evaluation is perceived by the school administration as an on-going improvement process. The formative purpose is satisfied through evaluation procedures conducted all along the academic year, while the summative purpose is approached through a final and comprehensive outcome in the form of a matrix or a report that reflects teacher's performance as a whole.

Area (2): The use of effective teaching standards in policy design

The interviewed policymaker assured that the current evaluation system was based on effective teaching standards that were formulated through deep and sufficient discussions between key members of the school. Moreover, these standards were inspired by the common practice in the educational field in UAE, in addition to various studies and practices conducted worldwide. The policymaker's own experience and the collective experiences of the school's educational team represented the main source for such standards. The interviewed policymaker stated: *"One of the main sources of these standards was my project for the EMBA degree titled: Developing a balanced score card for Al-Shola Private school"*.

The policymaker then described the methodology used by school to educate teachers about these standards, where he referred to their professional training policy that had successfully covered more than 150 topics related to teaching standards over the past 25 years. Such training included: seminars, workshops, questionnaires, and worksheets. Teachers themselves were encouraged to

conduct research about teaching standards, and were rewarded for that by the school administration.

Area (3): Systematic communication during evaluation

The interviewee provided examples on evaluation communication at the early stage of policy design and implementation. The school administration communicated a selected management team to provide suggestions on evaluation tools and the weight of each tool as well as the standards and sources of evaluation. This communication continued for more than six months before the first policy implementation, and various changes were introduced based on actual practice. The head of quality control department was the leader of this process. Communication expanded to include teachers, where many meetings were held with all teachers to illustrate the components of the matrix and their feedback was taken in consideration.

Area (4): The school climate established to facilitate evaluation

The school leadership attempted to promote teachers' acceptance of the matrix by providing individual feedback reports and sufficient classifications to teachers. In addition, all queries and objectives were patiently addressed. According to the experience of the school leader, these attempts had a positive impact; he stated: "*In principle, all teachers accepted the concept of the matrix*".

Area (5): Data sources employed to reflect teachers' performance

According to the policymaker's answers, the school referred to many data sources to reflect teachers' performance, including: examination department, training and development department, educational supervisors, section supervisors, human resources department, quality control department, and the IT department. These sources provided the necessary information for the twelve evaluation areas.

Moreover, he stated that most of those sources are applicable to all teachers; however, some of the matrix tools are not applicable to certain subjects and

levels. For instance, students' questionnaire could not be used in the primary stages. Certain subjects don't have an examination process; thus, students' results could not be used as a tool of evaluation.

Area (6): Qualification and credibility of evaluators and evaluation tools

The schools' leader described how evaluators went through an extensive training process over teaching standards and evaluation criteria since the beginning of the initial policy design efforts back in 2010. By time, they developed enough experience to deal with all issues and difficulties they faced during evaluation. Credibility and reliability issues related to evaluators and evaluation tools were addressed by the school administration through assigning the whole evaluation scheme to the quality control department, which reported directly to the school principal. This process guaranteed enough degree of credibility in general; however, the schools' leader admitted that certain tools still rely greatly on the judgment of evaluators and could include a margin of personal views. The feedback from teachers themselves could reveal any discrepancies.

Area (7): Resources employed to support the teachers' evaluation system

In relation to the sufficiency of evaluation resources, the schools' leader explained how most resources already existed inside the schools. Additional efforts were provided to support the data collection process. This was derived from the school leadership's belief in the significance of their evaluation policy and its entitlement of adequate time and money allocation.

Area (8): Feedback and professional development opportunities offered through the teachers' evaluation system

Taking in consideration that this area represents a main research question of this study, four questions were addressed to the schools' leader to explore policies and perspectives related to evaluation-based professional development.

The first two questions were related to individual performance feedback provided for teachers and the employment of such feedback in identifying

strengths and weaknesses. The interviewee explained the school's policy used to deal with evaluation results. Before granting the final approval of evaluation results, a copy of the matrix is given to every teacher in order to be aware of his/her grade in each area. They are given the right to express their views and submit any supporting evidence. Necessary alterations are done and then the final evaluation is approved and circulated among teachers. Moreover, the policymaker expressed: *"Teachers are offered the opportunity to identify strong and weak performance areas by knowing the grade allocated for each area of work and by referring to the standards of each point"*.

Other questions were related to the schools' efforts in employing group evaluation results, and to school leadership policies used to link evaluation with training and professional development.

According to the interviewee, the school administration is not only interested in the individual evaluation of teachers; they also aim at knowing which areas achieved high performance and which areas need to be dealt with. The average performance grade of each field of the twelve matrix fields is studied and compared with other fields and with grades from previous years to see whether the school collectively did well enough.

Evaluation-based professional development policies were explained by the schools' leader, where the final matrix results are used in planning for training courses and other professional development plans. Through special- purpose surveys; teachers are encouraged to select certain areas that cover topics revealed by the matrix such as: Building good relationships with students and parents, classroom management, employments of technology in teaching, teachers' social role, and analysis of students' results.

The training courses related to the matrix are well-documented in the school and offered to any interested organization to benefit from.

▪ Discussion and interpretations

Stake (2010) defined three main approaches to interview interpretation:

(1) Obtaining unique data from the interviewee, (2) Collecting aggregated numerical data from multiple interviewees, or (3) Finding out about an item or area only fully observable by the interviewee. The first and third approaches were followed during this interview to fully-investigate the teacher evaluation policy.

The interview answers were found to reflect a high experience level possessed by the schools' leader. They also reflected positive perspectives about the targeted policy, where almost all the eight evaluated areas yielded positive responses.

In the area of evaluation purpose, interpretation of interview answers shows a deep understanding of common teacher evaluation purposes recognized by literature (Hinchey, 2010; Goe *et al.*, 2014). In addition to other unique purposes defined by the school; where teachers' ethics and professional development represent a great priority to the school leadership. Moreover, the school leadership believes in the role of formative (improvement) and summative (accountability) characteristics in forming a meaningful evaluation. However, interview answers did not explain how the setting of these purposes was shared between all school members including teachers; moreover, the answers implied that these purposes were specified only according to school leadership's vision and objectives.

As for the area of basing evaluation on effective teaching standards, the school leadership along with several policymakers contributed to the process. The school used UAE educational standards and internationally recommended standards, in addition to own experience formed over more than thirty years. This was coherent with the suggestions of previous literature (Department of Accreditation, MOE, 2010; Goe, Holdheide & Miller, 2014).

The next area was evaluation systematic communication. The interviewee's statements reveal that communication with school members, including teachers, was mostly evident during the policy design process. However, the continuity of communication through later policy stages was seemingly inconsistent. This judgment requires further investigation through teachers' questionnaire results.

The schools' leader described a friendly evaluation climate where the school facilitated teachers' acceptance of the new system by providing enough explanations and performance feedback. However, he did not explain how a less friendly climate was established with low performance teachers as advised by previous educational experts (Stronge, 2005). Moreover, the overall good acceptance rate judged by the leader's experience requires further validation through teachers' perspectives expressed in the questionnaire.

At least seven examples on performance data sources were listed by the interviewee. This can be related to one of the school's evaluation purposes in using multiple perspectives. Moreover, it applies to literature recommendations (Stronge, 2005; Weisberg *et al.*, 2009; Hinchey, 2010; Glazerman *et al.*, 2011; Dretzke, Sheldon & Lim, 2015). However, the school leader admitted that not all these sources were applicable to all teachers. This indicates that teachers of certain subjects or grades require an adjusted form of the matrix, with perhaps different performance sources or measurements' allocation.

In relation to the qualifications and credibility of evaluators and evaluation measurements; the schools' leader did not demonstrate any selection policy set by the school in hiring supervisors and evaluators; instead, he only referred to the extensive training devoted to existing evaluators about standards and evaluation. The schools' leader highly trusted in the experience of evaluators; however, he monitored their work through the supervision of the quality control department which is directly linked to school leadership. Teachers also participated in evaluating the credibility and reliability of evaluators and

evaluation tools through feedback. On the other hand, the schools' leader was realistic enough to admit that some tools will always be subjected to evaluators' judgments and personalities, and this is an issue that the school attempts to minimize under the supervision of the quality control department.

Information provided during the interview reflected adequate allocation of time and expense resources. This interpretation was concluded based on the intense procedures devoted to evaluation which was tangible throughout the whole interview. Such procedures are time consuming and require a large team of personnel to cover all the targeted teachers.

The reported answers in addition to further discussions during the interview revealed a marked orientation of school leadership towards evaluation-based professional development. Evaluation results were used in individual performance improvement through comprehensive evaluation reports offered to teachers and administrators. Moreover, the school leadership designed policies to interpret group evaluation results and diagnose general weak areas, which were later addressed through training courses that covered all school teachers. To facilitate democratic professional development; teachers were allowed to participate in selecting training topics in relation to their evaluation results. These leadership approaches came in coherence with several literature recommendations (White, Cowhy, Stevens & Sporte, 2012; Al-Bustami, 2014).

4.2.2 Study Phase II: Quantitative questionnaire- Perspectives of evaluated teachers

- **Results and data analysis**

Since all the three schools share the same teacher evaluation system, and the purpose of this phase was to evaluate this system rather than comparing between schools; data analysis was performed collectively for all schools.

The teachers' questionnaires can be divided into **four main areas** with different purposes and answering scales.

(A) Questionnaire analysis: Likert-answering scale questions

Two questionnaire areas included a Likert-answering scale:

- 1- Evaluating the teacher evaluation matrix according to effectiveness criteria.
- 2- Evaluating the use of evaluation data in teachers' professional development.

For areas that follow a Likert-answering scale, the proportional percentage approach was selected for analysis because the researcher is interested in identifying teachers' satisfaction about each character related to the evaluation system; that would be more meaningful than measuring the overall satisfaction of teachers about the evaluation system as a whole using other statistical methods (Bell, 2014). For each statement, percentages of positive answers' frequencies (Strongly agree and Agree), and negative answers' frequencies (Strongly disagree and disagree) were calculated to facilitate comparison.

Data analyses of likert-answering scale areas are demonstrated numerically as **Appendix.7** and **Appendix.8**, and proportional wise in **Tables 4-1** and **Table 4-2**, in addition to a demonstrative graphical data analysis through **Figure 4-1** and **Figure 4-2**.

Table 4-1**Teachers' questionnaire results– Area (1): Evaluation of the targeted teacher evaluation system according to effectiveness criteria (Proportional percentile)**

No.	Area	Statement	+ Response	+ - Neutral	- Response
1.	Evaluation Purpose	School administration adequately informed teachers about the purpose of evaluation	98.9%	1.1%	0%
2.		School administration has a clear, justified purpose for teachers evaluation	92.6%	5.3%	2.1%
3.	Teaching standards	I received enough information about effective teaching standards set by school	94.7%	2.1%	3.2%
4.		The teaching standards set by school are clear and appropriate	90.3%	5.4%	4.3%
5.		I feel that the teacher evaluation matrix is highly related to good teaching standards set by school	71.4%	19.8%	8.8%
6.	Systematic Communication	School administration/Supervisors offered enough explanation to me about the new evaluation system	87.2%	9.6%	3.2%
7.		School administration/Supervisors maintained effective communication with me during evaluation procedures	59.1%	12.9%	28%
8.		I am encouraged to communicate with school administration/supervisors about my performance evaluation	61.5%	18.8%	19.8%
9.	Climate for high-quality evaluation	I feel that teacher evaluation efforts by the school are supportive to teachers	72.3%	14.9%	12.8%
10.		I generally support the teacher evaluation system implemented by the school	83.2%	10.5%	6.3%
11.	Qualified evaluators	My academic supervisor possesses knowledge and experience required for teacher evaluation	86.2%	11.7%	2.1%
12.		My academic supervisor is highly credible as a source of evaluation	85.1%	9.6%	5.3%
13.		My section supervisor possesses knowledge and experience required for teacher evaluation	87.4%	9.6%	3.2%
14.		My section supervisor is highly credible as a source of evaluation	86.3%	9.5%	4.2%

Figure 4-1

Teachers' questionnaire results– Area (1): Evaluation of the targeted teacher evaluation system according to effectiveness criteria (Graphical analysis)

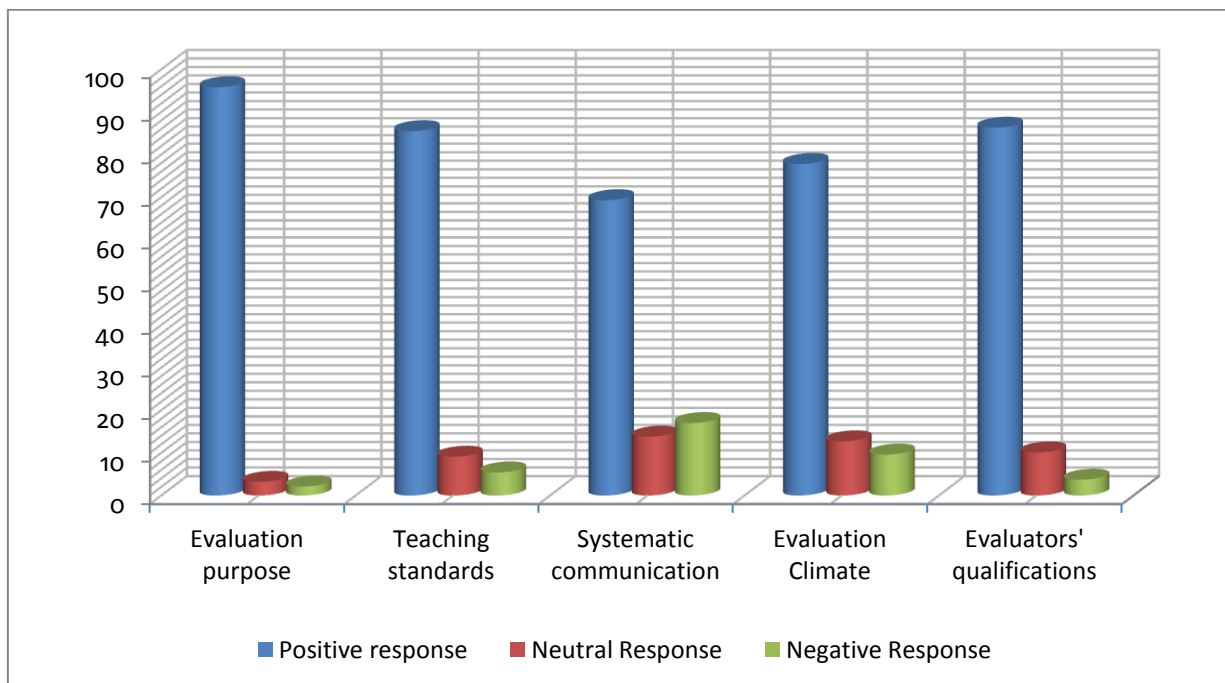


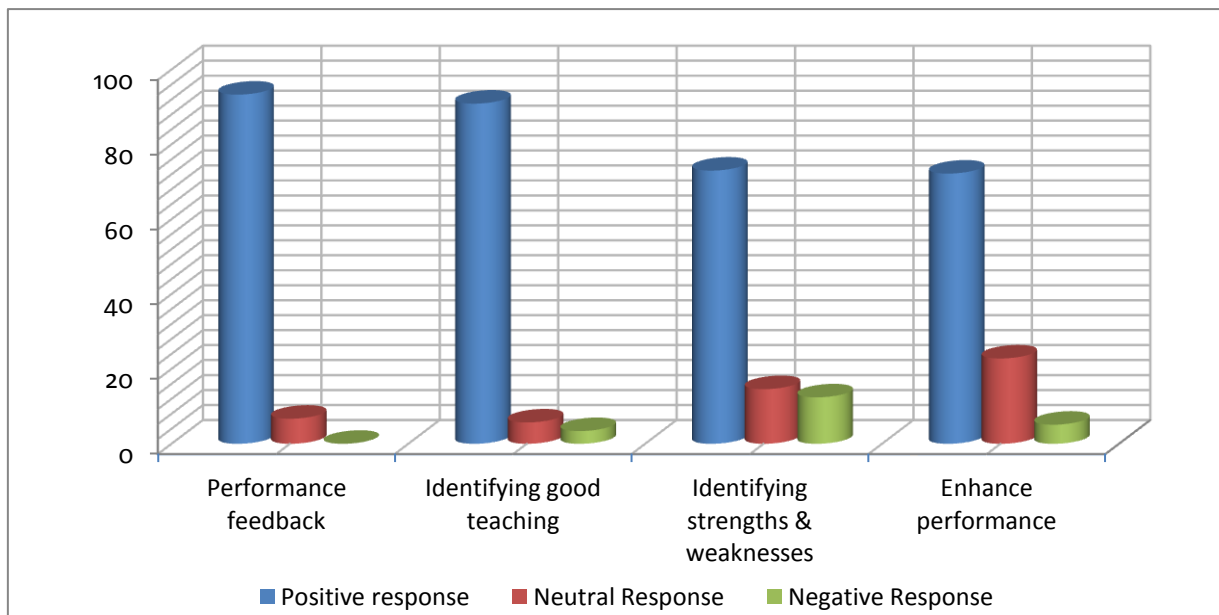
Table 4-2

Teachers' questionnaire results– Area (2): Evaluation of school's use of evaluation data in professional development of teachers (Proportional Percentile)

No.	Area	Statement	+ Response	+ - Neutral	- Response
1.	Performance feedback	I receive information about my performance evaluation results obtained through the teacher evaluation matrix	93.3%	6.7%	0%
2.	Professional development	The teacher evaluation matrix helps me in identifying good teaching criteria	90.9%	5.7%	3.4%
3.		The teacher evaluation matrix helps me in identifying strong and weak areas of my performance	73%	14.6%	12.4%
4.		My administrator/supervisor uses my performance evaluation data to enhance my performance	72.2%	22.8%	5.1%

Figure 4-2

Teachers' questionnaire results– Area (2): Evaluation of school's use of evaluation data in professional development of teachers (Graphical analyses)



(B) Questionnaire analysis: Teacher evaluation tools

Data analysis was performed to display the frequencies of importance ratings for each evaluation tool. The range of importance ratings descended from (1) to (12), with (1) being the most important and (12) the least. In order to compare the importance of each tool, a points' scale was designed to read the results. This scale grants points, from 1 to 12, to each frequency of importance ratings. The level of importance for each evaluation tool was estimated by calculating all the points earned by each evaluation tool as demonstrated in **Tables 4-3 and 4-4**.

The last part of the teachers' questionnaire was also related to teacher evaluation tools used in the targeted matrix. Teachers were asked to evaluate the reliability of each tool by selecting the tools that mostly succeeded in reflecting their true performance, and the tools that failed to do so according to their own evaluation results received post-evaluation. The researcher summarized the frequencies of answers to identify the evaluation tools that teachers thought were most reliable

in measuring actual performance, and the least reliable ones as displayed in **Table 4-5.**

Table 4-3

Teachers' questionnaire results– Area (3): Importance of each teacher evaluation tool used within the matrix according to teachers (Part I)

Evaluation tools	Frequencies of teachers' rankings											
	R. (1)	R. (2)	R. (3)	R. (4)	R. (5)	R. (6)	R. (7)	R. (8)	R. (9)	R. (10)	R. (11)	R. (12)
1. Students Results	23	5	4	7	12	7	13	4	3	6	6	0
2. Lesson planning evaluation	10	10	18	6	11	10	4	5	7	6	2	3
3. Classroom visits	10	12	7	16	12	24	3	10	4	3	1	0
4. Students assignments reports	0	3	4	5	6	8	10	13	11	12	14	4
5. Professional development report	4	0	3	4	7	8	17	15	8	12	8	7
6. Subject coordinator report	7	15	5	20	7	8	9	9	8	1	2	0
7. Section supervisor report	4	6	15	7	21	9	5	3	10	6	2	1
8. Commitment to laws and regulations report	16	17	11	10	1	6	11	5	7	2	2	1
9. Attendance report	9	14	9	9	7	15	5	7	5	6	1	1
10. Resources & Technology report	0	4	6	6	3	3	12	8	15	16	14	4
11. Students surveys	6	2	6	2	0	3	0	3	6	4	21	33
12. Extra-curricular activities Report	0	1	2	1	4	1	3	9	6	12	16	34

* (R.1) to (R.12) stand for the importance rankings of each evaluation tool according to teachers, with (R.1) being the most important and (R.12) the least.

Table 4-4

Teachers' questionnaire results– Area (3): Importance of each teacher evaluation tool used within the matrix according to teachers (Part II)

Evaluation tools	Level of importance
1. Classroom visits	825
2. Commitment to laws and regulations Report.	759
3. Subject coordinator report	729
4. Lesson planning evaluation	724
5. Students Results	719
6. Attendance report	700
7. Section supervisor report	665
8. Professional development report	494
9. Resources & Technology Report	455
10. Students assignments reports	394
11. Students surveys	319
12. Extracurricular activities Report	268

* The levels of importance in Table 4-4 were calculated by analyzing the data of Table 4-3 using the following scale:

Teachers' rating of the tool	Earned Points
(1)	12
(2)	11
(3)	10
(4)	9
(5)	8
(6)	7
(7)	6
(8)	5
(9)	4
(10)	3
(11)	2
(12)	1

Table 4-5

Teachers' questionnaire results– Area (4): Reliability of teacher evaluation tools from the perspectives of teachers

Teacher evaluation tools	Number of teachers who considered the tool reliable to measure actual performance	Number of teachers who considered the tool unreliable and must be excluded from evaluation
1. Subject coordinator report	33	0
2. Section supervisor report	28	1
3. Laws and regulations report	26	0
4. Students Results	25	3
5. Classroom visits	16	0
6. Attendance report	14	1
7. Students surveys	13	16
8. Lesson planning evaluation	12	1
9. Extra-curricular activities Report	11	23
10. Resources & Technology report	8	6
11. Students assignments reports	7	1
12. Professional development report	6	0

▪ **Discussion and interpretations**

One of the main objectives of this research was to evaluate the teacher evaluation matrix in Al-Shola schools. Analysis of the questionnaire results focused on defining strong and weak aspects in the design, implementation and use of the targeted matrix. The teachers' questionnaire included four main areas. Discussion and interpretations of the results will be introduced for each area individually.

Area 1: Evaluation of the targeted teacher evaluation system according to general effectiveness criteria

Participating teachers responded to fourteen questions related to criteria that define effective teacher evaluation systems.

The proportional percentile analysis reveals that the highest proportion of teachers responded positively to criteria related to ***the purpose of evaluation;***

with an average positive response of **95.8%**. Teachers admitted to receiving adequate information about the purpose of teachers' evaluation from their schools. Moreover, they justified the school's administration purpose for performing teacher evaluation.

Second on a raw was the area related to the ***qualifications of evaluators'***. Proportional percentiles for this area's sub-domains were very approximate with an average of **86.3%** positive responses. Teachers considered their supervisors to be highly qualified evaluators; however, an insignificant increase in negative responses is recorded in describing evaluators' credibility.

The teaching standards criteria came third in order of teachers' satisfaction with an average positive response of **85.5%**. A comparison between this area's sub-domains reveals that most teachers agreed on being informed sufficiently about effective teaching standards, and that they consider these standards to be clear and appropriate. However, the level of positive responses declined for the relation between these standards and the current evaluation system (71.4%).

The fourth area was the ***climate provided by school to establish high-quality evaluation***, which attracted an average of **77.8%** positive responses. A considerably high proportion of teachers (83.2%) supported the teacher evaluation matrix in general. The proportion decreased when it came to perceiving schools' evaluation efforts as being supportive. Discussions between the researcher and teachers during the questionnaire phase revealed that they believe their schools should pay more focus to support rather than accountability of teachers through evaluation.

The least satisfactory area of the evaluation matrix according to teachers was reported to be the ***systematic communication***. Although 87.2% of teachers admitted that their schools offered explicit explanation about the teacher evaluation system, a significant portion of teachers thought that this communication was one-directed and inconsistent. A portion of 28% of teachers

thought that their schools failed to communicate with them sufficiently during evaluation procedures. Additionally, 19.8% of teachers felt discouraged to communicate with their school's administration and supervisors regarding their evaluation. Overall, the percentage of positive responses in the communication criteria was the lowest among all other areas with an average of **69.3%**.

In general, positive teachers' perceptions about the effectiveness of the matrix were noticed through the percentile and graphical analysis; with positive responses exceeding 50% in all evaluation criteria.

Area 2: Evaluation of school's use of evaluation data in professional development of teachers

A high positive response (**93.3%**) and a zero negative response were detected in the area of ***receiving performance feedback following evaluation***, which indicates that the school administration was careful about informing teachers about their performance results.

Similar positive responses were noticed for the second domain, where **90.9%** of teachers thought the evaluation matrix helped them ***identify good teaching standards***. These results came in relation with the results of questionnaire area (1), where the information and training provided for teachers about the evaluation matrix acquired them knowledge about effective teaching standards. The detailed outline of teacher evaluation tools contained within the matrix could explain the high awareness level of teachers about several unperceived teachers' roles.

While most teachers admitted that they had received performance feedback through evaluation, fewer teachers could relate this feedback to their own performance diagnosis. ***Identification of good/weak own performance areas*** through evaluation feedback achieved a portion of **73%** positive responses.

Although this percentage is not considerably low, when compared to other criteria; it could be interpreted as a low response. The gap between teachers' agreement on receiving performance feedback and the lower satisfaction about

using that feedback for self diagnosis could be related to teachers' dissatisfaction of aspects related to the teacher evaluation methodology or tools. This outcome could not be fully explained until discussing questionnaire results related to teachers' perspectives on evaluation tools in the following section.

The fourth and final domain of this area was evaluating ***school's efforts in using evaluation data for teachers' professional development***. It was noticeable that the agreement level in this domain is lower than the others within the same area (**72.2%**), which indicates that the targeted schools should reconsider their evaluation-based professional development plans. Interestingly, the neutral responses for this domain were the highest among all questionnaire's domains (22.8%). This reveals that many teachers perceive professional development as a school leadership-level policy which teachers are not fully aware of; therefore, they were incapable of evaluating it positively nor negatively.

Area 3: Importance of each teacher evaluation tool used within the matrix according to teachers

As mentioned earlier, the level of importance was calculated for each teacher evaluation instrument using a points-scale that depends on teachers' ratings of importance. Results show that the most important tool according to teachers and with a major level difference was the ***classrooms visits***. This interestingly reveals teachers' tendencies to believe in the old school of teacher evaluation which depended solely on classroom visits as cited previously in Chapter 2. This implies that the level of teachers' awareness about the benefits of other teacher evaluation tools requires reinforcement.

The second most important evaluation tool was ***commitment to laws and regulations*** of the school. Teachers considered professional commitment to be more important than other roles related to academic responsibilities including, lesson planning and students' results. This finding could be supported by previous studies which linked professional commitment to teachers' attitudes

towards teaching through a significant positive correlation (Kumar, 2011). This indicates that professional and ethical commitment of teachers is a major requirement for the mastery of any academic role.

Analysis of the current research questionnaire show that most teachers considered **students' surveys** and **extra-curricular activities** to be the least important teacher evaluation tools. Post-questionnaire discussions with teachers revealed that most teachers did not consider students qualified enough to evaluate their teachers. This contradicts several literature recommendations that supported students' ratings of teachers as discussed in Chapter 2.

Additionally, many teachers during discussions stated that they consider any evaluation based on participation in extra-curricular activities to be injustice. Their rationale behind that was the inequality of teaching burdens between different subjects, where many subjects (e.g. languages, math, science..) require more extensive academic duties when compared to other subjects (e.g. social studies, physical and arts education..), which in turn limits the time available for extra-curricular activities for heavily-burdened teachers.

The importance of other evaluation tools varies as demonstrated in Table 4-4. Justification and explanation of such variations require additional research to explore teachers' intentions behind their selections, which could include unobservable reasons (e.g. dissatisfaction about the nature of the measured area or about the design of an evaluation instrument or certain evaluators...).

Area 4: Reliability of teacher evaluation tools according to teachers

Thirty three teachers considered **subject coordinators' reports** to be the most reliable evaluation tool. Moreover, twenty eight teachers perceived **section supervisors' reports** as highly reliable. Such results indicate a high-level of trust between teachers and their coordinators and supervisors. This does not come as a surprise when compared with the interview interpretations and the results of

area (1) of the questionnaire, where the majority of teachers agreed that their coordinators and supervisors are highly qualified.

Moreover, a large number of teachers considered ***commitment to laws and regulations'*** reports to be reliable in reflecting actual performance. This was similar to the answers in of questionnaire area (3), where teachers considered these reports to be highly important.

The least number of teachers considered ***professional development reports*** a reliable measurement. This could indicate setbacks in the design or application of this measurement. Further discussions with teachers revealed that many of them questioned the reliability of these reports in their schools, because they are used to evaluate a large number of teachers by a single professional development supervisor; thus, comprehensive evaluation of teachers is very difficult under such heavy burden on the supervisor.

On the other hand, the highest number of teachers considered that ***extra-curricular activities participation reports*** and ***students' surveys*** were the most un-reliable tools and should be excluded from evaluation. This result is compatible with questionnaire area (3) results, where most teachers perceived these two tools to be the least important.

Only six teachers thought that ***resources and technology reports*** were unreliable, and three had similar thoughts about ***students' results***. Almost none of the teachers raised unreliability concerns about other tools. Overall, the evaluation tools used in the matrix were considered as highly-reliable according to teachers, with the exception of extra-curricular activities reports and students' surveys.

This study phase reveals positive perspectives of teachers about most of the evaluation matrix's areas; however, some areas related to evaluation procedures and measurements achieved less favoring responses by teachers. Such areas represent a field for recommendations as will be discussed in Chapter (5).

4.2.3 Study Phase III: Quantitative analysis and correlation coefficient- The relationship between teachers' performance evaluation results and students' results

- **Results and data analysis**

To measure the validity of the targeted evaluation system, students' results were incorporated in this research as a second source of evaluation. Comparison between the three targeted schools was applied using two sources of evaluation:

- 1. The latest teacher evaluation matrix's results:**

The teacher evaluation matrix provides a numeric grade for each teacher out of 1000. To facilitate comparison; these grades were converted into values out of 100 using proportionality. As a final step, the average of teachers' grades of the same subject was calculated. Differentiation between the three targeted schools yielded three averages of teachers' evaluation results for each subject in each school.

- 2. Students' end-year results of the most recent academic year**

According to study sampling, the calculations included results of students from grade six to twelve. The average of students' end-year results (out of 100) at the same academic year for each of subject was calculated. Once again, differentiation between the three targeted schools produced three averages of students' results for each subject in each school.

Calculations described above resulted in a summary of teachers' evaluations resulting from the evaluation matrix, and students' end-year results of the same subjects and grades at the same academic year. Analysis of the collected data is demonstrated numerically in **Table 4-6** and graphically in **Figure 4-3**.

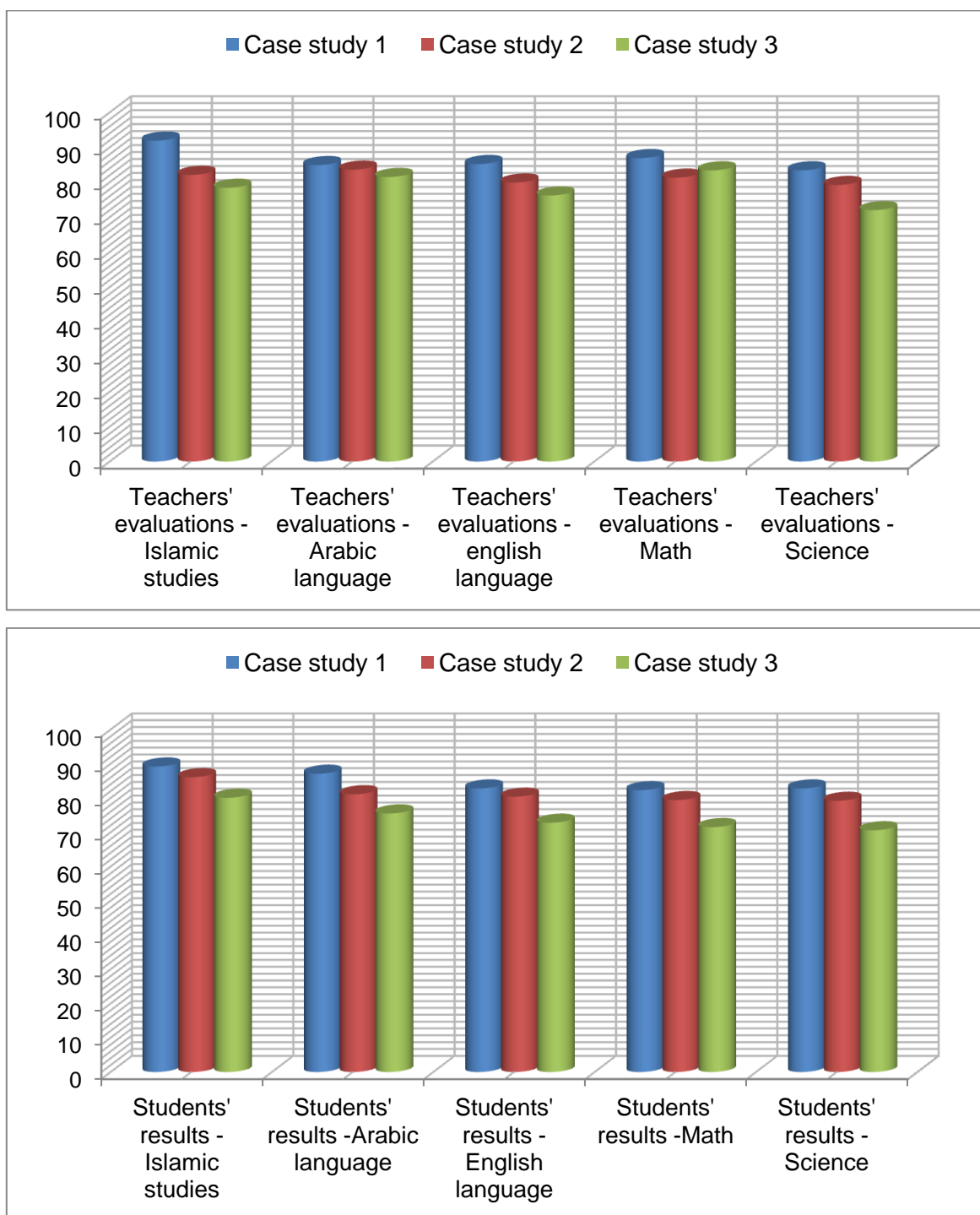
Table 4-6

A comparison sheet between teacher evaluations obtained through the targeted evaluation matrix and the end-results of students taught by the evaluated teachers at the academic year: 2014 / 2015

Subject	Area	Targeted schools		
		Case study1	Case study2	Case study3
		Al-Shola School for girls	Al-Shola School for boys	Manarat Al-Sharjah School
Islamic studies	Average of teachers evaluations through the matrix	91.9	82.1	78.4
	Average of students' end-year results (grade 6 to 12)	89.2	86.1	80.1
Arabic language	Average of teachers evaluations through the matrix	84.9	83.6	81.5
	Average of students' end-year results (grade 6 to 12)	87.1	81.1	75.5
English language	Average of teachers evaluations through the matrix	85.2	79.9	76.2
	Average of students' end-year results (grade 6 to 12)	82.8	80.4	72.8
Math	Average of teachers evaluations through the matrix	86.9	81.3	83.4
	Average of students' end-year results (grade 6 to 12)	82.3	79.4	71.5
Science	Average of teachers evaluations through the matrix	83.4	79.2	72
	Average of students' end-year results (grade 6 to 12)	82.8	79.2	70.6

Figure 4-3

Graphical comparison between teacher evaluations obtained through the targeted matrix and the end-results of students taught by the evaluated teachers at the academic year: 2014 / 2015



To examine the nature and strength of the relationship between teachers' evaluation results and students' results, the Pearson correlation coefficient of the two variables was calculated for a random sample of (40) teachers out of the 123 teachers who were evaluated on the targeted year. For each teacher, the evaluation grade obtained through the matrix was compared with the average end-year score of his/her students.

Statistical calculations and results are described in **Table 4-7** and **Figure 4-4**.

Table 4-7

Pearson correlation coefficient (r) between teacher evaluations obtained through the targeted evaluation matrix and end-results of students taught by the evaluated teachers at the academic year: 2014 / 2015

X		Y				
Teachers Matrix Evaluations	$X - M_x$	Average students' end-year results	$Y - M_y$	$(X - M_x)^2$	$(Y - M_y)^2$	$(X - M_x)(Y - M_y)$
92.1	10.09	89.5	8.65	101.86	74.87	87.33
93.6	11.59	90.5	9.65	134.39	93.17	111.90
90	7.99	91	10.15	63.88	103.07	81.14
86.6	4.59	87.5	6.65	21.09	44.26	30.55
90.1	8.09	89.6	8.75	65.49	76.61	70.83
82.3	0.29	84.5	3.65	0.09	13.34	1.07
89.8	7.79	88.1	7.25	60.72	52.60	56.52
81.7	-0.31	83.4	2.55	0.10	6.52	-0.79
78	-4.01	76.5	-4.35	16.06	18.90	17.42
81.2	-0.81	77.9	-2.95	0.65	8.69	2.38
94.7	12.69	87.7	6.85	161.10	46.96	86.98
85.9	3.89	79.8	-1.05	15.15	1.10	-4.08
76.6	-5.41	74.4	-6.45	29.24	41.57	34.87
85.5	3.49	81.9	1.05	12.20	1.11	3.68
87	4.99	82.5	1.65	24.93	2.73	8.25
85.9	3.89	88.7	7.85	15.15	61.66	30.57
79.6	-2.41	79.4	-1.45	5.80	2.10	3.49
89.6	7.59	86.7	5.85	57.65	34.25	44.44

80.4	-1.61	81.1	0.25	2.58	0.06	-0.41
82.2	0.19	85	4.15	0.04	17.24	0.80
81.6	-0.41	80.5	-0.35	0.17	0.12	0.14
73.5	-8.51	77.9	-2.95	72.38	8.69	25.08
82.7	0.69	81.8	0.95	0.48	0.91	0.66
79.9	-2.11	81.2	0.35	4.44	0.12	-0.74
82.4	0.39	80.9	0.05	0.16	0.003	0.02
75.8	-6.21	78.9	-1.95	38.53	3.80	12.09
84.2	2.19	82.5	1.65	4.81	2.73	3.62
82.4	0.39	80.9	0.05	0.15	0.003	0.02
75.8	-6.21	78.9	-1.95	38.53	3.79	12.09
74.7	-7.31	76.6	-4.25	53.40	18.04	31.04
84.8	2.79	74.5	-6.35	7.80	40.29	-17.73
87.3	5.29	85.9	5.05	28.01	25.53	26.74
69.5	-12.51	71.1	-9.75	156.44	95.01	121.92
87.3	5.29	85.9	5.05	28.01	25.53	26.74
82	-0.01	75.6	-5.25	0.00	27.54	0.04
78.4	-3.61	76	-4.85	13.01	23.50	17.49
64.4	-17.61	65.5	-15.35	310.02	235.55	270.23
67.1	-14.91	71.1	-9.75	222.23	95.01	145.31
78.5	-3.51	73.2	-7.65	12.30	58.48	26.82
75.2	-6.81	69.3	-11.55	46.34	133.35	78.61

Item	Value
N	40
$\sum (X)$	3280.3
$\sum (Y)$	3233.9
M_x	82.01
M_y	80.85
$\sum (X - M_x)^2$	1825.37
$\sum (Y - M_y)^2$	1498.78
$\sum (X - M_x)(Y - M_y)$	1447.11

Keys:

M_x : Mean of X Values.

M_y : Mean of Y Values.

$X - M_x$ & $Y - M_y$: Deviation scores.

$(X - M_x)^2$ & $(Y - M_y)^2$: Deviation Squared,

$(X - M_x)(Y - M_y)$: Product of Deviation Scores

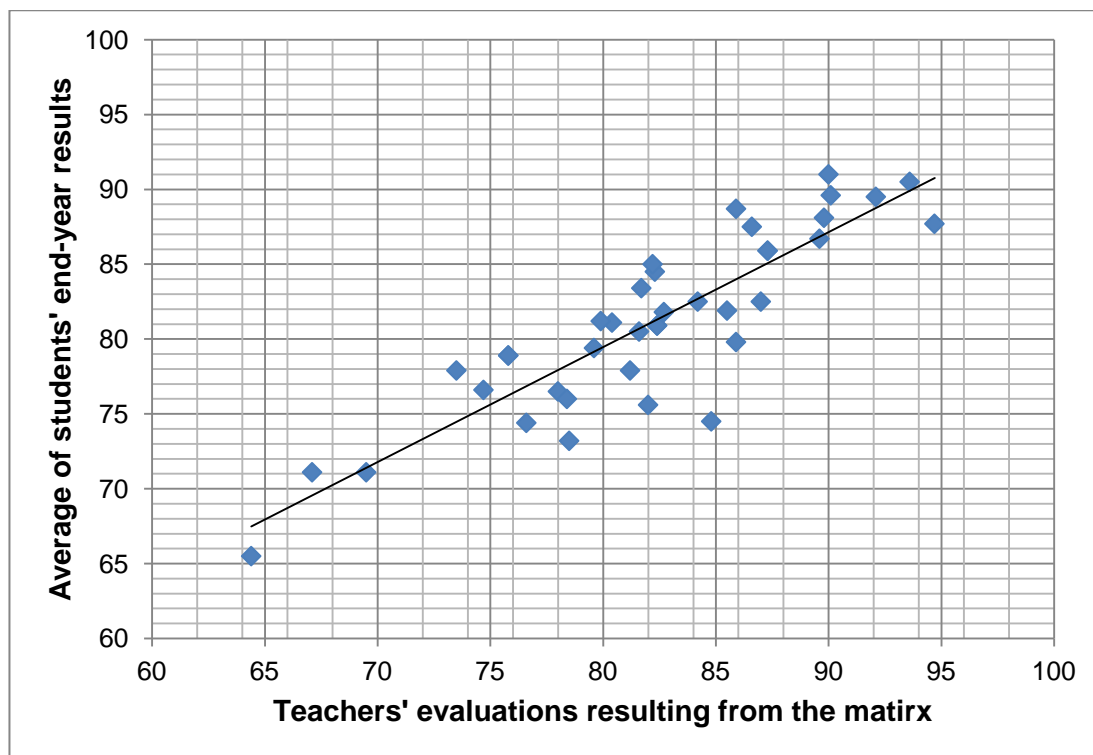
Pearson correlation coefficient (r):

$$r = \frac{\sum ((X - M_y)(Y - M_x))}{\sqrt{((SS_x)(SS_y))}}$$

$$r = 1447.106 / \sqrt{((1825.368)(1498.78))} = \mathbf{0.8749}$$

Figure 4-4

Pearson correlation coefficient chart between teacher evaluations obtained through the targeted matrix and end-results of students taught by the evaluated teachers at the academic year: 2014/2015



▪ **Discussion and interpretations**

The purpose of this research phase and all the comparisons described above was to evaluate the effectiveness and validity of the targeted teacher evaluation system. Students' achievement results were used as an indicator. Comparison results revealed the following:

- In all five subjects, case study (1); Al-Shola school for elementary stages and girls, ranked the highest in teacher evaluations resulting from the teacher evaluation matrix. Similar ranking was evident in the area of students' end-year results. Case study (1) had the highest average of students' results in all five subjects.

- Following was case study (2), Al-Shola school for boys; which ranked second in order when comparing teachers' evaluations, except for Mathematics teachers who ranked third in order with a slight difference when compared with case study (3). In the area of students' results, case study (2) ranked second in all five subjects.
- The last in order was case study (3). In both fields; teachers' evaluations and students' results; it ranked third in place, with an exception of mathematics where it ranked second in teachers' evaluations.

Although differences between comparables are slightly low in value, these differences indicate marked variations considering that the comparables are averages, thus; they reflect high proportions of teachers' evaluations or students' results.

The outcomes show that in almost all subjects there is a positive proportional relationship between teachers' evaluations obtained through the matrix with the end-year results of students taught by those teachers.

Moreover, the value of Pearson correlation coefficient between teachers' matrix evaluations and students' results was: **0.8749**, which indicates a strong linear positive relationship between the two variables (Taylor, 1990).

This relationship indicates that the targeted evaluation matrix has high reliability by measuring what is claimed and fulfilling the purpose of its design and implementation. Connecting teaching quality to students' achievement and vice versa had been concluded numerous through research. Teaching quality was found to be the most influential variable among other variables related to students' outcomes, such as student demographic variables, class size, resources, and teachers' incomes (Darling-Hammond, 2000).

Chapter 5: Summary, Conclusions and recommendations

5.1 Summary

The main objective of this research was to introduce and evaluate a new teacher evaluation system which the researcher had experienced as a part of the leadership team in the targeted organization. Despite the affiliation relationship between the researcher and the targeted system; being part of the policymakers who designed and implemented the evaluation system, she was committed to professional objectivity during this study. Judgments and interpretations were based on data collected through several sources in the school other than the researcher's own professional experience.

A series of procedures were followed to evaluate the targeted evaluation matrix:

- (1) Summarizing effective standards that define teacher evaluation systems through literature review.
- (2) Employing these standards in designing the research tools.
- (3) Accreditation of research tools by the dissertation coordinator.
- (4) Using the research tools to collect data about the targeted policy from different school members.
- (5) Adding the area of students' results as a quantitative approach to validate the results of the study, by comparing teachers' evaluation results obtained through the matrix to their students' results over three case studies.
- (6) Analyzing the data obtained the three research tools (Interview, questionnaire, comparison with students' results).
- (7) Interpreting the data to form a holistic judgment about the effectiveness of Al-Shola schools' teacher evaluation matrix.

5.2 Conclusions

The teacher evaluation matrix discussed in this research represents an innovative initiative by Al-Shola schools in Sharjah- UAE. The practices

described and validated reflect a high-quality teacher evaluation system; however, this system had only recently been developed and requires further research.

5.2.1 School members' perspectives

The evident outcome of this study is the shared agreement between school leadership and teachers on the effectiveness of the evaluation matrix. This agreement does not eliminate the fact that certain setbacks were diagnosed as will be discussed below; however, these setbacks represent minor aspects and are not related to the policy foundation.

The following final evaluation of the teacher evaluation matrix from the perspectives of schools members was concluded:

(1) Purpose of evaluation

School leadership efficiency in this area was reflected through interview and questionnaire responses; where formative and summative purposes were approached. Although teachers did not participate directly in defining these purposes; they were adequately educated about them by their school administration. Moreover, teachers believed that their school administration had appropriate and justified purposes for evaluation. The researcher based judgments over this area according to the collected perspectives, and in relation to the recommendations of Scriven's (1981) and Papay (2012).

(2) Basing evaluation on effective teaching standards

According to the perspectives of both school leadership and teachers, the evaluation system is highly related to effective teaching standards to which teachers were trained for. School leadership used several external and internal sources for such standards. In conclusion, this area can be rated as highly effective (Department of Accreditation, Ministry of education of UAE, 2010; Goe, Holdheide & Miller, 2014). However, further validation for such judgment requires

closer analysis of the detailed standards used in each evaluation tool and comparison of such standards with credible sources.

(3) Evaluation communication

A noticeable gap was noticed between the perspectives of school leadership and teachers regarding this area. The facilitating systematic communication described by the schools' leader was not highly admitted by teachers, where this area ranked the lowest among teachers' questionnaire's results. Communication during evaluation procedures was identified as the missing link.

(4) Evaluation climate

The amiable evaluation climate established by school leadership led to high acceptance of teachers to the evaluation matrix in general (83.2%). This was reported through two research tools; the interview and the questionnaire. Despite this general acceptance some teachers thought that the evaluation system needs to be more supportive to teachers.

(5) Evaluation data sources

The design of the matrix explained in Chapter1 and the interview responses reveal more than seven sources of data that the school uses to reflect teachers' performance. While numerous studies recommended multiple evaluation data sources (Stronge, 2005; Weisberg *et al.*, 2009; Hinchey, 2010; Glazerman *et al.*, 2011; Dretzke, Sheldon & Lim, 2015); no study was found to suggest as many and diverse sources as the ones included in the matrix. However, the school leadership admits that some tools are inapplicable to teachers of certain subjects or grades, thus; the matrix's data sources require adaptation and adjustments in such cases.

(6) Qualifications and credibility of evaluators

The high trust invested by school leadership in their evaluators' qualifications and credibility was validated through teachers' perspectives, where an average of 86.3% teachers agreed that their evaluators are highly qualified and credible.

Evaluators' bias risk admitted by school leadership responses could be minimized through the multi-evaluator scheme described previously, and through the efforts of the quality control department in the school.

(7) Resources employed to facilitate evaluation

Considering that teachers had no information about the resources allocated by school leadership to support evaluation; this area was only evaluated from the perspectives of school leadership; which reflected high understanding of the resources demanded to apply effective teacher evaluation.

However, this was only concluded through qualitative interpretation of the leader's interview responses; therefore, validation of such responses could be provided through the researcher's long professional experience in the targeted organization. The researcher has been working in Al-Shola private schools for more than seven years and clearly perceives the high amount of time and cost devoted by school administration to facilitate teacher evaluation.

(8) Teacher evaluation measurements

Representing the fundamental components of the matrix; teacher evaluation measurements were extensively analyzed during this study. Each measurement was individually evaluated by teachers in terms of importance and reliability. According to the school leadership interview, all evaluation tools are important; however, a closer look at the design of the evaluation matrix shows that the highest grades were allocated for two tools: Students' results and students' surveys. On the other hand; it was interesting to find that teachers considered classroom visits to be the most important evaluation tool. This comes as a surprise taking in account teachers' criticisms against using classroom observations as a comprehensive evaluation tool reported as early as the 1980's and 1990's (Scriven 1981; Brandt, 1996; Weiss & Gary, 1998; Danielson & McGreal, 2000; Little, Goe & Bell, 2009).

In conclusion, the comparison between the perspectives of school leadership and teachers regarding evaluation measurements reveals variant attitudes. While school leadership allocated the highest evaluation grades to the tools that measure teacher's impact on students (students' results and surveys); teachers provided high ratings to students' results and low ratings to students' surveys. Discussions with teachers during the study provided further clarification where many teachers questioned students' reliability as evaluators. Another measurement that attracted different perspectives was the extra-curricular activities report. The school leadership's emphasis on the social and proactive role of teachers in activities was not shared among many teachers. This evaluation area was considered the least important by teachers. Suggestions to address such attitudes and clear this contradiction will be discussed later in recommendations.

(9) Evaluation-based professional development

The school approaches to employ evaluation results in professional development was highly satisfactory from the perspectives of leadership and teachers. The approaches described through the leadership interview reflected high experience and marked adaptation to modern evaluation approaches (Danielson and McGreal, 2000; Milanowski, Kimball & White, 2004; Scherrer, 2009).

5.2.2 Teachers' evaluations-Students' results relationship

The researcher compared teachers' evaluation results obtained through the matrix with students' results to validate the findings. In the comparison that included five different subjects and three case studies, the Pearson correlation coefficient value (>0.8) between the two factors has confirmed the hypothesis of this research which suggested a positive relationship between the teacher evaluation matrix and students' results. This result conforms to several literature findings that related teachers' effectiveness to students' results (Darling-Hammond, 2000; Rockoff, 2004; Kimball, White, Milanowski & Borman, 2004; Kane, Staiger, 2008), thus; leads to a profound conclusion about the validity and reliability of the targeted evaluation matrix.

5.3 Limitations of the study

There were two main limitations in this research. The first one was the difficulty faced in finding related literature in the local region of the middle east and specifically in UAE, where the local research body lacked studies that targeted the area of evaluating teacher evaluation. Moreover, most of the local research found in this area was prepared as thesis and dissertations, with the lack of large scale or official studies.

The other main limitation was the nature of the targeted system itself. The complexity of the evaluation matrix requires further extensive research to examine all the components and deeply analyze each instrument used within the matrix. Such analysis would target each instrument individually to evaluate the appropriateness of standards, grades allocation, and numerical operations used. These procedures in addition to expanding the qualitative interview approach, hence; only one candidate was interviewed during this research as a representative of school leadership, are expected to help in attaining a more holistic and conclusive evaluation of the teacher evaluation matrix and expanding its benefits to the general educational society.

5.4 Recommendations

The findings of the current research suggest recommendations for the targeted organization as well as the general educational field; these can be summarized as the following:

❑ Recommendations for the general educational field:

The quality race in the present time makes teacher evaluation an inevitable requirement of school effectiveness. It is important for school leaderships to recognize that traditional observational-based teacher evaluation can no longer be adopted, and that the modern focus on standardized learning requires compatible standardized teacher evaluation.

According to this study, multi-measurement evaluation systems gained both, school leaders' and teachers' preference. These systems provide opportunities for school administrators to evaluate teachers in all the roles they are assigned to perform, which would in turn facilitate achieving professional development and accountability. As for teachers; they provide justice, minimize evaluation bias risk, and establish a more comprehensive improvement path.

Moreover, when implementing multi-lateral evaluation systems; educational leaders should keep in mind all the diverse roles assigned to teachers, where teachers are expected to become active members of their school community. This requires evaluation systems that not only target teaching areas, but extend to cover other important areas such as: professional commitment, professional development, and participation in extra-curricular activities.

Furthermore, this research recommends the use of numerically-scaled teacher evaluation systems. Although descriptive evaluations can provide teachers with more detailed information about their performance; assigning numerical scales to evaluation tools provides more accurate judgments about overall and specific performance. In an approach to facilitate professional development, the schools targeted by this research incorporate both descriptive and numerical reports in the evaluation process, and each teacher receives a set of detailed reports about his/her performance in each area. In addition to the model discussed in this research; other numerically-based models were successfully implemented including the recent model of New Jersey education department (2015).

❑ Recommendations for Al-Shola educational organization:

- Invest high trust in the current teacher evaluation matrix, as it is:
 - (1) Highly satisfactory to leadership's and teachers' needs.
 - (2) Supported by previous literature approaches.
 - (3) Positively related to students' outcomes.

- Increase teachers' participation in defining evaluation purposes and standards.
- Build up further communication routes with teachers about evaluation, especially during and after the evaluation process.
- Consider designing an adjusted form of the matrix for teachers of certain grades or subjects, where not all evaluation tools are applicable.
- Reconsider the design and application of two evaluation tools used in the matrix: Students' surveys and extra-curricular activities report. If no issues are found to be related to the design or application of such tools, the next step could be to educate teachers more about the importance of areas related to these two tools.

References

- Aaronson, D., Barrow, L. and Sander, W. (2007). Teachers and student achievement in the Chicago public high schools. *Journal of labor Economics*, 25(1), pp. 95-135.
- Al Bustami, G. (2014) "Improving the Teacher's Evaluation Methods and Tools in Abu Dhabi schools - Case Study", Athens: ATINER'S Conference Paper Series, No: SOS2014-1309. [Online]. [Accessed 26 February 2016]. Available at: <http://www.atiner.gr/papers/SOS2014-1309.pdf>
- Anderson, G. (1998). Fundamentals of educational research. 2nd ed. London; New york: Falmer, pp.152-153.
- Bell, J. (2014). Doing Your Research Project: A guide for first-time researchers. London. McGraw-Hill Education (UK)
- Bobbitt, J.F. (1912). The elimination of waste in education. *The Elementary School Teacher*, 12(6), pp.259-271.
- Brandt, R. (1996). On a New Direction for Teacher Evaluation: A Conversation with Tom McGreal. *Educational Leadership*, 53(6), pp.30-33.
- Clark, D. (1993). Teacher evaluation: A review of the literature with implications for educators. (ERIC Document Reproduction Service No. ED 359 174)
- Clotfelter, C.T., Ladd, H.F. and Vigdor, J.L. (2007). *Teacher credentials and student achievement in high school: A cross-subject analysis with student fixed effects* (No. w13617). National Bureau of Economic Research.
- Collins English dictionary. January 2016. Harper Collins Publishers. [Accessed 15 February 2016]. Available at: <http://www.collinsdictionary.com/dictionary/english/teacher-evaluation>
- Creswell, J. W. (2009), Research design: Qualitative, quantitative and mixed methods approaches. 3rd ed. Sage publications India Pvt. Ltd
- Danielson, C. (1996). Enhancing professional practice: A framework for teaching. Alexandria, VA: Association for Supervision and Curriculum Development.

- Danielson, C., & McGreal, T. L. (2000). Teacher evaluation to enhance professional learning. Princeton, NJ: Educational Testing Service.
- Darling-Hammond, L., Wise, A.E. and Pease, S.R. (1983). Teacher evaluation in the organizational context: A review of the literature. *Review of educational research*, 53(3), pp.285-328.
- Darling-Hammond, L. (2000). Teacher quality and student achievement. *Education policy analysis archives*, 8, p.1.
- Darling-Hammond, L. (2010). Evaluating teacher effectiveness: How teacher performance assessments can measure and improve teaching. *Center for American Progress*.
- Darling-Hammond, L., Amrein-Beardsley, A., Haertel, E. and Rothstein, J. (2012). Evaluating teacher evaluation. *The Phi Delta Kappan*, 93(6), pp.8-15.
- Davey, B. (1991). Evaluating Teacher Competence Through the Use of Performance Assessment Tasks: An Overview. *Journal of Personnel Evaluation in Education*. n5, (pp.121-132)
- Department of Accreditation, Ministry of education of the United Arab Emirates, (2010). *School evaluation guide 2010-2011*. Dubai, pp.23-24.
- Dretzke, B.J., Sheldon, T.D. and Lim, A. (2015). What Do K-12 Teachers Think About Including Student Surveys in Their Performance Ratings?. *Mid-Western Educational Researcher*, 27(3), p.185.
- Driscoll, D.L., Appiah-Yeboah, A., Salib, P. and Rupert, D.J. (2007). Merging qualitative and quantitative data in mixed methods research: How to and why not. *Ecological and Environmental Anthropology (University of Georgia)*, p.18.
- Duke, D. and Stiggins, R. (1986). Five keys to growth through teacher evaluation. Portland, OR: Northwest Regional Educational Laboratory.
- Dunkleberger, G. (1982). Classroom Observations--What Should Principals Look For?. *NASSP Bulletin*, 66(458), pp.9-15.
- Ellett, C. D., Teddlie, C. (2003). Teacher Evaluation, Teacher Effectiveness and School Effectiveness: Perspectives from the USA. *Journal of Personnel Evaluation in Education*, 17(1), 101-128.

- Farah, S., Ridge, N. (2009). Challenges to curriculum development in the UAE. Dubai School of Government Policy Brief No. 16.
- Follman, J. (1992). Secondary school students' ratings of teacher effectiveness. *The High School Journal*, 75(3), pp.168-178.
- Fraenkel, J., Wallen, N. and Hyun, H. (1993). How to design and evaluate research in education (Vol. 7). New York: McGraw-Hill.
- Glazerman, S., Goldhaber, D., Loeb, S., Raudenbush, S., Staiger, D. O. and Whitehurst, G. J. (2011). Passing muster: Evaluating teacher evaluation systems. Washington, DC: Brown Center on Education Policy at Brookings. Retrieved from: <http://www.brookings.edu/research/reports/2011/04/26-evaluating-teachers>
- Goe, L., Bell, C. and Little, O. (2008). *Approaches to Evaluating Teacher Effectiveness: A Research Synthesis*. Washington: National Comprehensive Center for Teacher Quality
- Goe, L., Holdheide, L. and Miller, T. (2014). Practical Guide to Designing Comprehensive Teacher Evaluation Systems: A Tool to Assist in the Development of Teacher Evaluation Systems. Center on Great Teachers and Leaders.
- Harris, D. and Sass, T.R. (2006). Value-added models and the measurement of teacher quality. *Unpublished manuscript*. Florida State University.
- Hinchey, P.H. (2010). Getting Teacher Assessment Right: What Policymakers Can Learn from Research. National Education Policy Center.
- Hopkins, W. G. (2000). Quantitative research design. *Journal of Sports sciences* 4(1) [Online]. [Accessed: 4 March 2016]. Available at: <http://gmar.sportsci.org/jour/0001/wghdesign.html>
- Hunter, M. (1983). Mastery teaching. El Segundo, CA: TIP Publications.
- Hussein, A. (2015). The use of triangulation in social sciences research: Can qualitative and quantitative methods be combined?. *Journal of Comparative Social Work*, 4(1).
- Iwanicki, E.F. (1989). 10. Teacher Evaluation for School Improvement. *The new handbook of teacher evaluation: Assessing elementary and secondary school teachers*, p.158. California: Corwin Press

- Kane, T. and Staiger, D. (2008). *Estimating teacher impacts on student achievement: An experimental evaluation* (No. w14607). National Bureau of Economic Research.
- Kimball, S.M., White, B., Milanowski, A.T. and Borman, G. (2004). Examining the relationship between teacher evaluation and student assessment results in Washoe County. *Peabody Journal of Education*, 79(4), pp.54-78.
- Kothari, C.R. (2004). *Research methodology: Methods and techniques*. New Age International.
- Kumar, S. (2011). Relationship between Professional Commitment and Attitude towards Teaching among Secondary School Teachers. *International Research Journal of Management Sociology & Humanities* (Vol.2). [Online]. [Accessed: 25 March 2016]. Available at: http://www.irjmsh.com/Artical_details.aspx?id=666
- Law, P. (2002). No child left behind act of 2001. *Public Law*, 107, p.110.
- Little, O., Goe, L. and Bell, C. (2009). A Practical Guide to Evaluating Teacher Effectiveness. *National Comprehensive Center for Teacher Quality*
- Mielke, P. and Frontier, T. (2012). Keeping improvement in mind. Educational Leadership. [Online]. [Accessed 4 February 2016]. Available at: <http://www.ascd.org/publications/educational-leadership/nov12/vol70/num03/Keeping-Improvement-in-Mind.aspx>
- Milanowski, A., Kimball, S. and White, B. (2004). The relationship between standards-based teacher evaluation scores and student achievement: Replication and extensions at three sites. TC, 4(01). *The American Educational Research Association annual meeting* [Online]. Wisconsin Center for Education Research, University of Wisconsin-Madison. Potsdam. 12-16 April. San Diego, California. [Accessed 25 March 2016]. Available at: http://www.cpre-wisconsin.org/papers/3site_long_TE_SA_AERA04TE.pdf
- Million, S.K. (1987). *Demystifying Teacher Evaluation: The Multiple-Strategies Model Used as an Assessment Device*. Washington, DC: ERIC
- New Jersey Department of Education (2015). *Multiple Measures of Student Achievement and Teacher Practice* [Accessed 25 March 2016]. Retrieved: <http://www.state.nj.us/education/AchieveNJ/intro/1PagerTeachers.pdf>

- New Jersey Department of Education (2015). AchieveNJ: Teacher Evaluation Scoring Guide. [Accessed 25 March 2016]. Retrieved: <http://www.state.nj.us/education/AchieveNJ/resources/TeacherEvaluationScoringGuide.pdf>
- OED Online. January 2016. Oxford University Press. [Accessed 1 April 2016]. Available at: <http://www.oxforddictionaries.com/definition/english/multilateral>
- OED Online. January 2016. Oxford University Press. [Accessed 1 April 2016]. Available at: <http://www.oxforddictionaries.com/definition/english/matrix>
- Papay, J. (2012). Refocusing the debate: Assessing the purposes and tools of teacher evaluation. *Harvard Educational Review*, 82(1), pp.123-141.
- Parfitt, J. (1997). Questionnaire design and sampling. *Methods in human geography: A guide for students doing a research project*, pp.76-109.
- Peterson, K. (2004). Research on school teacher evaluation. *NASSP Bulletin*, 88(639), 60-79.
- Rivkin, Steven G., Eric A. Hanushek, and John Kain. (2005). "Teachers, Schools, and Academic Achievement." *Econometrica*, 73(2): 417–458.
- Rockoff, J.E. (2004). The impact of individual teachers on student achievement: Evidence from panel data. *The American Economic Review*, 94(2), pp.247-252.
- Roulston, K. (2010). *Reflective interviewing: A guide to theory and practice*. Sage.
- Scherrer, R.J. (2009). *Evaluating public school teachers in western Pennsylvania: How and why has it changed since NCLB?*. Doctoral dissertation. University of Pittsburgh.
- Sclan, E. M. (1994). Performance evaluation for experienced teachers: an overview of state policies. Washington, DC: ERIC Clearinghouse on Teaching and Teacher Education. ED 373 054
- Scriven, M. (1981), Summative teacher evaluation, In J. Millman (Ed.), *Handbook of teacher evaluation*, pp.244-271, Beverly Hills, Ca: Sage
- Shinkfield, A. J., & Stufflebeam, D. L. (1995). *Teacher Evaluation: Guide to Effective Practice* (Vol. 41). Springer Science & Business Media.

- Stake, R.E. (2010). *Qualitative research: Studying how things work*. Guilford Press.
- Stronge, J.H. (2005). *Evaluating teaching: A guide to current thinking and best practice*. Corwin Press.
- Stufflebeam, D., (1988). *The personnel evaluation standards*. Report by the joint committee on standards for educational evaluation, Daniel L. Stufflebeam, Committee Chair. Newbury Park, CA: Sage Publications.
- Taylor, R. (1990). Interpretation of the correlation coefficient: a basic review. *Journal of diagnostic medical sonography*, 6(1), pp.35-39.
- Thorne, C. (2015). *An investigation of the professional identity of teacher educators in the UAE*. PhD. Thesis. [Online]. University of Exeter. [Accessed 10 March 2016]. Available at: <https://ore.exeter.ac.uk/repository/bitstream/handle/10871/20366/ThorneC.pdf?sequence=1>
- Tuytens, M. and Devos, G. (2010). The influence of school leadership on teachers' perception of teacher evaluation policy. *Educational Studies*, 36(5), pp.521-536. [Online]. [Accessed 20 March 2016]. Available at: <http://www.informaworld.com/smpp/content~db=all~content=a921983968~frm=titlelink>
- Weisberg, D., Sexton, S., Mulhern, J., & Keeling, D. (2009). The widget effect: Our national failure to acknowledge and act on differences in teacher effectiveness. [Online]. New York: The New Teacher Project. [Accessed 15 March 2016]. Available at: <http://tntp.org/publications/view/evaluation-anddevelopment/the-widget-effect-failure-to-act-on-differences-in-teacher-effectiveness>
- Weiss, E. M. W., & Gary, S. (1998). *New Directions in Teacher Evaluation*. Washington, DC: ERIC Clearinghouse on Teaching and Teacher Education.
- White, B.R., Cowhy, J., Stevens, W.D. and Sporte, S.E. (2012). *Designing and Implementing the Next Generation of Teacher Evaluation Systems: Lessons Learned from Case Studies in Five Illinois Districts*. Research Brief. Consortium on Chicago School Research. Chicago.
- Yin, R. (1989). *Case study research: Design and methods*, revised edition. *Applied Social Research Methods Series*, 5.

Appendices

Appendix.1
The teacher evaluation Matrix of Al-Shola Educational Organization

Total	Social Performance		Professional Performance				Educational Performance						Evaluation Tools Evaluation Areas	No.
	.12 Participation in Activities Report	.11 Students' Surveys	.10 Resources & Technology Report	.9 Attendance Report	8. Commitment to laws and regulations report	.7 Section Supervisor's Report	.6 Coordinator's Report	5. Professional development Reports	.4 Students assignments review reports	.3 Classroom Visits	2. Lesson Planning Evaluation	1. Students' Results		
230	10	20				15	10		10	15		150	Impact on students	1.
180	5	10		30	50	25	15	10	25		10		Professional Commitment	2.
110		15					10		30	40	15		Scientific Adequacy	3.
100	25	60				15							Social Role	4.
80	5		35			10	5	5	10		10		Employing Resources	5.
75		10	10			10	15			30			Classroom Management	6.
70	10	25		10		10	5			10			Personal traits	7.
55			5				5	35			10		Professional Development	8.
50	15	10		10		15							Professional Relationships	9.
50	5						10			5	30		Planning	10.
1000	75	150	50	50	50	100	75	50	75	100	75	150	Total	

Appendix.2

The school leadership interview questions

A Study to evaluate the teacher evaluation matrix of Al-Shola schools in Sharjah – UAE from the perspectives of school members and in relation to students' results

Interview with the leader of the targeted educational organization

Participant's information: Name, Functional Position, Years of experience in the educational field.

Based on your experience in designing and implementing the current teacher evaluation matrix, and derived from your perceptions as the school leader, please answer the following questions:

1. What were your purposes for developing and implementing the current teachers' performance evaluation matrix?
2. To which form of teacher evaluation purposes does you and other school's policymakers pay more focus on: Formative or summative?
3. During the policy design process, was the current teacher evaluation system based on effective teaching standards? And, what was the source of such standards?
4. Did the school administration educate teachers about effective teaching standards? If yes, please explain the methodology of doing that.
5. How would you describe communication between different school members during the implementation of teacher evaluation? Please provide examples of practical communication existing in the school.
6. How does the school promote acceptance of teacher evaluation by members of the teaching body?
7. What kinds of data sources are employed to reflect teachers' performance during the evaluation process?

8. Are all of these sources available for all teachers from different subjects and levels?
9. How do you evaluate the qualifications of evaluators participating in the teacher evaluation matrix?
10. How do school policymakers assure credibility and validity of evaluators' and evaluation tools used in the teacher evaluation matrix? Please support your answer with evidence.
11. Do you believe that the resources provided for teachers' evaluation in your schools in terms of time and expenses are sufficient?
12. Do teachers targeted by the teacher evaluation matrix receive information about their performance evaluation results? If yes, in what manner is that applied?
13. How are the teacher evaluation matrix's results employed in identifying strong and weak performance of areas for each teacher individually?
14. How are the teacher evaluation matrix's results employed in identifying strong and weak performance of areas for all teachers as a whole?
15. Is the teacher evaluation matrix employed in training and professional development practices? If yes, please provide examples to support your answers.

Appendix.3 The teachers' questionnaire

**A Study to evaluate the teacher evaluation matrix of Al-Shola schools in Sharjah –
UAE from the perspectives of school members and in relation to students' results**

Teachers' questionnaire

Teacher's information:

Teacher's name (Optional): Gender: (Male-Female)

Years of experience in the school:

1. General criteria:

**Please express your stand from each of the following statements within
the stated range: (Strongly Agree, Agree, Disagree, Strongly Disagree):**

No.	Area	Statement	Strongly Agree	Agree	Neutral	Disagree	Strongly Disagree
1.	Evaluation purpose	School administration adequately informed teachers about the purpose of evaluation					
2.		School administration has a clear, justified purpose for teachers evaluation					
3.	Teaching standards	I received enough information about effective teaching standards set by school					
4.		Teaching standards set by school are clear and appropriate					
5.		I feel that the teacher evaluation matrix is highly related to good teaching standards set by school					
6.	Systematic Communication	School administration/Supervisors offered enough explanation to me about the new evaluation system					
7.		School administration/Supervisors maintained effective communication with me during evaluation procedures					
8.		I am encouraged to communicate with school administration/supervisors about my performance evaluation					
9.	Climate for high-quality evaluation	I feel that teacher evaluation efforts by the school are supportive to teachers					
10.		I generally support the teacher evaluation system implemented by the school					
11.	Qualified evaluators	My academic supervisor possesses knowledge and experience required for teacher evaluation					
12.		My academic supervisor is highly credible as a source of evaluation					
13.		My section supervisor possesses knowledge and experience required for teacher evaluation					
14.		My section supervisor is highly credible as a source of evaluation					

2. Evaluation tools:

A. Please rank the following teacher evaluation tools used in the Matrix according to its significance to you:

No.	Evaluation tool	Your rank
1.	Student's Results	
2.	Lesson planning evaluation	
3.	Classroom visits	
4.	Students assignment review reports	
5.	Professional development report	
6.	Subject's coordinator report	
7.	Section supervisor report	
8.	Commitment to laws and regulations Report.	
9.	Attendance report	
10.	Resources & Technology Report	
11.	Student's surveys	
12.	Participation in extracurricular activities Report	

B. According to your own evaluation matrix reports, which tools do you believe have mostly succeeded in reflecting your true performance?

.....

C. Do you believe that any of the evaluation tools mentioned above should be excluded from evaluation? If yes, please explain your stand.

.....

3. Feedback and Professional development:

Please express your stand from each of the following statements within the stated range: (Strongly Agree, Agree, Disagree, Strongly Disagree):

No.	Area	Statement	Strongly Agree	Agree	Neutral	Disagree	Strongly Disagree
1.	Performance feedback	I receive information about my performance evaluation results obtained through the teacher evaluation matrix					
2.	Professional development	The teacher evaluation matrix helps me in identifying good teaching criteria					
3.		The teacher evaluation matrix helps me in identifying strong and weak areas of my performance					
4.		My administrator/supervisor uses my performance evaluation data to enhance my performance					

Thank you for your time,

Appendix.4-A The teachers' questionnaire - Translated

(تقييم مصفوفة تقييم أداء المعلمين المطبقة في مدارس مؤسسة الشعلة التربوية وفقاً
لوجهات نظر الأطراف المعنية ومقارنةً بنتائج تحصيل الطلبة)

استبانة المعلمين

بيانات المعلم:

اسم المعلم (اختياري): الجنس: (ذكر - أنثى)
سنوات الخبرة في المدرسة: التخصص:

1. معايير عامة للتقييم

يرجى التعبير عن موقفك من العبارات الواردة في الجدول أدناه باستخدام إحدى الخيارات التالية: (أوافق بشدة، أوافق، محايد، لا أوافق، لا أوافق بشدة)

م	المجال	العبارات	أوافق بشدة	أوافق	محايد	لا أوافق	لا أوافق بشدة
1.	أهداف تقييم المعلمين	قامت إدارة المدرسة بتعريف المعلمين أهداف تقييم أداء المعلمين من خلال الاجتماعات والدورات التدريبية					
2.		اعتقد بأن إدارة المدرسة عليها أن تقوم بتقييم أداء المعلمين لتحقيق من جودة أدائهم، ولديها أهداف مبررة لذلك					
3.	معايير التدريس	تقوم إدارة المدرسة بتعريف معايير التدريس الفعال من خلال الأدلة الإرشادية والدورات التدريبية					
4.	الفعال	معايير التدريس الفعال التي تحددها المدرسة مناسبة وواضحة بالنسبة لي					
5.		أشعر بأن مصفوفة تقييم المعلم مرتبطة بمعايير التدريس الفعال التي حددتها المدرسة					
6.	التواصل بين الأطراف المعنية	قامت إدارة المدرسة بتوفير الشرح الكافي حول مصفوفة تقييم أداء المعلمين					
7.		حافظت إدارة المدرسة على تواصل مناسب معي حول مصفوفة تقييم أداء المعلمين أثناء وبعد تطبيقها					
8.		أشعر بأنه لدي الفرصة للتواصل مع إدارة المدرسة والعاملين فيها حول تقييمي من قبل مصفوفة تقييم أداء المعلمين					
9.	مناخ ملائم لعملية تقييم المعلمين	أشعر بأن جهود المدرسة لتقييم أداء المعلمين توفر دعماً مناسباً للمعلمين لتطوير أدائهم					
10.		بصورة عامة، أنا أؤيد استخدام مصفوفة تقييم أداء المعلمين (تقييم أداء المعلم باستخدام عدة أدوات وتقارير)					
11.	مؤهلات المشاركين في عملية تقييم المعلمين	أعتقد بأن المشرف التربوي للمادة يمتلك المؤهلات والخبرة الكافية لتقييم أداء المعلمين تربوياً					
12.		أعتقد بأن المشرف التربوي للمادة يمتلك المصادقية الكافية كمصدر لتقييم أداء المعلمين تربوياً					
13.		أعتقد بأن المشرف الإداري لقسمي يمتلك المؤهلات والخبرة الكافية لتقييم أداء المعلمين إدارياً					
14.		أعتقد بأن المشرف الإداري لقسمي يمتلك المصادقية الكافية كمصدر لتقييم أداء المعلمين إدارياً					

2. أدوات التقييم :

أ. يرجى ترتيب أدوات المصفوفة المستخدمة في تقييم المعلمين والواردة في الجدول أدناه بحسب أهميتها بالنسبة إليك:

م	أداة التقييم	الترتيب
1.	نتائج الطلاب	
2.	تقييم التحضير اليومي	
3.	الزيارات الصفية	
4.	تقارير تدقيق الأعمال الكتابية للطلاب	
5.	تقارير التنمية المهنية	
6.	تقرير المشرف التربوي	
7.	تقرير المشرف الإداري	
8.	الالتزام بأنظمة وقوانين المدرسة	
9.	الالتزام بالدوام	
10.	توظيف التقنيات (إحصائيات استخدام الداتاشو)	
11.	استبانات الطلاب	
12.	المشاركة في الأنشطة اللاصفية	

ب. بناءً على درجاتك في التقرير الذي استلمته حول مصفوفة تقييم المعلم، ما هي برأيك أكثر الأدوات التي تمكنت من قياس أداءك الحقيقي من بين أدوات المصفوفة المذكورة في الجدول أعلاه؟

.....

ج. هل تعتقد بأن هناك أدوات من بين أدوات قياس أداء المعلم الواردة في الجدول المذكور أعلاه يجب استبعادها من التقييم؟ ولماذا تعتقد ذلك؟

.....

3. التغذية الراجعة والتنمية المهنية:

يرجى التعبير عن موقفك من العبارات الواردة في الجدول أدناه باستخدام إحدى الخيارات التالية: (أوافق بشدة، أوافق، محايد، لا أوافق، لا أوافق بشدة)

م	المجال	العبارات	أوافق بشدة	أوافق	محايد	لا أوافق	لا أوافق بشدة
1.	التغذية الراجعة حول الأداء	تزودني إدارة المدرسة بنتائج في مصفوفة تقييم أداء المعلم					
2.		تساعدني المصفوفة في التعرف على الأدوار والمهام المطلوبة من كل معلم في المدرسة					
3.	التنمية المهنية	تساعدني مصفوفة تقييم أداء المعلم في التعرف على جوانب القوة ومواطن الضعف في أدائي كمعلم					
4.		تستخدم إدارة المدرسة والمشرف التربوي نتائج المصفوفة في تطوير أدائي وفي التنمية المهنية					

شكراً على وقتك،،

Appendix.4-B

The teachers' questionnaire - Translated

(تقييم مصفوفة تقييم أداء المعلمين المطبقة في مدارس مؤسسة الشعلة التربوية وفقاً لوجهات نظر الأطراف المعنية ومقارنةً بنتائج تحصيل الطلبة)

استبانة المعلمين

بيانات المعلم:

اسم المعلم (اختياري): أحمد سليمان الجنس: (ذكر - أنثى)

سنوات الخبرة في المدرسة: 10 سنوات التخصص: الرياضة

1. معايير عامة للتقييم

يرجى التعبير عن موقفك من العبارات الواردة في الجدول أدناه باستخدام إحدى الخيارات التالية: (أوافق بشدة، أوافق، محايد، لا أوافق، لا أوافق بشدة)

م	المجال	العبرة	أوافق بشدة	أوافق	محايد	لا أوافق	لا أوافق بشدة
1.	أهداف تقييم المعلمين	قامت إدارة المدرسة بتعريف المعلمين أهداف تقييم أداء المعلمين من خلال الاجتماعات والدورات التدريبية	✓				
2.		أعتقد بأن إدارة المدرسة عليها أن تقوم بتقييم أداء المعلمين لتحقيق من جودة أدائهم، ولديها أهداف مبررة لذلك	✓				
3.	معايير التدريس الفعال	تقوم إدارة المدرسة بتعريفها بمعايير التدريس الفعال من خلال الأدلة الإرشادية والدورات التدريبية	✓				
4.		معايير التدريس الفعال التي تحددها المدرسة مناسبة وواضحة بالنسبة لي	✓				
5.		أشعر بأن مصفوفة تقييم المعلم مرتبطة بمعايير التدريس الفعال التي حددتها المدرسة		✓			
6.	التواصل بين الأطراف المعنية	قامت إدارة المدرسة بتوفير الشرح الكاف لي حول مصفوفة تقييم أداء المعلمين	✓				
7.		حافظت إدارة المدرسة على تواصل مناسب معي حول مصفوفة تقييم أداء المعلمين أثناء وبعد تطبيقها		✓			
8.		أشعر بأنه لدي الفرصة للتواصل مع إدارة المدرسة والعاملين فيها حول تقييمي من قبل مصفوفة تقييم أداء المعلمين		✓			
9.	مناخ ملائم لعملية تقييم المعلمين	أشعر بأن جهود المدرسة لتقييم أداء المعلمين توفر دعماً مناسباً للمعلمين لتطوير أدائهم	✓				
10.		بصورة عامة، أنا أؤيد استخدام مصفوفة تقييم أداء المعلمين (تقييم أداء المعلم باستخدام عدة أدوات وتقارير)	✓				
11.	مؤهلات المشاركين في عملية تقييم المعلمين	أعتقد بأن المشرف التربوي للمادة يمتلك المؤهلات والخبرة الكافية لتقييم أداء المعلمين تربوياً		✓			
12.		أعتقد بأن المشرف التربوي للمادة يمتلك المصداقية الكافية كمصدر لتقييم أداء المعلمين تربوياً		✓			
13.		أعتقد بأن المشرف الإداري لقسمي يمتلك المؤهلات والخبرة الكافية لتقييم أداء المعلمين إدارياً		✓			
14.		أعتقد بأن المشرف الإداري لقسمي يمتلك المصداقية الكافية كمصدر لتقييم أداء المعلمين إدارياً		✓			

2. أدوات التقييم :

أ. يرجى ترتيب أدوات المصفوفة المستخدمة في تقييم المعلمين والواردة في الجدول أدناه بحسب أهميتها بالنسبة إليك:

م	أداة التقييم	الترتيب
1.	نتائج الطلاب	2
2.	تقييم التحضير اليومي	10
3.	الزيارات الصفية	6
4.	تقارير تدقيق الأعمال الكتابية للطلاب	12
5.	تقارير التنمية المهنية	7
6.	تقرير المشرف التربوي	8
7.	تقرير المشرف الإداري	9
8.	الالتزام بأنظمة وقوانين المدرسة	3
9.	الالتزام بالدوام	4
10.	توظيف التقنيات (إحصائيات استخدام الداتاشو)	11
11.	استبيانات الطلاب	1
12.	المشاركة في الأنشطة اللاصفية	5

ب. بناءً على درجاتك في التقرير الذي استلمته حول مصفوفة تقييم المعلم، ما هي برأيك أكثر الأدوات التي تمكنت من قياس أداك الحقيقي من بين أدوات المصفوفة المذكورة في الجدول أعلاه؟

1. استبيانات الطلاب
2. استبيانات الطلاب
3. استبيانات الطلاب
4. استبيانات الطلاب
5. استبيانات الطلاب
6. استبيانات الطلاب
7. استبيانات الطلاب
8. استبيانات الطلاب
9. استبيانات الطلاب
10. استبيانات الطلاب
11. استبيانات الطلاب
12. استبيانات الطلاب

ج. هل تعتقد بأن هناك أدوات من بين أدوات قياس أداء المعلم الواردة في الجدول المذكور أعلاه يجب استبعادها من التقييم؟ ولماذا تعتقد ذلك؟

1. استبيانات الطلاب
2. استبيانات الطلاب
3. استبيانات الطلاب
4. استبيانات الطلاب
5. استبيانات الطلاب
6. استبيانات الطلاب
7. استبيانات الطلاب
8. استبيانات الطلاب
9. استبيانات الطلاب
10. استبيانات الطلاب
11. استبيانات الطلاب
12. استبيانات الطلاب

3. التغذية الراجعة والتنمية المهنية:

يرجى التعبير عن موقفك من العبارات الواردة في الجدول أدناه باستخدام إحدى الخيارات التالية: (أوافق بشدة، أوافق، محايد، لا أوافق، لا أوافق بشدة)

م	المجال	العبرة	أوافق بشدة	أوافق	محايد	لا أوافق	لا أوافق بشدة
1.	التغذية الراجعة حول الأداء	تزودني إدارة المدرسة بنتائج في مصفوفة تقييم أداء المعلم	✓				
2.		تساعدني المصفوفة في التعرف على الأدوار والمهام المطلوبة من كل معلم في المدرسة		✓			
3.	التنمية المهنية	تساعدني مصفوفة تقييم أداء المعلم في التعرف على جوانب القوة ومواطن الضعف في أدائي كمعلم	✓				
4.		تستخدم إدارة المدرسة والمشرف التربوي نتائج المصفوفة في تطوير أدائي وفي التنمية المهنية		✓			

شكراً على وقتك،،

Appendix.5-A
Ethical forms – Informed consent (School Leader)

Informed Consent

Researcher's name: Maha Ibrahim Barakeh

Research Title: A Study to evaluate the teacher evaluation matrix of Al-Shola schools in Sharjah – UAE from the perspectives of school members and in relation to students' results

This research was conducted in part fulfillment of the requirements for the degree of Masters in Education (Management, Leadership and Policy) from the British University in Dubai

Dear Chief executive officer of Al-Shola Educational Institution,,

I am looking forward for your voluntary participation in a study that I am conducting to evaluate the effectiveness of the Teacher Evaluation Matrix applied in Al-Shola educational institution schools. The main aim of this study is to evaluate two areas related to the system: (1) Design and implementation. (2) Employing evaluation in professional development of teachers.

Through this interview I would like you to share your experience as the key policymaker of the targeted teacher evaluation system. Moreover, I would like to investigate your perceptions and opinions about the evaluation system as the leader of schools that apply the referred to system.

If you accept to be a part of this study, please sign at the designated area at the bottom of this form, taking in consideration the following clarifications:

- The interview would explore your experience and opinions about the evaluation policy in several areas, including:
 - The purpose of the system.
 - Alignment of the system to effective teaching standards.
 - The nature of communication associated with the evaluation system.
 - The school climate established to facilitate teachers' evaluation.
 - Data sources employed to reflect teachers' performance.
 - Qualifications of evaluators and evaluation measurements.
 - Resources employed to support the teachers' evaluation system.
 - Feedback and professional development opportunities offered through the teachers' evaluation system.
- Participation at this study is optional, and you may withdraw from the research at any time.
- Data obtained through this interview will exclusively be employed in scientific research.

By signing this form, I declare my willingness to participate in the mentioned study after reading all the terms above.

Participant's signature

Appendix.5-B

Ethical forms – Informed consent (School Leader)

Informed Consent

Researcher's name: Maha Ibrahim Barakeh

Research Title: (The effectiveness of a new teachers' evaluation policy used in Sharjah private schools: The Multi-lateral Teachers Evaluation Matrix)

This research was conducted in part fulfillment of the requirements for the degree of Masters in Education (Management, Leadership and Policy) from the British University in Dubai

Dear Chief executive officer of Al-Shola Educational Institution,,

I am looking forward for your voluntary participation in a study that I am conducting to evaluate the effectiveness of the Teacher Evaluation Matrix applied in Al-Shola educational institution schools. The main aim of this study is to evaluate two areas related to the system: (1) Design and implementation. (2) Employing evaluation in professional development of teachers. Through this interview I would like you to share your experience as the key policymaker of the targeted teacher evaluation system. Moreover, I would like to investigate your perceptions and opinions about the evaluation system as the leader of schools that apply the referred to system.

If you accept to be a part of this study, please sign at the designated area at the bottom of this form, taking in consideration the following clarifications:

- The interview would explore your experience and opinions about the evaluation policy in several areas, including:
 - The purpose of the system.
 - Alignment of the system to effective teaching standards.
 - The nature of communication associated with the evaluation system.
 - The school climate established to facilitate teachers' evaluation.
 - Data sources employed to reflect teachers' performance.
 - Qualifications of evaluators.
 - Resources employed to support the teachers' evaluation system.
 - Teacher evaluation measurements used in the system.
 - Feedback and professional development opportunities offered through the teachers' evaluation system.
- Participation at this study is optional, and you may withdraw from the research at any time.
- Data obtained through this interview will exclusively be employed in scientific research.

By signing this form, I declare my willingness to participate in the mentioned study after reading all the terms above.

Participant's signature



Appendix.6-A
Ethical forms – Informed consent (Teachers)

Informed Consent

Researcher's name: Maha Ibrahim Barakeh

Research Title: A Study to evaluate the teacher evaluation matrix of Al-Shola schools in Sharjah – UAE from the perspectives of school members and in relation to students' results

This research was conducted in part fulfillment of the requirements for the degree of Masters in Education (Management, Leadership and Policy) from the British University in Dubai

Dear Participant,,

I am looking forward for your voluntary participation in a study that I am conducting to evaluate the effectiveness of the Teacher Evaluation Matrix applied in Al-Shola private schools. The study mainly aims to evaluate two areas related to the evaluation system: (1) Design and implementation. (2) Employing evaluation in professional development of teachers.

Since you have experienced the evaluation process through the targeted evaluation system as a member of the teachers' body of the school, you were selected to share your experience as a sample for this study. If you accept to be a part of this study, please sign at the designated area at the bottom of this form, taking in consideration the following clarifications:

- The enclosed survey aims at exploring your experience and opinions about:
 - The general design of the teacher evaluation system.
 - Teacher evaluation measurements used in the system.
 - Feedback and professional development opportunities offered to you through evaluation.
- Participation at this study is optional, and you may withdraw from the research at any time without loss of benefit or penalty.
- Data obtained through this survey will exclusively be employed in scientific research.
- In order to allow teachers to express their opinions free from any pressure or embarrassment, all the information obtained from the interview will be treated with secrecy, and no personal data is obligated from the teachers in answering the survey questions.
- A demonstration form of the evaluation matrix in addition to a brief summary of each of the evaluation instruments are enclosed with the survey.

By signing this form, I declare my willingness to participate in the mentioned study after reading all the terms above.

Participant's signature

Appendix.6-B

Ethical forms – Informed consent (Teachers)

استمارة موافقة على المشاركة في البحث

اسم الباحث: مها إبراهيم بركة

عنوان البحث: (تقييم مصفوفة تقييم أداء المعلمين المطبقة في مدارس مؤسسة الشعلة التربوية وفقاً لوجهات نظر الأطراف المعنية ومقارنةً بنتائج تحصيل الطلبة)

يتم إجراء هذا البحث استيفاءً لشروط منح درجة الماجستير في القيادة التربوية من الجامعة البريطانية في دبي

عزيزي المشارك،،

أدعوك للمشاركة الطوعية في دراسة إجرائية أقوم بتنفيذها لتقييم مصفوفة تقييم أداء المعلمين المطبق في مدارس مؤسسة الشعلة التربوية. تهدف الدراسة إلى تقييم جانبين متعلقين بمصفوفة تقييم أداء المعلمين:

(1) تصميم وتطبيق المصفوفة.

(2) توظيف مصفوفة التقييم في التنمية المهنية لأداء المعلمين.

انطلاقاً من خوضك لتجربة تقييم الأداء من قبل المصفوفة المستهدفة باعتبارك عضواً من أعضاء الهيئة التدريسية في المدارس المذكورة، فقد تم اختيارك لمشاركة تجربتك كعينة لهذه الدراسة. في حال موافقتك على المشاركة في هذه الدراسة فإنه يرجى التوقيع في المكان المخصص أسفل هذه الاستمارة، آخذاً بعين الاعتبار النقاط التالية:

■ تهدف الاستبانة المرفقة إلى استطلاع آرائك وتجربتك حول:

1. الفكرة والتصميم العام لمصفوفة تقييم المعلمين في مدارس الشعلة التربوية.
 2. أدوات تقييم أداء المعلم المستخدمة في المصفوفة.
 3. التغذية الراجعة وفرص التنمية المهنية التي تقدمها لك مصفوفة تقييم الأداء.
- المشاركة في هذه الدراسة اختيارية، وبإمكانك الانسحاب منها في أي وقت دون عواقب.
- المعلومات الواردة ضمن الاستبانة المرفقة سوف تستخدم لغايات البحث العلمي فقط.
- لإتاحة المجال أمام المعلمين للتعبير عن آرائهم دون ضغوط أو إحراج فإن جميع البيانات الواردة ضمن الاستبانة سوف تعامل بسرية، ولن يطلب من المعلم ذكر بياناته الشخصية أثناء الإجابة على الاستبانة.
- مرفق مع الاستبانة نموذج توضيحي لمصفوفة تقييم المعلم بالإضافة إلى شرح موجز حول أدوات المصفوفة وكيفية احتساب درجاتها.

من خلال توقيعك على هذه الاستمارة أقر بأنني قد قرأت الشروط الواردة أعلاه وأوافق على المشاركة في الدراسة المذكورة.

توقيع المشارك



Appendix.7

Teachers' questionnaire results– Area (1): Evaluation of the targeted teacher evaluation system according to effectiveness criteria (Numerical)

No.	Area	Statement	S. Agree	Agree	Neutral	Disagree	S. Disagree
1.	Evaluation purpose	School administration adequately informed teachers about the purpose of evaluation	51	43	1	0	0
2.		School administration has a clear, justified purpose for teachers evaluation	47	41	5	0	2
3.	Teaching standards	I received enough information about effective teaching standards set by school	44	45	2	3	0
4.		The teaching standards set by school are clear and appropriate	26	58	5	2	2
5.		I feel that the teacher evaluation matrix is highly related to good teaching standards set by school	15	50	18	2	6
6.	Systematic Communication	School administration/Supervisors offered enough explanation to me about the new evaluation system	43	39	9	3	0
7.		School administration/Supervisors maintained effective communication with me during evaluation procedures	16	39	12	21	5
8.		I am encouraged to communicate with school administration/supervisors about my performance evaluation	19	40	18	14	5
9.	Climate for high-quality evaluation	I feel that teacher evaluation efforts by the school are supportive to teachers	28	40	14	8	4
10.		I generally support the teacher evaluation system implemented by the school	38	41	10	1	5
11.	Qualified evaluators	My academic supervisor possesses knowledge and experience required for teacher evaluation	49	32	11	2	0
12.		My academic supervisor is highly credible as a source of evaluation	54	26	9	4	1
13.		My section supervisor possesses knowledge and experience required for teacher evaluation	54	29	9	1	2
14.		My section supervisor is highly credible as a source of evaluation	59	23	9	1	3

Appendix.8

Teachers' questionnaire results– Area (2): Evaluation of school's use of evaluation data in professional development of teachers (Numerical)

No.	Area	Statement	Strongly Agree	Agree	Neutral	Disagree	Strongly Disagree
1.	Performance feedback	I receive information about my performance evaluation results obtained through the teacher evaluation matrix	49	34	6	0	0
2.	Professional development	The teacher evaluation matrix helps me in identifying good teaching criteria	42	38	5	1	2
3.		The teacher evaluation matrix helps me in identifying strong and weak areas of my performance	37	28	13	9	2
4.		My administrator/supervisor uses my performance evaluation data to enhance my performance	11	46	18	2	2