



“It’s quite ruthless here, but it often is the case in Saudi ... it’s the way they operate”,  
(a participant in this study)

**Challenges to Teacher Evaluation at Saudi Universities and Suggestions for  
Solutions: A Perspective of English Language Teachers**

Submitted by

**Khadija Abdullah Ahmed Alamoudi**

To

The University of Exeter as a thesis for the degree of  
Doctor of Philosophy in Education

April 2018

This thesis is available for Library use on the understanding that it is copyright material and that no quotation from the thesis may be published without proper acknowledgement.

I certify that all material in this thesis which is not my own work has been identified and that no material has previously been submitted and approved for the award of a degree by this or any other University.

Signature:.....

## **Abstract**

This thesis explores EFL teachers' views concerning teacher evaluation practices and policies at five public universities in the Kingdom of Saudi Arabia. The research is informed by the interpretive paradigm due to its exploratory nature. With relativism as the ontological stance and constructionism as the epistemological stance, a mixed method sequential design was utilized to collect the required data. Quantitative and qualitative approaches were applied consecutively with more emphasis on the qualitative phase. The study utilised an online questionnaire in the quantitative stage in addition to one-to-one semi-structured interviews and focus group discussions in the qualitative stage. The numbers of participants were 249 for the questionnaire, 21 for the one-to-one interviews and 9 for the focus group discussions. Descriptive statistics were conducted for quantitative strands, and thematic analysis for the qualitative data. Both types of data were analysed with the help of SPSS and NVivo, respectively. The analysis of both sets of data culminated in the emergence of three main themes, 10 categories, and 33 subcategories. The three emergent themes are the importance of EFL teacher evaluation, challenges to EFL teacher evaluation, and suggested solutions for better teacher evaluation. Following the abundance of ideas grounded in the data, a suggested participatory teacher evaluation model informed by EFL teachers' voice is proposed. Based on the findings which provided evidence of challenges facing the current teacher evaluation in the Saudi higher education context, this study concludes by presenting recommendations for policymakers, administrators and teachers. For instance, a comprehensive framework needs to be developed for the professional development of teacher evaluators including initial training and continuous training programmes. Also, the management of higher education institutions needs to add a formative lens to the scheme of teacher evaluation at their workplace to help EFL teachers develop their instructional practice. Finally, suggestions for further research are also mentioned towards the end of the study.

## **Acknowledgment**

First and foremost, all praise and thanks are due to Allah the Almighty for His mercy and blessings on me to accomplish this thesis. I am indebted to my first supervisor, Professor Salah Troudi, for his constant guidance, encouragement and continuous support throughout my studies. As he is an expert in the field of research in TESOL, I always find his constructive critical feedback enlightening and informing. I would also like to express my sincere appreciation to my second supervisor, Dr Esmael Abdollahzadeh, for his support, guidance and insightful feedback. I owe my deepest gratitude to my beloved family. Special thanks go to my mother and father for their encouragement and heartfelt prayers, my husband for his understanding and support, and to my children for their love and priceless smiles during my hardest time. I shall always be indebted to all of them because their presence made it possible for me to carry out my work to a flourishing end.

# Table of Contents

## Table of Contents

TITLE PAGE.....	1
ABSTRACT .....	2
ACKNOWLEDGMENT .....	3
TABLE OF CONTENTS .....	4
<b>CHAPTER ONE: INTRODUCTION TO THE STUDY .....</b>	<b>13</b>
1. 1. INTRODUCTION .....	13
1. 2. BACKGROUND OF THE STUDY .....	13
1. 3. STATEMENT OF THE PROBLEM.....	14
1. 4. RATIONALE OF THE STUDY .....	17
1. 5. SIGNIFICANCE OF THE STUDY .....	19
1. 6. RESEARCH AIM AND OBJECTIVES .....	20
1. 7. RESEARCH QUESTIONS.....	21
1. 8. ORGANIZATION OF THE STUDY.....	21
1. 9. SUMMARY OF THE CHAPTER.....	23
<b>CHAPTER TWO: BACKGROUND OF THE STUDY.....</b>	<b>24</b>
2. 1. INTRODUCTION .....	24
2. 2. OVERVIEW OF THE SAUDI CONTEXT .....	24
2. 3. EDUCATIONAL SYSTEM IN SAUDI ARABIA.....	26
2. 4. ENGLISH LANGUAGE STATUS IN SAUDI ARABIA.....	27
2. 4. 1. <i>English Language in Saudi Higher Education</i> .....	28
2. 4. 2. <i>Preparatory Year English Language Programme</i> .....	29
2. 5. EFL TEACHER EDUCATION IN SAUDI ARABIA .....	31
2. 6. EFL TEACHER AS A FACULTY MEMBER IN SAUDI UNIVERSITIES .....	33
2. 6. 1. <i>Recruitment Standards</i> .....	35
2. 6. 2. <i>Duties and Responsibilities</i> .....	36
2. 6. 3. <i>Promotion Requirements</i> .....	37
2. 7. EFL TEACHER EVALUATION IN SAUDI UNIVERSITIES .....	41

2. 8. SUMMARY .....	44
<b>CHAPTER THREE: REVIEW OF THE LITERATURE .....</b>	<b>45</b>
3. 1. INTRODUCTION .....	45
3. 2. TEACHER EVALUATION .....	45
3. 2. 1. <i>Operational Definitions of “Evaluation” in Education</i> .....	46
3. 2. 2. <i>Introducing the Concept of “Evaluation” in the Field of Education</i> .....	52
3. 3. THE CHANGING LANDSCAPE OF TEACHER EVALUATION.....	53
3. 3. 1. <i>Why Educational Evaluation for Teachers and Faculties?</i> .....	56
3. 3. 1. 1. Formative Evaluation: Evaluation to Improve Performance .....	57
3. 3. 1. 2. Summative Evaluation: Evaluation to Make Decisions .....	61
3. 3. 1. 3. Diagnostic Evaluation: Evaluation to Provide Data to Outsiders.....	64
3. 3. 2. <i>How to Evaluate Teachers?</i> .....	68
3. 3. 2. 1. Student Rating.....	68
3. 3. 2. 2. Classroom Observation.....	71
3. 3. 2. 3. Self-Evaluation .....	73
3. 3. 2. 4. Teaching Portfolio .....	75
3. 3. 3. <i>EFL teachers: Evaluation and Special Concerns</i> .....	78
3. 4. EFL TEACHER EVALUATION: DIFFERENTIATED MODELS.....	84
3. 4. 1. <i>Technocratic Approach to Teacher Evaluation: Accountability and Teacher Evaluation.</i> 84	
3. 4. 2. <i>Democratic Approach to Teacher Evaluation</i> .....	90
3. 5. THE CONCEPTUAL FRAMEWORK OF THE STUDY.....	95
3. 6. SUMMARY OF THE CHAPTER.....	97
<b>CHAPTER FOUR: METHODOLOGY .....</b>	<b>98</b>
4. 1. INTRODUCTION .....	98
4. 2. PHILOSOPHICAL ASSUMPTIONS AND THEORETICAL FRAMEWORK.....	99
4. 2. 1. <i>Ontological Assumptions</i> .....	101
4. 2. 2. <i>Epistemological Assumptions</i> .....	101
4. 2. 3. <i>Methodological Assumptions</i> .....	103
4. 3. RESEARCH DESIGN.....	104

4. 3. 1. <i>Mixed Methods Research</i> .....	104
4. 3. 2. <i>Sequential Mixed Methods Design</i> .....	106
4. 3. 2. 1. Implementation .....	109
4. 3. 2. 2. Priority .....	110
4. 3. 2. 3. Integration of the Data .....	111
4. 3. 3. <i>Participants and Sampling Procedures</i> .....	112
4. 3. 3. 1. An Overview of the Participants Involved in the Questionnaire.....	118
4. 4. DATA COLLECTION AND ANALYSIS .....	119
4. 4. 1. <i>First Stage (Quantitative): Self-Completion Online Questionnaire</i> .....	120
4. 4. 1. 1. Developing the Questionnaire .....	122
4. 4. 1. 2. Examining Validity of the Questionnaire.....	125
4. 4. 1. 3. Piloting the Questionnaire .....	127
4. 4. 1. 4. Reliability of the Questionnaire.....	129
4. 4. 1. 5. Analysis of the Questionnaire Data .....	131
4. 4. 2. <i>Second Stage (Qualitative): Semi-Structured Interviews and Focus Group Discussions</i> .....	131
4. 4. 2. 1. Semi-Structured Interview Guide .....	134
4. 4. 2. 2. Piloting the Semi-Structured Interview .....	135
4. 4. 2. 3. Focus Group Discussion .....	136
4. 4. 2. 4. Establishing Dependability of the Qualitative Phase .....	138
4. 4. 2. 5. Establishing Credibility of the qualitative Phase.....	139
4. 4. 2. 6. Data Analysis .....	140
4. 5. RESEARCH ETHICS .....	142
4. 5. 1. <i>Gaining Access</i> .....	143
4. 5. 2. <i>Participants' Related Issues</i> .....	144
4. 5. 3. <i>Data Protection</i> .....	145
4. 6. CHALLENGES AND LIMITATIONS .....	145
4. 7. SUMMARY OF THE CHAPTER .....	147
<b>CHAPTER FIVE: FINDINGS AND ANALYSIS</b> .....	148
5. 1. INTRODUCTION .....	148

Table 5. 1. Themes, categories, and sub-categories .....	149
Table 5. 2. Sources and frequency of the three major themes.....	151
5. 2. IMPORTANCE OF EFL TEACHER EVALUATION IN SAUDI HIGHER EDUCATION .....	151
5. 2. 1. <i>Providing an Authoritative Body</i> .....	152
5. 2. 1. 1. Employment Decisions .....	154
5. 2. 1. 2. Contract Renewal and Salary Increment.....	155
5. 2. 2. <i>Improving the Teacher’s Performance</i> .....	158
5. 2. 2. 1. Reflection on Teaching Practices .....	160
5. 2. 2. 2. Show Differences between Teachers.....	161
5. 2. 2. 3. Measure Students’ Satisfaction .....	163
5. 2. 3. <i>Not Important and “Should Go Completely”</i> .....	164
5. 2. 4. <i>Effects of Teacher Evaluation?</i> .....	167
5. 2. 4. 1. Instructional Practices .....	167
5. 2. 4. 2. Professional Development .....	170
5. 2. 4. 3. Different Forms of Recognition and Rewards .....	172
5. 2. 4. 4. Teachers’ Self-Esteem .....	176
5. 2. 4. 5. Sense of Belonging and Job Security .....	177
5. 3. CHALLENGES TO EFL TEACHER EVALUATION AT SAUDI UNIVERSITIES .....	180
5. 3. 1. <i>“The Biggest Challenge is Classroom Observation”</i> .....	181
5. 3. 1. 1. Large Student Numbers .....	183
5. 3. 1. 2. Time Can Be a Challenge .....	186
5. 3. 2. <i>Are our Evaluators Capable of Judging Us?</i> .....	187
5. 3. 2. 1. Required Qualifications and Skills.....	189
5. 3. 2. 2. Personality and Exposure to Different Environments .....	191
5. 3. 2. 3. Associated Issues of Power .....	193
5. 3. 2. 4. Feedback Concerns .....	196
5. 3. 3. <i>Students Can Be a Challenge Too</i> .....	198
5. 3. 3. 1. Ability to Evaluate EFL Teachers .....	200
5. 3. 3. 2. “The Student Has a Punitive Ability to Be Honest or Not” .....	202

5. 3. 4. <i>EFL Teacher-Related Difficulties</i> .....	205
5. 3. 4. 1. "It is Quite Ruthless Here": Evaluation and Psychological Stress .....	206
5. 3. 4. 2. Fairness and Equality .....	209
5. 3. 4. 3. Teacher-Related Procedural Drawbacks .....	212
5. 3. 4. 4. Teachers' Preparation and Readiness .....	215
5. 3. 4. 5. Roles and Voice of Teachers .....	218
5. 3. 5. <i>Current Teacher Evaluation Systems</i> .....	225
5. 3. 5. 1. Transparency and Confidentiality .....	226
5. 3. 5. 2. Evaluation Criteria .....	229
5. 3. 5. 3. Modular System .....	231
5. 4. SUGGESTED SOLUTIONS FOR BETTER EFL TEACHER EVALUATION.....	234
5. 4. 1. <i>Micro-Level Suggested Solutions</i> .....	235
5. 4. 1. 1. Introduce Teacher Evaluation .....	235
5. 4. 1. 2. Involve Teachers, Students and Build Trust.....	237
5. 4. 1. 3. Appoint Qualified Evaluators .....	239
5. 4. 1. 4. Advance the Methods, Improve the Criteria and Scoring Techniques .....	241
5. 4. 1. 5. Fix the Existing Problems and Plan well .....	244
5. 4. 1. 6. Introduce a Different Evaluation System .....	245
5. 4. 2. <i>Macro-Level Suggested Solutions</i> .....	246
5. 4. 2. 1. Invest in Teachers .....	247
5. 4. 2. 2. Collaborate with Other Departments and Universities .....	250
5. 5. SUMMARY OF THE CHAPTER.....	253
<b>CHAPTER SIX: DISCUSSION OF THE FINDINGS</b> .....	254
6. 1. INTRODUCTION .....	254
6. 2. EFL TEACHER EVALUATION: VALUES VS DEMANDS .....	254
6. 2. 1. <i>The Demands of Accountability and Quality Assurance</i> .....	255
6. 2. 2. <i>The Values of Improving Instructional Practices</i> .....	261
6. 3. WHAT MAKES A QUALIFIED EVALUATOR? .....	266
6. 4. EFL TEACHERS' AVERSION TO TEACHER EVALUATION.....	272



6. 4. 1. <i>Teacher Evaluation: Stress and Threat Aspects</i> .....	272
6. 4. 2. <i>Alienated EFL Teachers</i> .....	277
6. 5. STUDENTS AND THE INTELLECTUAL LEAP BETWEEN GENERAL AND HIGHER EDUCATION.....	283
6. 6. SUMMARY OF THE CHAPTER .....	288
<b>CHAPTER SEVEN: CONCLUSION, RECOMMENDATIONS, AND CONTRIBUTIONS</b> .....	290
7. 1. INTRODUCTION .....	290
7. 2. THEORETICAL AND PEDAGOGICAL RECOMMENDATIONS FOR STAKEHOLDERS .....	290
7. 2. 1. <i>General Recommendations: Relationship between stakeholders</i> .....	291
7. 2. 2. <i>Recommendations for Policymakers</i> .....	293
7. 2. 2. 1. The Ministry Deputy of Planning, and Development .....	293
7. 2. 2. 2. The Ministry Deputy of Curricula and Educational Programs.....	295
7. 2. 2. 3. The NCAAA.....	296
7. 2. 3. <i>Recommendations for Administrators</i> .....	296
7. 2. 3. 1. University/Institution Management Teams .....	297
7. 2. 3. 2. Evaluators .....	300
7. 2. 4. <i>Recommendations for Teachers</i> .....	301
7. 3. CONTEXTUAL, METHODOLOGICAL AND THEORETICAL CONTRIBUTIONS.....	303
7. 3. 1. <i>Contributions to the Saudi Context</i> .....	303
7. 3. 2. <i>Methodological Contributions</i> .....	304
7. 3. 3. <i>Theoretical Contributions: Participatory Teacher Evaluation Model</i> .....	305
7. 3. 3. 1. The Model: A General Idea .....	307
7. 3. 3. 2. The Model: Who and What? Stakeholders Involved and Actions Required.....	308
7. 3. 3. 3. The Model: Stages of a Participatory Development Project.....	311
7. 3. 3. 4. The Model: Concerns and Complexities .....	312
7. 4. SUGGESTIONS FOR FURTHER RESEARCH.....	313
7. 5. MY PHD JOURNEY: REFLECTIONS ON AN ACADEMIC AND SOCIO-CULTURAL ADVENTURE.....	315
7. 6. FINAL THOUGHTS.....	317
<b>Appendices</b>	
Appendix 1: The Questionnaire of the Study.....	319
Appendix 2: One-to-One Interview Schedule .....	325

Appendix 3: Focus Group Discussion Schedule .....	327
Appendix 4: Ethical Approval Form .....	329
Appendix 5: Research Information and Consent Sheet .....	333
Appendix 6: Data Analysis / Questionnaire: Categorical Variables.....	335
Appendix 7: Data Analysis / Questionnaire: Continuous Variables.....	341
Appendix 8: Data Analysis / Questionnaire: Frequencies and Frequency Tables .....	348
Appendix 9: Data Analysis / One-to-One Interview: Transcription Sample .....	370
Appendix 10: Data Analysis / Qualitative Thematic Analysis: Nvivo Sample Nodes.....	385
Appendix 11: Data Analysis / Qualitative Thematic Analysis: Nvivo Interview Coding Sample .....	386
Appendix 12: Faculty Recruitment Standards .....	387
Appendix 13: Teacher Evaluation Form Sample/ Other Qualifications Holders.....	391
Appendix 14: Teacher Evaluation Form Sample/ Doctorate Holders.....	397
<b>List of References</b> .....	<b>410</b>

## List of Tables

Table 4. 1. Maximum variation of the qualitative stage sample .....	116
Table 4. 2. Number of participants in each stage .....	117
Table 4. 3. Nationalities of the participants .....	118
Table 4. 4. Length of the participants' contracts .....	119
Table 4. 5. Internal consistency of questionnaire items/1 .....	130
Table 4. 6. Internal consistency of questionnaire items/2 .....	131
Table 5. 1. Themes, categories, and sub-categories .....	149
Table 5. 2. Sources and frequency of the three major themes .....	151

## List of Figures

Figure 3. 1. The process of evaluation, Genesee, 2001 .....	48
Figure 3. 2. Evaluation theory tree, Christie & Alkin, 2013.....	51
Figure 3. 3. Foreign language teacher evaluation, Brown and Crumpler's Model, 2013 .....	80
Figure 3. 4. English language teacher evaluation, Akbari and Yazdanmehr, 2012 .....	81
Figure 3. 5. Ohio Teacher Evaluation System (OTES) .....	86
Figure 3. 6. Lee (2016) of Hanberger's democratic evaluation orientations .....	92
Figure 4. 1. The sequential mixed-methods design of the study .....	108
Figure 4. 2. Implementation and integration of the qualitative and quantitative strands.....	110
Figure 4. 3. Sequence and prioritization of the two strands .....	111
Figure 5. 1. First theme and categories .....	152
Figure 5. 2. First category and sub-categories, theme 1 .....	154
Figure 5. 3. Second category and sub-categories, theme 1 .....	159
Figure 5. 4. Fourth category and sub-categories, theme 1 .....	167
Figure 5. 5. Second theme and categories .....	181
Figure 5. 6. First category and sub-categories, theme 2 .....	183
Figure 5. 7. Second category and sub-categories, theme 2 .....	188
Figure 5. 8. Third category and sub-categories, theme 2 .....	199
Figure 5. 9. Fourth category and sub-categories, theme 2 .....	206
Figure 5. 10. Fifth category and sub-categories, theme 2 .....	226
Figure 5. 11. Third theme and categories .....	235
Figure 5. 12. First category and sub-categories, theme 3 .....	235
Figure 5. 13. Second category and sub-categories, theme 3 .....	247
Figure 7. 1. Macro- and micro- level stakeholders .....	290
Figure 7. 2. Relationship amongst stakeholders .....	292
Figure 7. 3. Proposed participatory teacher evaluation model .....	306

## List of Acronyms

<b>EFL</b>	English as a Foreign Language
<b>TESOL</b>	Teaching English to Speakers of Other Languages
<b>ELT</b>	English Language Teaching
<b>TE</b>	Teacher Evaluation
<b>MOE</b>	Ministry of Education
<b>NCAAA</b>	National Commission for Academic Accreditation and Assessment
<b>NTP</b>	National Transformation Programme
<b>EQAR</b>	European Quality Assurance Register for Higher Education
<b>CELTA</b>	Certificate in English Language Teaching to Adults
<b>DELTA</b>	Diploma in English Language Teaching to Adults
<b>CPD</b>	Continuous Professional Development

# **Chapter One**

## **Introduction to the Study**

### **1. 1. Introduction**

This chapter presents an introduction to this research. It commences with the background of the study where I introduce the topic. The statement of the problem and the rationale of the study then follow to provide the readers with my standpoint as a researcher. The following section highlights the importance of the study to the Saudi context and also to the existing literature. In the two sections that follow, the objectives of the research are identified and the research questions are stated. The chapter ends with a general overview of the chapters, which provides the readers with an idea of the organization of the thesis.

### **1. 2. Background of the Study**

In schools, universities and educational organizations, teachers are known to be at the centre of any educational enterprise (McCaffrey, Lockwood, Koretz, & Hamilton, 2003). Regardless of the educational programme, how strong it could be or how well it is planned, it can only be as effective as those people who run it. It might be true that effective teachers are usually encouraging and inspiring for their students; they know well how to stimulate their pupils' ideas, critical thinking about the subject matter, acceptance of challenging tasks, and even the choice of future careers. Therefore, the overall improvement-restructuring work in the field of education is continuously related to teachers and teachers' development and growth (Stronge & Xu, 2017), and this is contingent on the evaluation of teachers as its starting point.

Due to its significance, teacher evaluation attracts considerable attention from researchers and developers in the field of education. Its importance may rely on the idea that supporting innovation in the evaluation of teachers has the potential to touch upon everything else related to education. For instance, students' success in school may be highly related to the quality of their teachers (Danielson, 2011). Furthermore, according to Darling-Hammond (2000), the successful implementation of a curriculum along with the use of different teaching strategies that adequately serve students' learning needs are attributed to those teachers who are recognized as effective educators. Accordingly, educational institutions should ensure that they appoint efficacious staff members. It is worth mentioning that teacher evaluation can be considered as one method to ensure and maintain quality teachers in classrooms.

During the last few decades, research on teacher evaluation has received particular attention in education and social sciences studies, with a substantial number of articles in the literature in the area of teacher evaluation (see Bailey, 2016; Danielson & McGreal, 2000; Downing, 2016; Gullatt & Ballard, 1998; Gutmann, 2014; Kyriakides & Demetriou, 2007; Nagel, 2012; Peterson, 2000). However, Tovar (2011) claims that sufficient attention has not been given to the evaluation of English as a Foreign Language (EFL) teachers. With this in mind, it seems that those conducting research on the evaluation of teachers need to pay close attention to EFL teachers' evaluation and to explore and suggest more subject-oriented practices.

### **1. 3. Statement of the Problem**

The medium of instruction in all levels of education up to the tertiary level is Arabic, the official language in Saudi Arabia, and English is not widely used either by Saudi

people in their daily lives or in schools. Individuals are generally introduced to the language at school as a compulsory subject and treat it as a requirement to pass an exam, with the result that students rarely show interest in becoming competent users of English. Despite the fact that the Kingdom of Saudi Arabia has rapidly progressed in the field of education, the outcomes of teaching English as a foreign language are generally considered unsatisfactory (Rahman & Alhaisoni, 2013). In the same vein, Al-Shumaimeri (2003) earlier argues that students graduating from Saudi secondary schools are hardly able to carry out even a short conversation in good English. Thus, when these students enter university and have to study English as a compulsory subject, teaching the language as such may be challenging for the teachers.

Besides the lack of exposure to English, the students' attitude towards English language and their learning aptitude, further challenges in teaching English in the Arab world may be related to the teachers themselves. Inappropriate teaching methodologies, inadequate teachers' training, and the use of teacher-centred rather than student-centred activities have been noted as significant challenges to English language teaching in the Arab context in general (Fareh, 2010). From a personal perspective, I believe that the Saudi context is not an exception. Although being certified or holding credentials might make individuals eligible for jobs in the field of foreign language education, this may not be enough to ultimately make a good teacher. English language teacher preparation programmes are designed to help in training those teachers to become well qualified and effective. However, what has been done by the Ministry of Education to develop the curricula of English language in Saudi Arabia since 1991 may seem far superior to what has been done to improve

the programmes of English language teacher education (Al-Hazmi, 2003). Improvement not only requires efficient training for teachers; it entails effective and reflective supervision as well as a number of well-developed and planned professional development programmes (Habbash, 2011), which seem to be lacking in Saudi universities.

Arguably, teacher evaluation can be seen not only as an essential responsibility of leaders and supervisors in educational institutions but also as a significant means of inquiry to inform the needed professional development programmes that are intended to develop instructional practices and lead to a better language teaching in Saudi Arabia. Based on this assumption, every effort should be made to make the process of teacher evaluation successful and fruitful. Accordingly, EFL teachers' perceptions and ideas about teacher evaluation would appear to be essential in leading to more successful, effective, productive, and practical teacher evaluation. This is currently a problem in Saudi Arabia, given that policymakers and researchers interested in EFL teacher evaluation in the Saudi context do not pay sufficient attention to the teachers' views (Shah, 2016). It is worth noting that in the context of this research listening to the teachers' voice is not only about helping researchers and policymakers to make decisions for improvements. Listening to them further helps in understanding how their experiences in this specific context have shaped their opinions and informed the suggestions they propose.

Troudi (2005) in line with the above argument affirms that "if any element is to be the core of a teacher education programme it should be the teacher's view(s) of what



language education is about and what he/she considers teaching to be” (p. 118). In other words, teachers’ views, which tend to generally be neglected, are an essential element in teacher evaluation programmes in a particular context. Being involved in EFL teacher evaluation in a Saudi university, I noticed that teachers were rarely asked about their views on the practices of teacher evaluation. Hence, I think involving them would enhance solutions that are tailored to many context-specific situations. I also found that teacher evaluation was not always welcomed by some teachers and its system was characterised by ambiguity and confusion. Thus, I believe that teachers’ perceptions need to be investigated, analysed, and critically discussed in order to have a better understanding of the current evaluation system, which, in turn, should help towards having better English language teaching in Saudi universities,

#### **1. 4. Rationale of the Study**

This study explores EFL teachers’ views with regard to teacher evaluation, its importance as it is currently practiced in Saudi universities, challenges to the success of the evaluation of EFL teachers, and suggestions to improve its practices and policies. It also attempts to provide a base for further investigations in relation to the evaluation of teachers in the TESOL context and within the field of critical applied linguistics. Studies of this type that suggest possible alternatives and develop practices assessment will ultimately “inform educational policies to improve approaches to learning” (Klinger, DeLuca, & Merchant, 2016, p. 145).

Furthermore, the current study aims at reviewing the current status of EFL teacher evaluation at Saudi universities to identify positive and negative views of the teachers. This will help researchers and policymakers to understand the effects teacher

evaluation has on teachers and its general impact on the processes of learning. I hope to build a basis for better future practices and for correcting current mistakes by problematizing the given situation and critically questioning the taken-for-granted assumptions (Peeters & Robinson, 2015) and notions of EFL teacher evaluation.

It may seem unwise to call for a change in or termination of the current practices and/or policies of teacher evaluation in Saudi higher education. What seems to be needed is to involve teachers and policymakers in adopting a critical lens as one form of reflective critique (Percy, 2015, p. 881) which can help in making sense of the current practices of teacher evaluation. Thus, it is essential to note that this study is not against the evaluation of EFL teachers and does not discuss teacher evaluation as an obstacle towards the progress of EFL teachers. Rather, it discusses how those teachers perceive the current practices of teacher evaluation and challenges to the successes of the system of teacher evaluation, and seeks suggestions from the teachers for solutions.

This research simply attempts to provide a better understanding of the situation by identifying critical issues related to the quality of the current teacher evaluation system in Saudi universities. Critical issues such as evaluators' qualifications, teachers' involvement, teachers' rights and job security will be raised and discussed. Moreover, I hope to prompt different stakeholders to engage in debates about the possible consequences of the current teacher evaluation system and search for feasible solutions to any challenges that are pinpointed. I believe that one important step can be to challenge people's ideologies and raise their awareness.

It is worth mentioning that actions derived from endeavors to change educational or social practices based on other societies' experiences indicate that following the globalization movement blindly, such as applying readymade forms or models of teacher evaluation, is likely to be the least effective approach due to the "cultural resistance in a globalized world" (Fominaya, 2014, p. 81). Thus, I believe that this research can be seen as one of early introduction to the concept of a critical approach to research in EFL teacher evaluation at a higher education level in Saudi Arabia in that it is tailored to the context of the study. This study focuses solely on teachers' viewpoints as they are at the heart of teacher evaluation where they are in direct contact with evaluators and affected by the practices and policies of the system. Furthermore, their opinions and ideas have not received much attention in the teacher evaluation literature in Saudi Arabia. As it is teachers who are the stakeholders most influenced by the teacher evaluation policies, it seems significant to explore their views on teacher evaluation and to understand how they perceive it. It is my belief that, by focusing exclusively on their viewpoints and investigating them in more depth, more attention can be given to their opinions.

## **1. 5. Significance of the Study**

The idea that teacher evaluation plays a crucial role in improving instructional practices is well established in the literature of educational evaluation (Spillane, 2017). This can be achieved by diagnosing the strengths and weaknesses of each instructor based on a systematic approach, and then making appropriate decisions to elevate the level of performance in order to achieve quality in teaching. However, it can be argued that a limited number of studies have been conducted in evaluating EFL teachers'

performance based on their evaluation both generally (Tovar, 2011) and in the Saudi context in particular (Alamoudi & Troudi, 2017).

This study is significant for a number of reasons. First, the subject matter, i.e. teacher evaluation, is important in that it is one of the main concerns of those people working on quality assurance and trying to achieve academic accreditation as a global quest (Barqawi & Abuelrub, 2016). Second, Brown and Crumpler (2013) justifiably argue that there is little research that tackles context-specific areas such as foreign language teaching and teacher evaluation practices. Since this study focuses on teacher evaluation in the context of EFL, it sheds light on some of the neglected areas in this context. Third, according to Tovar (2011), despite the fact that teacher evaluation has been the focus of much research recently, due attention has not been given to EFL teachers in particular. In an attempt to fill this gap in the literature, this study is intended to explore EFL teachers' perceptions on the current evaluation systems in some Saudi public universities. To the best of my knowledge, EFL teacher evaluation has been the focus of a very limited number of research studies in Saudi Arabia. The only study I found is Al-Hammad (2011) on teaching performance assessment for English teachers at intermediate schools in the city of Hail. Accordingly, and as far as I know, the current study can be seen as one of the first studies that examines EFL teacher evaluation in higher education in Saudi Arabia.

## **1. 6. Research Aim and Objectives**

The aim of this study is to explore EFL teachers' perceptions about teacher evaluation in Saudi public universities. In order to achieve this aim, the researcher will focus on the following objectives:

- Understanding how EFL teachers perceive the importance of teacher evaluation in Saudi higher education,
- Investigating the perceptions of EFL teachers in higher education about the purpose of teacher evaluation in Saudi public universities,
- Exploring the methods that are used to evaluate EFL teachers in Saudi universities,
- Defining the existing challenges to the current evaluation system in Saudi public universities as described by EFL teachers, and
- Reporting some plausible solutions to these challenges as suggested by the teachers themselves.

## 1. 7. Research Questions

The current research is designed to address the following questions:

1. What are the purposes of EFL teacher evaluation in Saudi public universities?
2. What are the methods used to evaluate EFL teachers in Saudi public universities?
3. Why is teacher evaluation important (or not important) to EFL teachers in Saudi universities?
4. What are the challenges to the current teacher evaluation system in Saudi universities as described by EFL teachers?
5. How can we address these challenges to have a better teacher evaluation system?

## 1. 8. Organization of the Study

This exploratory interpretive research that utilizes a mixed method sequential design is divided into seven chapters as follows: **Chapter One**, the introduction, presents the study and states the research problem and the rationale of the research. It also addresses the significance of the study in general and how it makes an original contribution to the body of knowledge in the discipline of teacher evaluation in

particular. The main objectives of the thesis along with the research questions are also included in this chapter. **Chapter Two**, the background, provides the reader with background information about the context of the study to familiarize the readers with some key aspects about Saudi education. It introduces the educational system in Saudi Arabia and how English language is taught within that particular context. English language teaching in higher education and how English language teachers are educated and trained to teach English in the Saudi context are briefly addressed in this chapter. **Chapter Three**, the literature review, outlines the conceptual framework for the study after describing and analysing previous related research. This chapter critically discusses and analyses the body of relevant knowledge with the ultimate goal of determining the gap in the literature that the current study is attempting to fill. **Chapter Four**, the methodology, identifies the theoretical framework. It clarifies the philosophical standpoints along with the methodological assumptions underpinning the study. It also presents the description of and justification for the methods that are used to collect the data. This chapter also outlines how data is analysed and analysis techniques are revealed. The design of the study is explained thoroughly in this chapter and limitations are discussed. The following are also addressed: population description, justification for the sampling, instruments development, coding of data, and, most importantly, the ethical considerations that are taken into account while conducting the study. **Chapter Five**, the data analysis, addresses the results from analysing the data. It outlines the exploratory findings resulting from both the quantitative and qualitative analysis. These findings will be supported by some data that was gathered for the purpose of this research. The findings are also linked to findings from other relevant research in the literature. **Chapter Six**, the discussion,

considers what the research findings may mean in relation to both the context of the study and the literature. The outcome of the study will be linked to the theoretical body of existing knowledge on the topic of EFL teacher evaluation reported in chapter three. The main purpose of this chapter is to answer the "So what?" question and to elaborate on and move beyond the findings. **Chapter Seven**, the conclusion, brings the thesis to its end with conclusive remarks that provide closure. It addresses both theoretical and practical implications of the study and recommendations based on its findings. Furthermore, it presents a suggested model for EFL teacher evaluation that may help in overcoming the challenges reported by the teachers as well as in implementing their suggestions for solutions. It also raises current needs and proposes areas for future research.

## **1. 9. Summary of the Chapter**

The aim of this chapter is to introduce the study and to set the scene for the readers. It highlights the importance of the research for both the research context and the wider world. The problem of the study is then stated and followed by the rationale and the significance of the study. This chapter is brought to an end by presenting the objectives along with the questions of the research that will help in achieving them, and, finally, an overview of the organization of the study. The following chapter provides the background for the context of this study, the Kingdom of Saudi Arabia.

## **Chapter Two**

### **Background of the Study**

#### **2. 1. Introduction**

This chapter offers fundamental background information related to the Kingdom of Saudi Arabia as the context in which this study took place. The first section introduces the reader to the kingdom and how the English language is used there. The attention then is shifted to a general view on the educational system followed by a special focus on the status of the English language in Saudi Arabia. As this study is focusing on higher education, an outline of the teaching of English in higher education is presented. Furthermore, and as part of the concern of EFL teacher evaluation research in Saudi Arabia, EFL teacher education in this context is outlined in the fifth section. The sixth part is devoted to a consideration of the EFL teacher as a faculty member in Saudi universities, with a focus on recruitment standards, promotion requirements, duties and responsibilities. Finally, the chapter is brought to an end by presenting the teacher evaluation system currently employed in Saudi universities, followed by a brief summary.

#### **2. 2. Overview of the Saudi Context**

Since the establishment of Saudi Arabia as a kingdom in 1932, the country has been through exceptional development in almost all fields of life such as health, economy, and science; and more recently in technology (Rahman, 2011). The educational sector, in particular, has seen a rapid growth in all parts of the country. English language teaching has also been taken in a new direction. The Saudi government,



people in authority, policymakers, Saudi thinkers, teachers and pupils are sensitive to the significance of the English language in the modern world and, accordingly, a considerable amount of work has been done at all levels of education in Saudi Arabia in order to improve the proficiency of English as a foreign language (EFL) amongst teachers and students.

English has been taught as a foreign language in all Saudi public and private schools, universities, and many governmental and industrial institutions. In Saudi universities, English has been taught as both a major field of study or as an elective course; it is also used as a medium of instruction in a large number of Saudi public and private universities. In some public institutions and various private organizations, English language courses have been introduced into their training centres in order to improve the employees' proficiency in English. The King Abdullah bin Abdulaziz Programme for Scholarship, which gives Saudi students the chance to learn the language in a wide range of English-speaking countries, also reflects some of the efforts to improve the status of English language learning in the Kingdom. Furthermore, applicants who speak English well usually have a greater chance to compete in the job market in Saudi Arabia.

Despite all the previously mentioned efforts, the status of English in the kingdom remains in a continuous state of change. This change has been the consequence of the low achievement of Saudi students generally in terms of proficiency after learning English as a foreign language for many years. In fact, it may be unfair to accuse one particular aspect of the system of Saudi education, such as textbooks, students, or their teachers, for the unsatisfactory outcome of teaching English in the country, as

there are a number of factors that have led to this result and teachers are aware that, as instructors and communicators of knowledge, they may play a significant role in this problematic situation.

### **2. 3. Educational System in Saudi Arabia**

The system of education in Saudi Arabia is divided into two main stages: general and higher education, both of which are offered freely to Saudi and non-Saudi students. In the primary, elementary and secondary phases of general education, students' progress depends on passing monthly and annual comprehensive exams. The General Secondary Education Certificate is awarded upon the completion of the third phase and students can then be enrolled in higher education.

Three different types of educational systems can be found in Saudi schools: public, private, and foreign international systems. Public education is a government-funded sector that offers free education in all its stages. However, students need to pay if they opt to join the private or foreign international learning systems. The Ministry of Education oversees public and private education but foreign international schools, which are intended to educate diplomatic mission children, are not usually administered or supervised by the Ministry.

Higher education includes about 36 universities in different specializations and more than 51 technical and vocational training colleges and institutions (Ministry of Education, 2017). Higher education has experienced remarkable progress in Saudi Arabia over the last decade. Recently, the Saudi higher education system has expanded to comprise 26 government universities (174 undergraduate programmes

for education, 438 for health and social care, 245 for engineering and industries and 547 for natural sciences) and 33 private universities and colleges (ibid).

This rapid expansion is an expected outcome of the increase in the budget allocated to higher education by the Saudi government. Statistically, government spending on higher education rose considerably during the period 2010-2014 to reach its peak in 2014, when the government spent eighty billion Saudi riyals (around 14 billion British Pounds) on higher education (Ministry of Education, 2015). This has also led to a rapid increase in the number of registered students and, accordingly, to an increase in the number of faculty staff members. The number of registered students in public higher education jumped from 757.770 in 2010 to 1.358.312 in 2014 and in the private sector from 26.416 in 2010 to 193.221 in 2014. Over the same period of time the number of faculty members also rose from 41.589 to 64.689 (ibid).

## **2. 4. English Language Status in Saudi Arabia**

English was first introduced into the newly-established public secondary school system in 1937 (Al-Abdulkader, 1979) as a compulsory subject with just little emphasis given to efficiency in both teaching and learning the language (Rahman & Alhaisoni, 2013). At this early stage, and according to Al-Seghayer (2005), English was taught without any explicit goal. Since then, English has been taught to Saudi students not as a second or an official language but rather as a foreign language. However, the status of English in Saudi Arabia has been growing rapidly due to worldwide demands.

English has become the language of technology, science, business, economy and commerce all over the globe. Thus, English language has become one of the major subjects in the system of education in Saudi Arabia. Accordingly, English language as

a subject is being now introduced from year four in the elementary stage in public education continuing up to college level. The main objectives of English language teaching are to improve the socio-economic state of the kingdom, to advance individuals' careers, to interconnect with speakers of other languages, to develop international interaction, to spread Islam, and to obtain and provide knowledge (Alamri, 2008).

#### **2. 4. 1. English Language in Saudi Higher Education**

In higher education policy, Arabic is the medium of instruction in Saudi universities (Ministry of Higher Education, 1999). However, English has increasingly become more dominant in a considerable number of universities and higher education institutions across the Kingdom. For many Saudis, English is needed for the country's prosperity in different domains in the future (Aburizaizah, 2014). Aburizaizah (2016) suggests that all Saudi universities place considerable emphasis on English language learning for freshman students. In fact, it seems that the Saudi Ministry of Education has dictated this policy of language teaching.

To that end, according to McMullen (2014), the Ministry of Education in Saudi Arabia has recognized eleven goals for the teaching of English within the country. These goals include enabling students to acquire basic language skills, developing student awareness of the importance of English as a means of international communication, developing students' positive attitude towards learning English, and enabling students to acquire the necessary linguistic competence required in different professions (Ministry of Education, 2013).

In some higher education institutions, English has been taught as part of the preparatory programme while in others it has been used as a medium of instruction in all subjects. Nevertheless, only a few institutions fully transferred to the use of English as the only medium of instruction; these include King Fahad Petroleum and Minerals University and King Abdullah University of Science and Technology. Students in all universities are expected and prepared to read books and journals in English and to write theses in English. Moreover, colleges of technology, vocational training institutions and military colleges provide intensive courses of English language as well as teaching science-related courses in English (Habbash, 2011).

#### **2. 4. 2. Preparatory Year English Language Programme**

The introduction of a preparatory year in Saudi universities is recent and accordingly only a few studies exploring various aspects related to this preparatory year can be found in the literature (Al-Shehri, 2017). It is worth mentioning that the preparatory year programme is the first foundation year and that it is intended to narrow the gap between tertiary education and high school. Its main objective is to prepare students with the language and academic skills that they will need during their university education.

According to Al-Maliki (2013), preparatory year programmes are critical as they represent a middle/transforming stage between higher education and public education. In line with this, Al-Anzi (2015) summarizes the advantages of these programmes as follows:

- They present a significant educational leap given that they help improve the students' basic study and English language skills.
- They develop students' self-esteem, self-confidence, and improve their job skills.
- They advance communication skills, students' independence, and self-learning skills.

However, preparatory year programmes in Saudi universities have been criticized for the lack of clear objectives and standards together with poor quality or presentation. For instance, Al-Otaibi (2015) links the lack of objectives to the fact that newly established programmes do not profit from the experience of other universities in the country. In the same vein, Al-Maliki (2013) believes that the lack of educational quality standards in these programmes has contributed to the failure of Saudi universities to achieve their intended goals.

Al-Shehri (2017) claims that English language programmes in the preparatory year suffer from the “lack of vision, poor educational outcomes, poor academic guidance and student support programs, and the reliance on private companies to operate most of these programs” (p. 435). He argues that, with the poor English language skills students bring in from high schools, it is not possible for them to acquire sufficient language skills in only two semesters. This is likely to be true even at those Saudi universities where English language programmes are taught in four modules during the preparatory year and where each module lasts for seven weeks. Despite the fact that the average number of teaching hours for English language is 18 hours per week, which is quite overwhelming for both students and teachers, such programmes seem to be unable to develop students' language skills to any great extent.

Given the long teaching hours required for English language courses and the high volume of students enrolled every academic year in Saudi universities, and due to the shortage of preparatory year faculty members, Saudi universities frequently sign employment contracts with companies from the private sector. This situation has raised concerns for many educators in Saudi Arabia about the quality of education. Al-Sameti and Al-Seraji (2016) warn that the preparatory year programmes have been converted into commercial programmes, where private companies have a tendency to hire faculty members with low qualifications. Thus, it can be argued that it is critical for English language institutions to directly supervise and evaluate teachers assigned by such companies as part of their teaching teams.

It is worth noting that language institutions need to develop strategies in order to ensure that the courses they offer to students in preparatory year programmes are of sufficiently high quality. This is essential as it has been proven that one of the main reasons behind students' withdrawal or failure is their inability to retain sufficient English language skills upon completion of their preparatory year programme (Al-Juhani, 2012). Given that quality teachers are amongst the key aspects that can elevate the quality of any educational programme, their education and preparation seem to be essential steps in improving that quality. The following section explores teacher education in the country.

## **2. 5. EFL Teacher Education in Saudi Arabia**

In all levels of education in Saudi Arabia, the minimum requirement for teaching is the four-year Bachelor degree. Therefore, universities in the Kingdom are fully responsible for teaching, preparing, and training prospective teachers and staff members. Colleges

of education at these universities have objectives that reflect their roles in preparing teachers in cooperation with the Ministry of Education. These objectives include preparing well-qualified staff as well as upgrading the professional and educational standards of the current teachers, principals and administrators through providing different training programmes.

Schools of education place considerable emphasis on educational aspects in their programmes to prepare teachers and university staff (Zaid, 1994). They offer an extensive curriculum in various educational aspects, theories and methods. They also have distinct departments for different subjects such as Biology, Physics, Mathematics, Arabic, English, Geography, History and Islamic Studies. Graduates of education colleges are trained in learners' psychology, teaching methodology and educational measurement. They are also enrolled in practicum courses where they practise teaching in intermediate or secondary schools. It is noteworthy that "many Saudi EFL teachers who are serving in schools around the country are graduates of faculties of arts and had never had any training" (Al-Saadat, 2004, p. 1) in educational practices and therefore they are not sufficiently prepared to be qualified language teachers.

In Saudi universities, Saudi and non-Saudi teachers can occupy different kinds of position. Apart from being language instructors, they may be appointed as teaching assistant, lecturer, assistant professor, associate professor and professor, subject to vacancy and type of contract. With regard to lecturers and professors, they are usually assigned to teach according to their academic qualifications. Thus, while holders of doctoral degrees can teach any course in their specialization, those with a Masters



can only teach undergraduate courses, and holders of a Bachelor degree only teach introductory courses. Unfortunately, teacher training is not considered a compulsory requirement for teachers at higher education level. Accordingly, many lecturers are found to be insufficiently qualified and they frequently lack some essential teaching skills.

In-service training programmes are offered on a very limited scale via some education departments scattered throughout the Kingdom. Furthermore, and according to Al-Seghayer (2014, p. 21), these programmes are conducted in “a poor manner” and on a limited scale. Some English language teachers have taken few if any of the in-service training courses on offer despite the fact that they have been teaching English for a considerable period of time. Apart from promotion to supervisory positions (ibid), career advancement is not assured for English teachers. By the same token, incentives are not provided for English language teachers to engage in professional self-development and resources for teacher training are very limited. Nevertheless, developmental courses provided by professional development departments are highly recommended for EFL teachers in Saudi universities.

## **2. 6. EFL Teacher as a Faculty Member in Saudi Universities**

EFL teachers who are working in Saudi universities are generally treated as faculty members and they adhere to and abide by rules and regulations from the “*Regulations Document governing the affairs of faculty members in Saudi universities*” taken at the sixth meeting of the Council of Higher Education held on 26/08/1417 AH (06/01/1997) approved by the Royal Order No. 7/ B/12457 on 08/22/1418 AH (22/12/1997). Thus,

it is important to look at some aspects of the previously mentioned document when we investigate EFL teacher evaluation in higher education in the Kingdom of Saudi Arabia.

In Article 1 of the document, it is stated that Faculty members in Saudi universities comprise Professors, Associate Professors and Assistant Professors. Article 2 clarifies who those of equivalence can be, namely Lecturers, Teaching Assistants, Language Teachers, and Research Assistants. Accordingly, seven different types of EFL teachers can be found in Saudi universities.

In the following section, I will present a summary of the relevant Articles relating to three main areas: firstly, how faculty members are selected in Saudi universities and the standards for their recruitments; secondly, the responsibilities and duties they are required to fulfill as faculty members; and thirdly, the criteria and conditions for their promotion.

Recruitment standards, responsibilities, and promotion criteria are closely related to EFL teacher evaluation for many reasons. Concerning recruitment standards, they give evaluators some idea of the type of qualification and education EFL teachers should be expected to have before embarking on any evaluation of EFL faculty members. The area of responsibilities and duties, on the other hand, restricts and defines the borders for the criteria of evaluation. With regards to promotion, an evaluation containing 100 points for faculty promotion is formally considered as a type of faculty evaluation.

## **2. 6. 1. Recruitment Standards**

In each Saudi university, a Standing Committee for the affairs of teaching assistants, lecturers, language teachers and research assistants is composed, with the Vice President for Graduate Studies and Scientific Research as its head, where appointment of its members is based upon a decision of the University Council. The responsibilities of this committee are listed in Article 3 as follows:

1. Proposing the general policy for the selection of teaching assistants, lecturers, language teachers and research assistants and distributing it to departments and colleges.
2. Approving recommendations issued by the college boards on the appointment of teaching assistants, lecturers, language teachers and research assistants according to the following criteria:
  - A. The number of current Saudi faculty members and their percentage of the total number of faculty members in the department,
  - B. The number of lecturers, teaching assistants, language teachers and researcher assistants in the department,
  - C. The number of scholarships in the department and expected dates for their return along with their specialties.
3. Proposing the distribution of job vacancies for teaching assistants, lecturers, language teachers and research assistants as needed for current and future departments.
4. Determining the transferability of lecturers and teaching assistants to administrative jobs within the university or sending them to the Ministry of Civil Service.

It is worth mentioning that there are different recruitment criteria depending on the position applied for. More information about these recruitment standards is available in Appendix 12 (p. 387).

## **2. 6. 2. Duties and Responsibilities**

In the following three Articles, the characteristics and responsibilities of staff members or those of equivalence in Saudi universities are outlined in detail. In Article 38, it is indicated that faculty members should show the following characteristics:

1. Demonstrating honesty and morals besides adhering to the regulations, instructions and rules of moral behaviour that rises above all and reflects respect for the dignity of the job.
2. Following the latest developments in his/her field, and contributing through various academic activities in order to foster the progress of his/her specialization.
3. Conveying up-to-date academic findings in his/her field to the students, and nourishing the love of and the striving for knowledge and critical academic thinking.
4. Participating actively in the Department Council work as well as in other boards and committees in which s/he is a member, whether at department, college or university level. Being actively engaged in different community service events.
5. Devoting her/himself to fully working at the university, and, in accordance with the university rules and regulations, obtaining prior approval to work outside the university.

In Article 39, it is stated that the faculty member should keep order and system within teaching halls and laboratories. If any incident disrupts order, it must be reported to the Head of Department.

Teaching loads are specified in Article 40. It is stated that the following should be the upper limit for the shares of the faculty members or those of equivalence:

1. Professor: 10 Teaching Units
2. Associate Professor: 12 Teaching Units
3. Assistant Professor: 14 Teaching Units

4. Lecturer: 16 Teaching Units (reduced during study period)
5. Teaching Assistant: 16 Teaching Units (reduced during study period)
6. Language teacher: 18 Teaching Units

A teaching unit comprises the weekly theoretical lecture of not less than fifty minutes, or the weekly field practice of at least one hundred minutes in an academic semester. These figures are also known as teaching loads and they are used as one of the categories under the personal information section in the questionnaire designed for this study.

Article 41 indicates that faculty members or those of equivalence are expected to work thirty-five hours a week, which is subject to an increase to forty hours upon a decision from the Council of the League. Those hours are spent in teaching, research, academic counseling, office hours, committees and other work assigned to them by competent authorities of the university.

In Article 42, it is stated that the teaching load can be reduced for faculty members or those of equivalence to three units of teaching if they are assigned to administrative positions such as acting as Dean, Vice-Dean, Chair of the Department,.

### **2. 6. 3. Promotion Requirements**

For EFL teachers in Saudi universities it is well known that performance evaluation results inform decisions on promotion. Promotion requirements are consistent across universities as they are set up by the Ministry of Education. Furthermore, promotion for an Assistant Professor to the grade of Associate Professor is open for all enrolled faculty members regardless of their nationality. It is worth mentioning that, in Saudi universities, such positions are open to any qualified individual subject to availability.

Article 21 clarifies that for an Assistant Professor to be promoted to the grade of Associate Professor the following criteria are required when a staff member is to be evaluated:

1. Serving for at least four years as an assistant professor at a Saudi university or other recognized universities with a minimum of one year in a Saudi university
2. Meeting the minimum promotion requirements of research output according to Article 32 of this document
3. Publishing the subject-specific output while serving as an assistant professor

As indicated in Article 22, for an Associate Professor to be promoted to the grade of Professor, the following criteria are required:

1. Serving for at least four years as an associate professor at a Saudi university or other recognized universities with a minimum of one year in a Saudi university
2. Meeting the minimum promotion requirements of subject-specific production according to Article 33 of this document
3. Publishing the subject-specific production while serving as an associate professor

Article 25 states that the promotion of faculty members is carried out according to the following evaluative criteria:

1. Subject-specific production
2. Teaching
3. University and community service

Article 27 is directly related to faculty member's evaluation for promotion. The article indicates that promotion evaluation contains one hundred points divided as follows:

60 points for subject-specific production

25 points for teaching

15 points for serving the university and the community

The University Council sets the evaluation standards in teaching or serving the university and the community based on the recommendation of the Scientific Council.

For a faculty member to be promoted, as stated in Article 28, the sum of points obtained should not be less than 60, with at least 35 points in the field of subject-specific production for promotion to the grade of Associate Professor and 40 points for promotion to the grade of Professor. Promotion to the grade of Associate Professor requires agreement of the majority of the three arbitrators, and promotion to the grade of Professor requires agreement of all three arbitrators; should one of the arbitrators disagree, the research output can then be given to a fourth arbitrator for a final decision.

In Article 29 it is explained that the minimum subject-specific production required for faculty member promotion includes the following:

1. Research published or accepted for publication in peer-reviewed journals, and the Scientific Council sets acceptance standards;
2. Peer-reviewed research presented in specialized subject-specific conferences and symposia whether published or accepted for publication, and only one can be accepted as a single unit;
3. Peer-reviewed research published or accepted for publication by specialized university research centres;
4. Peer-reviewed subject-specific references or books, and only one can be accepted as a single unit;

5. Peer-reviewed reviews of a subject-specific book and only one can be accepted as a single unit;
6. Peer-reviewed translation of specialized subject-specific books and only one can be accepted as a single unit;
7. Research and books published by academic bodies approved by the Scientific Council and subject to arbitration, and only one can be accepted as a single unit;
8. Inventions and discoveries that have been issued with copyrights by bodies recognized by the Scientific Council;
9. Outstanding creative activity in accordance with the rules set by the Council of the University on the recommendation of the Scientific Council, and only one can be accepted as a single unit.

For promotion to the grade of Associate Professor, according to Article 30, publication in academic journals must not be less than one unit. However, for promotion to a Professor grade it must not be less than two units.

Article 31 clarifies that, for subject-specific productions to be counted, they should be published or accepted for publication in more than one body, and these bodies should belong to different universities or academic institutions.

Article 32 states that the minimum subject-specific production required for promotion to the grade of Associate Professor is four units published or accepted for publication, two of which, at least, must be individual work. The University Council upon the recommendation of the Scientific Council has the right of exemption from this condition for some disciplines; however, at least one unit should already have been published.

It is indicated in Article 33 that the minimum subject-specific production required for promotion to the grade of Professor is six units published or accepted for publication,



three of which, at least, must be individual work. The University Council upon the recommendation of the Scientific Council has the right of exemption from this condition for some disciplines; however, at least three units should already have been published.

Article 34 assures that any subject-specific production is counted as one unit if it is single-authored, and only a half unit if it is co-authored by two, and if the joint research is between more than two it is calculated as a half unit for the principal researcher and a quarter unit for each of the remaining researchers, and if another joint research was conducted by more than two it is counted as a quarter unit for each of them.

Article 35 explains that, for a subject-specific production to be counted, it must not be part of a Masters or doctoral thesis, or a previous work of the applicant. If this is the case, the applicant will not be promoted and will be prohibited from applying for promotion for a whole year starting from the date of this decision by the Scientific Council.

## **2. 7. EFL Teacher Evaluation in Saudi Universities**

There appear to be few resources that provide information about policies and practices related to the evaluation of EFL teacher in Saudi universities. Some researchers have focused on a specific aspect of EFL teacher evaluation in the Saudi higher education context. For instance, classroom observation has attracted the attention of researchers such as Shukri (2014), who investigated the perception of EFL teachers towards classroom observation and its challenges. She found that the participants held positive viewpoints about classroom observation, pre- and post-observation meetings, and called for further training and workshops. In the same vein Shah and AlHarthi (2014) raised a number of critical concerns associated with classroom observation.

Some of these issues are the trust deficit between observers and observed teachers, observers' subjectivity, the inadequate training of observers, and the insecurity and threat linked to classroom observation.

Albaiz (2016) took the issue further and investigated the use of peer-observation as a method for evaluating the performance of EFL teachers at Saudi universities. The findings of her study revealed overwhelming support from the teachers; however, she argues that the implementation of such an approach needs the support of the accreditation bodies in Saudi Arabia. Aburizaizah (2016), on the other hand, found that developing EFL teachers' performance and raising awareness of good teaching practices can contribute to assessing the performance of language teachers in Saudi universities.

It is worth mentioning that only a very limited number of studies give a holistic view about the practices of teacher evaluation in Saudi higher education and to the best of my knowledge none of them has tackled EFL teacher evaluation. Given that the evaluation of language teachers in Saudi universities does not essentially differ greatly from that of teachers of other subjects, I will outline those attempts that tackle teacher evaluation holistically in the Saudi higher education. It is well known that Saudi public universities have different systems for teachers' evaluation. A comprehensive overview of these systems has been provided in the work of Alaidarous (2011). In general, these systems include three main criteria: teaching, academic research and community services. However, I believe that all these approaches to teacher evaluation tend to overlook administrative and professional development which frequently form part of teachers' duties at university level. According to Alaidarous

(2011), the most comprehensive evaluative system for teachers is found in King Fahad University of Petroleum and Minerals and King Abdulaziz University, where three aspects are considered: self-evaluation, evaluation of head of department, and college evaluation (a sample of the evaluation form employed at King Abdulaziz University is available in Appendix 13 & 14, p. 391 & 379). Academic profile is used in two universities: King Abdulaziz and Al Taief University. The evaluation of the head of department is considered important in four universities: King Saud, King Fahad, King Abdulaziz, and King Faisal University. Students' evaluation is more public in Saudi universities and is employed in five universities: King Saud, King Abdulaziz, Um Al Qura, Al Taief, and Noura University; however, it is not included in the academic profile at King Abdulaziz University and merely used as feedback. College evaluation is utilized in four universities: King Fahad, King Faisal, and Taiba University. Self-evaluation is found useful in only two universities: King Fahad and King Abdulaziz. Alaidarous (2011) argues that despite its significance, peer evaluation is not considered fundamental in Saudi Universities.

To examine the quality of teacher evaluation in Saudi universities, Alharbi (2015) conducted a case study in Al Qassim University. He found that the majority of the participants believe the following:

- The university pinpoints staff weaknesses in order for them to be able to overcome them;
- The university impels staff members to make use of modern technology in their teaching;
- The university attempts to encourage communication between students and staff.

However, the participants also believe that the university needs to evaluate staff members in a periodic manner, and to appoint an external member as an evaluator of

the staff. The study eventually recommends developing the regular evaluation of the faculty members and considering not only the teaching process but also research and other activities.

The above outline suggests that different sources for evaluation have been used in Saudi universities; however, special attention for EFL Faculty cannot be found within those general guidelines for evaluation. Given the current situation, it seems that more studies are needed in the field of EFL teacher evaluation in Saudi Higher education.

## **2. 8. Summary**

In conclusion, this chapter gives the reader an account of the role of English in the Saudi context and explores some essential aspects that are closely related to the issue of evaluating EFL teachers in Saudi Higher Education. The educational system in Saudi Arabia has been introduced and the status of English in the kingdom has also been discussed. As the focus of this research is on higher education, the use and teaching of English in that particular sector are given special attention. EFL teacher education in the Saudi context has been outlined in order to give some background information on the preparation to become an EFL teacher. Since this research focuses on higher education, some light has been shed on the EFL teacher as a faculty member in Saudi universities and issues such as recruitment standards, promotion requirements, duties and responsibilities have also been explored. Finally, a summary of the teacher evaluation system in Saudi universities has been presented. In the following chapter, the literature on the phenomenon of EFL teacher evaluation will be reviewed.

## **Chapter Three**

### **Review of the Literature**

#### **3. 1. Introduction**

In this chapter, I attempt to conceptualize the issue of evaluating English language teachers at universities in general and in the Kingdom of Saudi Arabia in particular. Of all the structural elements in higher education institutions, teachers seem to receive the least amount of attention, support, opportunities to evaluate their proficiency or to value their experience and teaching skills in the Saudi context (Alamoudi & Troudi, 2017). This situation impacts significantly on foreign language teachers with regards to their professional development and academic growth, especially when demands on teaching quality in higher education are to be considered. The main objectives of this chapter are, first, to offer a review of the current body of knowledge on the purposes and methods of evaluating language teachers, and, second, to reflect on the philosophies and models developed for the purpose of evaluating teachers or faculties in tertiary education. The models are carefully analyzed, discussed and eventually critiqued in light of the researcher's reading in the wider literature along with the researcher's experience as an English language teacher who has been working in a Saudi University and has participated in the EFL Teacher Evaluation Committee prior to developing an interest in research of teacher evaluation.

#### **3. 2. Teacher Evaluation**

It has previously been observed that many attempts to define the concept of evaluation in the field of education have been made. Some of these attempts were to distinguish

between evaluation and other closely related conceptions such as measurement, assessment or appraisal. However, the comparison between these terms is not part of the objectives of this research and hence these terms are not discussed here. In the following section, I will provide some of the well-established definitions for the term “*evaluation*” and how it has been developed in the field of education in order to introduce and conceptualize the evaluation of language teachers’ performance.

### **3. 2. 1. Operational Definitions of “*Evaluation*” in Education**

As a significant educational concept, evaluation has received much attention in the literature. Various distinctive definitions have been provided in order to help people conceptualize this important yet sensitive issue to value and develop educators’ efforts and contributions.

In the context of evaluation research, Ralph W. Tyler entered the field as one of the prominent figures in educational evaluation. He is well known for his work in associating evaluation with the concept of learning objectives in educational programmes (Alamoudi & Troudi, 2017). According to Tyler (1949), evaluation is “*the process of determining to what extent educational objectives are actually being realized*” (p. 69). His model of objectives has had a lasting impact on the current conceptualization of evaluation, especially when the achievement of students, educational programmes, or institutions is to be taken into account. However, the model has been criticized in the literature for its inability to present a method for assessing educational objectives themselves. Yet, and for the purpose of evaluating contributions made by teachers of English as a foreign language in higher education, Tyler’s definition may be regarded as adequate, considering the standard-based and

students-learning-objectives-oriented approaches to classroom observation which are largely adapted by higher education institutions in the Saudi foreign language teaching context.

Cronbach (1960, cited in Azike, 2014), on the other hand, links evaluation in education to decision-making instead of objectives, and defines it as "*the collection and use of information to make decisions about an educational programme*" (p. 65). His work involves evaluation in three different areas in educational decisions: administrative regulation, course improvement, and decisions about individuals. I believe that, although Cronbach's definition seems effective in guiding decision-making, it could also be criticized for equating evaluation to only one of its various roles and aspects besides focusing on only summative evaluation (for more information about Summative Teacher Evaluation see page 61, Section 3. 3. 1. 2.).

Another definition is provided by Rossi, Lipsey, and Freeman (2003), who identify the concept of evaluation as the "*use of social research methods to systematically investigate the effectiveness of social intervention programs in ways that are adapted to their political and organizational environments and are designed to inform social action to improve social conditions*" (p. 16). Employing a systematic investigation for the purpose of evaluation is an important element of their definition. Another strength of the definition can be seen as the link between evaluation and political environments and social action in order to improve social conditions.

The idea of a systematic approach has been taken further by Patton (2008) who defines evaluation as the "*systematic collection of information about the activities,*

*characteristics, and results of programs to make judgments about the program, improve or further develop program effectiveness, inform decisions about future programming, and/or increase understanding”* (p. 38). Although this definition can be seen as a definition of evaluation in general, it can be extended to educational contexts. Patton has included not only a systematic approach in his definition but the definition can be perceived as embodying an inclusive description for various purposes of evaluation such as educational activities, characteristics and effectiveness of the programme. Yet, the question remains as to whether or not this definition could be applicable in evaluating the performance of EFL teachers within the Saudi higher education context in the sense that it can also be relevant to evaluating the performance of teachers of other subjects. .

When Teaching English to Speakers of Other Languages (TESOL) is taken into account, various definitions of evaluation can be found in the literature. For instance, Genesee (2001) defines evaluation in TESOL settings as “*a process of collecting, analysing, and interpreting information about teaching and learning in order to make informed decisions that enhance student achievement and the success of educational programmes*” (p. 144). He also provides different evaluative examples and argues that evaluation, from that perspective, can have various aspects to focus on in different pedagogical contexts. In his opinion, the examples provided lead to four fundamental components of evaluation in TESOL as follows: articulating the purpose of evaluation, identifying and collecting relevant information, analyzing and interpreting information, and finally making decisions. The four components are presented in four sequentially connected circles in figure 3.1.



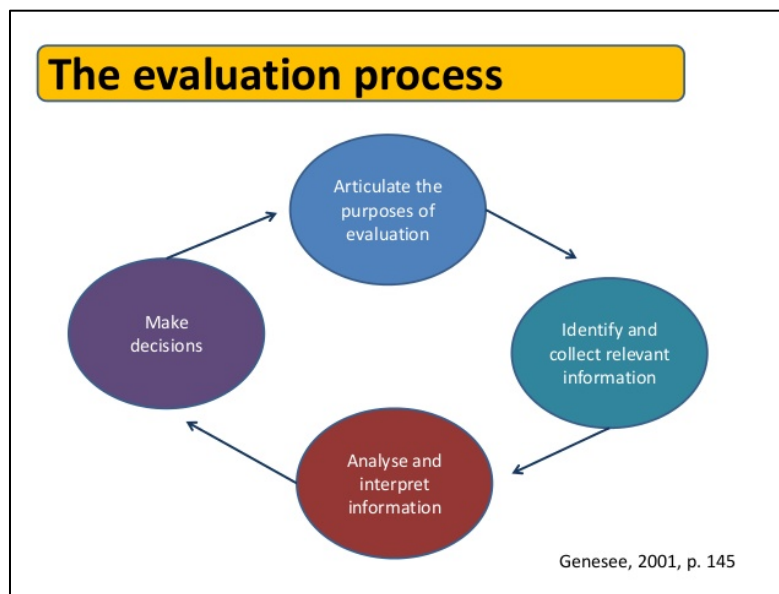


Figure 3. 1. The process of evaluation, Genesee, 2001

Given that the context of this study is higher education, it might be worth looking, briefly, at how researchers define evaluation of faculty performance. In the wider literature of educational evaluation, faculty evaluation has been tackled from different perspectives. From a rather practical perspective and according to Seldin (2006), *“faculty evaluation has many facets. It is an exercise in observation and judgement. It is a measurement and feedback process. It is an inexact, human method that must meet key requirements if it is to succeed (p. 1)”*. The second sentence in Seldin’s definition might be taken as an appropriate reflection of the spirit and practice of EFL teacher evaluation in some Saudi higher education institutions. In these institutions, teacher evaluation schemes are mainly employed as a means of observing and judging the performance of EFL teachers. However, the sentence that concludes the definition appears to reinforce the importance of perceiving faculty evaluation as an “inexact human method” which is a concomitant of the necessity for criteria.

While all the previously mentioned definitions differ in their details and the way they help researchers conceptualize the term “evaluation”, the decision to choose one of those definitions depends on a number of significant research-relevant factors. In my opinion, these factors include, but are not limited to, the following: the context of the study, the research question(s), and the issues to be addressed by the research *per se*. It is worth mentioning that the definition of evaluation for the purpose of this research is informed by the aforementioned Patton’s (2008) definition for a number of logical, philosophical and practical reasons. First, his definition seems to be comprehensive (in purpose) in the sense that it includes a variety of purposes such as making judgment, improving effectiveness, informing decisions, and increasing understanding, which works well with the research questions of the current study and the context of higher education. Second, Patton considers evaluation as a way to collect information about different aspects that include activities, characteristics and results, which allows different stakeholders’ perspectives to be involved. This does not contradict the constructivist philosophy of multiple realities underpinning this study. Third, Patton has included the notions of development and improvement in his definition. This presumably allows for the definition of a practical outcome that will help improve instructional practices. However, evaluation as such needs to be used with a considerable amount of caution in higher education contexts, as Patton’s (2008) definition can be seen as focusing more on the use of evaluation outcomes. It may thus overlook what Christie and Alkin (2013, figure 3.2. below) called “methods” for evaluating teachers’ performance, whether inside or outside the classroom, and valuing faculty staff members’ effort and other activities such as non-academic community service (for more information about Christie and Alkin’s evaluation tree see

Section 3.3.1 p. 15). With reference to Christie and Alkin's classification in their evaluation tree, another caution may be that Patton's definition can be viewed as having more social accountability roots, thus giving it a control-oriented approach towards teacher evaluation rather than a social-inquiry or epistemology orientation.

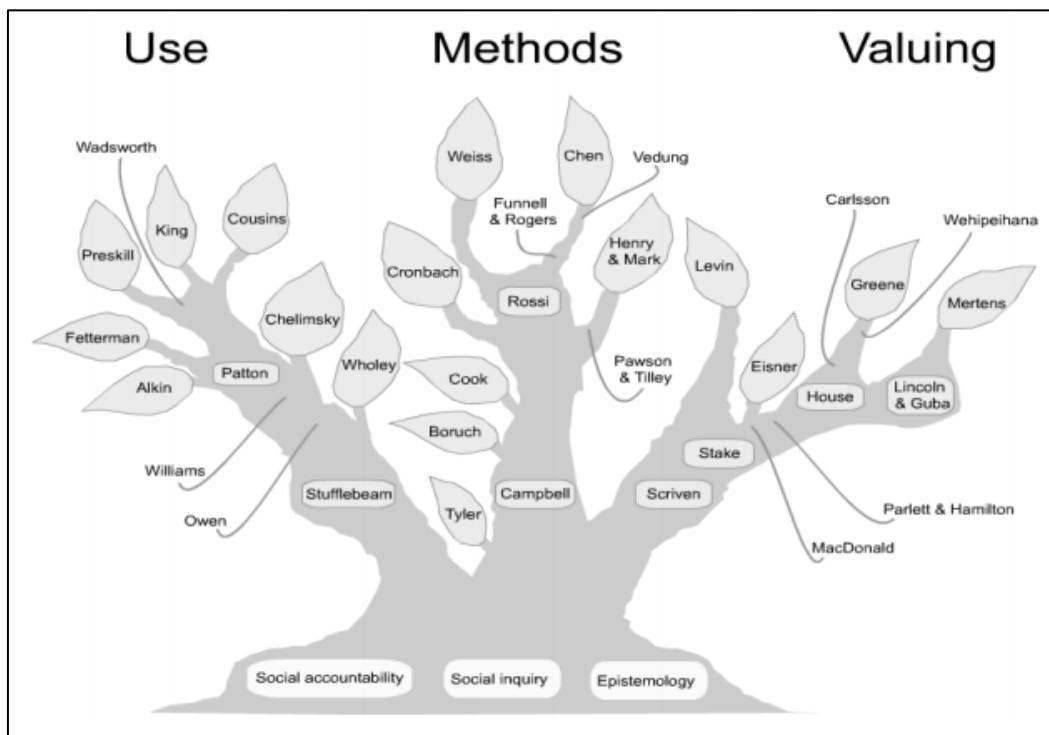


Figure 3. 2. Evaluation theory tree, Christie & Alkin, 2013

In order to make Patton's definition more appropriate to EFL teacher evaluation in the Saudi higher educational context, and drawing on my previous research and experience in EFL teacher evaluation in that context, I have developed the following proposed definition for EFL teacher evaluation for the sake of the current study. EFL teacher evaluation in higher education can be defined as:

*The systematic collaborative collection and analysis of information about the activities, characteristics, and results of EFL teachers' performance in order to increase understanding of teachers' needs, develop their teaching practices, and to inform personnel decisions or decisions about future developmental programmes that are adapted to their social and organizational environments.*

The use of this definition helps conceptualize the term evaluation to suit the context and enquiry of this study.

### **3. 2. 2. Introducing the Concept of “Evaluation” in the Field of Education**

In the field of education, there appears to be agreement that the history of evaluation began before the turn of the 20<sup>th</sup> Century (Glasman, 1986; Guba & Lincoln, 1981; Norris, 1990). Glasman (1986) argues that the history of educational evaluation can be divided into three distinct phases: the first continued until 1930s, the second lasted until 1960s, and the third is still going on. It seems that expansion rather than substitution of the old ideas is the main characteristic of the development of educational evaluation throughout those three periods.

Evaluation in education was seen first as measurement and the focus was on the measurement of the level of intelligence of children and their ability to learn a specific subject (Glasman, 1986). Glasman claims that educational evaluation before the 1930s was used widely in the life and physical sciences. On the other hand, Guba and Lincoln (1981) argue that during the last decade of the nineteenth century, Joseph Rice, who is known as the father of educational research, devised a number of achievement tests supporting his debate about the inadequate use of school time. His

test published in 1904 has become the base for almost all tests that measure intelligence since then. However, the publication of Fredrick Taylor's *The Principles of Scientific Management* can be considered as having the greatest impact on ideas about the standardization and systematization of the industry of education in that it offers a systematic methodology for educational administration (Norris, 1990).

Despite the fact that Tyler's contribution to the field of educational evaluation in the 1930s maintained that evaluation was merely synonymous with measurement, he is regarded by many researchers as the father of educational evaluation and the invention of the term "evaluation" was attributed to him (Norris, 1990). This idea was opposed by Guba and Lincoln (1981), who argue that Tyler's method of evaluation was a distinct improvement on the measurement-directed methods that were popular at that time. The reasoning in Tyler's approach is the systematic nature of evaluation, which seems justified given that Tyler's focus was on refining programmes and curricula in particular by means of examining educational objectives that can be considered an essential impetus for evaluation in the field of education.

### **3. 3. The Changing Landscape of Teacher Evaluation**

Medley, Coker, and Soar (1984) offer a concise summary of the change in teacher evaluation in the twentieth century. They divide it into three main phases: (1) the quest for great teachers; (2) determining the quality of teachers by students' learning; and (3) observing teaching performance. The three phases are discussed in the following sections. In 1896, the issue of 'Great Teachers' started with the study conducted by Kratz, who asked 2,411 students in Iowa to define the features of the best teachers (Medley, 1979). According to Medley, Kratz was hoping to establish a benchmark that

all teachers could be judged against. In his study, "helpfulness" was labeled as the most significant characteristic of a great teacher and "personal appearance" was reported as the next important feature. This result seems understandable if one considers just the students and disregards other factors or measures when evaluating teachers.

That idea was not accepted by Barr (1948), who claimed that supervisors' assessment of teachers was the actual metric of choice. However, some researchers at that time started to examine student achievement and used students' learning to reflect teacher quality, assuming that supervisors' opinions of teachers do not reveal anything about students' learning. Domas and Tiedeman's (1950) review of more than 1,000 studies of teacher characteristics indicated that for evaluators there was no clear direction. The notion of using students' achievement to evaluate their teachers was also rejected by Getzels and Jackson (1963), who argued that many of the students' tests were inappropriate to address the effectiveness of teachers. Medley, Coker, and Soar (1984) supported this opinion, claiming that students' achievement may vary and achievement tests can be poor measures of the success of the students themselves, let alone their teachers. This is true, especially considering that students' achievement can be affected by a number of other factors and may be linked to a wide range of distinct considerations.

The second era, Examining Teaching Performance, started with a focus on detecting effective teachers' behaviours that result in student learning. Brophy and Good (1986) argued that learners who receive quality instruction from their teachers achieve more

than those who work independently or receive poor instruction. Clark and Peterson (1986) concurred with this and went further, claiming that good teachers tend to adapt their instruction to their students' needs. However, Powell and Beard (1984) argued that there is a risk of subjective judgment when comparing one domain in teacher performance to another. Their bibliography of teacher evaluation research between 1965 and 1980 remains a valuable reference over time.

The third phase, teacher evaluation based on teacher performance, has experienced a number of changes over time. Many challenges to this approach in valuing the performance of teachers have been detected, "including evaluation inflation, highly subjective instruments, and a lack of objective measures" (Nagel, 2012, p. 33). Noticeably, despite the fact that there are many methods of assessing the quality of teachers, each has its own strengths and limitations. Notwithstanding the restraining factors, the notion that better student learning is a result of effective instruction remains strong. Darling-Hammond's (2000) attempt to find a link between teacher quality and student achievement can be seen as support for this argument. Although her study is based on a large collection of data from a 50-state survey of policies, it can be also argued that the study is based on some out-of-date estimates. What really makes her review distinctive is that it examines different student-related variables that may affect their achievements, other than the quality of the teachers.

Taking the value-added model in math and reading into account and in a more recent study, Cowan and Goldhaber (2015) attempt to examine the effectiveness of NBPTS-certified (National Board for Professional Teaching Standards) teachers by comparing

them to NBCTs ( National Board Certified Teachers). Although Cowan and Goldhaber conclude by stating that NBPTS-certified teachers are more effective than NBCTs, it can be argued that the rationale of their study is to use “individual assessment exercise scores to estimate optimal weights for value-added prediction” (Cowan & Goldhaber, 2015, p. 1). However, they end the study by calling for the creation of more professional development opportunities, and a better licensing system. This may not be relevant to higher education in Saudi Arabia, especially where TESOL is concerned; however, a well-designed teacher evaluation system should identify the features of effective teaching and locate those effective teachers in order to improve the value of current professional development training programmes.

### **3. 3. 1. Why Educational Evaluation for Teachers and Faculties?**

Bearing in mind the global concern with ranking, standards and sustainable development for education, and considering the research which has been carried out so far within different educational fields in higher education, teacher evaluation is seen as an essential part of educational institutions and/or language programmes around the world. However, when teaching in higher education is taken into consideration, the issue of how to sustain an effective teacher evaluation that contributes to continuous professional development for teachers has caused considerable debate over recent decades (Disterheft, Caeiro, Azeiteiro, & Filho, 2013). In the following section, I review some of the debates on the importance of teacher evaluation and the main purposes for evaluating the performance of teachers.

Christie and Alkin (2013) argue that the evaluation theory tree has three main roots: “social accountability, systematic social enquiry, and epistemology” (p. 11). These



three roots can be seen as the foundation from which evaluation emerged and hence they feed into the fundamental purposes of evaluating teachers (figure 3.2. p. 51). According to McGreal (1983), evaluation in general is expected to serve two fundamental needs: accountability to satisfy summative purposes of evaluation (for more details about summative evaluation see page 61, Section 3.3.1.2) and improvement to satisfy formative purposes of evaluation (see page 57, Section 3.3.1.1). It seems that the push for both accountability and improvement has resulted in supervision relying on integrated models of formative and summative evaluation (Gullatt & Ballard, 1998). However, and according to Towe (2012), both purposes for teacher evaluation cannot be satisfied by only one evaluative system. Additionally, some researchers in the field of educational evaluation and supervision such as Daresh (2007) and Bailey (2006) suggest further purposes for language teacher evaluation. They name it the diagnostic type of evaluation (see page 64, Section 3.3.1.3). Given those three purposes (formative, summative and diagnostic), one could argue that, when an evaluation scheme is claimed to satisfy more than one purpose, one of the intended purposes may have more weight than the other two. In such cases, a clear evaluation system that utilizes multi purposes should be clarified for both teachers and evaluators. In the following section, the three purposes for teacher evaluation will be discussed in detail.

### **3. 3. 1. 1. Formative Evaluation: Evaluation to Improve Performance**

Formative evaluation can be considered a distinct form of teacher evaluation from different perspectives. From the teachers' perspective, formative evaluation might be seen as a method for supporting individual growth. Danielson and McGreal (2000)

argue that formative evaluation places importance on teacher improvement, growth, and development. In line with this, Bailey (2006) argues that formative evaluation is conducted mainly to offer feedback for the purpose of improvement. Peterson (2000) goes further and claims that formative assessment data may be used as a type of feedback to shape future performance, to build new practices or to alter existing practices. Peterson's standpoint suggests that teachers' perspectives on teacher evaluation might differ from institutional administrative perspectives. Thus, institutions need to consider teachers' perspectives about teacher evaluation when they carry out formative evaluation. In line with this argument and in an attempt to ascertain the reasons why EFL teachers feel disconnected from professional endeavours, Glisan (2005) makes a distinction between three different types of EFL teachers: those who are interested in remaining on the periphery of the profession, those who are in isolated settings such as rural areas, and the vast majority who do not believe in the value of professional development plans. From a personal perspective, I believe that a number of EFL teachers strongly believe in the importance and value of professional development programmes. These teachers can help develop a better individual and institutional growth while practising formative EFL teacher evaluation in higher education; their voices can feed decisions on workshops that are needed by the majority of teachers in their context. Nevertheless, the involvement of any EFL teacher in professional development activities carried out within the educational programmes on offer may remain a personal decision made by the teachers themselves.

Hughes (2007) carried out an empirical study to investigate the indicators of quality for English language teachers by examining the importance attributed to formative self-

evaluations done by English language teachers. Findings from the questionnaire show that teachers involved in processes of continuous improvement plans found them useful. The results of that study can be seen as supporting the call for EFL teacher engagement with the institutional formative development programme. However, the situation in higher education may be more challenging. In support of this argument, and taking the performance of university teachers into account, Seldin (2006) states that evaluation carried out to improve performance can be seen as only one of the reasons for the evaluation of faculty members. Where faculties are concerned with formative evaluation for their own performance as individuals within higher educational institutions, approaches such as continuous professional development and informal peer observation can be useful.

In the TESOL context, Shukri (2014) conducted a study to investigate EFL teachers' perceptions towards classroom observation, a teacher evaluation method, as a form of professional development. The participants included 50 female teachers from King Abdulaziz University in Saudi Arabia who received a 5-point Likert-scale questionnaire in the first stage and open questions in the second. The findings revealed that the majority of the participants agreed and strongly agreed that classroom observation provides feedback for teachers that allows them to improve their teaching practices. In my opinion teacher evaluation as such can play a significant role in developing teaching and learning. In addition, 75% of the participants agreed and strongly agreed that training opportunities based on teacher evaluation and observers' feedback help to enhance their teaching skills. Shukri claims that this high tendency of agreement is a result of the fact that EFL teachers view professional development workshops as

essential in keeping them up to date in terms of teaching styles and methods. Amongst the study's recommendation is the implementation of "informal" peer observation and microteaching sessions in order to enhance teachers' instructional practices and contribute to their professional development. The findings and recommendations of the study support the argument that formative evaluation can be a method of supporting teacher growth.

In line with Shukri's (2014) claim, Albaiz (2016) suggests that peer-observation as a teacher evaluation tool can improve student-learning outcomes and ensure quality of teaching. The 107 respondents from the University of Jeddah in Saudi Arabia involved in her study had different qualifications and work experience as EFL teachers in different Saudi higher education institutions in the Kingdom of Saudi Arabia. The research instrument was a survey with 37 close-ended questions. The findings indicate that the majority of the participants perceive peer-observation as a way to support EFL teacher professional development that enhances their teaching skills and knowledge of the subject matter, and to develop better attitudes about teaching. Moreover, 44.16% of the participants believe that classroom visits by colleagues encourage language teachers to identify different strategies that can help promote the success of their students in the process of learning English language at tertiary level. Furthermore, only 21.62% of the participants believe that teacher evaluation by colleagues is biased, while 30.01% believe it is not biased, and the overwhelming majority were not sure. The study recommends that the application of peer-observation requires relevant authorities to cooperate with institutions of higher.

### **3. 3. 1. 2. Summative Evaluation: Evaluation to Make Decisions**

Summative evaluation, on the other hand, can be described as the summary of evaluation that serves decision-making in different educational programmes. According to Danielson and McGreal (2000), its focus might be on ranking, rating, and making judgments about the adequacy of teachers' performance. Bailey (2006) claims that results from summative evaluation help to determine when funding is to be continued: it is "*a final assessment, a make-or-break decision at the end of a project or funding period*" (p. 184). However, and prior to Bailey's opinion, teachers' evaluators have long considered the difficulties and challenges involved in evaluating work done by professionals (teachers). Wragg, Wikeley, Wragg, and Haynes (1996) argue that those in "receipt of public funds must be accountable" (p. 1) and consider it a society's right to know; yet, they acknowledge that valuing educational performance and procedures can be sensitive and might lead to rejection instead of cooperation. In their view, the performance of teachers can also be negatively affected when teacher evaluation is carried out with a focus on effective performance. In higher education, the case may become more controversial, given that faculty members and teachers in higher education are expected to provide high quality education to their students.

In higher education, summative evaluation can be linked, but not limited, to decisions for faculty members' tenure and routines for individual promotion. Youn and Price (2009) claim that people involved in processes related to routines of promotion and tenure decision-making in *comprehensive* universities can be faced with ambiguity when the probable outcome is to be considered, given that tenure and promotion are known to be rule-based actions. Before proceeding with the issue of evaluation for

personal faculty decision, it seems wise to explain what the term “*comprehensive*” university means. According to Henderson and Buchanan (2007), universities that are labeled comprehensive are those universities that “grew rapidly in size, structure and mission over the past 40 years, struggling to attain a consistent character that distinguishes them from the doctoral universities and liberal arts college” (p. 523). Accordingly, it seems that struggle with identity is a main feature amongst these universities. Henderson and Buchanan argue that members of faculty at such universities, often with heavy teaching loads, may be unaware or confused about their own roles, when they can be required to provide different kinds of services for their communities. Unfortunately, this may also be the case with some faculty members involved in EFL teaching at a number of Saudi universities. Hence, the faculty roles must be taken into consideration before making decisions based on the summative type of evaluation.

Routines, promotion, and tenure may be issues related to the summative evaluation of faculty members. These academic actions should be based on a set of clear and well-regulated rules rather than the subjective views of the evaluator or decision-maker. Monin, Noorderhaven, Vaara, and Kroon (2013) mention this issue while calling for justice, equity, and equality in organizational decision-making. These issues can be crucial in any academic organization, and the case may become more serious when EFL teacher summative evaluation is to be considered in Saudi universities, given that some teachers feel uneasy about the rather opaque system of teacher evaluation. According to Youn and Price (2009), people involved in decision-making processes are faced with ambiguity where probable outcomes are to be predicted from

beginning to end. Accordingly, organizational actions that are based on rules, as opposed to those based on choices, should require the involvement of more than one actor in the decision making, operating in a cyclical, complex, temporal sequence of actions. Institutionally, it may be difficult to predict the outcome of each step in the decision-making procedure for an individual faculty member, given the dilemma mentioned above. Such an academic working environment has been criticized for assuming that actors in decision-making procedures may rely on looking at appropriateness instead of making a sensible calculation based on data to make decisions on tenure or promotion for faculty members in order to deal with the ambiguity they may encounter.

From what has been discussed above, what is crucial to this study is what Youn and Price (2009) call the loss of “maximizing the value of certain outcomes” (p. 211). This should be one of the main targets for summative evaluation, where the value of a faculty member’s performance and academic/social contributions is expected to be under the evaluators’ magnifying-glass. From a teacher’s perspective rather than an academic intuitional/organizational one, Seldin (2006) argues that in some higher education institutions, and in haste, administrators tend to collect, not to select, inaccurate data or use poorly designed methods for gathering information that are implemented in a “get-it-done-quickly” (p. 6) manner to conduct evaluation of faculty members for personnel decisions. This, in fact, may be the case in some summative evaluation of EFL teachers in Saudi universities, especially where teacher evaluation is a new trend for accreditation or university/institutional ranking purposes.

Accordingly, and taking into consideration the importance of the summative type of evaluation, faculty members seem to be in general agreement with their universities' norms; however, they may balk at what looks like administrative screening with zero value objectives (Seldin, 2006) and could end up rejecting the evaluation system along with its unexpected outcomes. Ultimately, teachers need to be aware of the idea that promotion or tenure decisions are significant acts yet they can be costly decisions. Youn and Price (2009) arguably suggest a termly contract that is linked to increases in salary in a compensatory manner as a prospective alternative for tenure system. This might be accepted by some EFL teachers in higher Saudi education, although faculty members may look at different aspects such as job security and sustainability while considering evaluation for tenure promotion, given that they may see themselves as struggling with the previously mentioned teaching loads, research, and community services. In the following section, an additional purpose for the evaluation of teachers or faculty members, namely diagnostic evaluation, will be presented to give a complete picture of the purposes for teacher evaluation.

### **3. 3. 1. 3. Diagnostic Evaluation: Evaluation to Provide Data to Outsiders**

Daresh (2007) argues that evaluation is diagnostic when determinations in educational programmes are to be considered. According to Daresh, this type of evaluation is used to “determine the beginning status or condition (...) prior to the application or intervention or treatment” (p. 281). In line with this, Bailey (2006) claims that before any attempt to change, and in order to provide data about the current status, diagnostic evaluation should be carried out. Bailey goes further and argues that it seems sensible to start with diagnostic evaluation, followed by systematic formative evaluation, and



then summative evaluation can be conducted after an extended period of formative evaluation.

The term “decision” in the context of EFL teacher evaluation in higher education might be perceived as vague or misleading *per se*. Concerns such as promotion versus retention constantly arise. When teachers’ understanding and conception are to be considered and in an attempt to investigate how administrators influence retention decisions made by teachers, Boyd et al. (2011) conducted a study based on a survey completed by 4,360 first-year teachers to examine the relationship between the reported working conditions and teacher turnover (teachers leaving or being replaced). They found that administrations have “the greatest influence on teacher retention decisions” (p. 303). However, Boyd and his co-researchers argue that the teachers’ characteristics and experience can also predict teachers’ turnover. They also claim that old and young teachers in their study seem to have a higher retention rate than middle-aged teachers. Building on this argument, and given that educational programmes are expected to seek continuous development for teachers in their profession, one could claim that diagnostic evaluation is only needed when change decisions are to be carried out within educational structures. This can be easily solved for new-comers to the profession through recruitment-based decisions. However, this may raise other concerns, such as how to conduct diagnostic evaluation of those EFL teachers who have been involved for years in the profession of teaching. Furthermore, the same concern may apply with faculty members who are part of the academic field at their universities, involved in research activities and expected to support social enquiries in their communities.

In a relevant study, Henderson and Buchanan (2007) investigated faculty members' involvement in publishing activities related to teaching and learning scholarship at comprehensive universities (for a definition of comprehensive universities see page 62, Section 3.3.1.2). Their findings illustrate that those faculty members at universities which are labeled comprehensive show more involvement in articles publication focusing on pedagogical interests rather than in research-oriented journals. What matters in their study is that the faculty members' roles as academics seem to be confusing and unclear and hence it may not be wise to report the results of any diagnostic evaluation to an outsider. They use Coser's (1974 cited in Henderson and Buchanan, 2007) adjective "greedy" (p. 524) to describe those post-secondary educational institutions that place multiple academic and community service demands on faculty members.

Moreover, when diagnostic evaluation for teachers in higher education is to be considered, issues such as academic role, research duties and professional identity need to be taken into account when collecting or reporting the necessary data. In support of this argument, Seldin (2006) claims that academic institutions need to gather and propagate information on student services, student satisfaction and learning ratings, drop-out, employment of alumni, patterns for achievement, and contributions to the community and society. This can be carried out for a number of reasons such as providing faculty hiring, retention, promotion, or tenure data that may be required by government legislation, trustees' boards, or other departments to check on faculty efficiency. Nevertheless, Seldin does not report on faculty publication,

which seems essential as part of faculty members' contributions that need to be reported as well.

Henderson and Buchanan (2007), in contrast, argue that publication remains the central and most significant factor for a faculty member to attain academic status in higher education. They argue that putting some effort into publishing by scholars on teaching and learning might be seen as a challenge for those involved in low status institutions, where the contribution of publications may not count towards faculty members' tenure, promotion, or increase of merit-based salary. This would appear to be the case in many contexts, especially where EFL teachers are involved in higher education yet not treated as faculty members but rather as service teachers. Although a number of Saudi and non-Saudi EFL teachers who are appointed as assistant or associate professors are promoted based on their publications, other language teachers who are interested and involved in research activities dearly wish that such activities were counted for when they are being evaluated.

Having outlined above the various debates on the main purposes for teacher evaluation, I believe that utilizing all the three types, namely formative, summative and diagnostic evaluation, is the logical way forward. However, the three purposes can be given a different amount of attention and significance depending on the context, objectives, and the rationale for the teacher evaluation programme adopted in a specific educational institution.

### **3. 3. 2. How to Evaluate Teachers?**

In the wide range of literature on teacher evaluation, there has been discussion of different methods and various sources that can be used, such as student rating, classroom observation, self-evaluation and teaching portfolio (Arreola, 2006; Poster, Poster, & Benington, 1991; Seldin, 2006; Towe, 2012). Quirke (2010) suggests that, in order for any teacher evaluation system to be balanced and effective, it should allow for the collection of data from multiple sources. In line with this, Brown and Crumpler (2013) argue that, due to the complexity of evaluating the performance of foreign language teachers, there should be a number of peer observations in addition to the classroom visits carried out by administrators, as well as other forms of teacher evaluation.

In the following, I present a number of teacher evaluation methods, namely student ratings, classroom observation, self-evaluation, and teaching portfolio. However, this does not suggest that any one of them should be utilized in a specific context; rather, the educational institution should adopt one or more of them to satisfy the purpose(s) for EFL teacher evaluation in that particular context.

#### **3. 3. 2. 1. Student Rating**

Student rating is used extensively to reflect on the performance of teachers and to provide colleges and educational institutions with their perspectives and feedback. Seldin (2006) argues that it is widely thought that the ratings of students are all that is needed to evaluate the effectiveness of teaching. However, there are ongoing concerns about utilizing this approach in evaluating the performance of teachers. Students, as a key element in educational systems, have a very close and extended

interaction with the teacher; hence, their judgment can be valuable and genuine. However, their ratings should not be the only source. Despite the perspective that students may be seen as a significant source to evaluate the teacher's performance in a wide range of educational institutions, including the Saudi context, their ratings as a tool can have their own limitations.

In this respect, it can be argued that students may neither be well prepared nor have enough experience to be able to evaluate their teachers (Al-Seghayer, 2016). They may focus on the teacher's personality and give it more attention than academic and teaching skills. Arreola (2007) reports that two-thousand articles or more in the literature have been devoted to students' ratings of faculties in various subjects over the period of more than eighty years of research. This reflects the considerable attention in the literature given to the rating of students for faculty members. In his chapter on student ratings, Arreola (2007) challenges the validity of what he calls "Homemade Faculty Evaluation Tools" (p. 99).

Bearing in mind higher education contexts and cultural aspects, and according to Arreola, there is doubt about the usefulness of the ratings of students. For instance, Burden (2010) conducted an interpretive study in order to understand EFL teachers' views on the usefulness of students' ratings of ELT teacher faculty in Japanese universities. He claims that students tend to take evaluation as a revenge tool and to show their frustration at having to take compulsory English language classes, which they consider a boring requirement. Thus, I believe that student ratings are more beneficial for professional development purposes and programmes rather than making decisions based on them. Furthermore, any inclusive evaluation system may need to

consider research findings before relying on student ratings in order to avoid any expected limitations.

Davidson (2010) argues that student rating is potentially an effective tool to help administrations to evaluate EFL teachers only if both students and teachers value teacher evaluation. He claims that teachers frequently have negative attitudes towards student ratings. Davidson investigated the attitudes of 94 EFL teachers towards students' ratings of teachers at two campuses of Zayed University in the United Arab Emirates. He found that 68.5% of the participants believe that students do not read questions on the evaluation survey carefully and 62.8% think that evaluation surveys do not generate valid data about their teaching. He concludes his study by confirming that students' evaluation of teaching is the least effective way to evaluate the performance of EFL teachers.

In a comparative study in a Saudi context, Al-Qahtani (2010) asserts that a considerable number of Saudi university students do have the critical thinking abilities that might enable them to participate in staff evaluation compared to other university students in the Gulf area. This might be considered as an encouraging aspect for more student ratings for faculty in Saudi higher education. However, Al-Seghayer (2016) claims that the language used by individuals in Saudi society, including students, can be used as a tool to evaluate the society in which they live or the culture they come from. Thus, researchers or practitioners who are developing teacher evaluation tools need to be aware of, sensitive to and reflective on the language of the forms that are used by students while commenting on their teachers' performance.

### **3. 3. 2. 2. Classroom Observation**

The literature on teacher evaluation has acknowledged the fundamental role of classroom observation in enhancing the quality of teaching (Howells & Meiring, 2015), yet teachers tend to perceive classroom observation as a fear-inducing practice in that it is commonly thought that the observer puts them on the spot and is there to 'get them'.

Shukri (2014) argues that there are two main approaches to classroom observation as a method of evaluating EFL teachers, namely supervisory and non-directive approaches. In supervisory classroom observation, the observer is an administrator or a supervisor who usually pays visits to classrooms in order to give overall feedback and/or judgement on the teacher's performance and the lesson s/he observes. The relationship between the observed teacher and the observer involves a kind of power relationship as the observer occupies a superior role. This type of classroom observation tends to be judgmental since it mainly relies on the observer's judgement, which can be subjective, especially where the observer does not receive sufficient training. The non-directive approach to classroom observation, on the other hand, focuses on motivating the observed teachers to think critically about their performance and providing alternatives to what the observers have seen in the classroom during their class visit(s). It can be argued that the ultimate aim here is developmental rather than judgmental. As such, the relationship between the observer and the observed teacher is one of equality, where no superior role is taking place.

Seldin (2006) suggests three phases for classroom observation: first, the pre-visit consultation session, where the observer generally reviews the syllabus and other

relevant materials with the observed teacher; second, the visit itself, where the observer visits the classroom and observes the performance of the teacher; and, third, the follow-up session, where the observed teacher meets with the observer in order to discuss ideas after the class visit to help improve the performance of the teacher. For the purpose of easing what is to be observed and discussed throughout these three phases, basic checklists are usually considered and a wealth of literature in this area can be found such as Weimer, Parrett, and Kerns (1988), Braskamp and Ory (1994), and Chism (1999). Despite its ease of use, Zepeda (2014a) claims that it is essentially challenging to reduce all actions and words down to a predetermined list and categorize them in one form. I believe that checklists may be useful to look for patterns of occurrence, but tend not to offer sufficient space to describe details of the lesson observed.

Weaknesses that are associated with classroom observation as a method for evaluating teachers' performance have received attention in the literature. For instance, it is not an easy task to make sure that the piece of teaching that is being observed is a proper representation of everyday practice. Teachers' performance inside the classroom may also be altered by the presence of the observer within the classroom environment and disturb learning accordingly. Arreola (2006) suggests that scheduling multiple visits, training a special team to carry out classroom observation, preparing the students, preparing the instructor, and scheduling a post-observation conference can help in addressing some of these drawbacks.

Another weakness of classroom observation, which observers do not usually pay attention to, is that students' behaviours inside the classroom may be affected by the



observer's presence in the classroom. In line with this argument, Skaalvik and Skaalvik (2015) claim that reduced accomplishment and loss of self-efficacy have been reported inside classrooms by class-teachers in their study. Interestingly, some of the 30 teachers in their study have reported that they are highly satisfied with their jobs as teachers, yet they suffer severe stress and exhaustion because of negative feedback after class visits. In the context of TESOL in higher education, Shah and AlHarthi (2014) argue that EFL teachers perceive post-observation feedback as being threatening and causing undesirable stress due to the observers' lack of training and relevant qualifications. I believe that, with such an unpleasant atmosphere of threat in the work place, teachers are likely to suffer different levels of job insecurity. To this end, Al Boqami (2014), in his investigation of 3603 participants from King Abdulaziz University, asserts the importance of job security in promoting creativity for academic staff members, which will ultimately enhance students' learning.

### **3. 3. 2. 3. Self-Evaluation**

Though not widely used as a pattern for teacher evaluation in many educational programmes, when compared to the previous methods, self-evaluation can be a worthwhile source for evaluating teachers. The teachers themselves seem to perceive the lack of self-evaluation as a weakness in any teacher evaluation system (Towe, 2012). Self-evaluation can be significant, given that teachers critically analyze their own instructional practices, which will help towards their professional growth. Doherty (2009) argues that teachers have different perceptions from administrators about the teacher assessment process (TAP). In fact, the teachers in his study perceived teacher evaluation as a tool to reflect on their own performance for developmental purposes, while the administrators tended to see it as a tool to inform decision-making.

Both views, however, may be acceptable where teacher evaluation serves both purposes in educational institutions.

A major criticism of self-evaluation as a method for teacher evaluation is that teachers tend to give themselves higher ratings than they deserve (Centra, 1977). In line with this and in a more recent study, King (2014) argues that more clarity is needed when teachers' development is to be considered to follow any teacher evaluation scheme. The study raises questions on how to ensure that teacher development programmes will make a difference to teachers' growth. King's study reached the conclusion that "teachers' retrospective recollections" help when a school review needs to gather evidence, prioritise what they want to accomplish, set targets, and finally devise plans in order to get there. This may confirm the usefulness of teacher self-evaluation as an effective tool towards the improvement of instructional practices. With regard to this point, Centra (1977) made an early claim that teacher self-evaluation cannot be used for making certain decisions such as promotion. In higher education, Abu Arrab and Gadadah (2008) argue that faculty members play a significant role in enhancing the teaching quality in educational institutions. In light of their argument, self-evaluation can be more effectively applied with other methods to have a more effective and comprehensive evaluation scheme.

As an example of the suggested methods for teachers' self-evaluation, Robles-Gómez et al. (2015) suggest using an online continuous assessment and self-evaluation academic users system. By encouraging the use of internet-based platforms in the Spanish University for Distance Education, they claim that they have developed an online tool that can enable facilities to implement a new learning and teaching service

system. From their point of view, their proposed model enables faculty members to track the progress of their students as well as to assess their practical activities but not their theoretical knowledge. Their findings are significant in supporting the claim that faculty feel more confident with the use of that system, namely the Technology Acceptance Model, TAM (Robles-Gómez et al., 2015, p. 99), which is used to measure the acceptance of the model. However, it seems that there are a number of justified limitations surrounding the study. First of all, their study gives more weight to students' points of view than to faculties' opinions; the study includes 52 students and only 12 faculty members as its sample. Second, the system is designed to satisfy the needs of postgraduates in the College of Computer Science at their university; yet, it could be applicable in other contexts to enhance the investigation of faculty statements, by using questions adapted to measure faculties' response in terms of usefulness, ease of use, or any relevant aspects that can be connected to the needs of faculty members.

To conclude the discussion on faculty self-evaluation, linking self-evaluation to students' feedback using the available technological support and services seems to be a worthwhile conclusive judgement from different perspectives. However, whether teachers accept the use of technology to have more ideas about their own performance and to facilitate their own professional development remains a debatable issue.

### **3. 3. 2. 4. Teaching Portfolio**

Teachers frequently create teaching portfolios as a means of collecting educational documents and to provide practical evidence, instructional materials, and both teaching and learning records. Portfolios can also be used to reflect on the individuality

of teaching and are often more effective where they can be shared with other teachers working in a relevant field. Seldin (2006) claims that developing teaching portfolios can allow educational contexts to connect intended pedagogical theories with current teaching practice, where teachers are able to provide a “natural outcome” for improvement to take place (p. 114). This suggests that teachers’ portfolios are more useful for developmental purposes. In line with this, Davison and Leung (2009) argue that portfolios can be used as one of the methods for formative assessment; however, this may be a misconception, as teaching portfolios can also be used in an external larger-scale set of standards to assess the performance of language teachers. This, however, can be a challenge in those educational institutions that have a large number of teachers.

One of the key concerns with this approach to assessing the extent to which language teachers are teaching effectively is that it is highly dependent on how teachers are able to present their work in the portfolio. Teaching portfolios as such can be part of the faculty members’ identity and the way they exhibit themselves. Some researchers who have tackled the issue of teacher evaluation in the literature, such as Seldin, Miller, and Seldin (2010) and Davison and Leung (2009), agree on this point. However, Seldin et al. (2010) provide more endorsement for the role of portfolios in teachers’ identity and how teachers perceive their roles in educational systems. Based on these views, a very high level of trust between teachers and principals is required, as claimed by Arreola (2006); it seems reasonable to suggest that teachers need to be trusted, especially in reporting and documenting their own work for performance evaluation purposes in order to have better practices.

It can be argued that no one is more aware of their practice, preparation and readiness for teaching than teachers and faculty members themselves. Accordingly, I believe that faculties and teachers involved in teaching EFL in higher education need to be clear, transparent, and honest in conducting their teaching portfolio. By doing so, they will help to make the portfolio useful and serve its purposes. Part of this honesty and transparency involves including results from formal evaluation where that practice is established in educational institutions. In line with this and within the context of TESOL, Troudi and Rich (2012) argue that “while taking different formats such as direct classroom observation, forms with a set of items and criteria, and self-observation summaries, it is important to keep a record of the reports written by one’s evaluators” in their teaching portfolios (p. 61).

Besides reports written by others, language teachers may also need to include their own reflection as an important part of the portfolio. According to a study by Korenev, Westbrook, Merry, and Ershova (2016) on a language teachers’ target language project, which was based on responses from 670 participants, more than 60% of them reported that they perform reflection on their teaching. They claim that the materials of English language courses should help teachers to develop “interactive communicative competence”, part of which involves reflections that can be incorporated with other evidence in their teaching portfolios.

In higher education, there may be certain areas where the productivity of teachers cannot be easily measured. For instance, Kim, Wolf-Wendel, and Twombly (2011) claim that it is not easy to measure or keep records on faculty member research productivity; yet measures like conference presentations, grants, and publications can

be considered and included in teaching portfolios. Seldin (2006) also admits the complexity of reporting on teaching practices and the individuality of faculty members' teaching. Thus, the purposes of teaching portfolios may need to receive considerable attention before implementing them. Personal improvement programmes, evidence for personnel decisions, application for awards, applications for new positions, or other purposes towards senior faculties' retirement are the main purposes for teaching portfolios of faculty members, according to Seldin, which brings us back to the main issue of how to evaluate teachers, bearing in mind the main purposes behind not only the teaching portfolios of faculty members but also the reasons for creating any teaching record.

### **3. 3. 3. EFL teachers: Evaluation and Special Concerns**

Research in the field of teaching English as a foreign language deals with different concerns related to language learning, curricular aspects, teaching and teachers' performance. In their theoretical article, Brown and Crumpler (2013) claim that there is little agreement as to what makes a method of language teachers' assessment effective. This problematic issue can affect the evaluation of teachers of foreign languages. Brown and Crumpler argue that this evaluation faces many challenges, especially for evaluators who do not have sufficient knowledge about second language acquisition. This issue becomes more complex when those evaluators do not speak the target language that is used inside the classroom. Despite the fact that in such cases it may be challenging for classroom observers to reflect on the content knowledge of the teacher, it becomes more challenging to observe the degree to which students comprehend the information provided. It is worth mentioning that senior

teachers can observe teachers' performance in foreign language classrooms in different ways and using different methods, for instance, classroom observation checklists, which include the content knowledge of foreign language teachers as merely one of the criteria for teacher performance assessment. This might be considered insufficient.

Brown and Crumpler (2013, figure 3.3., p. 80) have developed a model that positions the assessment of foreign language teacher portfolios at the heart of teacher evaluation in order to shift teacher evaluation towards a more learning and progression model. Their assessment portfolio model offers an inclusive and wide-ranging approach to instructor's performance assessment that is informed by "multiple sources of evidence, which leads to a more complete and authentic" (p. 145) evaluation. They argue that administrators, due to their busy schedules, cannot supervise and evaluate foreign language teachers in a proper sufficient way. Hence, they pay more attention to the teacher portfolio, where teachers themselves keep records of their teaching evidence and materials. In fact, the model proposed by Brown and Crumpler tends to overlook "self-evaluation" as an approach for a better system of foreign language teacher evaluation. If it were added to the model, teachers would be able to diagnose and reflect on their teaching as a whole, which can be seen as a better attempt to improve the quality of the teachers' performance.

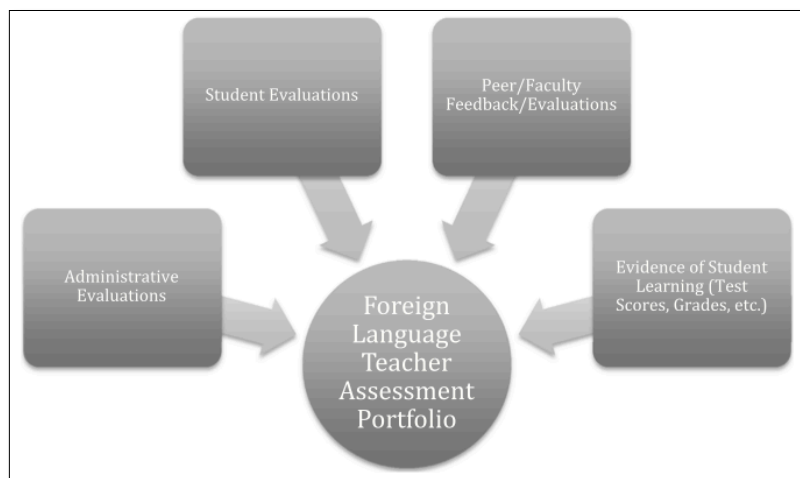


Figure 3. 3. Foreign language teacher evaluation, Brown and Crumpler's Model, 2013

In an attempt to examine the main criteria of in-service English language teachers' evaluation, Akbari and Yazdanmehr (2012) conducted an exploratory study in five private language institutions in Iran. Interviews with the supervisors along with analysis of application forms, observation sheets and other relevant documents illuminated the procedures and criteria of teacher assessment in the target setting. Their suggested procedures in assessing in-service English language teachers' performance are categorized into four groups: teacher's command of English, teaching skills, compliance with the syllabus and personal/affective features. The model they developed exclusively for English language teacher evaluation is presented in figure 3.4 as follows:



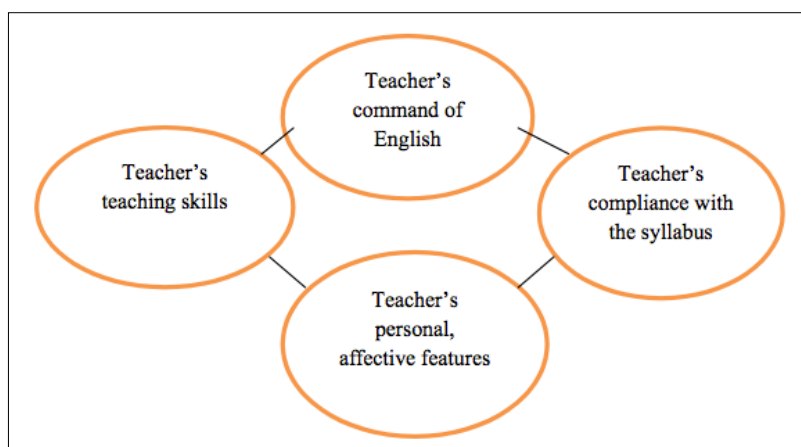


Figure 3. 4. English language teacher evaluation, Akbari and Yazdanmehr, 2012

In Akbari and Yazdanmehr's (2012) model, the teacher's command of English involves accuracy of speech, structure, pronunciation, and performance in discourse along with fluency in speech. Personal/affective features include punctuality, rapport with learners, tolerance in error treatment, enthusiasm and dynamism in involving learners. Teacher's compliance with the syllabus comprises expected content to be covered, educational goals to be achieved, and the way of presenting the material to be followed. Teaching skills involve communication skills, class management techniques and task management. Their model might be seen as distinctive and uniquely designed for EFL teachers but it fails to consider other social and administrative duties/skills as well as the community services and research activities that frequently form an essential part of EFL faculty members' activities and which should not be overlooked.

In a survey conducted with 457 post-secondary foreign language teachers, Bell (2005) examined their perceptions on the teaching attitudes and behaviours contributing to effective foreign language teaching. The participants were French, German, and Spanish language teachers at the American Council on the Teaching of Foreign Languages in New York. Her study demonstrated a strong positive agreement on all five standards for foreign language teaching. Other categories in which the teachers agreed with the majority of items include teachers' qualifications, general theories related to the communicative approach to foreign language teaching, the significance of small group activities, and negotiation of meaning and strategies in foreign language classes. In fact, the study was more concerned with the teachers' behaviour and attitude towards aspects highly related to language acquisition rather than contributing to the effectiveness of foreign language teachers and teaching.

Brown (2006) investigated language teachers' and students' perceptions of effective teaching in foreign language classrooms, which, he argues, are distinctive from the classrooms of other subjects. People involved in EFL teacher evaluation may find such studies on effective language teaching useful in giving them insights on teaching qualities they need to pay attention to when they evaluate teachers. Brown's (2006) findings are the result of analysing a questionnaire distributed to 49 university language teachers and 1400 of their students at the University of Arizona. From the teachers' perspective, engaging students in information-gap activities, assessing group tasks, being as knowledgeable about culture as language, and having students respond to physical commands are the main characteristics of effective foreign language teachers. Concerning students' opinions, correcting oral errors indirectly,

being as knowledgeable about culture as language, having students respond to physical commands, addressing errors with immediate explanation, presenting grammar with real-world contexts, speaking with native-like control of language, using real-life materials in teaching language culture, and engaging students in information-gap activities are the most prominent features of effective foreign language teaching. Arguably, Brown's study can be seen as being more concerned with instructional practices and disregarding other areas that can be utilized to evaluate language teachers.

Al-Hammad (2011) conducted a study aiming at examining the teaching performance level of 18 English language teachers from intermediate-level schools in the city of Hail, Saudi Arabia, according to the teaching quality standards. By employing an analytical and descriptive approach, Al-Hammad applied a controlled observation method on teaching standards and found that teachers were competent in the use of teaching aids and class management skills. In that study and within her sample, students' assessment and lesson delivery were satisfactorily accomplished. Lesson planning, however, was achieved by her participants to only the lowest quality standard. Al-Hammad's study reinforces the importance of conducting in-service training sessions on teaching quality standards for English language teachers mainly in the three dimensions of planning, implementation and assessment. Despite the fact that the sample was limited to English language teachers, the dimension and the criteria were not subject-content oriented and could be applicable to teachers of other subjects. In the Gulf context Al-Mahrooqi, Denman, Al-Siyabi, and Al-Maamari (2015) compared Omani school students and teachers' perceptions of the characteristics of

good EFL teachers. 171 Omani students and 233 English teachers took part in the study, which showed general agreement between students and teachers about the importance of all characteristic categories, with special importance given to English language proficiency and equality in treating students.

To conclude, a number of studies with different focuses and approaches to teacher evaluation can be found in the literature, each of which has its own interest and philosophy under the umbrella of a specific paradigm and school of thought. In the following section, the two main approaches to teacher evaluation that are informed by different schools of philosophy are discussed.

### **3. 4. EFL Teacher Evaluation: Differentiated Models**

The issue of teacher evaluation has received considerable critical attention and different models related to the various philosophical and theoretical evaluative rationales behind them can be found in the literature. In the following section, models that can be utilized to evaluate teachers of English as a foreign language are divided into two main categories: first, the technocratic accountability-oriented model, and second, the democratic model.

#### **3. 4. 1. Technocratic Approach to Teacher Evaluation: Accountability and Teacher Evaluation**

Towards the end of the last century, the reflections of New Public Management continued to be intensified by audit cultures (Power, 1997). Ryan (2004) claims that educational accountability “reflects this audit intensification, particularly since the passage of the No Child Left Behind Act of 2001” (p. 443). In fact, such legislations

may rely heavily on performance indicators as their principal strategy in order to institutionalize the audit practice and to improve students' learning and achievements. It is through monitoring that these performance indicators assist stakeholders and the public. However, and within such audit cultures of accountability, teacher evaluation, which is closely related to classroom observation and supervision, has been given the "negative connotations of the word *inspection* in the EFL contexts" (Ashour, 2011, p. 43).

According to Power (1997), the main characteristic of educational accountability is that it is a hierarchical top-down mechanism that reflects an atmosphere of efficiency and control in an attempt to improve educational achievements. Power (1997) claims that this accountability approach can be labeled the "new rationality of governance" (p. 10). In the same vein, Ashour (2011) adds that when teacher evaluation tends to be top-down in the context of ELT, it becomes hierarchical, power-based and bureaucratic. In the context of education, Glanz (1998) argues that bureaucracy consists of "a hierarchy of authority, prescribed rules, centralized decision-making and procedural specifications" (p. 45). Robinson (2009), likewise, confirms that when we think of teacher evaluation within bureaucratic educational systems, we consider it in terms of hierarchy, authority and control. As such, teacher evaluation might be seen as a *pro forma* with a negative touch of its artificial practices. Educators, under such bureaucratic teacher evaluation schemes, will tend to view themselves as subjected to vigilance by a central authority whose ultimate goal is to judge them against a set of predetermined standards.

It is worth noting that an example of the attempts of teacher evaluation within the trend

of accountability in education was found in 2004 in Ohio through the report of the Governor’s Commission on Teaching Success (Downing, 2016). That report was later followed by a call for an Educator Standards Board of Education. In 2009, the Educator Standards Board was directed by House Bill 1 to recommend “model evaluation systems for teachers and principals” for their adoption and review. In response to that, The Ohio Teacher Evaluation System (OTES) was created and designed by the Ohio Department of Education (ODE), to assess the Ohio teachers’ performance (Ohio Department of Education, 2017). The Original OTES Framework for 2015–2016 includes two components: teacher performance on standards and student growth measures, each accounting for 50% of the overall structure (Figure 3.5. below).

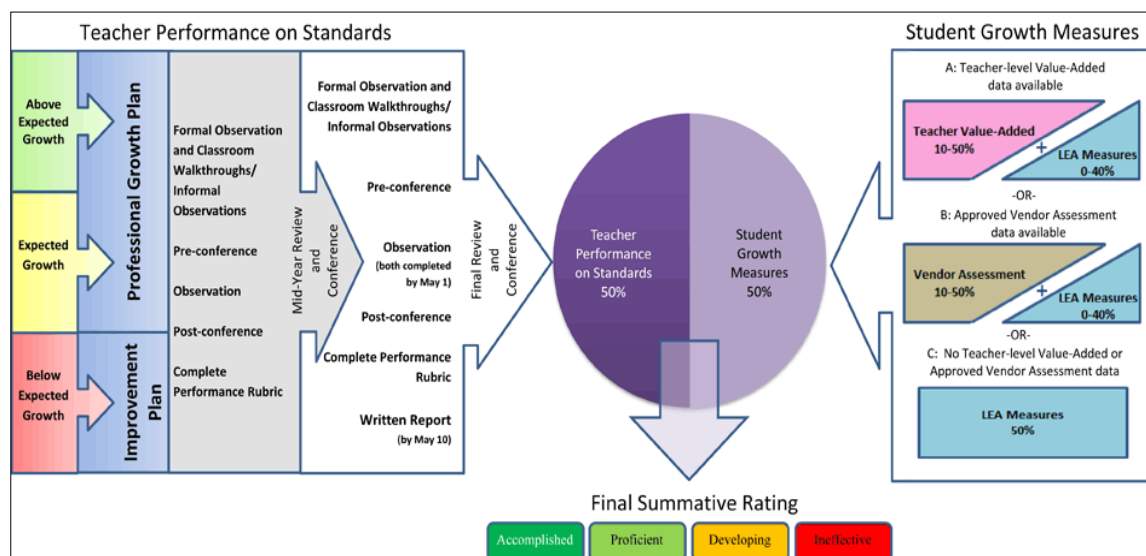


Figure 3. 5. Ohio Teacher Evaluation System (OTES)

Teacher performance on standards, which is the first component, can be determined by three aspects: a professional growth plan, a mid-year review of formal and informal observations, and a final review of formal and informal observations. The processes

require the evaluator to use data collected from multiple sources as evidence. These involve professional improvement or growth plans, formal or informal observations, walkthroughs, conferences and/or any other artifacts and evidence the teacher is happy to share with the evaluator. Pre- and post-observation conferences are not compulsory but are contemplated as rich sources of evidence for best practice. Furthermore, the post conference is seen as a reflection opportunity for both the teacher and the evaluator to decide on areas for reinforcement and refinement. Refinement areas are usually anticipated to drive decisions regarding professional development for that instructor (Ohio Department of Education, 2016).

Student growth measures, which form the second component, also carry a weighting of 50% of the overall structure of the system. Student growth measures are a mechanism for determining the degree of academic gains students have made (Ohio Department of Education, 2017). This can be done through the calculation of student growth between two points in time. There are three methods, each of which can be used to measure student growth: teacher-level value added, approved vendor assessment data, or Local Education Authority (LEA) measures. The LEA measures can only be used when data from the other two methods are not available (Ohio Department of Education, 2015) and they include Shared Attribution or Student Learning Objectives (SLOs). Shared attribution is a student growth measure that is shared by or attributed to a group. It can be helpful in supporting collaboration for meeting school goals. SLOs are another way to measure the influence of a teacher on student learning; they are typically intended to cover an interval of instruction that is long-term and includes a target for academic growth for each and every student (Ohio Department of Education, 2015). Finally, data from both components, 50% each, are

used to determine a teacher performance rating of Accomplished, Skilled, Developing, or Ineffective (figure 3.5, p. 86.).

It is worth noting that the movement towards accountability was one that grew in momentum and vastly increased in volume. Downing (2016) claims that the nationwide movement to call for accountability was technocratic and efficiency-oriented in nature. Kyriakides and Demetriou (2007) earlier argue that, within this accountability movement, a teacher evaluation system is not only a kind of technocratic affair but is also “determined by political influences” (p. 45). Accordingly, and with the political influence in mind, many schools and educational institutions in the US developed closely identical programmes in order to ensure accountability in education, with nearly all of them having employed data and results from large-scale assessment (Popham, 2000).

Despite the fact that there were opposing premises of democratically oriented teacher evaluation, many evaluation commissioners rejected the democratic “value-committed stance for evaluation” and favoured what House and Howe (1999) call ‘stripped of values technocratic evaluation’. These critics, clearly, have confidence in the impartial objectivity of the standards, which were needed for the support of the concepts of contemporary accountability such as results-based management, evidence-based decision-making, and teacher performance indicators (Greene, 2007). All these reinforce the climate of control associated with technocratic evaluation within the accountability movement. On the other hand, there have been different movements against this stream of accountability. For instance, Ravitch (2016) fears that results-based teacher evaluation systems could create a toxic atmosphere amongst teaching



staff and eventually colleagues will tend to compete with each other. She also warns that the accountability movement is designed to improve the welfare of companies, which profit from the movement, rather than the welfare of the students.

So far, however, there has been little discussion about technocratic teacher evaluation in the context of TESOL. It seems that, to date, no large-scale studies have been performed to investigate the prevalence of technocratic approaches to evaluate EFL teachers. A qualitative study was conducted by Mahamadou (2010) in the Republic of Niger to explore supervisory practices in the evaluation processes of EFL teachers in Salma city. The sample included 9 EFL teachers and 3 supervisors and the researcher employed observations, interviews, group discussions and journal logs to collect the data. Mahamadou's study revealed that EFL teacher evaluation in Niger is summative in nature and technocratic in approach. Accordingly, it is performed for administrative purposes in a bureaucratic way. The findings indicate that the only method for EFL teachers to obtain feedback after classroom observation was the directive controlled method, irrespective of the teachers' teaching experience. In that method, the supervisor mainly controls the feedback, which puts supervisors in a power position, giving the impression of them being the expert, and the teachers need to follow the supervisor's guidance closely. Mahamadou (2010) claims that this clarifies the tension between teachers and supervisors as that method may only suit novice teachers. Responding to the participants' request for a democratic supervisory approach, the study recommends listening to the teachers' desire for democracy, freedom, collegiality and dialogue in the supervisory processes in the city of Salma.

These ideas amongst others help towards the development of contrasting movements

and approaches to teacher evaluation, one of which, the democratic way, will be discussed in the following section.

### **3. 4. 2. Democratic Approach to Teacher Evaluation**

Within the realm of evaluation, the democratic approach can be seen as empowering people to initiate, take corrective action, and control. According to Bailey (2016), “reliance on principles of scientific management as a means of regulating pedagogical practice undermines the democratic and liberating aims of education through the application of business-oriented systems” (p. 143). I believe that the use of a predetermined set of checklists makes teachers more dependent on evidence-based practices instead of being involved in making decisions. In line with this, Trujillo (2014) argues that removing professionalism from teachers frequently results in changing instructional rules dramatically and conceptualizing them as a merely managerial practice. As such, designating the roles of the teachers by virtue of such a limited paradigm may be pedagogically unsound.

It seems that lapses within technocratic approaches to evaluation resulted in resistance from within the field of education, which eventually sparked teachers’ call for involvement in various ways. In this respect, Gutmann (2014) argues that technocratic teacher evaluation flaws in the USA set the stage for contests against management and encouraged teachers to call for systems that are more democratic. The doctrine of technocratic evaluation was thought to set teachers as cogs in the complex factory of education at a time where the principles of scientific management “operate vis-à-vis the technical control and de-skilling of teachers” (Au, 2011, p. 36). As such, it was thought that democratic evaluation could introduce ways of engaging

teachers and encouraging them to think critically about their instructional practices, to empower them, and in particular to strengthen social justice. However, perspectives, interests and values of other stakeholders may remain as core issues in the procedures. Accordingly, one of the main characteristics of democratic evaluation is that it is dialogical, as an evaluator should not presume that he/she knows how stakeholders think without engaging them in intensive dialogue. This will eventually help the evaluator to define the situation and stimulate reasonable solutions.

One of the first theorists to call for and write about democratic evaluation was the British evaluator Barry MacDonald, who published on the concept of democratic evaluation in the 1970s. Podems (2017) claims that MacDonald's ideas about democratic evaluation were guided by the premise that the public has the right to get access to enough information, which will enable them to have meaningful and thoughtful engagement. This suggests that, for an evaluation to be democratic, it should provide appropriate information to diverse and various types of audiences and stakeholders with assorted interests. Kushner (2016) goes further and suggests that, for evaluation systems to be democratic, they should involve the voice of the marginalized population. Accordingly, it seems that Kushner recognizes the complexities of political and social dynamics that an evaluator needs to pay attention to in order to ensure the engagement of those stakeholders who are marginalized.

The issue of democratic evaluation has been a focus of attention for researchers and practitioners during much of the last decade. Like MacDonald, Earnest House campaigned for a deliberative democratic approach to evaluation that addresses social class inequities, involves minority cultures, and advances social justice. House's

work has engaged strongly with the political spirit of evaluation. House and Howe (1999) claim that values, facts, and methods of evaluation are intertwined with politics. Thus, it can be argued that, besides informing institutions, evaluation can contribute to shaping their content and morale.

In collaboration with philosopher colleagues, Hanberger (2004) goes further and argues that, besides the idea that democratic evaluation places informed stakeholders at its core, it is expected to evolve around confidentiality, negotiation and criticality. However, different democratic evaluation systems may tackle these issues differently. According to Hanberger, there are three different democratic evaluation orientations: elitist, participatory, and discursive democracy-oriented evaluation, which I will discuss in the following section. Lee (2016) offers a visual representation of the three democratic orientations and identifies their positions in terms of inclusion, dialogue, and deliberation (Figure 3.6. below).

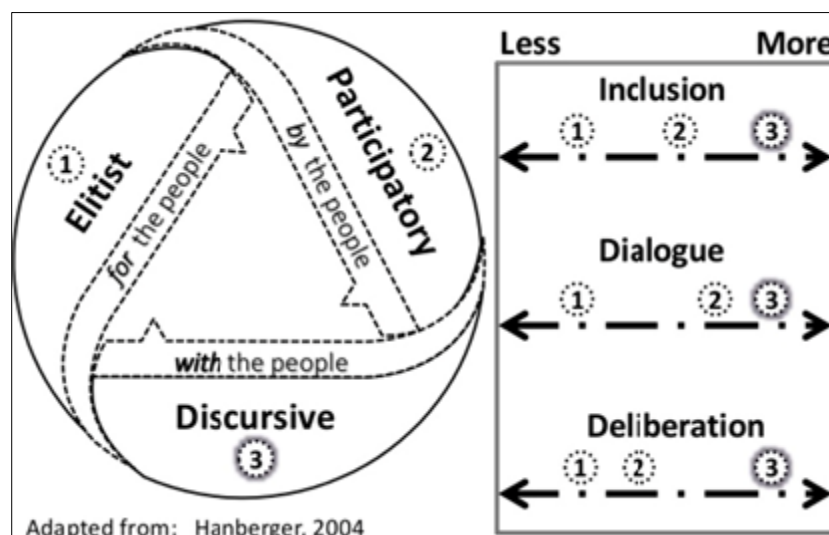


Figure 3. 6. Lee (2016) of Hanberger’s democratic evaluation orientations

According to Schumpeter (2013), political elites grapple with one another for power in open societies within the theory of *elitist democracy*. In the literature, this theory is sometimes referred to as the Lockean or the liberal view and it assumes that people have the ability to control those who rule over them by means of choosing amongst competing elites (Habermas, 1994). The main principle in elitist democracy is that of *representative democracy* (Hanberger, 2004), which is embedded in technocratic policy and/or expert-oriented practice (Hanberger, 2001).

The participatory democratic evaluation is derived from the *participatory* theory of *democracy*, which assumes that people's participation in making decisions is the most significant value of democracy. Groen (2017) claims that participatory evaluation is one of the forms of collaborative social inquiry that is based on problem-solving. Stakeholders' participation in the process of policy-making frequently involves political aims that intend to endorse freedom of choice and to "promote fairness through the involvement of individuals associated with all groups" (Weaver & Cousins, 2007, p. 21).

Discursive democratic evaluation, sometimes called "*deliberative democracy*" (Van Aaken & List, 2017), is established in a form of democracy linked with public commitment to an approach of reasoning on concerns of public policy (Hanberger, 2004). Peters (2017) argues that it is through deliberation that discursive democratic evaluation can help people to change their beliefs, habits and actions. It can be argued then that this change can take place by means of communication without domination.

Surprisingly, teacher evaluation within the TESOL context has not yet been closely investigated. As far as I am aware, there is only one study, looking at theoretical

aspects of EFL teacher supervision in Syria that discusses the oppression of teachers and calls for a democratic approach to reviewing EFL teachers' performance. Ashour (2011) argues that "dialogue, trust, mutuality, democratization, and love are unfortunately absent, or made absent by those in power, from teacher-supervisor relationships in the Syrian educational system" (p. 78). In his opinion, educational systems should abolish the dichotomy of the 'teachers-do not-know and the-supervisor-knows' autocratic approaches to the evaluation of teachers. Ashour (2011) adopted Freire's ideas of critical pedagogy: that it is the role of education to empower and free the whole of society. Freire (2014) believes that teachers should implant hope in their educands (the term Freire used for students), that there is always a better tomorrow, and that democracy is always possible, no matter how tough the oppressors, who they are and how much they are deprived of their humanity.

Prior to that, Freire (2000) argues that "those who have been denied their primordial right to speak their word must first reclaim this right and prevent the continuation of this dehumanizing aggression" (P. 88). Thus, Ashour (2011) claims that dialogical and collaborative relationships between EFL teachers and supervisors are needed to replace the existing power relationships. Accordingly, his study recommends that, in practical terms, liberation needs the involvement of all people, and their determination and willingness, irrespective of any differences amongst them. The study also recommends that teachers and supervisors should get involved in special training workshops, programmes, and regular meetings where ideas can be shared respectfully.

It is worth mentioning that critics have reported different challenges to democratic

approaches to teacher evaluation. For instance, Groen (2017) argues that this type of evaluation requires considerable time and planning, given the involvement of various stakeholders. Thus, the commitment of the participating stakeholders seems to be a key to the success of this type of evaluation. Moreover, Weaver and Cousins (2007) claim that selecting stakeholders in the democratic type of evaluation can be problematic. The degree of authority and control that different stakeholders have over the data obtained can be another issue and there may be a risk of information misuse. Cousins (2004) warns that, whether intentionally or unintentionally, there is always the hazard of misuse of data, considering the involvement of stakeholders in interpreting and reporting the findings of this type of evaluation.

### **3. 5. The Conceptual Framework of the Study**

This study aims at exploring and examining current practices of EFL teacher evaluation implemented in Saudi Arabia at tertiary level with some elements of criticality that challenge the contemporary ideologies, whether selected by or imposed on the participants. The body of research discussed above helps in providing different models to feed into the conceptual framework, which is also informed by the constructive theoretical framework, as will be discussed in the forthcoming chapter.

In light of the objectives and questions of this study, a critical approach to EFL teacher evaluation seems appropriate. This stance is mainly concerned with questioning policies and practices related to teacher evaluation and problematizing the taken-for-granted givens. Pennycook (2004) highlights that, in a critical approach to TESOL, making things socially relevant and raising awareness are fundamental steps in the desired change processes. It aims not only to describe what is wrong but also to

propose various ways for change. The conceptual framework has been informed by some of the reviewed literature such as MacDonald's 1970s ideas about the meaningful engagement of stakeholders and Hanberger's (2004) opinions on criticality of this engagement. As discussed in the literature review, critical approaches to teacher evaluation can help empowering teachers to raise their voice and advance taking corrective actions. I believe that critical notions help moving teachers from being seen as merely gears in the complicated factory of education to empower them as agents of change and to engage them to think critically about their teaching practices. This becomes a necessity in contexts where teachers, amongst other stakeholders, are being marginalized given the complexities of political and social dynamics.

Moreover, this critical vision helps the TESOL researcher to focus on the extent to which dominant ideologies enforce the creation of meanings and understanding that results in privileging some and oppressing others (Hawkins & Norton, 2009). Thus, in putting emphasis on awareness of what a critical approach might be, researchers and teachers become transformative agents who can support the engagement of different partners collaboratively in educational activities, making them aware of the social construction of the givens and accordingly questioning them. It is worth mentioning that, within the boundaries of this research, criticality is an approach that evolves through exploring the experience of EFL teachers (the participants) rather than a prescriptive theory. The ultimate goal of this critical approach is to naturally appeal to EFL teachers' awareness of policies and questioning practices around teacher evaluation and other critical issues such as marginalization and discrimination. This critical approach is also utilized in an attempt to maximize EFL teachers' representation and legalize the role of teachers in order to empower them. This can



eventually lead to authentic learning (Martin, Breunig, Wagstaff, & Goldenberg, 2017) and agency (Akbari & Tavassoli, 2014).

On that account, from a critical point of view, the success of EFL teacher evaluation can be determined by observing how those who either are using the policies or being affected by them react to them in a certain context. I believe it is not sufficient to examine the policies away from the context in which they are applied and the people who are in direct contact with and affected by them. Thus, the current practice of EFL teacher evaluation and its policy should be critiqued based on the Saudi context in order to see how successful it is when applied in that specific context and to be in an appropriate position to seek suggestions and solutions to the existing challenges.

### **3. 6. Summary of the Chapter**

In this chapter a body of research in the literature on teacher evaluation have been reviewed, in order to show how evaluation has been conceptualized in education in general and higher education in particular. It was my intention to link it to the context of the study in a reasonable logical approach to help the development of tools aimed at finding answers for the research questions. The chapter moves gradually through discussion of concepts, theories and models for teacher evaluation in order to inform an appropriate conceptual framework to facilitate a proposed teacher evaluation model for faculty members who are involved in EFL teaching in Saudi higher education. The model is to be proposed based on the results of this study to enhance pedagogical aspects of EFL teaching in Saudi universities after the data collection and analysis are complete. In the next chapter, the theoretical framework, research design, methods utilized, and tools designed to collect data for this study are outlined.

## **Chapter Four**

### **Methodology**

#### **4. 1. Introduction**

In order to capture the participants' unique experience to help in achieving the objective of the study, namely to explore the challenges faced by EFL teachers at Saudi public universities under the current evaluation systems, a mixed method sequential design is applied. Quantitative followed by qualitative methods of data collection and analysis are utilised following a participant–selection variant of the mixed method sequential design. Both quantitative and qualitative data are important to this research and expected to contribute equally to the study. Nevertheless, the emphasis is on the qualitative phase, and this has greater priority since the study is exploratory in nature and qualitatively oriented since it is informed by the interpretive paradigm. The quantitative phase employs a cross sectional survey to answer the first and second research questions. It also provides some information which is helpful in answering the third question. The items of the questionnaire were developed partly from the related literature while other items were driven from my personal experience with teacher evaluation in the Saudi context. The main purpose of the first quantitative stage is to generate data as a starting point for the second qualitative phase. The qualitative phase employs semi-structured interviews and focus group discussions to answer research questions number 3, 4 and 5. This chapter is divided into four main sections: the philosophical assumptions, the research design, the data collection and analysis, and finally the research ethics.

## **4. 2. Philosophical Assumptions and Theoretical Framework**

The study is informed by the interpretive paradigm. Interpretive research supporters believe that the objective claims and the methodological practices of scientific research reflect a number of theoretical limitations that cannot give researchers the chance to understand social phenomena and interpret human actions (Sandberg, 2005). Human beings can be distinguished from physical objects in the way they attach meanings to the world around them (Pring, 2017). As Creswell states, the meaning individuals attach to a social action is mainly grounded in their knowledge and experience, which can be considered the most prominent feature of interpretive enquiry (Creswell, 2013). When experience is to be considered, the interpretive approach is progressively applied in order to justify “experience-based methods” (Garrick, 1999, p. 147). In this study, teacher evaluation as experienced by EFL teachers in Saudi universities was interpreted in order to explore the meaning or multiple meanings they attach to this phenomenon; thus, the philosophical assumptions of the interpretive approach seem to be the most suitable philosophy to underpin the study.

Experienced realities, which might be equated with subjective states, are likely to be different from one individual to another (Johnson & Onwuegbuzie, 2004). The subjective meaning of the participants and the various ways of understanding a social phenomenon are what an interpretive researcher is looking for and aiming to explore, which is precisely the case in this study. The basics for the subjective knowledge of interpretivists’ claims are the interpretive assumptions, which are based on the distinctive and subjective stories of the participants, the language they utilise, their

descriptions of the phenomenon, and even the metaphors they use (Garrick, 1999). The subjective meanings and individual interpretation of teacher evaluation as perceived by EFL teachers in Saudi universities form the core argument of this study. From this standpoint, the interpretive paradigm fits within the philosophical assumptions of the proposed research.

Since the way people perceive the world differs from one individual to another within a social group and between social groups as well (Pring, 2017), it is the interpretivist's intention to explore these various ways and to reflect on the expected differences. This view is also held by Neuman (2002), who claims that the most prominent concern of interpretivists is the "systematic analysis of socially meaningful action" (p. 62). Therefore, reality can be seen as socially constructed (Mertens, 2005 cited in Mackenzie & Knipe, 2006, p. 3). Accordingly, new knowledge perceived daily forces individuals to form and reform how things can be understood and how social phenomena are constantly reconstructed, which reinforces interpretive subjectivity. According to Burr and Hooser (1995), in social constructionism, there is a focus on interaction, social practices, and process. Explanations for a social phenomenon can be found in the "interactive processes that take place routinely between people" (p. 8) instead of in individuals or social structures. Accordingly, the emphasis is more on the process rather than the structure of social interactions and knowledge, which can be seen as what people do together and not only what a person has experienced individually. This will be further discussed in the next section.

#### **4. 2. 1. Ontological Assumptions**

According to Crotty (1998), ontology can be defined as “the study of being. It is concerned with “what is”, with the nature of existence” (p. 10). By exploring the perceptions of EFL teachers at Saudi universities on teacher evaluation, I hold the assumption that social phenomena and their meanings differ from one individual to another. This view suggests that human beings attach different meanings to the same phenomenon (Pring, 2017), resulting in a world with multiple, possibly relative, ways of perceiving reality. This implies that there can be multiple realities that are experienced by the same person about the same phenomenon, depending on the type of social interaction or process which may result in constructing these relatively different realities. Accordingly, relativist ontology will serve as the underpinning stance to understand reality. In the present research, participants hold their own perspectives about the investigated phenomenon, developing a variety of meaningful relative realities. On the other hand, the researcher is expected to construct meaningful realities while interpreting their perceptions. All these realities are “contingent upon human practices, being constructed in and out of interaction between human beings and their world, developed and transmitted within an essentially social context” (Crotty, 1998, p. 42). In other words, teachers may have the same experience and have been working under the same conditions but, since they interact differently with their social context, they do not constitute the same reality and they develop, rather, some relative truth.

#### **4. 2. 2. Epistemological Assumptions**

Understanding what can be involved in knowing, in other words “how we know what we know” (p. 8), is the definition of epistemology (Crotty, 1998). In this research, a

constructivist approach will be the epistemological stance underpinning the study. Accordingly, an objective truth is 'not out' there, waiting to be uncovered by the researcher. When we engage with the realities in our world, truth consequently exists. Truth cannot be discovered; instead, different people construct meanings or "truth" differently.

In conducting this study, there are two levels of interpretation: the participants will provide their interpretation about the phenomenon of teacher evaluation and I, as the researcher, will provide my interpretation of their interpretation. Bryman (2015) has added a third level of interpretation, given that "the researcher's interpretations have to be further interpreted in terms of the concepts, theories, and literature of the discipline" (p. 31). As a result, one can see how reality is constructed and affected by the many levels of social interactions it goes through.

Given the exploratory nature of this research, relativism as the ontological stance and constructionism as the epistemological stance seem to be appropriate theoretical perspectives. The purpose of this study is to explore and to deeply understand EFL teachers' perceptions on teacher evaluation in Saudi universities, which assumes that participants are engaged with a reality or a phenomenon and make sense of it in various ways. In this sense, there will be no "true or valid interpretation" (Crotty, 1998, p. 47) and all diverse interpretations are appreciated and valid. Thus, the interpretive rather than the positivist approach will allow me to communicate with the participants and the collected data effectively in order to understand their interpretation of the investigated phenomenon.

### **4. 2. 3. Methodological Assumptions**

Any discussion of the ontological and epistemological assumptions of research necessitates a consideration also of the methodological assumptions, as these philosophical constructs not only affect the quality of the research, but also the appropriateness of the methods that are employed to collect data. According to Sapsford (2006), methodology can be seen as the “philosophy of methods. It encompasses, first, an epistemology, which is the ‘rules of truth’ for warranting the validity of conclusions; and secondly, an ontology establishing the ‘objects’ about which questions may validly be asked and conclusions may be drawn” (p. 176).

As discussed in earlier sections, this research is exploratory in nature and the rationale behind all the quantitative and qualitative methods applied is to explore and understand how the participants view teacher evaluation in their context. In line with this, Stebbins (2001) argues that exploration cannot be a synonym of qualitative research; he claims that exploration indicates the development of theory from the collected data, whereas qualitative research emphasises the actual collection of the data by which a researcher develops a theory. This is the case of the current research, where the gathered data helps the researcher to explore the targeted phenomenon regardless of the types of methods that have been used. It is worth highlighting that the use of the quantitative elements in this study does not indicate any positivistic intention or confirmatory purposes. As Shannon-Baker (2016) puts it, “‘quantitative’ and ‘qualitative’ should not be used as synonymous with paradigms. These qualifiers are instead more about our approaches to data and methods rather than signaling a singular worldview” (p. 320). Accordingly, both quantitative and qualitative methods were integrated to help in establishing the research findings rather than influencing the

ontological or the epistemological assumptions of this research. Thus, a systematic design for utilising both types of method was developed before commencing the data collection, as outlined in the following section.

### **4. 3. Research Design**

Creswell (2013) defines research design as “the specific procedures involved in the research process: data collection, data analysis, and report writing” (p. 20). This suggests that the researcher needs to plan a clear, well-constructed and developed research design before commencing the research procedures. This view is supported by Bryman (2015), who reminds us that for researchers, the research design serves as a “framework” for the two main stages of the research: namely data collection and data analysis. As discussed earlier, the philosophical assumptions underpinning this study have fed into an exploratory approach to answer the research questions. Given the fact that there are different variations when research design is to be considered, making a decision on a specific design can be challenging. Yet, one needs to remember that the central aim of any research is to address the predetermined set of research questions (Tashakkori & Teddlie, 1998).

#### **4. 3. 1. Mixed Methods Research**

Mixed method research is defined as “collecting, analysing, and “mixing” both quantitative and qualitative methods in a single study or a series of studies to understand a research problem” (Creswell, 2013, p. 623). As discussed in the previous section, the research aim and questions has informed the research design. As the study includes a number of Saudi universities in different geographical regions of the country and, given the fact that each of the universities has its own system for teacher



evaluation, commencing with a questionnaire to have a general view about the trends seems the most appropriate approach to collect the required information. Since this research is exploratory in nature, the qualitative stage then serves to provide a deep understanding of the explored phenomenon.

The mixed methods design employed in the current research includes two phases: a quantitative phase followed by a qualitative phase. The rationale for mixing both quantitative and qualitative data is that neither the qualitative nor the quantitative approach alone is sufficient to capture the participants' experience and/or perception on EFL teacher evaluation in the five Saudi universities selected. To enhance the breadth of this study and to give a wider understanding of EFL teachers' evaluation in Saudi higher education, a quantitative approach is suitable to cover a considerably larger number of participants. As for the depth of the data that is needed to address the predetermined research questions and to satisfy the exploratory claim of the research, a qualitative approach to collect the data is suitable. Brown and Crumpler (2013) claim that researchers who do both qualitative and quantitative research, particularly in TESOL, have a considerable number of advantages compared to other researchers considering only a qualitative or quantitative approach. In the same vein Tashakkori and Teddlie (1998) remind us that quantitative and qualitative methods complement one another when they are used in combination and provide a more comprehensive picture about the social phenomenon. Given these advantages of mixed methods research, the subsequent step after selecting the design is to determine which variant to select that could best fit the research philosophy and objectives. The following section gives more details on this issue.

### 4. 3. 2. Sequential Mixed Methods Design

Creswell, Gutmann, and Hanson (2010) claim that most of the writings about mixed methods design tend to place more emphasis on sequential two-phase approaches. According to Creswell (2017), the explanatory sequential mixed methods design is:

a mixed methods strategy that involves a two-phase project in which the researcher first collects quantitative data and then follows up or builds on this database with a second qualitative data collection and analysis (p. 243).

The design chosen for this study is a customised version of one that is labelled by Creswell (2017) “*explanatory sequential*” (p. 243), where the qualitative data is used to explain the quantitative findings. However, this is not the formula used in this research; instead, the quantitative data is intended to give more breadth to the study, a better understanding of the explored phenomena and its practices in different variants of the Saudi context, and to provide valuable insights into the issues that require more attention by researchers. For the sake of this research, the less common *participant-selection* variant is applied and not the more common *follow-up explanation* variant. Unlike the follow-up explanation variant, in the participant-selection design, the priority is on the second “qualitative phase instead of the initial quantitative phase” (Creswell, 2017, p. 86). Having considered the philosophical assumptions underpinning this research discussed earlier, I came to the conclusion that the participant-selection design would fit better with my study. Creswell emphasizes that this variant may produce more accurate results than the other variant for studies that have a focus on the qualitative inquiries to investigate a phenomenon. However, an initial quantitative data stage is added, based on the fact that it is needed to inform the

second qualitative phase and to identify and to purposefully select the appropriate sample of participants.

The aforementioned method is used in this study where the quantitative phase is used for the purpose of informing the qualitative main phase. The main qualitative approach of the current study is compatible with the philosophical assumptions for the research's main purpose, which is exploring EFL teachers' perceptions, while the other set of data can be seen as subservient within the mixed methods sequential design (Creswell, 2014). Furthermore, the quantitative method of the first phase serves an exploratory purpose as it gives an idea about the general trends in a number of institutions and what people think about teacher evaluation. It is worth mentioning that the main approach of this research is qualitative in nature, where the study will be informed by a constructivist approach to knowledge, and this design fits well for that purpose. A representation of the sequential mixed-methods design can be found in the following figure 4.1.

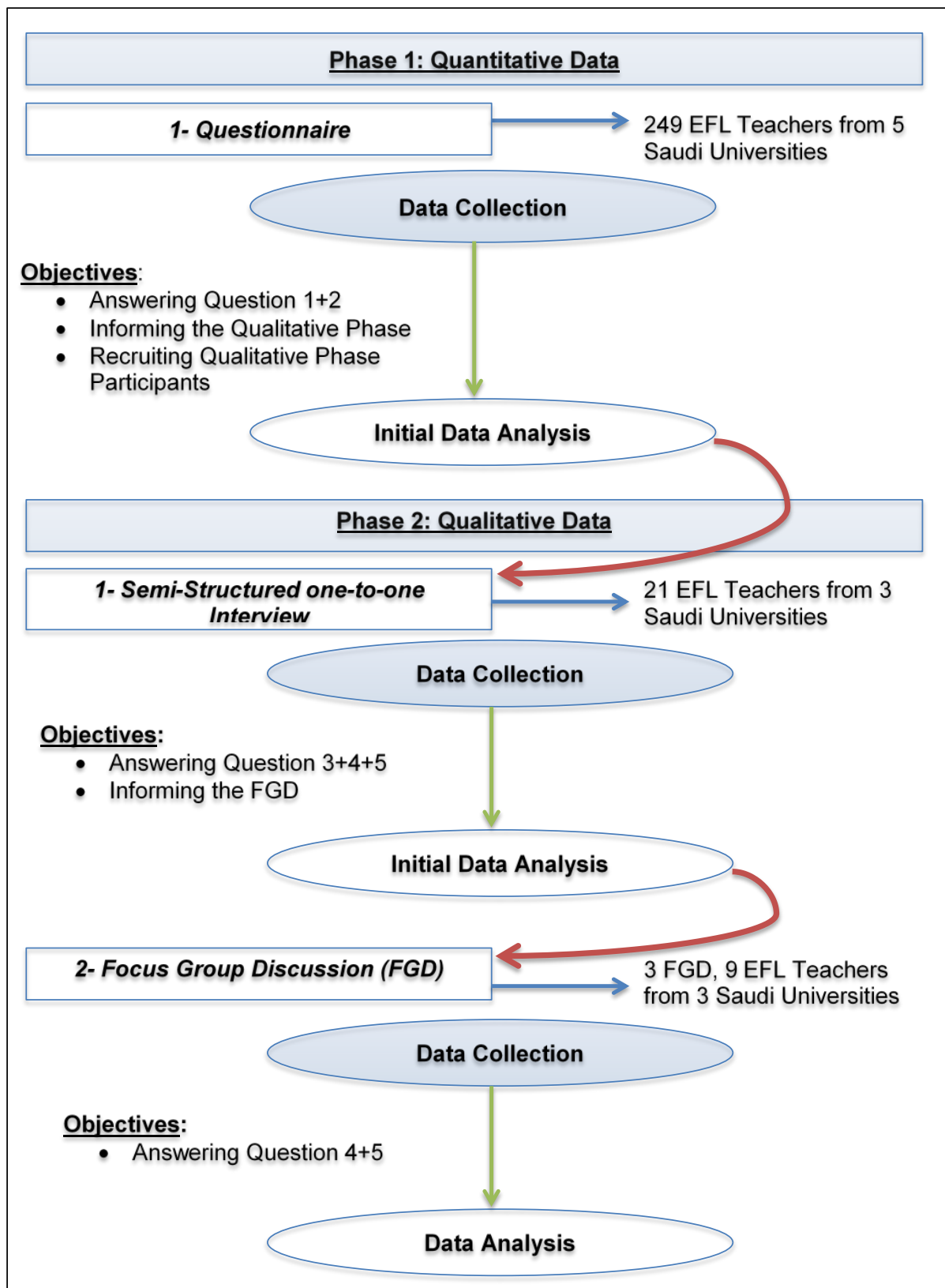


Figure 4. 1. The sequential mixed-methods design of the study

#### **4. 3. 2. 1. Implementation**

The sequential mixed methods design allows the researcher to combine qualitative and quantitative data, and as Bryman (2015) puts it, research methods that cross both qualitative and quantitative research strategies are invited. The implementation assumption, according to Creswell et al. (2003), refers to the sequences of qualitative and quantitative data collection. Most importantly, implementation, or 'timing' as Creswell and Clark (2011) label it, "describes the order in which the researchers use the results from the two sets of data within a study - that is, timing relates to the entire quantitative and qualitative strands, not just data collection" (p. 65). They suggest that implementation can be divided into three main types: concurrent, sequential, and multiphase combination. The present study implements three methods sequentially in the two phases: one method in the quantitative stage and two in the qualitative stage.

In the first phase, I implemented a quantitative strand that involved collecting and analysing purely quantitative data using a quantitative instrument, namely a questionnaire, the main purpose being to explore the investigated phenomenon and to feed into the next qualitative stage. With regard to the second stage, I employed a qualitative strand that included two different methods, namely one-to-one interviews and focus group discussions. The two phases are illustrated in figure 4.2. below. The data in the second qualitative phase was collected using two methods: semi-structured one-to-one interviews and focus group discussions (FGD). The semi-structured one-to-one interview was conducted to discuss the main findings of the questionnaire, to have a better understanding of the results and to find answers for research questions 3 to 5. While carrying out the semi-structured one-to-one interviews, answers to research question 4, namely on evaluation challenges facing EFL teachers, were

obtained. These one-to-one interview outcomes also informed the FGD phase and guided the construction of the discussion guide. All the phases are clearly illustrated in figure 4. 2.

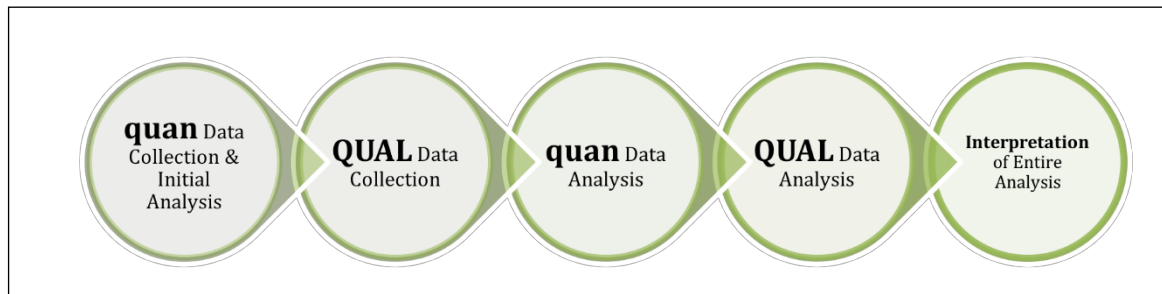


Figure 4. 2. Implementation and integration of the qualitative and quantitative strands

In the following section, I will present the way in which both quantitative and qualitative strands are implemented in this study and which one has priority.

#### **4. 3. 2. 2. Priority**

Priority can be defined as the relative weight or importance given to the quantitative or the qualitative methods while addressing the problem of the research in mixed methods deigned research (Creswell & Clark, 2016, p. 415). There can be three different options for researchers for a weighting decision: equal priority, quantitative priority, or qualitative priority. Bearing this in mind and given the philosophical assumptions underpinning this study, it seems more appropriate to prioritize the qualitative strand regardless of the implementation or sequence.

For the purpose of typological organization and using Johnson and Onwuegbuzie's (2004) visual representation, Dörnyei (2007) artfully draws nine different combinations

for mixed methods varieties that clearly indicate both the sequence ‘implementation’ and dominance ‘priority’. According to his classification, this research might best fit with the ninth combination, which is **quan**  $\longrightarrow$  **QUAL**, where the arrow represents a sequential collection of the data, small letters indicate lower priority or weight, and capital letters indicate priority or increased weight. This can be further illustrated in figure 4.3. below.

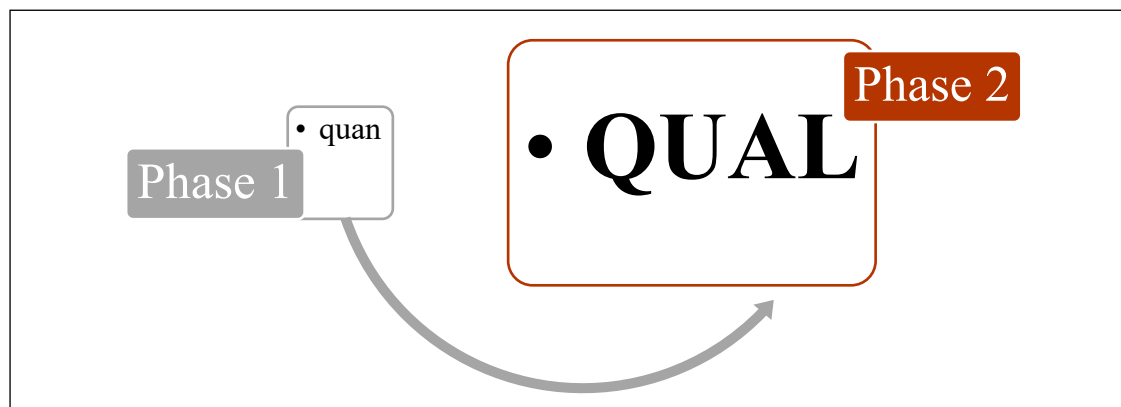


Figure 4. 3. Sequence and prioritization of the two strands

#### 4. 3. 2. 3. *Integration of the Data*

In mixed methods research, combining or integrating the data can be used to refer to how the quantitative and qualitative results are brought together in the research. It can be argued that integration shows the interactive or independent relationship of both strands when they are mixed together in a single study. This depends to a large extent on the design of the mixed methods study and the type of integration that is included. According to Creswell (2017), “the way the researcher combines the data needs to relate to the type of mixed methods design used” (p. 123). Before deciding on which type might properly fit with this research, I should mention that Creswell identifies four integration types: merging, explaining, building, and embedding. However, Creswell and Clark (2016) argue that there are four possible points of interface where qualitative

and quantitative strands can be mixed: at the level of design, during data collection, during data analysis, or during interpretation. This four-point conceptualisation appears more convincing, especially when practicality of application is to be considered, as the first classification of types may seem more of a theoretical approach related to the chosen design.

Creswell (2003) identifies two stages where mixing both strands and data integration can become effective, namely the data analysis and interpretation. It is worth mentioning that mixing both types of data of this study comes was at the level of interpretation, as the analysis tools used for both types of data are different, with SPSS being used for the quantitative data and NVivo for the qualitative data.

#### **4. 3. 3. Participants and Sampling Procedures**

Participants, who are also known as the *sample* of the study, can be defined as “the segment of the population that is selected for research. It is a subset of the population” (Bryman, 2015, p. 715). Before discussing the sampling strategies applied, I will give a brief summary of the participating universities and sampling procedures. The five selected universities belong to the largest three and most significant regions in the county. Three universities are from three different cities that are situated in Makkah Region (where the Holy City is), the fourth is located in Riyadh Region (where the Capital City is) while the fifth is in the Eastern Region (where oil comes from). Three universities from different regions are also known to have the highest ranking in Saudi. In order to answer the research questions, my aim was to involve as participants of this study EFL teachers in five public Saudi universities from different regions in the country. A link to the questionnaire, with an invitation letter to participate, was sent to



the whole population. However, prior to this, a letter seeking permission was sent to deans and heads of departments, given that they act as gatekeepers who have access to the intended population, to ensure that the questionnaire would be sent to as many participants as possible and that I could reach a sufficient number of participants.

After the gatekeepers had given access to the participants by forwarding the invitation letter along with the questionnaire link to their EFL teachers, the intended population was successfully reached. Any individual from the population had an equal chance of being selected and all participants were seen as being representative of the whole population; thus, the sample should consist of what Dörnyei (2007) calls 'self-selected participants', where participants choose to participate without any systematic control by the researcher, as is the case with probability sampling. Given this situation, I should refer to non-probability sampling strategies, which may be seen as inappropriate to aim for in a psychometric sense. Bryman (2015) reminds us that, with probability sampling, researchers can make generalisations from findings they derive from a study sample for the whole population. However, this is not the case in this research and generalisation is negligible. As a researcher, I am not aiming at generalising the results of the quantitative phase but rather to give more breadth that helps to understand the phenomenon at this preliminary stage.

According to Creswell (2017), there are two major types of non-probability sampling: convenience and snowball sampling. In the former, those who are willing and available to participate will be members of the sample, while in the latter the researchers tend to ask participants to identify others to be members of the sample. As the selection of participants in this research was based on availability and willingness of the teachers,

it becomes apparent that convenience sampling can better fit with the data collection procedures of this research than snowball sampling. Dörnyei (2007) reminds us that convenience sample can seldom be totally 'convenience-based' but rather it tends to be partially purposeful. This suggests that, regardless of the ease of access to participants, they should have certain key features that are fundamentally related to the purpose of the study and which will enable them to contribute to the research questions which is the case in sampling for this research. In this research, it was stated in the invitation email and on first page of the online questionnaire that participants should be EFL teachers, teaching at a Saudi university, and should have been evaluated at least once in their current workplace. Therefore, a convenience purposeful sampling best describes sampling strategy in the first quantitative phase.

For the qualitative stage, 3 universities that permitted me as a researcher an access and EFL teachers from those universities who were interested in the topic and willing to continue to the next qualitative phase were identified. It is worth mentioning that there was a special section towards the end of the questionnaire which was allocated to recruit qualitative phase participants. Thus, a sampling frame was easily initiated. The sampling frame according to Bryman (2015) is "the listing of all units in the population from which a sample is selected" (p. 715). However, whether to include individuals from outside this list or not was a question that emerged at a certain point of selection time. As the design suggests that the qualitative phase is informed by the quantitative phase, it could be argued that participants in the qualitative phase should be expected to also have answered the survey questions. According to Creswell (2017), the sample of the qualitative phase in an explanatory sequential design is a

subset of the sample of the quantitative phase and should be drawn from the participant pool of the quantitative sample. As the questionnaire in this study was sent online, it could be the case that some of the prospective participants were not able to respond to it due to a number of technical issues. Therefore, I came to the conclusion that individuals who fulfilled the conditions that enabled them to participate in the first quantitative phase could also participate in the second qualitative phase even if they were not listed in the sampling frame.

Following the *purposive sampling* method, individuals who subsequently volunteered to take part in the one-to-one interviews were contacted. Concerning the purposive sampling, the goal is to sample the participants using a strategic method in a way that “those sampled are relevant to the research questions that are being posed” (Bryman, 2015, p. 418). Researchers who utilise purposive sampling usually aim at making sure that there will be a convenient variation in the resulting data so that each participant contributes differently in terms of significant elements in relation to the study questions. Dörnyei (2007) identifies nine different strategies for purposive sampling and argues that issues of feasibility, such as availability of respondents, time and budget need to be taken into consideration. He also reminds us that above all ‘saturation’, where interviewing adds very little or nothing to the data, needs to be considered. In his opinion saturation is easily reached when individuals of the samples tend to be homogenous. Bearing this in mind and given the exploratory nature of the study, amongst the nine different strategies Dörnyei refers to, I judged that maximum variation sampling could most benefit this research and help in addressing the research questions.

In *maximum variation sampling*, researchers select respondents with notably different types of experience and commonalities are usually underscored, which allows researchers the intended variations within the group of participants (Dörnyei & Taguchi, 2009). This approach to sampling will allow for more variability and diversity to be explored within the phenomenon under consideration in the present study. Creswell and Clark (2016) go further and indicate that this strategy can be seen as one of the most common sampling strategies in mixed methods research. There can be different criteria to maximise differences, such as demographic information (gender, level of education, age, etc.,) of the participants or other similar categorisations, which will enable us to assume that respondents who are purposefully selected can be seen as different individuals (see table 4. 1. below on the maximum variation of the sample). One of the strengths of this strategy is that the participants' variation can provide the researcher with different views, which indicates that variation may offer some good qualitative data and reflects the complex picture of the explored phenomenon.

Table 4. 1.

*Maximum variation of the qualitative stage sample*

No.	University	Pseudonym	Gender	Nationality	Highest Qualification	No. of working years
1	1st	Omaima	Female	Indian	MA	7
2	1st	Fatin	Female	Saudi	Doctorate	5
3	1st	Farah	Female	Egyptian	BA	8
4	1st	Ferdaws	Female	Algerian	MA	3
5	1st	Majeda	Female	Saudi	MA	9
6	1st	Mariam	Female	Saudi	MA	18
7	1st	Sundos	Female	Sudanese	MA	6
8	1st	Karam	Male	Jordanian	BA+Diploma in Education	6
9	1st	Hatim	Male	Tunisian	BA	6
10	1st	Hamdan	Male	British	MA+MSc	6
11	2nd	Joseph	Male	British	MA + MPhil	5
12	2nd	Daniel	Male	Irish	BA+Diploma in TESOL	10
13	2nd	Brian	Male	American	MA	3
14	2nd	Albert	Male	British	MA	5
15	3rd	Samar	Female	Pakistani	MA	2

16	3rd	Roza	Female	Indian	MA	7
17	3rd	Nadin	Female	Jordanian	MA	4
18	3rd	Mawadda	Female	Saudi	BA	2
19	3rd	Asher	Male	Zimbabwean	BA	5
20	3rd	Ayman	Male	Sudanese	MA	3
21	3rd	Abdulhameed	Male	Saudi	BA	1

The sample of this study in the piloting stage involved 20 participants who provided complete answers to the questionnaire. The one-to-one interview was piloted with 3 EFL teachers from 3 different Saudi universities. Later in the study, I received a total of 250 completely answered questionnaires from 5 different Saudi universities. However, I decided to exclude one participant who seemed not to take the questionnaire seriously. Thus, only 249 participations were taken further and analysed as the sample of the first quantitative stage. For the qualitative phase, 21 participants from 3 Saudi universities (whose gatekeepers allowed access) participated in the one-to-one semi-structured interviews. As for the 3 focus group discussions that were carried out in the same 3 universities, a total of 9 participants took place in the discussions. The following table (4.2.) summarizes all participations in the pilot and final study during both the quantitative and qualitative stage.

Table 4. 2.

*Number of participants in each stage*

<b>Method Stage</b>	<b>Quantitative Tools Questionnaire (5 universities)</b>	<b>Qualitative Tools (3 universities)</b>		<b>Total</b>
		<b>One-to-One Interviews</b>	<b>Focus Group Discussions</b>	
<b>The Pilot</b>	20	3	0	23
<b>The Study</b>	249	21	9	279
<b>Total</b>	269	24	9	302

**4. 3. 3. 1. An Overview of the Participants Involved in the Questionnaire**

As mentioned earlier, the total number of respondents who took part in the questionnaire was 249. Of the total number, only 65 (26.1%) were Saudis compared to 184 (73.9%) who were non-Saudis (table 4.3.). The difference in number of the two groups is due to the fact that the majority of the English language teachers in the preparatory year programme are of different nationalities and only a few are Saudis. This fact is likely to impact on the data as Saudis have more job security with long-term contracts.

Table 4. 3.

*Nationalities of the participants*

	Frequency	Percent	Valid Percent	Cumulative Percent
Valid American	16	6.4	6.4	6.4
Algerian	1	.4	.4	6.8
British	21	8.4	8.4	15.3
Canadian	5	2.0	2.0	17.3
Egyptian	17	6.8	6.8	24.5
Filipino	1	.4	.4	24.9
Indian	14	5.6	5.6	30.5
Iranian	1	.4	.4	30.9
Irish	1	.4	.4	31.3
Jordanian	22	8.8	8.8	40.2
New Zealand	1	.4	.4	40.6
Pakistani	38	15.3	15.3	55.8
Saudi	65	26.1	26.1	81.9
South African	5	2.0	2.0	83.9
Sudanese	11	4.4	4.4	88.4
Syrian	4	1.6	1.6	90.0
Tunisian	10	4.0	4.0	94.0
Yemeni	2	.8	.8	94.8
Zimbabwean	1	.4	.4	95.2
Not Specified	13	5.2	5.2	147.3
Total	249	100.0	100.0	

While the highest number of respondents, 117 (47%), are working on a one year type of contract, only 39 (15.7%) are working on a six year or more contract (table 4.4.). This reinforces the idea of lower job security for the majority of the sample and how

crucial results from teacher evaluation can be in affecting decisions on their contract renewal every year. This will be drawn upon in the coming chapters.

Table 4. 4.

*Length of the participants' contracts*

	Frequency	Percent	Valid Percent	Cumulative Percent
Valid One year	117	47.0	47.0	47.0
Two years	13	5.2	5.2	52.2
Three years	8	3.2	3.2	55.4
Four years	10	4.0	4.0	59.4
Five years	8	3.2	3.2	62.7
Six years or more	39	15.7	15.7	78.3
Others	54	21.7	21.7	100.0
Total	249	100.0	100.0	

#### 4. 4. Data Collection and Analysis

According to Bryman (2015), data collection can be perceived as the process where the researcher gathers data from the intended sample that would help in answering the research questions. As mentioned earlier, this research utilises and mixes three different methods to collect the required data and to create a corpus to answer the intended research enquiry. The mixed methods that are used for data collection in this research are, namely, questionnaire, semi-structured one-to-one interview and focus group discussion (FGD). As each method is intended to inform the following one, the methods will be applied separately in the same previously mentioned order. The data collected by each method was initially analysed to feed into the questions for the following method. The procedures for collecting and analysing the data for each method are explained separately in the following section.

#### **4. 4. 1. First Stage (Quantitative): Self-Completion Online Questionnaire**

The questionnaire is designed to mainly measure the perceptions and attitudes of EFL teachers about the phenomenon under investigation. Attitudinal measures can be constructed by writing the questions, finding the instrument(s) in the literature (Creswell, 2017), or by using both techniques, where the researcher consults the literature and adds more questions. In fact, the third approach was applied in this study by referring back to various studies on EFL teacher evaluation involving questionnaires, adapting them as needed, and adding other questions that suit the context of the study and help in addressing the research questions. The quantitative phase applies a cross-sectional strategy to collect data. In other words, the product should elicit what Cohen, Manion, and Morrison (2013) call a “snapshot of a population at a particular time” (p. 267).

The questionnaire is designed to include aspects that reflect the participants’ background (e.g., years of experience, age, academic qualifications) to ensure that the “snapshot” represents the intended population. I constructed a 5-point Likert scale (strongly agree, agree, neutral, disagree, and strongly disagree) questionnaire (appendix 1, p. 319) to include closed questions that are quick to complete and easy to analyse. In addition, the questionnaire included open-ended questions that are intended to avoid the limitations of questions with pre-set categories. Adding open-ended questions can be considered an advantageous feature, especially in a study with an exploratory approach, such as this study. This type of question gives the participants an opportunity to freely express their opinion, which helps researchers to better understand the phenomenon being explored.



It is worth stating that using questionnaires in educational research has its advantages and disadvantages. Self-completion questionnaires are widely known to be cheap and quick to administer (Bryman, 2015). In this study it should be noted that the sample was scattered in different regions of Saudi Arabia, while I was located in the UK: a self-completion questionnaire offered me what Bryman calls an '*unrestricted compass*', where the use of online surveys has almost no constraints when geographical coverage is to be considered, which is the very reason why an online questionnaire was chosen for this study. It is also an efficient use of time in the sense that it can be distributed to a relatively large number in a relatively short time. The researcher effect is also absent with self-completion questionnaires, which makes the outcomes more reliable. Given that it is an online questionnaire, better data accuracy is an advantage, as entering the data on the electronic platform is automated and the researcher does not run the risk of error by entering the data manually onto a spreadsheet.

Despite these strengths, self-completion questionnaires have some weaknesses that are discussed by Bryman (2015). Great attention must be paid to ensuring the questions are clear and the questionnaire can be easily completed, given that no probing can be carried out to get more information from the participants. Moreover, the respondents are able to read other parts of the questionnaire before responding to the first question, which cannot ensure that the questions are answered independently of one another. These challenges, in fact, may be of a great concern where studies are merely quantitative or rely heavily or solely on questionnaires. However, this is not the case in this study, where the questionnaire functions as an

initial source of data and the qualitative phase attempts to overcome the questionnaire's shortcomings. Finally, there is a considerable risk of lower response rates and/or missing data. Daniel (2011) claims that with electronic surveys response rates are lower than with paper surveys. This argument gains some support from the following factors: participants may use more than one email account; email and internet may not be frequently used by some people; emails can be blocked; there are issues related to anonymity and confidentiality; emails are occasionally ignored or filtered into junk mail, and email addresses frequently change. However, this is not always the case.

In contrast, Cobanglu, Ward and Moreo (2001, cited in Bryman 2015) claim that, compared to post and fax paper questionnaires, the online survey in their study, which combined all three types, proved to achieve a "higher response rate (26 per cent versus 44 per cent) and a faster response speed, and was cheaper" (674). As a researcher doing her research by means of an overseas scholarship granted by the Saudi Ministry of Education and having participants in the study who reside in my home country, I believed that an online web-based questionnaire would be the most convenient and appropriate option, given that the majority of the intended population have internet access at their workplace and they can complete the questionnaire easily.

#### **4.4.1.1. Developing the Questionnaire**

To develop questionnaires and as a general recommendation by questionnaire specialists, one should start writing items without being restricted to a specific number. Thus, researchers need to free their imagination and create as many questions as

they can. Dörnyei and Taguchi (2009) have labelled the “resulting collection of items” an *item pool* that “should include many more items than the final scale” (p. 40). In fact, different sources can be identified to generate items for the item pool, such as the gathered qualitative data, an established questionnaire, and the researchers’ own thoughts (Dörnyei, 2009). As the qualitative phase was designed to come after the quantitative one, the questionnaire developed for this research relied on the last two sources: some items were adopted from earlier questionnaires in the literature of teacher and faculty evaluation and occasionally modified while others were based on my own understanding of and experience with teacher evaluation in the Saudi context.

According to Punch (2013), questionnaires are frequently used to seek factual information such as background, demographic and knowledge information, and often contain ‘measures’ of attitudes, beliefs, opinions and values. Both aims are important and were considered while developing the questionnaire for this study. However, when the sequence of both types is to be taken into account, there are two options: to start with the factual and end with attitudinal questions or to start with the attitudinal and end with the factual questions. Oppenheim (2000) claims that placing factual questions that enquire about the participants’ demographic information at the top of first page is normally preferred by novice researchers. As this style may resemble what Oppenheim calls the many ‘*bureaucratic*’ forms that we have to fill out and as information regarding age and marital status may be perceived as private in many cultures, it seemed better to place these demographic questions at the end of the questionnaire designed for this study. With regard to number of questions, in the first draft, I started with 54 questions derived from an item pool of 123, with 10 factual

questions at the end preceded by 4 open-ended questions, while the rest were mainly attitudinal questions.

Two question formats can be found in questionnaires, namely closed questions and open-ended questions. According to Dörnyei and Taguchi (2009), theoreticians tend to discourage open-ended questions which are known to be exploratory and qualitative in nature. It has been claimed that “the desire to use open-ended questions appears to be almost universal in novice researchers, but is usually rapidly extinguished with experience” (Robson, 1993, p. 243). From a personal perspective, and given the exploratory nature of this research, it seemed reasonable to include 4 open-ended questions towards the end of the questionnaire. As the questionnaire was distributed among EFL teachers at 5 different locations, open-ended questions gave the participants the opportunity to communicate further experiences they have had and would like to share that I was not aware of when constructing the closed questions. This would eventually help me develop and construct the questions for the qualitative phase. Indeed, this encouraged me to make the decision to include open-ended questions, a decision supported by Dörnyei and Taguchi’s (2009) view that open-ended questions may offer a greater kind of ‘*richness*’ than entirely quantitative data.

Returning to my questionnaire, the closed part includes 35 questions, all of which are Likert scale in form, which Dörnyei and Taguchi (2009) claim to be the most popular format, with five response options indicating the extent to which participants agree or disagree with a given statement. The website used for presentation of the questionnaire offers both vertical and horizontal layouts and I selected one that would

work over all the items in order not to confuse participants. In this respect, Bryman (2015) argues that, if the questionnaire is being responded to in haste, with the horizontal alignment there could be a risk of misplacing the tick. However, Bryman seems to overlook how lengthy the questionnaire may look with the contrasting vertical layout, which may result in discouraging participation. With the issue of discouragement in mind, and given that Bryman's argument may make more sense with a paper-based questionnaire, I decided on the horizontal format. Wording is another issue that needs further attention by researchers: this issue will be discussed later in the piloting section (section 4.4.1.3.). It is worth mentioning that the questionnaire developed for this research went through four main stages: designing and developing the initial draft, revising, piloting the revised version, and finally refining the final draft that was sent to the study informants. In the following sections I will discuss how a final draft was developed.

#### **4. 4. 1. 2. Examining Validity of the Questionnaire**

According to Dörnyei and Taguchi (2009), a quantitative instrument can be termed valid when "it measures what it is supposed to measure" (p. 51). They remind us that there are three types of validity: 'criterion validity', which is the correlation of an instrument with another instrument that is similar; 'construct validity', which shows how the results of an instrument conform to a theory where the targeted construct is a part; and finally 'content validity', which concerns the judgement of experts about the content of an instrument. Since theory validation is not the aim of the questionnaire in this study and given that the questionnaire is a researcher-developed instrument, I focused my attention on the content analysis of the questionnaire in order to ensure

that content validity could be established. Field (2013) defines content validity as the “evidence that the content of a test corresponds to the content of the construct it was designed to cover” (p. 872).

In line with this but in a rather simpler approach, Bryman (2015) argues that a researcher who has developed a new questionnaire could seek the judgement of experienced individuals in determining “whether on the *face* of it, the measure seems to reflect the concept concerned” (p. 171), and hence determine what can also be called ‘*face validity*’. Before proceeding with experts to ensure the content validity of my questionnaire, I decided to examine the wording, clarity, length of questions and the format of the developed questionnaire by sending it to 4 research peers for consultation and feedback. Based on their feedback, the sequence of questions in some parts of the questionnaire was modified, the wording in a small number of items was improved, definitions for two technical terms (namely summative evaluation and teachers’ portfolio) were included and a closing question was added. As experts in the field of language teaching, both my supervisors were contacted to examine the content validity of the questionnaire. Based upon their feedback and judgment, the first draft of the questionnaire was finalised and the ready-for-piloting version produced. This strategy to assess the content validity is supported by the approach adopted by Creswell and Clark (2016) to validate instruments by employing content validity. They remind us that researchers need evidence of content validity by looking at how experienced judges examine whether the questions are representative of potential items. After ensuring the instrument was ready for piloting, it was sent to the piloting

phase participants. In the following section, how the questionnaire was piloted and how it was further examined will be discussed.

#### **4. 4. 1. 3. Piloting the Questionnaire**

Besides making sure the questions of the questionnaire work well, according to Bryman (2015), “piloting also has a role in ensuring that the research instrument as a whole functions well” (p. 263). Therefore, piloting is a very significant stage that a researcher should not overlook. Piloting may be even more important in quantitative approaches to collecting data. Sudman and Bradburn (1983) argue that, if researchers do not have the resources to pilot their questionnaires, they do not have to carry on with the study. In the questionnaire developed for this study, piloting is intended to verify the first draft in terms of clarity of items and consistency of constructs.

Twenty EFL teachers from two Saudi universities received the link to the online questionnaire for piloting purposes. The first draft they received contained a total of 54 statements, and was divided into four main sections: the first section was devoted to methods used for teacher evaluation with a total of ten items (from item number 1 to 10), where questions 2 and 5 were taken from Mazandarani's (2014) questionnaire, and the rest were developed by me based on my readings in the wider literature on teachers' evaluation. The second section was dedicated to the purposes for teacher evaluation and contained 8 statements (from item number 11 to 18), where question 13 and 14 were based on my readings and the rest from the aforementioned Mazandarani's questionnaire. The third section concentrated on teacher evaluation as experienced by EFL teachers at Saudi universities and constituted 25 elements (from item number 19 to 43), where only question 20 was from Mazandarani's (2014)

questionnaire; questions 19, 32, 38, 40, 42, and 43 were written by me, and the rest were from Shough's (2010) study. The fourth and final section was designed to provide me with the participants' demographic information with 11 items (from item number 44 to 54).

Based on the test-pilot results and after examining the internal consistency (which will be discussed in more detail in the following section 4.4.1.4), the finalised questionnaire ended up consisting of 52 items (see appendix 1, p. 319). In the first section, 3 items were deleted and 3 were added, which gives a total of 10 items (from item number 1 to 10), where only question 4 was from Mazandarani's (2014) questionnaire and the rest were developed from my readings together with the piloting stage feedback. With regard to the second section, it was modified to include 7 items (from item number 11 to 17), where questions 11, 12, 15, and 16 were from Mazandarani (2014) and the rest were developed from my readings and the piloting stage feedback. The third section remains with 24 statements (from item number 18 to 41), where question 19 was from Mazandarani's (2014) questionnaire; items 18, 30, 36, 38, 40, and 41 were based on my experience and readings and the rest were from Shough's (2010) study. Finally, the fourth section remains the same with 11 items (from item number 42 – 52).

It is worth mentioning that the numbers that show the sequence of statements were intentionally deleted from the participants' view in the final draft of the questionnaire. When I analysed the pilot questionnaire, the report of the platform of the online questionnaire showed that the completion rate was 38.89% with a total of **20 completed** responses. Further analysis indicated that there were 33 participants who dropped out after starting the questionnaire. The drop out started at the beginning of



the second section and reached its highest point at the beginning of the third section, with only one dropping out in the fourth section. When I asked for participants' feedback, some of them replied that a questionnaire with a total of 54 questions seemed rather long given the busy schedules of teachers. As a result, I deleted the numbers that show the sequence of the questions from the participants' view or layout. Arguably, showing the numbers was significant in the piloting stage where the informants needed to refer to a specific item number to give a particular comment when providing feedback. However, in the final draft and for the sake of collecting sufficient quantitative data for this study, I found the strategy of removing the numbering had a positive effect on the response rate as it increased from 38.89% in the pilot stage to 56.12% in the actual data collection stage. In the following section, the focus will be on the reliability of the instrument.

#### **4. 4. 1. 4. Reliability of the Questionnaire**

Field (2013) defines reliability as “the ability of a measure to produce consistent results when the same entities are measured under different conditions” (p. 882). It is well established in the literature that there are four different ways to ensure the reliability of an instrument, namely test-retest reliability, inter-rater reliability, inter-method reliability and internal consistency reliability. Since the instrument in this study is designed to include a number of items and is intended to be run only once, internal consistency seems to be the most appropriate measure for reliability (Cohen et al., 2013). Internal consistency indicates how homogeneous the items of a scale are (Muijs, 2010). Pallant (2013) claims that Cronbach's alpha coefficient is one of the most frequently used indicators to measure the internal consistency of a questionnaire.

In line with this, Bryman (2015) claims that the popularity of using Cronbach's alpha to test internal consistency is due to its incorporation into software to analyse quantitative data. This incorporation makes it easy and fast for researchers to calculate internal consistency and accordingly assess the reliability of the instruments they use. After piloting the questionnaire of this study, the data was made ready for this type of software incorporation. I entered the data into the IBM SPSS programme to check out its internal consistency in order to assess its reliability. Initial examination revealed that the Cronbach coefficient alpha value was 0.768 (table 4.5.). Nunnally (1978, cited in Pallant, 2013) reminds us that a minimum of 0.7 Cronbach alpha value is considered a high reliability and sufficient to claim an instrument is reliable.

Table 4. 5.

*Internal consistency of questionnaire items/1*

Cronbach's Alpha	Cronbach's Alpha Based on Standardized Items	N of Items
.768	.780	36

By eliminating 9 items from the scale, the Cronbach coefficient alpha value increased to reach 0.865 (table 4.6.). This level of internal consistency value, according to Pallant, can be considered a very good level of reliability. While values above 0.7 are seen as 'acceptable', values above 0.8 are 'preferable' Pallant (2013, p. 104). Therefore, and given that the questionnaire includes a relatively large number of questions, I decided to remove those 9 items from the final draft of the questionnaire rather than to find the faults with them and revise them.

Table 4. 6.

*Internal consistency of questionnaire items/2*

Cronbach's Alpha	Cronbach's Alpha Based on Standardized Items	N of Items
.865	.869	27

**4. 4. 1. 5. Analysis of the Questionnaire Data**

The data collected in the first stage of the study by means of a questionnaire was analysed using the IBM Statistical Package for the Social Sciences (SPSS) analysis software, Version 24. SPSS is known to be the most popular computer programme among social scientists for analysing quantitative data (Bryman, 2015). Furthermore, it is a powerful programme that enables researchers to handle any statistical procedure that seems complex (Pallant, 2013). SPSS allows users to analyse quantitative data in various ways and for different enquiries very quickly. With the help of SPSS, descriptive statistics for categorial and continuous variables were calculated (see Appendices 6 & 7, p. 350 & 357). Moreover, frequencies for each question and a frequency table were extracted through the software (Appendix 8, p. 348). The final part of the questionnaire with its open-ended questions was analysed qualitatively using thematic analysis techniques and Nvivo, which will be explained in detail in the qualitative data sections. It is worth mentioning, and given the qualitative focus of my study, that SPSS is used mainly to provide some descriptive analysis for this project.

**4. 4. 2. Second Stage (Qualitative): Semi-Structured Interviews and Focus Group Discussions**

Lichtman (2012) argues that interviewing is the most frequently used method for data collection in the qualitative approach. By interviewing, the researcher can explore

aspects that cannot be examined through observation and it offers the researcher the opportunity to investigate and probe the “interviewee’s thoughts, values, prejudices, perceptions, views, feelings and perspectives” (Wellington, 2015, p. 137). In the literature, three different types of interviewing can be identified: structured, semi-structured, and unstructured interviews. In the semi-structured type, the researcher prepares the main questions with prompts to stimulate participants’ replies and to make sure that they contribute as much as they wish.

Brinkman and Kvale (2015) define the semi-structured interview as “a planned and flexible interview with the purpose of obtaining descriptions of the life world of the interviewee with respect to interpreting the meaning of the described phenomenon” (p. 327). In their opinion, the semi-structured interview is flexible as opposed to the structured interview with its fixed and structured design. In this study the flexibility of the semi-structured interview was intended to give me the opportunity to follow up the elaboration of the participants’ unique ideas and to adjust the questions accordingly, which makes it suitable for the exploratory nature of the proposed study. Furthermore, Mann (2016) claims that one of the aims of the semi-structured interview is to achieve participants’ descriptions of their life world along with their interpretations of the investigated phenomenon. This concurs with constructivism, subjectivity, and the philosophical assumptions underpinning this research.

Interviewing, like any other method, has a number of advantages and disadvantages. Brinkman and Kvale (2015) claim that “interviewing is an active process” (p. 17); accordingly, and during that active process, interviewers have the opportunity to get

explanation, clarification and examples. In this case, this type of interviewing will enrich the data needed for this exploratory study. The interviewee, on the other hand, is able to clarify ambiguous questions and seek further explanation. Thus, I ensured that the participants fully understood the questions that were asked. Furthermore, non-verbal data such as facial expression and body language can be obtained to help and support interpretation of the outcomes. Finally, the flexibility of semi-structured interviews enabled me to adjust my interview schedule and questions to suit individual participants and to follow the flow of their ideas. The use of semi-structured interviews at the second stage of the research was also intended to give me the opportunity to fully and deeply understand the investigated phenomenon through the lenses of the informants.

Some of the disadvantages of the semi-structured interview include the need for a skilled interviewer, its high cost, and the extended time it requires. Skilled interviewers know when to probe and when to stop and listen; they are also capable of maintaining the required level of formality during the interview, while unskilled interviewers may find it challenging to deal with a style of interview which Drever (2006) describes as a “formal encounter, with a specific purpose” (p. 4), especially if the interviewee is a friend. To deal with this challenge, I excluded my close friends and attempted to have sufficient knowledge of conducting interviews to carry them out myself. Piloting the interview and seeking interviewees’ feedback at the piloting stage helped me to improve my interviewing skills, especially the skills discussed in this section. Furthermore, interviewing can be seen as both costly and time-consuming: travelling from one region of the country to another in order to conduct interviews with the

selected participants was an economic disadvantage of this method. I also had to spend a considerable period of time reaching my interviewees, given that they were scattered in three different universities, and carrying out the interviews with them myself. Finally, time is also needed for the transcription of the collected data and for the analysis. This challenge is indeed unavoidable; however, the degree of depth and understanding of the phenomenon resulting from interviews made it worth the challenge.

#### **4. 4. 2. 1. Semi-Structured Interview Guide**

Bryman (2015) has made a clear distinction between interview schedule and interview guide. He defines a schedule as a set of questions that are prepared by researchers to ask their participants, and this is what tends to be used in the structured type of interviewing. On the other hand, an interview guide can be described as a collection of prompts of areas to be addressed in an unstructured interview or the “more structured list” of issues to be covered in a semi-structured interview. Given this distinction, the flexibility assumed with the so-called ‘interview guide’ gives the interviewer the chance to explore how the interviewees view their own social world. It seems clear then that interview guides play an essential role in qualitative interviewing as they help interviewers ensure that all areas of enquiry are covered and they offer appropriate wording for the questions.

Dörnyei (2007) addresses issues related to question types and wording when designing an interview guide. The issues include, for example, starting with easy questions to help participants open up and relax and to end with an open question to invite the interviewee to have the final say about what has not been discussed at the

same time as offering the participants the opportunity to comment on their own interests and experience. Bryman (2015) mentions 9 different kinds of interview questions and makes a clear distinction between them as well as when and how to adopt each of them. Using this as a guide, and being informed by other interview guide templates suggested by Brinkman and Kvale (2015), I developed an interview guide for this study (see Appendix 2, p. 325). The guide starts with an introductory section containing general factual questions assumed to function as warm-up questions. Thematically formulated questions that address the main areas of research enquiry followed, along with probing questions. The guide ends with a final question that invites the interviewee to freely add aspects that are related to his/her experience with the phenomena that have not received attention or may not be addressed in the interview.

#### **4. 4. 2. 2. *Piloting the Semi-Structured Interview***

According to Dörnyei (2007), since quantitative studies rely heavily on the psychometric components of their instruments, piloting is more important in a quantitative approach than in a qualitative approach to data collection. However, this does not lessen the importance of piloting the interview guide before commencing the actual interviews. Just as issues such as ambiguity, complexity or duplication can be avoided by consulting experts in the field and peer researchers, aspects related to the researcher's interviewing skills and the estimated length of the interview can only be tested by piloting the guide beforehand. Accordingly, a total of **3** EFL teachers were interviewed to pilot the interview guide.

As I had anticipated, piloting did not result in changes to the questions but it increased my confidence in the guide and gave me the chance to develop my interviewing skills.

I became more sensitive to the proper use of probes and silence to encourage the interviewees to clarify or say more about their opinions. Piloting also gave me some ideas about the time needed to conduct the interview and to cover all the questions. It also gave me the chance to use the data obtained from the piloting stage at a later stage. In line with this, Richards (2014) argues that with methods of collecting qualitative data, there is no actual stage of piloting where researchers test their tools, as piloting and the actual data collection stage can be finely merged. However, I did not include the qualitative pilot data as the participants involved in the pilot sample were found to be from universities that I did not get access to and were eventually excluded.

#### **4. 4. 2. 3. Focus Group Discussion**

Focus Group Discussion (FGD) can be simply defined as “a group interview” (Lichtman, 2012, p. 129). However, it can be argued that an FGD is different from a group interview. Morgan (1988, cited in Cohen et al., 201) describes FGD as “the interaction within the group who discuss a topic supplied by the researcher” (p. 436). The idea of interaction between the participants has been reinforced by Barbour and Kitzinger (1999), who define an FGD as “the explicit use of group interaction to generate data” (p. 4). This interaction is expected given that the role of the FGD moderator is to stimulate the discussion so that the desired data emerge.

There are three main questioning strategies to moderate any FGD: open-ended, more structured, and the topic guide approach (Flick, 2014). In the first two types, the moderator has a list of questions that suit the approach being implemented. For the last approach, a list of topics is used to stimulate the discussion and questions are not



required. It was my intention to apply the open-ended questioning method for a number of reasons. First, the FGDs were conducted after initially analysing the outcome of the questionnaire and after conducting the one-to-one interviews. Accordingly, I had a general understanding about the phenomenon in the given context so there was no need for the topic guide approach which is appropriate where the researcher has no prior knowledge about the context. Second, in preparing a set of questions, I could expect some kind of comparability between the 3 FGDs that were conducted with participants from different universities, where each FGD would bring participants from the same university together. Finally, the open-ended approach helped to give the participants the opportunity to raise issues they perceived as significant with more freedom and flexibility since they were not restrained by a more structured method, and this suits the exploratory nature of the research.

FGDs can be seen as effective tools for collecting qualitative data. The moderator can probe the cognitive and emotional responses of the group members that were unattainable in the first quantitative phase. Thus, like individual interviews, FGDs will balance the drawbacks of the questionnaire and they will all work complementarily. FGDs also encourage multiple views of the phenomenon, as is required for a study with an interpretive focus. Given the participatory collective approach of this method of collecting qualitative data, it seemed reasonable to encourage participants to provide answers to research question number 5, related to solutions to overcome the challenges, an area of enquiry investigated beforehand in the individual interviews. Furthermore, and since FGDs can be adapted easily to different settings and cultures

(Krueger & Casey, 2014; Morgan, 1996), they should be beneficial for a study carried out at 5 universities scattered throughout the regions of Saudi Arabia.

Some of the expected challenges associated with FGDs include the required control over the proceedings of the sessions. Participants speaking at the same time, late comers, no-shows, expert informants, dominant speakers, and disruptive or quiet members are different types of participant-related issues that might hinder the effectiveness or smoothness of the sessions. However, a trained moderator can deal successfully with such challenges. It was my aim to seek a well-qualified moderator to lead FGDs or to lead them myself, given that I had practised moderating FGD sessions a couple of times after receiving some theoretical guidance during my work as a senior coordinator of the programme at my institution. Another challenge may arise during the analysis stage in that FGD data are difficult to analyse (Litosseliti, 2017). Recording the sessions, taking notes, and developing some strategies for analysis helped me overcome these expected obstacles. It is worth mentioning that I developed a guide for FGDs (see Appendix 3, p. 327) following all the procedures for the semi-structured interviews. However, given the larger number of participants in the FGDs compared to the one-to-one interviews, the number of questions was intended to be fewer considering that the participants needed to contribute equally to the discussion.

#### **4. 4. 2. 4. Establishing Dependability of the Qualitative Phase**

According to Brown (2008), the main characteristics of sound research are consistency and fidelity. Consistency in qualitative research is termed dependability and equates to reliability in quantitative research. On the other hand, fidelity is termed

credibility in qualitative research and corresponds to validity in quantitative research. Within the notion of *dependability*, the aim is to ensure the consistency of the process in the study and the stability across methods (Miles, Huberman, & Saldaña, 2014).

Shenton (2004) claims that the dependability of a qualitative research study can be achieved by using multi-methods such as the two qualitative methods that are used in this study, the focus groups and individual interviews. In line with this, Brown (2014) reminds us that dependability can be enhanced by triangulation. There are three types of triangulation that have been utilised in this mixed methods research, namely method, location and perspective triangulation. The three methods used to collect data for this research provide the research with the use of multiple data gathering procedures that enhance method triangulation. Gathering data from five different sites boosts what Brown (2014) refers to as location triangulation. Perspective triangulation was achieved through taking the negative perspectives that contradict the findings of the results and implying multiple perspectives. In addition, having different categories of EFL teachers in terms of nationality, experience, qualification and age group increases perspective triangulation.

#### **4. 4. 2. 5. Establishing Credibility of the qualitative Phase**

It is important to make sure that the findings are credible to the readers and the participants involved in the study (Shenton, 2004). As they did with dependability, Miles et al. (2014) identify a number of ways of ensuring credibility in qualitative enquiry; these include involving thick description, ensuring the account makes sense and is convincing for the readers, linking data to previous theories, reporting clear

findings that are coherently and systematically related, and showing the accuracy of predictions.

The credibility of this research was first achieved through engagement with the participants whether as individuals or in groups. I visited them at their universities and work sites, which gave me the opportunity for more engagement and involvement in the context of the study. In addition, member checks and peer debriefing were carried out in order to enhance the credibility of the qualitative data. I returned interview transcripts to the participants so that they could modify any parts of the interviews they thought needed to be changed. After these member checks, some minor grammatical, editorial and content changes were suggested. I also invited a colleague who was familiar with my research to look through the themes, categories and sub-categories for peer debriefing. This colleague subsequently challenged the logic behind my construction of the findings, which resulted in a much more insightful progression than would otherwise have been plausible.

#### **4. 4. 2. 6. Data Analysis**

After transcribing all the interviews (For a sample transcription, see Appendix 9, p. 370), the data obtained was imported into NVivo and analysed using a thematic analysis strategy, where I examined the data to extract the main themes. Bryman (2015) indicates that coding is one of the core elements in identifying the main themes in any transcript. After coding the separate parts of the transcripts, they were combined into themes. While doing this, I attempted to link the themes to the research questions as well as to the literature whenever possible to make sense of the thematic grouping. However, the thematic construction was mainly guided by the data and the

participants' ideas. I used the NVivo programme to store the interview transcriptions for analysis, to label segment codes, and to have the codes visibly displayed (Appendix 10, p. 385). This makes it easier to classify codes into themes and to finally interpret them (Appendix 11, p. 386).

Themes might be seen as the central product of the processes of data analysis that can yield some practical results in the research. Vaismoradi, Jones, Turunen, and Snelgrove (2016) argue that "the creativity of the researcher is an integral part of the analysis and in presenting the result in terms of a story line" (p. 101). Thus, identifying the themes and organizing them into categories and sub-categories reflects the interpretation of the researcher of the set of data s/he deals with. This eventually gives meaning and sense to the information collected, which helps to develop a story line from the data. It is essential to build up a convincing account for the reader while presenting the themes, categories and sub-categories. Some of the strategies utilized in this research to build up convincing themes are "conveying subthemes, or sub-categories; citing specific quotes; using different sources of data to cite multiple items of evidence; and providing multiple perspectives from individuals in a study to show the divergent views" (Creswell & Clark, 2016, p. 209).

As a software programme, NVivo is known to help researchers to code data around themes, which is a significant characteristic of qualitative data analysis. The programme does not do any coding for the researcher; rather, it helps them to import interviews from word documents and to label segments from stored interviews as "nodes". In NVivo, nodes "can represent anything that the researcher wants them to be and grants easy organization and reorganization of themes in the data"

(Castleberry, 2014, p. 1). A section of a text can be highlighted and moved easily into a node chosen by the researcher by a “drag and drop” feature in the programme. NVivo also allows users to create categories and sub-categories under each node and to move segments smoothly between them. All these features help researchers to create, organize and reorganize themes, categories and sub-categories. The use of NVivo thus facilitates the coding stage and makes it run easily in an organized way.

Brown (2014) identifies 7 different steps in analysing qualitative data as follows:

- Step 1: Get the data into a usable form.
- Step 2: Look for patterns.
- Step 3: Map out tentative patterns.
- Step 4: Organize and recognize the categories.
- Step 5: Search for connections.
- Step 6: Consider multiple perspectives.
- Step 7: Be skeptical. (p. 102)

After coding the data, I applied Brown's model in analysing the qualitative data, given that it is a recently developed model with a focus on TESOL mixed methods research, which is the case in my study. I followed all the steps he mentioned to analyze the data, in the same sequence. It is worth emphasizing that I did not separate the one-to-one interviews and FGDs while analyzing them but I made a distinction between both types of qualitative data while feeding into the themes. In addition, boundaries between both types were made clear in the NVivo programme by locating each type in a separate section in the software.

## **4. 5. Research Ethics**

In any educational research, ethical considerations must be taken into account in all stages. Research that is likely to cause any harm for participants and any other party

is always considered unacceptable. The three main areas that will be discussed in the following section are gaining access to the institutions where the data will be collected, issues related to the participants, and the data protection process. I have paid close attention to the established research ethics required in any stage of the proposed study.

#### **4. 5. 1. Gaining Access**

As I was dealing with different universities in the Kingdom of Saudi Arabia, different procedures were required at each university. In general, Saudi public universities allow and support research students who are sponsored by the government to gain access and to meet with participants to conduct their studies. These universities can be considered closed settings, where a researcher needs permission to access them. Gatekeepers of these closed settings are normally members of the top management and senior administration; thus, I needed to contact deans and vice-deans in order to secure a sufficient level of agreement and preferably to get clearance to ensure access to target participants.

Creswell and Poth (2017) claim that gaining local access permission is a type of ethical issue which can be challenging. They argue that contacting a gatekeeper to gain approval is required to address this issue. For this study, my only means for communication was through emails, given that I was in the UK at that stage. Out of the many universities I contacted, five gave me access and three of them required a formal letter from the Dean of Postgraduate Studies at my university to officially grant their approval. I requested the letters and sent them back to the deans of the three universities and I provided them with a clear explanation of my objectives and

methods, a copy of the participants' consent form (Appendix 5, p. 333), and the University of Exeter research ethics form (Appendix 4, p. 329), to help mitigate their worries and concerns. On gaining their approval for access I started to communicate with heads of language centres and institutions, in addition to the prospective participants, in order to arrange for interviews.

#### **4. 5. 2. Participants' Related Issues**

As far as participants are concerned, issues such as informed consent, harm, invasion of privacy, and deception need to be considered. Giving particular attention to the code of ethics to protect the identity of participants and locations (Denzin & Giardina, 2016) and in line with BERA's Ethical Guidelines for Educational Research (BERA, 2011), this study observed sensitively the confidentiality, privacy, and anonymity of the participants. Informed consent should include information about the research, voluntary participation, the possibility to withdraw without negative consequences, how confidentiality will be maintained, and whom to address for further questions; I made sure that these were included in the introductory section of the questionnaire designed for this research (see Appendix 1, p. 319). To ensure that there would be no harm for the informants or negative consequences for their teaching routines, the place and time for the interviews (FGD and semi-structured one-to-one) needed to be convenient for the informants (Creswell & Clark, 2016). Participants were also required to sign the informed consent document and they needed to be made aware that all sessions would be recorded, both the FGD and one-to-one interviews. Invasion of privacy is related to issues of anonymity and confidentiality (Bryman, 2015), and participants were assured of these in the informed consent document for this study in order to avoid refusal to answer certain questions that might invade their privacy. In



the transcription stage, anonymity was obtained by means of the use of pseudonyms to refer to each of the participants.

### **4. 5. 3. Data Protection**

As strongly advised in the Data Protection Act outlined by the British Educational Research Association (BERA, 2011), data protection was maintained in this study. One of the strategies I utilized in this respect was to keep the completed questionnaires in password-protected files. They were kept, along with the tape recordings of the interviews, in a secure location. The data was also not shared with other parties and it was indicated in the informed consent document that, if sharing was necessary, participants' permission would be asked for in advance. Furthermore, it was planned that the entire data resulting from all the research instruments and tools would be destroyed upon completion of the study. This remains an essential part of the researchers' responsibilities in order to protect the data obtained (Creswell, 2012).

### **4. 6. Challenges and Limitations**

The challenges I experienced while conducting this research were mainly of two kinds: topic-associated and methodological limitations. Limitations of a study can be defined as those characteristics of methodology or design that can influence the interpretation of the study findings (Price, 2004). According to Creswell and Poth (2017), it is the researchers' responsibility to reduce the likelihood of such limitations by predicting probable undesirable scenarios which might arise at any stage of the research. As a researcher, I anticipated some of these prior to the study. However, a number of unforeseen limitations may still occur, some of which may seem to be beyond the researcher's control.

One of the limitations in this study is that participation was restricted to teachers in public (state) universities in Saudi Arabia. It is worth pointing out that it was not my intention to exclude private universities or limit the participants to this restricted sample. In fact, I commenced by contacting universities from both public and private sectors; however, it seems that the research did not raise much interest in any of the private universities that I contacted. Unfortunately, I did not hear back from their gatekeepers to help me determine the reasons that made them hesitate to participate in this study. Including participants' perspectives from such different sites could have enriched the findings of this study and helped in developing a better understanding, as it is likely that they have different practices and policies for EFL teacher evaluation from those of the public sector. Thus, I decided to choose universities in a variety of regions in order to gain a varied picture of unique practices and policies of the investigated phenomenon. It is worth highlighting that this problematic aspect appertaining to the selection of the type of universities was inevitably beyond my control as a researcher; yet it can be seen as an opportunity to make suggestions for further research.

Other limitations and challenges that I came across were mostly due to the characteristics of the subject under investigation in this research study. These challenges generally emanate from the context and the focus of this study and are difficult to deal with, given the sensitivity of the topic *per se*. I found that some of the participants were hesitant to express their honest opinion or to criticize the existing teacher evaluation practices or policies. This became apparent during the FGDs, when the presence of other members in the group resulted in some of the participants being restrained and less communicative. For this reason, I needed to implement different

strategies in order to establish mutual trust such as reminding the participants at the beginning of each FGD that the issues discussed should remain confidential and real names of participants must not be released under any circumstances as per the informed consent. By adopting these strategies, I was able to minimize a cautious and guarded attitude and to encourage participants to freely reveal their opinions and address the drawbacks of the current systems of teacher evaluation.

#### **4. 7. Summary of the Chapter**

The chapter presents a thorough discussion of the methodology for the proposed study. It starts by discussing the philosophical assumptions underpinning the research, and then moves on to outline the design of the study. The chapter also explains in detail the methods for data collection and analysis. It concludes with the research ethics that were taken into account while conducting the study and, finally, a brief account of the limitations of the study is also presented.

## Chapter Five

### Findings and Analysis

#### 5. 1. Introduction

This chapter provides a detailed exploration of the issues investigated in the questionnaire, one-to-one interviews, and the focus group discussions. The participants shared their experiences of the investigated phenomenon in great and sufficient detail, which has resulted in rich data and findings. The findings include various quotes from the participants, employing the pseudonyms given to maintain the participants' confidentiality and privacy as discussed in the previous chapter. In an attempt to answer the research questions, I organised the findings into three major themes, under which categories and sub-categories are presented and supported by evidence extracted from the three research instruments mentioned earlier. In the first theme, the importance of EFL teacher evaluation in Saudi higher education from the teachers' perspective is displayed in three categories and ten sub-categories. The emergent categories and sub-categories are illustrated by data with one example from the datasets. Then contradictory perspectives from the data are provided whenever available to achieve perspective triangulation (Section 4.4.2.4, p. 135). All outcomes are interpreted in light of my experience of and knowledge about the context. The second theme presents findings on challenges to EFL teacher evaluation at Saudi universities and includes five categories and seventeen sub-categories. In the third and final theme, solutions suggested by the participants are displayed in two categories and eight sub-categories. The following table represents all themes, categories and sub-categories.

Table 5. 1.

*Themes, categories, and sub-categories*

Theme	Category	Sub- Category
<p style="text-align: center;"><b>Importance of EFL Teacher Evaluation in Saudi Higher Education</b></p>	<p style="text-align: center;"><b>1. Providing an Authoritative Body</b></p>	1. Employment Decisions
		2. Contract Renewal and Salary Increment
	<p style="text-align: center;"><b>2. Improving Teacher Performance</b></p>	1. Reflection on Practice
		2. Show Differences between Teachers
		3. Measure Students Satisfaction
	<p style="text-align: center;"><b>3. Not Important and “Should Go Completely”</b></p>	
	<p style="text-align: center;"><b>4. What Follows from Teacher Evaluation?</b></p>	1. Instructional Practices
		2. Professional Development
		3. Different Forms of Recognition and Rewards
		4. Teachers’ Self-Esteem
5. Sense of Belonging and Job Security		
<p style="text-align: center;"><b>Challenges to EFL Teacher Evaluation at Saudi Universities</b></p>	<p style="text-align: center;"><b>1. “The Biggest Challenge is Classroom Observation”</b></p>	1. Large Student Numbers
		2. Time Can Be a Challenge
	<p style="text-align: center;"><b>2. Are our Evaluators Capable of Judging Us?</b></p>	1. Required Qualifications and Skills
		2. Personality and Exposures to Different Environments
		3. Associated Issues of Power
		4. Feedback Concerns
	<p style="text-align: center;"><b>3. Students Can Be a Challenge Too</b></p>	1. Ability to Evaluate EFL Teachers

<b>Challenges to EFL Teacher Evaluation at Saudi Universities</b>		2. "The Student Has A Punitive Ability to Be Honest or Not"
	<b>4. EFL Teacher Related Difficulties</b>	1. "It is Quite Ruthless Here": Evaluation and Psychological Stress
		2. Fairness and Equality
		3. Teachers' Related Procedural Issues
		4. Teachers' preparation and Readiness
		5. Roles and Voice of Teachers
	<b>5. Current Teacher Evaluation Systems</b>	1. Transparency and Confidentiality
		2. Criteria
		3. Modular System
	<b>Suggested Solutions for Better EFL Teacher Evaluation</b>	<b>1. Micro-Level Suggested Solutions</b>
2. Involve Teachers, Students and Build Trust		
3. Appoint Qualified Evaluators		
4. Advance the Methods, Improve the criteria and Scoring Techniques		
5. Fix the Existing Problems and Plan well		
6. Introduce a Different Evaluation System		
<b>2. Macro-Level Suggested Solutions</b>		1. Invest in Teachers
		2. Collaborate with Other Departments and Universities

The themes, categories and sub-categories are supported by data from the qualitative research material, namely the 21 one-to-one interviews and the 3 focus group discussions, which I refer to as sources for the themes. The number of times a

particular theme was identified across the sources is referred to as frequency. The following table (5.2) shows the frequency of sources and references for the three major themes (including categories and sub-categories). It is worth noting that despite the variety of the participants' ideas, there seems to be a lack of variability in response based on their demographic characteristics (gender, nationality, institutions, region, etc.,) which reflects an overall homogeneity of the data. This is probably due to similarities of education policies in Saudi universities and the general philosophy behind the practice of teacher evaluation in that particular context.

Table 5. 2.

*Sources and frequency of the three major themes*

Theme		Source	Frequency
1.	Importance of EFL Teacher Evaluation in Saudi Higher Education	23	149
2.	Challenges to EFL Teacher Evaluation at Saudi Universities	24	365
3.	Suggested Solutions for Better EFL Teacher Evaluation	16	116
Total		<b>63</b>	<b>630</b>

## 5. 2. Importance of EFL Teacher Evaluation in Saudi Higher Education

The first theme that emerged from the data gives a general idea about the importance of teacher evaluation to EFL teachers in Saudi higher education. Under this theme, four main categories emerged from the data and were organised to explore different opinions of the participants. These four categories are providing an authoritative body, improving teacher performance, not important, and consequences of teacher evaluation, as presented in figure 5.1. below.

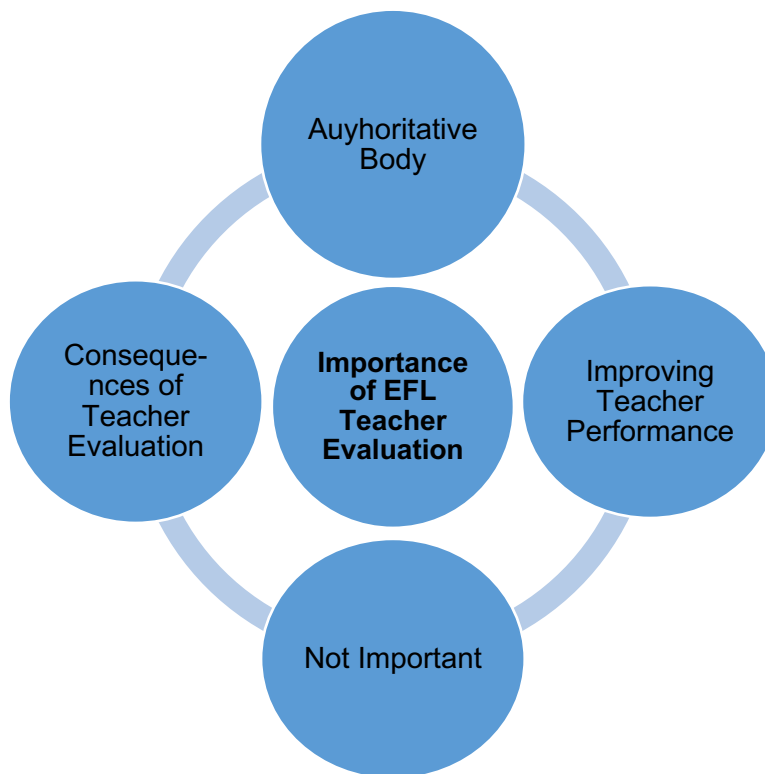


Figure 5. 1. First theme and categories

The first category shows that EFL teacher evaluation provides an authoritative body that is entitled to make decisions on employing new EFL teachers and/or renewal of ongoing contracts, which can be considered as two sub-categories. Both will be elaborated on in the following sections.

### 5. 2. 1. Providing an Authoritative Body

Teacher evaluation is important from the perspective of EFL teachers as it provides the university with an authoritative body that is in control over teaching profession practice. This idea is illustrated by one of the participants, Sondus:

*“I believe teacher evaluation is very important. Teachers need to know that there is an **authoritarian body** that controls **standards** of practice and ensures that ethical and professional teaching is taking place.”*



As the quote above suggests, EFL teachers perceive teacher evaluation as a means of authoritative practice. Mainly, the aim of these authorities is to check the quality of teaching by inspecting the performance of the teachers. One participant in the questionnaire, as a response to one of the open questions about the strength of the current teacher evaluation system at his/her university, commented on the importance of setting standards by such bodies as follows:

*“Teacher evaluation sets the **standards** for teachers.”*

The statistical results also give strong support to the importance of setting standards, as one of the statements suggests that working towards improving the teacher’s performance on the evaluation **standards** will help to improve the quality of his/her teaching, where 62.2% of the participants agreed and 14.9% strongly agreed.

As the above findings suggest, institutions of higher education need to adhere to a specific level of standards of practice to ensure that they provide a quality education to the students. It appears that in order for those institutions to achieve this goal, they need to have a body that monitors and advises on standards to work by, besides other related responsibilities. I believe that keeping an eye on standards is closely related to a number of important issues in higher education, one of which is quality assurance, which is required to fulfil public trust, given the fact that public universities spend public money to educate students. This issue will be discussed and linked to the demands and values of teacher evaluation further in the following chapter.

Despite the fact that some participants have a positive perception of authoritative bodies, as Norris and Mills (2014) claim, the top-down and authoritative orientation of

language programme evaluation can result in dire consequences in some cases. This idea will be discussed in the discussion chapter. Besides monitoring the quality of teaching and as an authoritative entity, teacher evaluation can help in making decisions on employment and contract renewal for university EFL teachers, which will be discussed in the following section. Those two sub-categories are illustrated in the following figure 5.2.

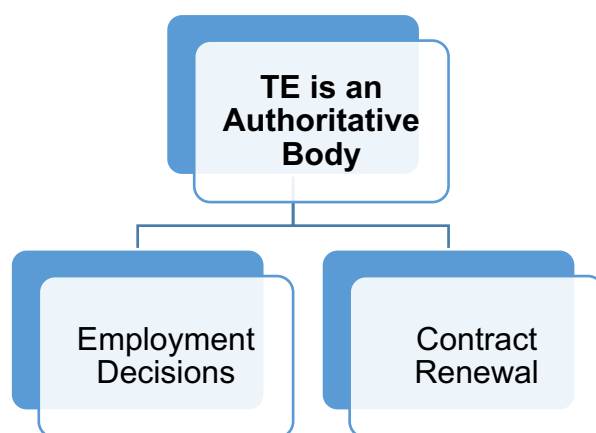


Figure 5. 2. First category and sub-categories, theme 1

### **5. 2. 1. 1. Employment Decisions**

After recruiting new teachers, making decisions on their employment is a significant part of authoritative bodies' decision-making responsibility. The data suggests that there may be a link between teacher evaluation and making decisions on EFL teachers' employment. For instance, in one of the focus group discussions and to reveal the key purpose for teacher evaluation at her university, Sameyah indicated that:

*“For new teachers, to **take decisions**, just for the first semester or a year joining the university.”*

However, the statistical analysis of the questionnaire shows contradicting results, where decision-making is found to be the least recognized purpose of teacher evaluation by EFL teachers themselves. Only 18.5% of the participants were of the opinion that teacher evaluation is carried out solely to make decisions. From a personal perspective, I believe that teachers may not be fully aware of managerial actions and administrative practices as such. Decision-makers can use teacher evaluation as a useful basis for their judgement before making any employment decision. This does not necessarily imply that teachers are to be involved in these procedures, nor may those practices have been made clear to them.

#### **5. 2. 1. 2. Contract Renewal and Salary Increment**

Connected to the previous sub-category but with more frequency, contract renewal was among the purposes acknowledged by the participants for EFL teacher evaluation. While the issue of employment decisions related to new teachers was brought up only three times in the interviews, contract renewal and salary increment were mentioned twenty times during different interviews. According to Albert:

*“The basic discomfort of people, I think, is not so much being observed. It’s what’s behind that and what it can lead to. Here, it’s completely obvious that’s what they’re doing because they will actually tell you if your evaluations are not high enough, then **you will be terminated.**”*

Daniel and Brian, in a similar vein, highlighted the idea of teacher evaluation influencing their salary in one of the focus group discussions and said:

*“Brian: These two things are always **correlated**. Is that correct [to Daniel]? Their grade and their--*

*Daniel: **Salary increase**. Always.*

*Brian: Always.*

*Daniel: Yeah.*

*Brian: The **grade** of the instructor when the **contract renewal** is happening.”*

As these findings suggest, EFL teachers are very concerned about the results of their evaluation, given the considerable influence it may have on personnel decisions such as salary increment or contract renewal. A possible contextual factor that may explain these findings is that there are many expatriate EFL teachers at Saudi universities whose contracts are renewed annually or every other year. Statistical analysis of the participants' demographic information shows that 47% of them, which forms the majority of the sample, have a one-year contract. When the qualitative data is considered, the issues of contract renewal and salary increment were mentioned by teachers of different nationalities but not by Saudis. Karen from the UK stated this clearly in one of the focus groups:

*“I think basically, it all comes down to the same thing that we are **contractees** and we feel worried about **renewing our contracts**, basically. And my colleague has said it several times; security.”*

The findings suggest that Saudi EFL teachers at Saudi universities may appear to be less worried about contract and increment issues as they have a permanent contract with stable annual increments and hence did not mention such issues in the interviews. However, a close analysis of the data reveals a contradiction to the previous findings.

One of the participants, namely Nadin from Jordan, believes that contract and salary may not be influenced by teacher evaluation results. According to Nadin:

*“Well, **contract and salary, no**, because I haven’t heard anyone talking about being affected with salary and contract.”*

The participant contended that she did not have any experience of having a colleague fired based on their teacher evaluation and justified her response on that. There may be other contextual reasons for her answer: Nadin and her colleagues talked about the introduction of a mentorship programme at their university, where their institution provides a peer mentoring scheme in order to support the new teachers, especially those with some teaching deficiency; furthermore, Nadin herself is one of the peer mentors and has spent more than 4 years in the institution and therefore may have a more secure job than the other participants, who linked contract renewal and salary increment to teacher evaluation.

Another factor that may mitigate the teacher evaluation effect on expatriate teachers’ salary was highlighted by Joseph, who is teaching at another university:

*“It **doesn’t affect mine** because any salary increments stop once you reach the **age of 60** at this university. So it’s the **pressure of ageism**. You could be the best teacher in the world, but after you reach the age of 60, you **don’t get any increments**. Before that, yes, I understand.”*

This finding suggests that the influence of teacher evaluation can be precluded at a certain age where teachers do not expect any raise in salary. Age, however, cannot stop the effect of teacher evaluation on contract termination, which is an unfortunate situation for any experienced professional. To conclude this category, it is worth

mentioning that the kind of teacher evaluation that leads to decision-making by authoritative bodies belongs in the category that is labelled summative evaluation in the literature. This type of evaluation appears to be significant in education, as Asher (2016) contends that those who receive public funds should prove to be accountable. This significant issue can be linked to the idea of public trust and will be elaborated on in the following chapter.

### 5. 2. 2. Improving the Teacher's Performance

The second category reveals how teacher evaluation can help improve teacher performance as perceived by the participants. When teacher evaluation takes place, many EFL teachers may place considerable importance on their improvement and growth as a result of the evaluation process, which could be categorised as formative evaluation. There is a wealth of literature on teacher evaluation (e.g. Bailey, 2016; Danielson & McGreal, 2000; Peterson, 2000) that illuminates the notion of formative evaluation. Some works assume that formative evaluation is essentially conducted to offer feedback for the purpose of improvement. This idea received strong support from the data in this study, as most of the participants mentioned the importance of evaluation in developing their practice. For instance Hatim stated:

*"It's through **evaluation** that teachers try to see their own way of teaching and to identify the **strong aspects** of their practice, as well as their **weaknesses** which may need to be changed and **improved**."*

As this finding suggests, teachers as practitioners need evaluation feedback to reinforce their strengths and to avoid weaknesses. Abdulhameed adopted the same

stance and added that, through evaluation, teachers can benefit from senior teachers.

He asserted that:

*“I think it [referring to teacher evaluation] is very **important** since teachers benefit from the accumulated **experiences of senior teachers.**”*

The statistical findings highly supported the significance of evaluation for improvement purposes, where 53% of the participants agreed that teacher evaluation should mostly focus on formative purposes and another 27.3% strongly agreed, totaling 80.3% of the sample. It seems that improving teacher performance is seen as the core target for teacher evaluation from the participants’ perspective, thus highlighting the significance of teacher evaluation and the high expectation teachers place on it. The analysis of the data resulted in the emergence of three sub-categories, namely reflection on practice, show differences between teachers, and measure students’ satisfaction. The three sub-categories are presented in the next section and displayed in figure 5.3. below.

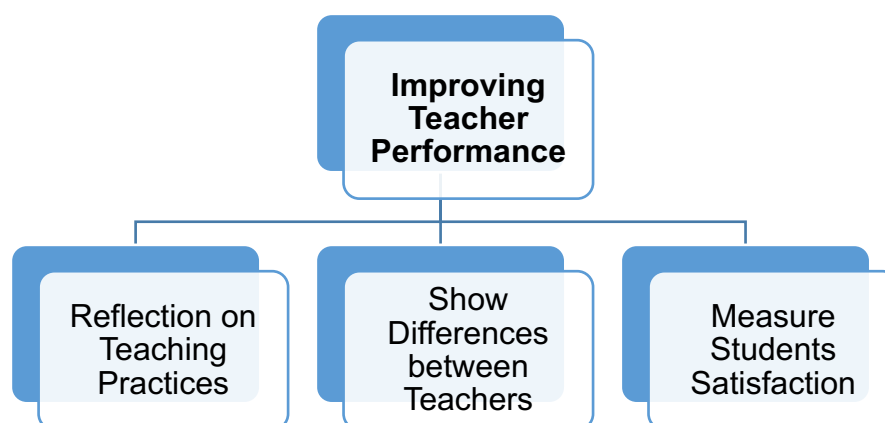


Figure 5. 3. Second category and sub-categories, theme 1

### 5. 2. 2. 1. Reflection on Teaching Practices

In order for a language teacher to improve her/his performance, evaluation should encourage reflection on teaching practices. According to Farrell (2014), reflection, or 'reflective practice', is one of the buzzwords in education and has become a popular term in language teachers' education and professional development. Reflection involves the process of analysing and evaluating what is happening inside the classroom (Zwozdiak-Myers, 2012). As Bailey (2006) claims, when teacher evaluation reflects on the teacher's practice, s/he can achieve the desired improvement. Explaining that evaluation helps her reflect on her professional practice, Omaima maintained that:

*"If you're asking my personal opinion, I think that it makes you a very **conscious teacher, a very self-reflective professional.** There have been things that have been pointed out to me during my observations which I **had not really paid attention to.**"*

As Omaima stated, there are areas for improvement in her practice which she had not recognized, as they had become routine over the course of time. Teacher evaluation should help bring these areas to the attention of the teacher as a kind of reflection and feedback, and hence information about teachers' practices can be useful. Statistical results showed that 52.6% of the participants agreed that teacher evaluation provides them with objective information about their teaching and another 8.8% strongly agreed. This suggests that more than half of the sample supported this finding. Similarly, the following quote from Ashir, one of the interviewees, illustrated the importance of the reflective feature of evaluation:



*“Evaluation is important as it gives both the teacher and supervisor an **indication of the strengths and weaknesses** of the teacher and where necessary, **corrective measures** taken.”*

Besides highlighting the significance of evaluation, he also took the issue somewhat further by asserting that corrective measures are to be taken when necessary. In a similar vein, Ayman pointed to identifying the areas that need improvement:

*“It's through evaluation that teachers try **to see** their own way of teaching and to **identify** the strong aspects of their practice, as well as their weaknesses which may need to be **changed and improved.**”*

As an expected consequence of reflection on practice, improvement and change need to take place after identifying strengths and weaknesses. Interestingly, analysis of the survey data showed that 58.8% of the participants agreed that teacher evaluation increases their reflection on choice of teaching strategies and another 12.4% strongly agreed. Such a relatively high level of consensus amongst EFL teachers about the role of evaluation on choice of teaching strategies underlined the significance of teacher evaluation in reflection on their practice. Waring's (2013) findings, however, highlight the significance of interaction in *post-observation meetings* to reflect on practice in the context of second language teaching.

#### **5. 2. 2. 2. Show Differences between Teachers**

In line with reflection on teachers' practices and connected to the previous sub-category, teacher evaluation serves to distinguish between teachers. As evaluation takes place, teachers get different results based on their evaluation and hence

differentiation will appear. Fatin linked this point to the importance of teacher evaluation stating:

*“Because it **shows the difference** between people; somebody who is working and somebody who is not working; somebody who is doing their job and somebody like this, just ignoring the tasks that she's given. It is **important.**”*

The quote suggests that differentiating between teachers is significant as teachers are not the same. Different actions might be taken according to what teacher evaluation brings out. Enrolment in different professional development programmes may be one of them; as posited by some participants, teachers can decide on different professional development plans based on their evaluation. This is illustrated in the following extract from Farah:

*“Those who want to have **more practice in some areas**, they need development; they prepare **presentations and workshops** for those teachers in order to **enhance the performance** of the teachers and to have a **better outcome** for them.”*

The data suggests that teacher evaluation helps EFL teachers to know more about the areas they need to plan for professional development activities. Knowing more about differences between teachers could help institutions to provide a variety of developmental workshops to match different professional needs. This may be the ultimate purpose for teacher evaluation. Farah made a clear link between teacher evaluation and improving teachers' performance in the following quote:

*“In case, if this teacher evaluation is done for my own sake for **improving the level of performance**, I totally agree.”*

The statistical results also support this finding. The survey analysis indicates that 51% of the participants agreed that the teacher evaluation process is helpful to their professional growth and 18.1% strongly agreed, giving a total of 69.1% agreement. In my opinion showing differences between teachers based on teacher evaluation and consequently planning various workshops may further highlight the importance of teacher evaluation.

### **5. 2. 2. 3. Measure Students' Satisfaction**

As Seldin (2006) states, academic institutions need to gather and propagate information on student service, student satisfaction and learning ratings. EFL teachers at Saudi universities appear to be aware of institutional demands to measure students' satisfaction and some teachers accordingly linked it to the importance of teacher evaluation. For instance, when Samar was asked about the importance of teacher evaluation, she acknowledged the significance of measuring students' satisfaction amongst other indicators in the following extract:

*"I know that it is **important** for the seniors, for the authorities what the teacher is doing inside the class and **whether the students are satisfied**, whether she is applying the techniques as she claimed to have learned previously and which are **necessary for effective teaching**."*

Roza also placed value on measuring students' satisfaction as an inevitable aspect of teacher evaluation and asserted:

*"Sometimes maybe I'm teaching in a way you're maybe not satisfied with (...). But if the **students are satisfied** or not, that is really **important**."*

Regarding results from the survey and when participants were asked to describe EFL teacher evaluation in two words only, one of the participants described it as “*customer satisfaction*”, referring to the student as a customer. The data suggests that student satisfaction may be significant in indicating quality teaching. In line with this, Carbone et al. (2015) in their study found that teaching quality was highly related to students’ satisfaction. It appears that creating an encouraging learning environment which can positively influence student satisfaction will accordingly help achieve better evaluation schemes and higher teaching quality. This sub-category brings the importance of EFL teacher evaluation category to an end. In the following category, I will present the results of a category which contradicts those shown above.

### **5. 2. 3. Not Important and “Should Go Completely”**

This category not only contrasts with the previous two categories in the ideas reported by EFL teachers at Saudi universities but also differs in structure of presentation, in that it has no sub-categories. The reason behind this is not because of lack of variation in the ideas but is rather a result of not having much support from the data to create sufficient sub-categories. Instead, the ideas were presented by the participants, introduced without much support and were discussed only in brief. Therefore, I will present them as various reasons to support this category. The insignificance of teacher evaluation was stated very clearly by some participants: for example, in one of the focus groups, Huda, who was highly ranked by her evaluator, made explicit reference to this idea:

*“The other thing is sometimes we feel **it’s unnecessary**, especially if I get something around 4 or 4.5 or 5 [out of 5] in the last year, and now I’m being observed again for the following year. So I find it **unnecessary** as a teacher. So if I*

*needed development in any sort of any part of my teaching, I would expect the evaluation, so why should you re-evaluate me if you already know my performance beforehand?"*

Daniel had a similar stance towards the importance of teacher evaluation:

*"Well, some teachers say it **should go** completely."*

He added that:

*"For me, I think it's fair enough (...) but unfortunately, because I'm almost 60 now, they don't link it anymore to my salary increase. So for me, I **couldn't care**."*

Analysis of the qualitative data revealed several reasons put forward by the participants to justify their opinions on the insignificance of teacher evaluation. Mariam, for instance, linked this issue to lack of aims:

*"But it's **aimless**, it's fruitless"*

The data here reveals that clarity of the aims and reasons behind evaluation make it important. It appears that institutions need to inform EFL teachers about the aims of the evaluation before proceeding with the evaluation so teachers can better understand the practice and its aims. It is likely that teachers will invent some trivial reasons for teacher evaluation when they have no clear explanation for it. A humble reason proposed by the participants was to categorize teachers, as stated by Ferdaws:

*"I don't know. I feel it's more into **sorting teachers into categories**. It's really more about scoring, if you get 5, if you get 4 or 3, etc. It's not about to develop. They just give you a score and then go home, and that's it."*

The extract clearly reveals that teachers know little about the reasons for teacher evaluation, so it appears to some of them that it is simply enacted to aimlessly sort teachers into separate categories and evaluate them to ascertain in which category they best fit. The data also showed a less insignificant yet unpleasant aim, which is documentation. Majedah referred to the notion of 'documentation' in the following extract:

*"You feel in a way, between the lines, **it's only for documentation** reasons. They just want to document that we've done evaluating as part of the **institute professional stuff**, and that's it."*

The data shows that teachers need to guess why teacher evaluation is taking place at their institution in order to make sense of the evaluative procedures carried out. This makes them feel that it has no value and that it is merely an administrative routine.

The final issue that emerged from the data is the lack of follow-up plans. This reason can be identified in the following quote from Mariam:

*"There's **no follow up**, and at the end of the day, she is graded for what she was supposed to develop in and **nobody followed up** on her development. And no real consequences."*

As the quote suggests, having a follow-up plan might help EFL teachers to conceive a better understanding of the goals of teacher evaluation.

To conclude this section, it is clear that teachers need to be convinced about the practice of teacher evaluation by providing persuasive reasons and explicit aims for it.

In the following category, the consequences of teacher evaluation as suggested by the participants will be outlined.

#### 5. 2. 4. Effects of Teacher Evaluation?

In this category, issues that follow as a result of teacher evaluation from the perspective of EFL teachers will be presented. Instructional practices, professional development, forms of recognition and rewards, along with teachers' self-esteem, sense of belonging and job security will be presented in the following as five different sub-categories (displayed in figure 5.4. below). Various and sometimes opposing opinions were held by teachers and they will all be presented in detail in the following section.

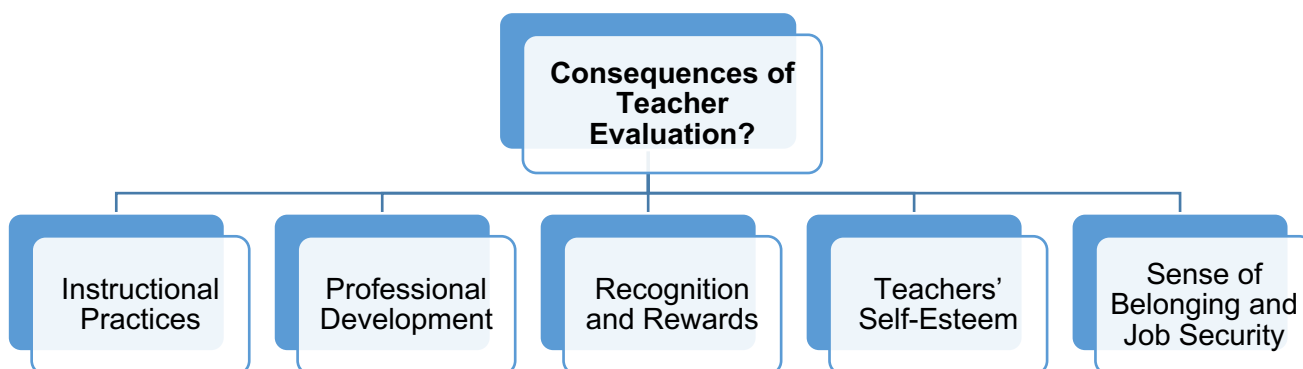


Figure 5. 4. Fourth category and sub-categories, theme 1

##### 5. 2. 4. 1. Instructional Practices

When instructional practices were discussed, the majority of the participants indicated that evaluation can positively affect their own practice. Thus, teacher evaluation can be perceived as offering different contributions to both teachers' performance and additional improvement plans. As Papay (2012) maintains, teacher evaluation can

afford a rich view of teachers' instructional practices when classroom observation implements a standard-based evaluation system and evaluators pay frequent visits to the classroom. The following extracts indicate how the participants consider teacher evaluation as a source by which they can develop their instructional practices. Ayman and Asher in their one-to-one interviews made explicit reference to this idea:

*“Ayman: The evaluation notes have a **positive effect** on my instructional practices.”*

*“Asher: Such results should **help improve** a teacher's instructional practices leading to professional growth and by extension benefit the students.”*

Interestingly, the analysis of the survey data showed that only 29.3% of the participants agreed that they regularly receive instructional support based on their performance evaluation and another 3.2% strongly agreed, while 35% remained neutral about the statement. In line with this, a considerable number of the participants, specifically 50%, agreed that the process used under the teacher evaluation system fosters a climate for instructional improvement and 8.48% strongly agreed, while 26.9% decided to remain neutral. It appears that teacher evaluation can provide EFL teachers with a useful source to decide on the quality of their instructional practices so teachers can recognize what practices might need more improvement; however, it seems that they do not receive sufficient instructional support as a consequence of their evaluation. Some participants link instructional improvement to constructive feedback. Hamdan in his interview referred to this link:

*“Where the teacher was constructively alerted to a particular point in the teaching process by the evaluator and the teacher being evaluated [has] **accepted** this point wholeheartedly. So, there could be a pathway for **constructive feedback** where*



*the teacher genuinely believes that the evaluator is fairly reflecting on his teaching and giving credit to the teacher where credit is due and alerting him in a professional manner to ways of **improving** in other **aspects of his teaching**.”*

It seems that teachers place a high value on the feedback they receive from their evaluators and improve their instructional practices accordingly. On the other hand, some participants reported that teacher evaluation does not affect their instructional practice. Brian, for instance, viewed this issue from a rather different perspective:

*“I normally forget about the student evaluations and the other evaluations throughout the year. I think that it **doesn't really affect the way that I teach** or the way I present myself as a professional in the field as it should, I believe.”*

In an extremely contradictory viewpoint, the data revealed that some teachers believed that teacher evaluation has a negative impact on their instructional practices. When Karam was asked in his interview about the ways in which teacher evaluation affects his instructional practices, he contended:

*“They usually **increase our frustration**. For example, inside the class you should be thinking all the time whether what you do pleases X and Y or not.”*

There is a lot to say about such an understanding of teacher evaluation and its effect on instructional practices. Saudi universities' enthusiasm for international recognition and accreditation is a hidden agenda which EFL teachers may feel obliged to follow which may cause them to perceive teacher evaluation as a source of frustration. It seems the link between teacher evaluation and its negative consequences may reinforce the idea that evaluation does not do much to improve instructional practices.

Instead, some people believe it is done to respond to universities' thrust for accreditation; an issue of demand that will be discussed in the forthcoming chapter.

#### **5. 2. 4. 2. Professional Development**

Unlike the previous sub-category, the majority of the participants expressed the view that teacher evaluation does not motivate their professional development. The core of these ideas may reside in the view that educational institutions do not provide proper developmental workshops. The need for teacher evaluation that feeds into institutional decisions on workshops for teachers is highlighted in the following extract from Abdulhameed:

*"I do not think that evaluation results are used to put on a specific **workshops** or fed into these workshops to be specifically tackled, rather they are workshops revolves around specific teaching topics have **no specific link** to anyone's performance."*

The extract above suggests that there is no link found between teacher evaluation and the workshops their institutions usually provide. In other words, institutions do not make any effort to design developmental plans based on the needs of their EFL teachers. As a result, teachers hardly find these workshops valuable. Some participants critically commented on the purpose of the workshops provided: that they do not support the type of professional development they need. For instance, Joseph levelled a serious criticism of the purpose for such workshops at his university:

*"Workshops are used rather **cynically** here. We often finish the semester before the holiday begins, and I think we have workshops then to ensure that people aren't leaving before the end of the semester."*

As the quote suggests, aimless workshops are sometimes offered just to keep teachers busy and to ensure that they do not leave before formal holidays begin. A plausible solution for such an issue would be to make professional development programmes more useful by linking them to the results of teacher evaluation practices. I believe that workshops should match the needs of the teachers in order to improve their instructional practices. In the same vein and reflecting on the link between teacher evaluation and professional development, one of the participants, Karam, attempted to articulate how teacher evaluation demotivates his professional development:

*“To me, they **demotivate**, frustrate, disappoint, demoralise me simply because they want me to please them in a fifty-minute-class without pleasing Allah and my students the whole year.”*

Although the participant was referring to classroom observation and how it is conducted, he asserted that teacher evaluation results demotivate teacher professional development, the main reason being that it seeks the evaluator’s satisfaction in a short period of time. This may be true, especially where students’ rating is not part of the overall evaluation of the teacher and classroom visits are the sole or main method of evaluating performance. Some participants, however, trust that teacher evaluation can provide them with a reliable source for motivating professional development and identifying training needs. Appreciating the value of teacher evaluation in providing opportunities to identify training needs, Sondus commented:

*“Yes. I believe teacher evaluation is a good opportunity to have the teacher’s **training needs** pointed out and so it*

*becomes easier for teachers to go and seek training in these areas.”*

It appears that the notion of identifying training needs partly relies on further actions taken by teachers themselves, given that institutions do not provide training opportunities that meet teachers' requirements. For instance, Mawaddah stated:

*“And I learned, and I was looking for information about that. Every point they told me after that, I **started researching.**”*

The extract suggests that EFL teachers need to depend on themselves to promote their performance and carry out individual professional plans. Furthermore, the statistical results revealed that teacher evaluation was found to be of paramount significance in guiding teachers in their professional development efforts. With regard to the survey findings, 56% of the participants agreed that they focus their professional development efforts on activities that directly help them achieve the evaluation standards and 16.1% strongly agreed, giving a total of 72.1%, while 18.5% were neutral.

#### **5. 2. 4. 3. Different Forms of Recognition and Rewards**

The data revealed that, when they do exist, rewards and recognitions may take different tangible and intangible forms. The issue of recognizing and rewarding highly-performing teachers seems to be a common trend in education. As Yuan (2015) states, the practice of rewarding teachers with high performance based on their evaluation score became increasingly common in the USA, especially those rewarding programmes that are supported by the Race to the Top Fund. Indeed, although

rewarding teachers may not be clearly linked to teacher evaluation outcomes in Saudi universities, teachers consider rewards an indication of a good performance. The following extract from an interview with one of the EFL teachers showed how recognition is perceived by participants of this study. Karam stated:

*“Recognition here means that you are **distinguished** from other teachers and so you can get **credits, bonus, overtime, vacations**....etc.”*

The statistical findings partly support the idea that teacher evaluation helps in recognizing highly-performing teachers in Saudi universities. The survey showed that 45.8% of the participants agreed that the teacher evaluation system recognizes teachers' contribution to the academic institution and 12% strongly agreed, totaling 57.8% of the entire sample, where 16.9% disagreed and 18.9% remained neutral. The data also revealed that there can be tangible and intangible rewards: tangible rewards may include certificates or shields, monetary compensation (bonuses or overtime) and salary increments; the intangible may include contract renewal. The following extracts from different one-to-one interviews show how EFL teachers perceived this aspect of rewarding. Fatin stated:

*“Yes we have always like **certificates**. They always have certificates and sometimes **shields**”*

In the same vein, Joseph commented:

*“The only reward, if it is a reward, was when I've been asked to teach on an evening class, which is **overtime**”*

Albert was less interested in the rewards offered at his university:

*“There’s only a **small bonus in salary terms**. It’s the only reward that we get.”*

Omaima seemed to be disappointed with recognition of her efforts and articulated her dissatisfaction as follows:

*“I suppose the **contract** that you can get to sign the next year is the only reward we know, and it just shows that, yes, probably we’re **still needed** and it’s okay for whatever we’ve done. Unfortunately, there is no other thing.”*

The data showed that some EFL teachers recognize the reward system taking place at their institutions, whether linked to teacher evaluation or not. They also seemed to admit the significance of the rewards in motivating teachers to perform better. The quotes may also suggest some ambiguity in the way institutions recognize and reward teachers’ high performance. In contrast to this group of teachers who reflected on how Saudi universities positively recognised their effort, another group of participants denied the existence of any reward or recognition systems at their universities altogether. When asked about rewards and recognition in their evaluation system, Daniel stated:

*“Never got that. We **never get it**. Not at all.”*

In line with the above quote and acknowledging the significance of having a reward system, Abdulhameed contended:

*“Unfortunately, we **do not have** that in our evaluation system as far as I know, nonetheless, I think this will **intensify** teachers to **run their limits** in **improving** their quality of teaching.”*

Moreover, in one of the open end questions in the survey where EFL teachers were asked about the weaknesses of current evaluation systems at their universities, one of the participants clearly stated:

*“There are **no incentives** for achieving a high evaluation”*

The data also showed that there is a third viewpoint, where teachers link high performance scores to getting an undesirable kind of recognition such as having extra work or duties. This may encourage teachers to show a low performance level in order to avoid high scores, even if it is within their capacity. The following extract from Hamdan illustrates his views on the disadvantages of having high evaluation:

*“They expressed their intentions to just get near the highest scores (4 or 4+ out of 5) **since getting a score of 5** will mean the compulsory listing that faculty’s name (who achieved the highest evaluation score) on a list of mentors who need to carry out the process of an ongoing mentoring and evaluations (without any incentive or compensation) which is an **arduous job** that takes at least 8 hours a week of **extra work** that the faculty members believe they could do without.”*

The data suggests that recognition of teachers’ performance and the subsequent reward system are not set in a way that is clear, desirable and motivating to teachers at universities in Saudi Arabia. There seems to be a need for a more professional approach to rewarding EFL teachers that is grounded in an effective evaluation scheme. Similar to many other contexts, Gosling and Hannan (2013) suggest that the best way to elevate the quality of teaching can be to reward highly performing teachers

in a selective and competitive approach. They claim that this idea was behind the foundation of the Centres for Excellence in Teaching and Learning (CETLs) initiative in the UK. This can also be linked to the issue of public trust and funding, which will be developed further in the following chapter.

#### **5. 2. 4. 4. Teachers' Self-Esteem**

One of the sub-categories that emerged from the data is how EFL teacher evaluation and self-esteem can be related to each other. As an academic endeavour, teacher evaluation can be seen by teachers as emotionally charged, which may influence the way they feel about themselves. As Reddy, Dudek, Kettler, Kurz, and Peters (2016) advise, teachers' perceptions of self-esteem depend on their performance and the way their performance is being evaluated, since teachers invest a great deal of their 'selves' in their work. Most of the participants linked high self-esteem to high evaluation scores and vice versa. For instance, the following extract from one of the interviews illustrates how Asher related them:

*"Self-esteem can either be low or high **depending** on how the evaluation process has been handled. If it's like a witch hunting exercise then **negative** results will lead to **low** self-esteem. And the converse is true"*

Samar, similarly, explained how her self-esteem declined when she once received low evaluation results:

*"Yes. I had explained it earlier that I felt like a **total failure**. When I got that evaluation, I knew what were the factors involved in that bad or **below average evaluation**. I know that very well, but still I was **distressed**."*



The extract above suggests that knowing the reasons behind getting low evaluation and acknowledging them does not help mitigate the negative effect of low evaluation on teachers' self-esteem. On the other hand, some participants confirmed that the process of teacher evaluation does not lead to any consequences for their self-esteem regardless of the result they get.

Interestingly, some of the participants attributed this state of being emotionally unaffected by evaluation to having sufficient qualifications and adequate experience.

Referring to these issues, Hamdan proudly stated:

*"I am an **experienced** and **well-qualified** teacher, and therefore, my self-esteem was **not affected** at all. However, some of the inexperienced and less qualified teachers may bear the brunt of teacher evaluation since the psychological context involved is very tense and stressful."*

The findings reveal that teacher evaluation can shape teachers' self-esteem and may affect them emotionally. In some cases, judgmental evaluative environments may result in tension and stress. In the following sub-category, how such an atmosphere may affect EFL teachers' sense of belonging and job security will be presented.

#### **5. 2. 4. 5. Sense of Belonging and Job Security**

Job security and sense of belonging have been extensively investigated in the literature. According to Albaqami (2015), job security can promote creativity for those who work in higher education. When teachers feel their jobs are secure, they can develop some sense of belonging to academic institutions, which eventually results in high performance and creativity. Teacher evaluation can create such an atmosphere.

For example, when Asher was asked about how evaluation affects his sense of belonging, he contended:

*“It does in the sense that it **fosters** a sense of a **shared goal/vision** and collegiality”*

The data indicates that having shared goals, through evaluation schemes, can encourage a sense of belonging. However, the data also shows that developing a sense of belonging may not be directly related to teacher evaluation but rather may be attributed to many various aspects of the participants' work environment. Amongst other issues that affect their sense of belonging, the participants mentioned relationships with students and/or with colleagues, satisfaction with the salary, quality of team work, commitments teachers have, and work space settings. In addition, some expatriate teachers added that their sense of belonging is affected by the way they are made to realize they are foreigners to the Saudi culture, whether by colleagues or students.

It appears that, when EFL teachers have more interaction with colleagues, they develop a better sense of belonging. For instance, one of the participants, prior to his current place of work, was working in another place at the same university where he was only rarely able to meet up with his colleagues. In his interview Joseph stated:

*“I do have a **better sense of belonging** to the prep year than I did in the other department, but that's because we're all in one building and I see my colleagues, whereas when I was in the department, we all used to go off and teach in different buildings, so you'd very seldom see your colleagues.”*

The data suggests that being physically close to colleagues helps to develop a stronger sense of belonging. It also appears that the quality of collegiality influences how teachers feel about their belonging to the institution. Although teacher evaluation appears not to have much influence on EFL teachers' sense of belonging to institutions of Saudi higher education, it seems to play a significant role in job security. This can be expected, given that authoritative bodies in Saudi universities tend to rely on teacher evaluation to make personnel decisions (for more details, see Section 5.2.1.1.). When Daniel was asked if evaluation affects his job security, he claimed:

*“Yes. Teachers are being **fired**. Actually, one teacher was just fired about two weeks ago in the beginning of the module because his **evaluations** were quite **low**.”*

Interestingly, the participants who associated job security with teacher evaluation were generally teachers of different nationalities but not Saudis. It seems that being a non-tenured faculty member can also affect the way EFL teachers feel about job security.

In one of the discussion groups, Brian contended:

*“**Tenured professors** have job security with their employer until they either decide to leave of their own or if there is some outside, like really bizarre event that happens (...) So that kind of thing really doesn't exist here probably because a lot of the faculty members teaching outside of universities are **foreigners**”*

As most of the expatriate EFL teachers at Saudi universities are generally contractees (73.9% of the sample are expatriate teachers), and it is rare to find tenured non-Saudi faculty, their residence in the country is directly related to their contracts. A number of political factors may contribute to this issue as most Gulf countries do not offer the

possibility for permanent residence or naturalization for non-citizens to become citizens. If expatriate teachers had better opportunities to become permanent residents or citizens, they might have more job security and more opportunities to become tenured professors. When job security is taken into consideration, teaching can be a predictable and stable occupation; however, and as Goldhaber (2015) states, the increased risks that are linked to teaching can lead to the perception of teaching as a less desirable job. The data also shows that some EFL teachers, including expatriates, still feel secure in their job even when they have to be evaluated. When asked in her interview about how teacher evaluation affects her job security, Farah stated:

*“What I have seen here as a contractee, **it didn’t**. The teachers who have shown low performance in their classes, they were given extensive workshops.”*

The data suggests that what follows from teacher evaluation is what indeed affects teachers’ job security. It is possible that teacher evaluation followed by professional development plans helps teachers feel more secure and at less risk than evaluation followed by an immediate decision.

### **5. 3. Challenges to EFL Teacher Evaluation at Saudi Universities**

Another theme that emerged from the data is the challenges to EFL teacher evaluation under current evaluation systems at Saudi universities. This theme comprises five main categories under which a number of sub-categories are involved. The five categories comprise challenges related to classroom observation, evaluators, EFL

students, EFL teachers, and teacher evaluation practices and they are shown in the following figure 5.5.

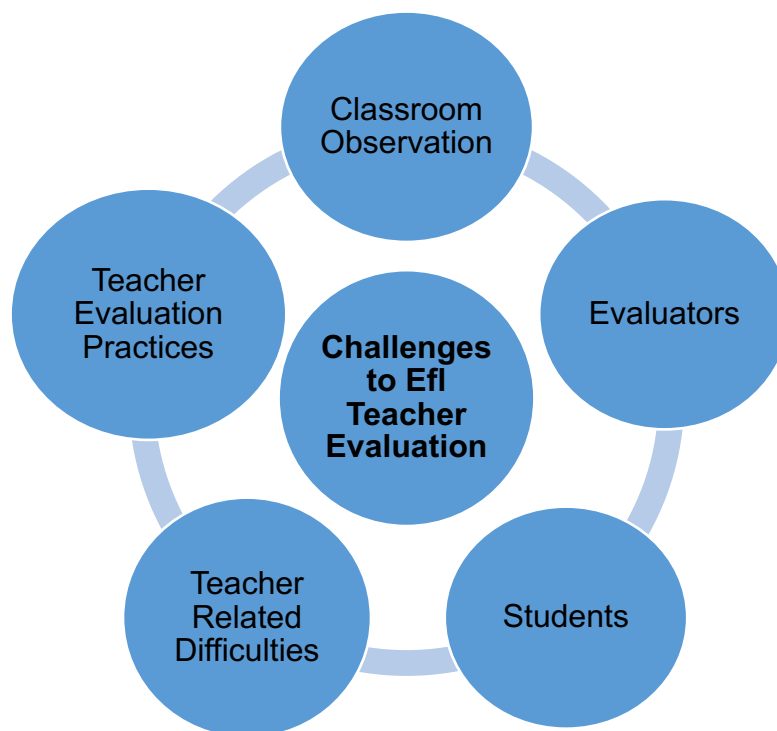


Figure 5. 5. Second theme and categories

### 5. 3. 1. “The Biggest Challenge is Classroom Observation”

The data shows that there are a number of issues related to classroom observation which hinder the success of teacher evaluation. There is a wealth of literature on classroom observation as one of the most significant evaluation measures (e.g. Arreola, 2006; Galbraith & Merrill, 2012; Seldin, 2006; Williams & Crates, 2015). Williams and Crates (2015) highlight the importance of classroom observation in providing meaningful feedback to teachers. Indeed classroom observation can be a good source for EFL teachers to reflect on their practices and how others might see their instructional practices, as the findings of this study indicate (Section 5.2.2.1. p.

160). Interestingly, some participants revealed that classroom observation can be perceived as a challenge, as indicated in the following extract from Omaima:

*“I think one of the **biggest challenge** now for us, to be honest, is that **classroom observation** itself.”*

She continued:

*“So I think that that is the biggest issue we have is just the classroom observation where we are probably **not seeing the value of classroom observation** for what it is (...) And also, the perception, the perceived notion of a classroom observation where we think that it’s a hire and fire mechanism.”*

The data suggests that classroom observation is not being used for formative evaluation and professional development. This contradicts what Philip, Martinez, Lopez, and Garcia (2016) suggest: that classroom observation offers a critical constructive anchor that can inform improvement plans for instructional practices and professional development. It seems that the participants are against its use in only forming decisions about their jobs, i.e. renewal decisions. EFL teachers seem to have concerns about intimidation and power that receive more attention than improving their instructional practices based on their evaluation. These issues will be discussed thoroughly in the discussion chapter where improving instructional practices are to be considered.

Given these factors, there may be some cases where EFL teachers manipulate their teaching materials and accordingly compromise the quality of their teaching. In her interview, Majedah hinted that some teachers ‘act’ teaching on the day they are observed:

*“The pacing guide is heavy. Sometimes I have to cover a whole unit in one session. So how come you come at the end of the year and evaluate me? Evaluate me based on what? I wasn’t teaching well throughout the year? So on that day, the teacher has sometimes **to fake a lesson**, and that’s it. So again, it’s not real. It’s not the real teaching. So the system here doesn’t help you.”*

The extract above suggests that, when teachers have a heavy teaching load, they may attempt to keep on track by just going through the materials. Yet, when an evaluator is coming in to the class, they will ‘fake’ the lesson in order to get a high evaluation. The data analysis reveals two relevant sub-categories to this category, namely large student numbers and time, which will be presented in the following section and is shown in figure 5.6. below.

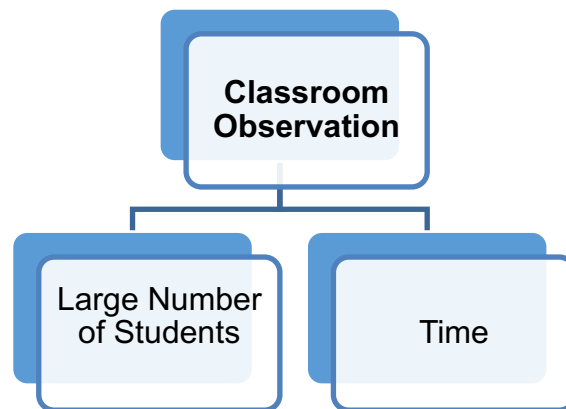


Figure 5. 6. First category and sub-categories, theme 2

### **5. 3. 1. 1. Large Number of Students**

Some of the participants indicated that the number of students inside the class can influence teachers’ performance and accordingly affect their evaluation when they are

being observed. Large class numbers were associated with different aspects that may lead to undesirable results, especially when classroom observation is to be considered. For instance, the participants in one of the focus group discussions agreed that large numbers of students can be seen as a challenge to the evaluation of EFL teachers. The following extract highlights this idea:

*“Tasneem: **Large class number.***

*Sameyah: Yes. The class.*

*Basmah: When I came here, I used to teach around **90 students** in my class. Ninety students means the whole school.”*

Basmah in the following extract raised a number of issues related to teaching classes with a large number of students:

*“That’s a very large number. And then there is **no teaching**. There is just **giving**. Teaching means interaction between me and my students, but here only me, no students (...) students are only **passive listeners**. Nobody points out anything. Just you have to prepare yourself, mug up the lecture, give the lecture, and come back. And there is **no interaction** between you and them.”*

The quote above suggests that a large number of students discourages both teachers and students from interacting effectively. Some studies in the English language learning literature have investigated teacher-student interaction in large classes at tertiary level. For instance, Abdelgadir (2012) argues that the interaction between students and their teacher in large classes can be limited, while Arora (2014) claims that, in large language classes, the “transformation of information in the form of feedback can be arduous and cumbersome” (p. 194).



In order to boost students' engagement in the classroom, researchers have suggested different actions. According to Van Uden, Ritzen, and Pieters (2014), a positive relationship with teachers enhances the engagement of the students inside the classroom. They add that students' interaction with the teacher can be fostered where students' autonomy is supported in learning environments. Also, Klem and Connell (2004) attribute the decrease of students' engagement not to the size of classroom but rather to age factors, where students become less engaged as they grow older. One of the participants, Roza, suggested that large numbers can inhibit her from paying sufficient attention to all students:

*“As a teacher, every class we have 85 to 100 students, so it is a little difficult to **give attention** to each student (...) so we have to sometimes change the way of teaching, divide them in groups, and be in contact with those students so that they can come to **know us really well** and we can be able to **know them well.**”*

Roza's statement shows that she found a link between large numbers and the difficulty of giving attention to students. As the data suggests, giving attention to students results in teacher and students reciprocally knowing each other, which could be seen as building a good rapport with students. Based on my experience both as a teacher and as an evaluator, teachers' rapport with students can influence teacher evaluation, as teachers are expected to maintain a good rapport with their students. In her interview, Samar linked large numbers to issues of control:

*“When you are new here, this is a different culture, and it is a big challenge to **control the students**, not to teach the students but to control. So with time, when you learn the techniques to control this type of students, evaluation is not a*

*big challenge itself. (...) and it was an extraordinarily **large group** comprising 86 or 88 students”*

Samar highlighted above a relevant issue to teacher evaluation related to classroom management, part of which is controlling students in the classroom. Large numbers can be an issue for new inexperienced teachers that might result in poor classroom management. However, gaining more familiarity with students and the context can lead to better management and the large number factor may become less problematic with the passage of time.

### **5. 3. 1. 2. Time Can Be a Challenge**

In all three focus groups that I carried out, issues relevant to time were brought up in the discussion. In the first focus group, Karen revealed that:

*“Taking students from A1 to B1 in such a **short time** is a real challenge as well, and especially because the institute has a modular system, so there are like **six or seven weeks** for each level, whereas the rest of the university have a semester system. So even like practically the exams for other departments which clash with teaching weeks for the institute, so we lose time there as well. So there are **time constraints**.”*

Working under a similar modular system, Brian, from another university, also stated:

*“There is not a lot of flexibility on what we can actually bring into the classroom from outside because we have so **much material** to cover in the **eight weeks of our module**. Some teachers like to bring in outside materials, but for the most part, we focus on the reading textbook, the writing skills that are being tested, and the grammar book that we use. There’s not a lot of flexibility.”*

Even at universities where a semester system is applied, the same time issue emerged from the data. Sameyah, in the third focus group, contended:

*“Sometimes, she [the teacher] tried her best, but I think **the time** for leading her, it’s not enough (...) I have to send her some links in WhatsApp. Then I point out some figures for her and arouse them.”*

It appears that, whether the institution is applying a modular or a semester-based system, time is always a challenge to EFL teachers. In other words, the total amount of time allocated to teaching a particular course is too short for the nature of the curriculum and its objectives. Some of the contextual factors that may have led to such findings are the fact that English is taught as a foreign language in Saudi Arabia, where it is not used in formal communication, and thus a large quantity of materials is needed to improve students’ performance levels. Furthermore, since students do not need to use English outside the class, their level is on average low, which was supported by the data in some of the interviews. For instance, Farah confirmed:

*“We have to present a course book plus other materials that complete the course book within six to seven weeks. The **students’ level** mostly is very **low**.”*

The following section presents data related to the evaluators of language teachers’ performance.

### **5. 3. 2. Are our Evaluators Capable of Judging Us?**

During the interviews, I noticed that the participants used the terms “evaluator” and “observer” interchangeably. It seems that, in the context of education, these two terms are frequently combined and accordingly used to refer to the same concept. For

example, Page, Thomas, and Marshall (1977, cited in Krishna, 2004) in their *International Dictionary of Education* merge the concepts of evaluation and observation; they define evaluation as the “value judgement on an **observation**, performance test, or indeed any data whether directly measured or inferred” (p. 198). Nevertheless, evaluation and observation are defined separately and distinctly in other educational literature.

In those cases where the main reason for classroom observation is to evaluate teachers’ performance and their instructional practices, EFL teachers might perceive the observer as an evaluator and hence use the two terms interchangeably. It is worth mentioning that I retained the participants’ preferences in choosing between the two terms and no further differentiation is to be made as long as the derived data feed into the category or sub-category they belong to. The evaluator-related challenges category includes the following four sub-categories: required qualifications and skills, personality and exposure to different environments, associated issues of power, and feedback concerns. They will be presented in detail in the following sections and are presented in figure 5.7.

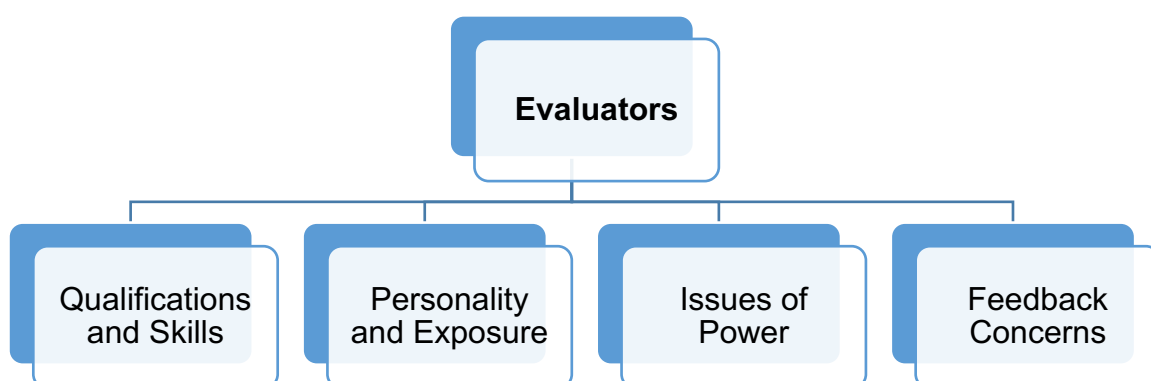


Figure 5. 7. Second category and sub-categories, theme 2

### 5. 3. 2. 1. Required Qualifications and Skills

It seems that, whenever the issue of evaluators was referred to in the interviews, the evaluator's qualifications and skills became the essence of the participants' discussion. Aspects such as knowledge, skills and experience were mentioned. Perceiving experienced evaluators as people more qualified to carry out teacher evaluation is a common notion in Saudi universities. Alaidarous (2011) contends that years of experience are a significant factor in the qualifications of evaluators. When evaluators with less experience or poor qualifications are assigned to evaluate EFL teachers, the participants of this study perceived them as one of the challenges facing teacher evaluation. For instance, Hatim stated:

*“Frankly speaking, most of the evaluators do not have any qualification, knowledge, or skills in doing this task. Some of them have **less classroom experience** than the teachers whom they evaluate.”*

Besides experience, some participants highlighted the importance of interaction skills and consider evaluators who lack them as another challenge. This idea is clearly presented in the following extract from Ashir:

*“The use and knowledge of **appropriate interpersonal supervisory** skills suited to the teacher's development level, expertise and commitment. Supervisors are meant to be aware of this in order to make any evaluation meaningful.”*

It also appears that, when evaluators of EFL teachers receive special training, it can help them become more qualified to perform their job appropriately and make their judgement sound more convincing to the teachers. Referring to this issue, Sundos emphasized the significance of evaluator training in the following quote:

*“Also, training for evaluators is essential; most evaluators have not received **specialised training**. Sometimes the evaluator is not qualified enough to make sound, informed and accurate **decisions** about teachers’ teaching.”*

The statistical analysis does not provide much support for the idea that the current evaluators at the participants’ universities are appropriately qualified, and reflects the participants’ concerns about this issue. Only 24.9% of the sample agreed that evaluators have been properly trained to consistently evaluate teaching and merely 3.2% strongly agreed, with a total of only 28.1%. Furthermore, 25.3% disagreed and 11.2% strongly disagreed, which is a total of 36% disagreement, and the rest (35.3%) remained neutral to the statement. The data suggests that challenges related to the qualifications of evaluators of EFL teachers’ performance are a multifaceted phenomenon. Having said that, EFL teachers may find themselves in awkward situations where they receive the results of evaluation carried out by unqualified evaluators. Hamdan, for instance, articulated in the following extract the doubt teachers feel about the outcomes communicated to them:

*“Some teachers may resent the idea that some equally or even **less qualified** peers are appointed as coordinators and evaluators where there is perhaps a feeling of **lack of integrity and trust** in the decisions as well as ever existence of **doubting** the authenticity of the evaluation **outcome**.”*

Interlinked to the previous aspect of the evaluator’s qualifications, it seems that not only the results of teacher evaluation may be questionable but also the overall effectiveness of the practice of teacher evaluation. A comprehensive evaluator training programme could help evaluation systems overcome such challenges. This issue will be explained in more detail in the following chapter.

### 5. 3. 2. 2. *Personality and Exposure to Different Environments*

In all five Saudi universities involved in this study, evaluators were found to be insiders and, more precisely, colleagues of the participants. Having a colleague from the same institution as evaluator has advantages and disadvantages. According to the participants, being evaluated by colleagues can allow evaluators' personality issues to interfere with evaluation results. Issues such as bias, subjectivity, penalization and grudges were reported by many participants. When Albert was asked about his reasons for considering evaluators a challenge to teacher evaluation, he claimed the following:

*“Because if you have a colleague, you always have the problem there of **personalities** entering in. **Subjectivity, bias, and then you can have grudges and personal conflicts.**”*

The data suggests that there is a link between having colleagues as evaluators and personality interference. However, Safta (2016) contends that, in any evaluation system, subjectivity is unavoidable. Arreola (2006) suggests controlling the impact of subjectivity and calls this process 'controlled subjectivity' (p. xviii). The data also shows other personal drawbacks of evaluators; for instance, Ferdaws stated in her interview:

*“They're not **emotionally and socially intelligent.**”*

The quote suggests that having emotional and social intelligence is a significant characteristic of evaluators. Although these aspects were mentioned in this study by two of the participants during their interviews, it seems that there is not much support for this issue in the literature. Seldin (2006), for instance, does not include emotional and social intelligence in the features an evaluator must have, but such intelligences

are linked to the psychological and physical health of the teacher in the literature. Some studies showed that emotional intelligence can help predict positive outcomes and educational competency for teachers (Austin, Saklofske, & Egan, 2005; Petrides & Furnham, 2001; Vesely, Saklofske, & Leschied, 2013). The data in this study also uncovered some of the desired features that participants would prefer to see in their evaluators. Apart from qualifications, a number of personal and ethical concerns were raised by the participants when evaluator's characteristics were under discussion. One example is the following extract from Mawadda, one of the interviewees, who described competent evaluators as follows:

*“They need to be **open-minded** because teacher techniques are different from one teacher to another.”*

Interestingly, Omaima not only contradicted Ferdaws but also described the evaluators she was in touch with as being open-minded, a feature that is not only desired by EFL teachers but also impacts on their personality:

*“I feel that they have a lot of **exposure**. They've been educated abroad and in the best of universities around the world (...). No seriously, this has an impact. I place a lot of value on people who have had **different exposures to different environments** because it opens up their minds and it gives them this **tolerance** and respect for many of many of us.”*

Omaima continued, placing more emphasis on ethics:

*“There will always be a general **ethical standard** and criteria that they're going by, but then every now and then, they have to make concessions for certain aspects that are foreign to the culture of the kingdom.”*



A plausible reason for such a contradiction in the data is that each institution has a number of evaluators whose personalities and experience are different. Although coming from the same university, Omaima and Ferdaws hold two opposing opinions about how open-minded their evaluators are. These opinions were directly derived from each participant's unique experience, which enriches the outcomes of this study. Such conviction on the part of the participants carries different meanings which are grounded in the context of the study. In both standpoints, the data reinforces the significance of being open-minded as an evaluator. Given that the cultural setting of Saudi Arabia is Arab and Islamic, people who have never been overseas stand almost no chance of different cultural exposure. Yet, when such people work at language institutions, they need to be in direct contact with expatriate EFL teachers who have come from different cultural backgrounds. Such socio-cultural aspects of the Saudi higher education context may play an important role in reinforcing the tolerance and open-mindedness of the Saudi people; an issue that can be linked to evaluators' quality and will be discussed in the following chapter. Bearing these issues in mind, and acknowledging that a considerable number of EFL teachers at Saudi universities are from a different cultural background, as are almost 74% of the participants, it appears that evaluators need to be tolerant of different ideas held by their colleagues from different educational and cultural backgrounds when evaluation takes place.

### **5. 3. 2. 3. Associated Issues of Power**

The data reveals that there are different issues related to power relationships between teachers and evaluators. These issues may affect the evaluation of EFL teachers in Saudi universities. There is a wealth of literature on issues of power in educational

contexts associated with different authoritative aspects. In higher education, tension might be found as a result of two different levels of power clash; according to Fulton (2013), the main sources may be authority, balance of power, and accountability. The first level is between academics as knowledge experts, the state as the funder in public systems, and other stakeholders as clients. The second level is ‘intra-institutional’ (p. 187) i.e. between working academics themselves, managers, governors and other stakeholders. Issues resulting from power clash in higher education can influence EFL teacher evaluation at Saudi universities. In his interview, Karam referred to that influence as follows:

*“Sometimes, the **observers are instructed** by their **superiors**, who sometimes do not have any qualifications or experience in English language teaching, to give X or Y a (Z) grade or rating. Sometimes these superiors are driven by **inhumane motives** like revenge, insulting, offending, or humiliation.”*

The data suggests that power relationships can affect decisions made upon observing teachers under evaluation schemes. The data also reflects the way in which educational institutions integrate and divide authority and responsibility and, more significantly, how managerial superiors may make decisions and force them upon their fellow subordinates. Hierarchy and the structure of power may be perceived in different ways by EFL teachers. In his interview Albert highlighted how EFL teacher evaluation is perceived as a sort of judgement as a result of power issues:

*“Because it’s such a delicate human issue, there’s kind of so much there psychologically in terms of **power structure**, and our whole issues as human beings with **power and authority**, and all that comes up. So it’s a huge thing really, and **it’s not taken that seriously**, except only in these very narrow terms.*

*Looked at in these very narrow terms, hire and fire issues. But it's really very big. It's like people sitting in **judgment** on another person."*

The data shows that the way evaluation of teachers is practised by authorities can be seen as a judgmental matter. Tension in higher education is likely to result from the clash between promoting efficiency, claimed by management as authoritative entities, and academic freedom, treasured by teachers. De Boer (2013) concludes that this type of tension is associated with any hierarchical structure that requires management to interfere with academics' responsibility. In the same vein, Joseph, one of the interviewees, mentioned this interference by authorities and linked it to what might be called 'filed complaints':

*"I think that the policy at this university is what I would call **laissez-faire** on the part of management. If there are no problems, (...) they assume that things are going well. So they kind of rely on, I suppose, students to complain."*

Interestingly, one of the participants in the open-ended questions of the questionnaire went further and highlighted the negative consequences of such power relationships and related it to the whole Saudi society structure. The participant contended:

*"Nothing will change evaluation systems here. All **social structures operate top-down**, where downward blame precludes upwards constructive criticism."*

This issue of social structure operation appears closely related to the issue of socio-cultural aspects of the Saudi society raised previously and will also be discussed in further detail in the following chapter in relation to stress aspects of teacher evaluation.

#### 5. 3. 2. 4. Feedback Concerns

One of the frequently reported challenges to EFL teacher evaluation relates to issues resulting from the post-evaluation/observation feedback. Some of the drawbacks linked to feedback were lack of feedback, delay, insufficiency, uselessness or questionability. In his interview, Asher claimed:

*“It actually depends on whether the evaluation or observation benefits the teacher. As it is, I'm afraid to say it doesn't. There's **no feedback** and all teachers might be under the false illusion that they're on the right track.”*

Similarly, the following quote from Hamdan reflects how delayed feedback may not be perceived as enhancing professional development:

*“My philosophy is that the feedback should be given right after the teaching session when things are hot and fresh. The last feedback on the observation and evaluation I had over a year ago was handed to me and discussed **three weeks after teaching observed**, which does not necessarily reflect a professional development process as such.”*

From Ferdaws's point of view, limited feedback is perceived as insufficient even if it is constructive:

*“It was somehow constructive but it **was not really that sufficient** because, as I told you, the aim was not continuous evaluation. It was just summative and that's it (...). I got some good feedback and then full stop, and that's it. So it was a little bit limited feedback.”*

Also, when feedback concentrates on certain criteria from the rubric instead of checking whether teaching/learning processes are taking place, teachers frequently

find contradiction amongst descriptors or evaluation team members. This idea is presented by Karam in the following extract:

*“The feedback always concentrates on whether teachers **follow the rubric** designed by the PDU, which **changes** every now and then, or **not** regardless the taking place of the **teaching learning process**. Sometimes there are many **contradictions** among the PDU people themselves.”*

In the same vein, the following quote from Huda in one of the focus group discussions illustrates the participant’s aversion to what she perceives as questionable and untrustworthy comments:

*“I will **not trust** her judgment (...) not disrespect, but let’s say I wouldn’t take her judgment or her **comments** for **trust or seriously**.”*

The statistical analysis, however, divulged some intriguing results in that almost half of the sample (48.6%) agreed that the feedback they receive from their evaluator is valuable and 17.3% of them strongly agreed, with a total of 65.9%; a total of 22.9% remained neutral and only 10% disagreed and 1.2% strongly disagreed with the statement. Moreover, analysis of the qualitative data shows that, in the view of the participants, follow-up plans should follow evaluation feedback in order to make it fruitful. For instance, the following extract from Ferdaws represents this idea:

*“It’s constructive, but it **doesn’t yield** anything. Why? Because as I said, it’s **not continuous**. You receive just a feedback, and that’s it. It has been two years now and nobody came for me to see if I developed.”*

Findings from the survey indicate that follow-up plans do not receive much attention under current evaluation systems at Saudi universities. When the participants were asked to respond to the following statement: “I regularly receive focused follow-up based on my evaluations”, only 28.1% of them agreed, while 24.1% disagreed and 34.1% decided to remain neutral. Furthermore, results from the open-ended section support the need for follow-up after evaluation: when participants were asked about the changes they wish to have in current teacher evaluation systems, one of them requested:

*“Mentorship programme to follow up the teachers’ performance.”*

It appears that the feedback teachers receive upon completion of the evaluation can be a challenge if it is not done properly or not followed up. This may be related to the evaluators’ qualifications and training, where they should be taught how to make teacher evaluation useful and give it more value. All the above mentioned evaluator-related challenges can be critically linked to the way in which evaluators are selected and make it questionable. Selecting evaluators in some Saudi universities is based on personal relationships with people in charge of selecting evaluators. In other words, favouritism tends to play a significant role, an issue that can be linked to the evaluators’ qualifications and will be discussed in more depth in the coming chapter. The following section focuses on challenges related to students.

### **5. 3. 3. Students Can Be a Challenge Too**

Students’ rating of teachers has received considerable attention in the literature of teacher evaluation (e.g. Al-Qahtani, 2010; Arreola, 2006; Benton & Cashin, 2014;

Brown, 2006; Mohan, 2011; Seldin, 2006). Analysis of the questionnaire and the interviews indicated that, where students' rating does not directly influence teacher evaluation as a measure of teacher evaluation, students may still affect the rating of the teacher. This usually occurs where students' performance level has a significant weight, especially under teacher evaluation schemes that consider students' performance as a factor in teacher evaluation. The participants of this study tend to perceive students' evaluation as an important measure. In the survey, 34.9% of the informants agreed that student rating is an effective method for evaluating EFL teachers and 13.3% strongly agreed, making a total of 48.2%, which is almost half of the sample. However, 17.3% disagreed with the statement and 8.4% strongly disagreed, with a total of only 25%, and the rest remained neutral. Analysis of the data revealed that two different sub-categories can be linked to this category of students as a challenge to teachers' evaluation as follows: ability to evaluate EFL teachers and punitive chances. I will elaborate on these sub-categories in the following section and they are presented in figure 5.8. below.

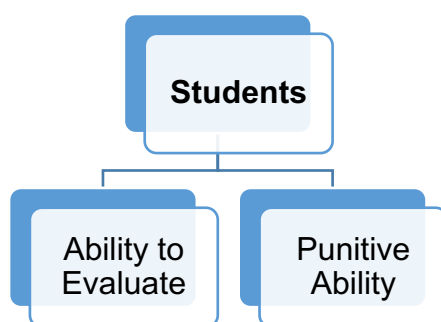


Figure 5. 8. Third category and sub-categories, theme 2

### 5. 3. 3. 1. Ability to Evaluate EFL Teachers

Despite the importance the participants attach to students' rating of teachers, the data shows that some EFL teachers have a number of concerns related to students' capability to evaluate their performance. The following extract from Samar reflects her view that student rating may not be fair given that students do not fully understand evaluation systems:

*“And that’s why I say the students either are **not fair or they don’t understand** what they are doing.”*

Interestingly, the data reveals that this lack of understanding can be linked to other relevant contextual factors. For instance, some participants make a reasonable connection between the type of students we have in Saudi higher education and the skills they typically develop in current pre-university education and consequently bring to higher education. Referring to drawbacks resulting from lack of reasoning skills and inability to produce analytical information, Brian posited:

*“And so because we’re seeing students right out of **high school** with no other instructional type other than we give you, you take what we have and put that down on your tests, then the **higher reasoning skills** and **production of analytical information** is just not present. And so if they are expected to evaluate the success and ability of their instructors, I really **don’t think** that these skills that they’re expected to produce are **present** in most of them.”*

Statistical results partially support the findings above. 27.3% of the sample disagreed that “students have the critical thinking skills that enable them to participate in staff evaluation” and 14.1% strongly disagreed, which gives a total of 41.4% of the participants. On the other hand, only 29.7% agreed with the statement and 14.1%



strongly agreed. The remaining participants, 23.7%, chose to remain neutral. These outcomes appear to contradict findings from Al-Qahtani (2010), who contended that students at Saudi universities have the required critical thinking skills that enable them to rate teachers, especially when they are compared to their Gulf-area counterparts. Apart from their skills, Roza went further in her interview and questioned students' honesty and lack of any training in how to produce valid evaluation. She stated:

*“Some teachers are not happy with their evaluation because students, all of them are **not really honest** and certainly **not qualified**. They are so young to evaluate teachers. They are not well-qualified and **not well-trained**. This is true. And if this evaluation is telling about our contexts and our future things, then it is really teachers are in **danger**.”*

The quote suggests that counting on students' rating of teachers may put teachers at risk when students are not qualified and lack training in how to evaluate teachers. This is an important factor, especially in contexts where students' ratings are used to inform decisions on keeping or firing EFL teachers. Some cases were reported where teachers were dismissed from their job as a result of evaluation carried out by students, as illustrated in the following quote from Daniel, one of the interviewees:

*“Is somebody who has never taught, at 18 years of age, **qualified** to tell somebody? We had an incident of a teacher, who was 65 years of age, had been teaching at the university all his life, and the students gave him a very, very negative evaluation, and he lost his job. Well, he changed job. **He lost his job in the university over it.**”*

In line with the previous findings, it appears that EFL teachers in the sample prefer students' rating to inform formative rather than summative decisions and plans. In response to the statement “student ratings should be used for formative purposes

only”, 45.8% agreed and 14.9% strongly agreed, making a total of 60.7% of the sample. Only 18.1% disagreed and 3.2% strongly disagreed. Where students are able to provide comments with their formal rating of teachers, which may mainly be used for developmental purposes, the language they use can be a challenge when students use their mother tongue instead of English. This issue was touched upon by some of the participants. For example, Brian stated:

*“I received comments that were in **Arabic** (...) and so I asked someone to read it for me, and there were some really pertinent comments that I appreciated but I **didn't understand** them. Most of my colleagues can't read Arabic at all.”*

The quote suggests that sometimes EFL teachers cannot benefit from students' feedback. This may be true especially given that a considerable number of EFL teachers do not speak Arabic; in this research, 42.1% of the participants were non-Arab (4.8% did not specify their nationality) and of 11 different nationalities. Apparently, this may result in useless comments where teachers cannot comprehend students' reflection on their teaching practices.

### **5.3.3.2. “The Student Has a Punitive Ability to Be Honest or Not”**

Given the limited pre-university evaluative experience of students, issues such as bias, subjectivity, rancour and other anomalies of personality might be expected when students are asked to evaluate their EFL teachers. Furthermore, as illustrated in the previous section (5.3.3.1. p. 200), students do not have training opportunities that can help them develop skills to give valid and reliable evaluation of teachers. As a result, teachers seem to be reluctant to take their students' comments seriously and tend to blame management when they give weight to students' ratings without questioning

them. Bias and subjectivity are two of the most frequently reported drawbacks EFL teachers associate with student rating. As indicated in the following quote from Tasneem in one of the group discussions, she explicitly related the lack of objectivity in students' rating to the students' feelings:

*“They’re young, 18 to 19, and they **don’t evaluate you objectively**, so that they have this kind of emotion. “If I like this teacher, I’m going to give her high marks.”*

In the same vein, Albert maintained that students' dislike of teachers or what they say and the subjectivity of their comments can cause a problem for teachers when students' comments are taken seriously:

*“Obviously, you’ve got to take into account the **subjectivity** of that (...). But you always have the danger of some odd student bringing in something **personal** or something completely beside the point and that being taken seriously. That’s the main danger of it. Could be they just **don’t like you**, or you might have **said something they don’t like.**”*

The data suggests that participants tend to link students' subjectivity to their feelings, especially when they form a negative opinion about the teacher. Samar, in her interview, put forward the view that rancour can be a reaction when students feel insulted and this, accordingly, precludes good evaluation of the teacher:

*“When we reprimand them for not maltreating the teacher or not having books or misbehaviour in the classroom, they **feel insulted** and they have some **rancour** against the teacher and they **don’t give good evaluation**. The teacher has to be here very diplomatic and very sensitive.”*

As the quote above suggests, EFL teachers may manipulate the way they deal with their students and become ‘diplomatic’ so that their rating will not be affected by students’ bias or subjectivity. The idea of adjusting the way teachers deal with their students to avoid undesirable evaluation results received support from various participants during the interviews. For instance, Brian raised the same issue in one of the focus groups:

*“**Brian:** So the student has like a punitive ability to be honest or not (...) we’ve had situations where instructors were **manipulated** by the entire class who had agreed upon certain behaviour. (...) They get together and **band together** and say, “Look, this guy is doing this. We don’t like it. **Let’s all give him a bad evaluation.**”*

A plausible reason behind EFL teachers’ diplomacy in the way they treat the students is that students’ ratings are given considerable weight in the teacher evaluation systems. It is worth mentioning that data from universities that do not utilize students’ ratings to make decisions did not produce the same findings. This also indicates that teachers are careful in the way they deal with students in order to keep their jobs and not to let personal issues affect their contract renewal. Moreover, some participants saw good ratings and good grades as conditional: in other words, if students receive high grades, they will give their teacher a good rating and vice versa. The following extract from Joseph clearly reflects the grade bias:

*“I think what you’ve got is students will give a **good rating** to a teacher who is enabling them to get **good grades**. Whether that teacher is a good language facilitator is often not necessarily the case. I believe you can be an excellent teacher of TOEFL and a pretty awful teacher of language.”*

The participants' standpoint was also confirmed by the statistical results as only 14.9% of them disagreed that "students tend to give high ranks for teachers who give high marks", while 39.4% agreed and 25.7% strongly agreed, totaling 65.1%, and the remaining 19.3% were neutral. In order to eliminate the grading bias effect, collecting students' ratings of teachers can take place mid-course instead of after releasing students' assessments grades. When asked about this, 39% of the participants agreed and 21.3% strongly agreed, which makes a total of 60.3%; 25% of the informants were neutral and only 14.4% disagreed. This may indicate that the timing of collecting students' ratings plays an important part in avoiding grading bias effect.

#### **5. 3. 4. EFL Teacher-Related Difficulties**

Under this category, challenges related to EFL teachers at Saudi universities are presented. Critical issues such as pressure, threat, trust, freedom, readiness, roles and voice of EFL teachers were frequently reported by participants and were strongly associated with challenges to current evaluation system difficulties. The way teachers perceive these concerns and relate them to their evaluation demonstrates the importance of paying attention to them in any system of teacher evaluation in higher education. In the following sections, these concerns are presented as sub-categories, of which there are five: psychological stress, fairness and equality, procedural glitches, teachers' readiness, and roles & voice of teachers (figure 5.9.).

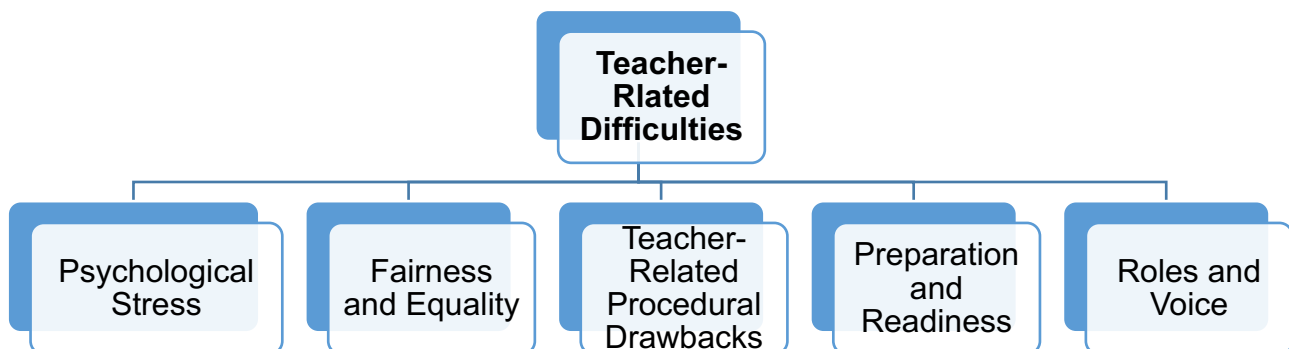


Figure 5. 9. Fourth category and sub-categories, theme 2

#### 5. 3. 4. 1. *“It is Quite Ruthless Here”*: Evaluation and Psychological Stress

Although teaching is frequently perceived as a rewarding career, given that teachers can witness the progress of their students, many teachers have reported a high degree of burnout and stress. Research reveals that teachers’ stress is highly related to their working conditions (Day, Sammons, & Stobart, 2007; Skaalvik & Skaalvik, 2015). These working conditions can be affected by several factors, one of which is the evaluation of teachers as part of the general schemes and practices of educational institutions. It seems that teachers have many concerns about teacher evaluation and this may eventually lead to stress, which in turn may lead to an uncomfortable working atmosphere. In one of the focus group discussions, Brian explained how the stress of teacher evaluation does not develop a good working environment:

*“And so by having regular evaluations and that in the **back of your mind** as punitive possibly, it doesn’t lead for a **good work environment**.”*

It is grounded in the data that the key purpose of evaluation can contribute to the stress caused by teacher evaluation. It appears that, where teacher evaluation has more of a summative than formative nature, it places teachers under considerable psychological stress. This is illustrated in the following extract from Farah, one of the participants:

*“And connecting the idea that it’s for **terminating** and the teacher who’s not going to be well-evaluated is going to be terminated from her place makes a kind of **freaking out** before the session. The teachers are **under stress**. They’re under too much **pressure** so that they don’t know are they doing well.”*

Holding the same opinion, Albert reported:

*“It’s quite **ruthless here**. But it often is the case in Saudi, if you don’t mind me saying. It’s the way they operate.”*

These critical words reflect how teachers perceive teacher evaluation, especially when the results psychologically affect them and lead to insecurity about their jobs. In addition, the data revealed that there are a number of threats associated with the purpose of teacher evaluation from the perspective of EFL teachers, for example, in the case of decision-making oriented evaluation, i.e., for hiring and firing purposes.

Majeda commented:

*“Some of the teachers here, when you come to their classroom, they see you as a **threat** because, seriously, the practicing of evaluation at this institute is a **threat**. **Hire and fire** reasons. It is a **threat**.”*

In the same vein, Albert referred to the fear caused by teacher evaluation:

*“Why some people in your institution would be reluctant to involve in evaluations because of **fear**. The fear because these are used as **tools**. You can say **weapons** really on the part of the administration **to hit you over the head**.”*

Analysis of the data also showed that some teachers relate evaluation to trust issues on the side of the educational institutions, which can contribute to any psychological stress they may already feel. According to Ayman in his interview:

*“I feel that the institute **not fully trust me** so it's through the teacher's **evaluation objectives** are achieved.”*

Having considered these aspects of psychological stress, it is worth referring to the purpose and practices of teacher evaluation. As stated earlier (in Section 5.2.4.1. p. 167) and amongst other aims, teacher evaluation in Saudi universities is carried out to accomplish accreditation goals and requirements. Fulfilling this aspect of accreditation drive, teacher evaluation schemes appear to focus on making decisions and may overlook some important resulting aspects such as the psychological impact on teachers. According to Evans, Lee, and Thompson (2016), policymakers in educational institutions should turn their attention to the ‘moral dimensions’ associated with teacher evaluation and consider the potential consequences of evaluation. In other words, and within the current reform movement in education, evaluators in Saudi higher education should consider the ethical elements of their practice. Acknowledging the pressure of evaluation, it seems that with time teachers develop some techniques to cope with the threat of evaluation. Where EFL teachers are being evaluated by



students, they may alter their behaviour to avoid low evaluation, as illustrated in the following comment from Brian in one of the focus group discussions:

*“A lot of them pay attention to their evaluations and possibly alter not their teaching style, but some of their **behaviours** in order **to increase future evaluations** from students. So things like not being very strict with attendance because students have marked them low because they marked them late when they are late.”*

The data also shows that teachers may alter their teaching practices not only to satisfy students to get better evaluation but also to please the evaluators. Karam, one of the participants, in the following extract showed how some teachers might act on the day of observation for evaluation:

*“Therefore, we have **to act** as clowns, practice **hypocrisy**, **make a show** on that day (...) Thus, we are free to teach our own way and do whatever we want to do the rest of the year.”*

This quote from Karam shows the effect of this style of observation/ evaluation on how teachers see themselves and is only one of many comments from the participants related to the psychological stress associated with the practices of teacher evaluation.

#### **5. 3. 4. 2. Fairness and Equality**

Another aspect closely related to the previous sub-category is the issue of fairness and equality in teacher evaluation. Acceptance of any teacher evaluation system depends to a great extent on the perceived notion of its fairness (Colby, Bradshaw, & Joyner, 2002; Delvaux et al., 2013; Santiago, Roseveare, Van Amelsvoort, Manzi, & Matthews, 2009). Accordingly, people in charge of teacher evaluation may need to improve the fairness of the system in order to make it more acceptable to teachers.

The issue of fairness and equality was frequently reported by the participants, as in the following extract from Karen in one of the focus group discussions:

*“Another thing, (...) a lot of the **contractees** feel it’s **unfair** that they’re observed while their **Saudi colleagues are not** observed. And I know the status is different between contractees and Saudis, but there should be a feeling of **fairness**, I think.”*

This quote suggests that some expatriate teachers have experienced the unjust or prejudicial treatment of different categories of EFL teachers in the same institution and especially on the grounds of their nationality as described above. In the same vein, Majeda in her interview reported that teacher evaluation at her institution usually targets a specific group of teachers in the following quote:

*“But however, even if it is for hire and fire reasons, it should be **fair for all teachers**, not only for a specific **group** of teachers.”*

The following extract from Karam, talking about the greatest challenge to teacher evaluation, illustrates different types of unfairness:

*“**Fairness**. That is to say, teachers should feel that their superiors are **fair and just** (...) when they evaluate them. For example, I was told by some observers that they **were asked** and instructed to give X teacher 3 out of 5 and Y teacher 4.5. I was also told by other observers that they gave Z teacher 2/5 then they **discovered** that that teacher got 4/5. (...) In short, your **relationship** with some administrators determines your evaluation mark most of the time.”*

Asher in his interview mentioned salary unfairness in some institutions:

*“Unfortunately, teachers are **paid** according to their **nationality** not their expertise, expertise etc. If you are from Europe, America, Australia, native speakers of English then your **salary is much higher** than those from other countries.”*

As the data suggests, issues such as unfairness, inequality, subjectivity, hypocrisy and racism may influence some systems of EFL teacher evaluation in Saudi higher education. When any of them occur, EFL teachers tend to perceive them as challenges to the teacher evaluation schemes at Saudi universities. Having said that and despite the significance of social justice leadership, previous research shows that issues related to justice in education seem to reside in the margins of preparation programmes in educational leadership (Karaxha, Agosto, & Bellara, 2014). However, fairness and equality should be an essential part of the evaluators' training programmes before they are assigned to evaluate EFL teachers. Some participants blamed inequality on the standards for teacher evaluation. In the open-ended question of the survey on changes to be made to the current evaluation system, one participant contended the following:

*“**Standards** should be **equal** for all nationalities”*

This finding, however, was not highly supported by statistical analysis of the survey, in which it was found that only 23.3% of the participants disagreed with the following statement: “The standards used in the teacher evaluation system are fair”, whereas 43% agreed and 33.7% remained neutral.

### 5.3.4.3. Teacher-Related Procedural Drawbacks

In this sub-category, and moving away from stress- and fairness-related issues, procedural issues will be explored and presented. It is apparent that, in order for any teacher evaluation scheme to be successful and effective, all parties involved in the process and/or outcomes of evaluation should be made aware of the scheme. Not only should policymakers evidently set out reasonable aims, clear procedures, probable findings and potential consequences but they should also communicate them clearly to evaluators, teachers and other related bodies in advance. The data shows that failure to do so may make teachers perceive the whole situation as a mess and therefore challenging. Some participants revealed that teacher evaluation had not been made clear to them. For instance, Fadwa and Karen, in the following quote from their focus group discussions, commented on the vagueness of their teacher evaluation and lack of clarity for evaluators:

***Karen:** There isn't really a lot of discourse around teacher evaluation. It **just happens**, and people are just told, 'You will be observed at this time on this day. You can have a pre-observation meeting if you like,' and that's it. I mean, the criteria are sent by email (...) but they're very vague. On short notice (...).*

***Fadwa:** Yes **very vague**.*

***Karen:** Contradicting between different criteria, between different descriptors. There are contradictions. And we very often find that even when the **observers** come to the classroom, they're **not really aware** of what they should be looking for. They themselves are not aware of the criteria."*

As the quotes suggest, both teachers and evaluators need to be fully aware of the criteria and the teacher evaluation system. This may also provide support for the need for teacher evaluator training programmes (section 5.3.2.1.) before launching any

teacher evaluation scheme in any educational institution. With regard to the quantitative data from the questionnaire, the analysis of the statement “the teacher evaluation instrument includes clear explanations for each performance descriptor” revealed contradicting results, where 57.8% of the participants agreed and 26.1% were neutral. The need for information about teacher evaluation was, however, stated very clearly in another focus group discussion. The following extract from Brian raised the issue of lack of information:

*“One of the challenges also is knowing the importance of evaluations. There’s been **no information**. In the faculty handbook, there’s no description of the importance of student evaluations or **how much weight** is given or **what it’s used for**. (...) As far as the self-evaluations, it’s the same. We know that we have to do it at the end of each year (...), but we don’t know exactly **how much weight** is put on that self-evaluation or **what it’s used for**.”*

In the same vein, the following quote from Samar illustrated how establishing clear standards may motivate EFL teachers to manage their performance in order to achieve those standards:

*“I think they should tell the teachers **why**, or what are the **good things** they saw in the teacher who they have selected or who they have chosen for the award of “Best Teacher.” Others would **be motivated**, and we will be able to prove ourselves **according to their lines**.”*

The data suggests that teachers need to be informed about all details related to teacher evaluation schemes such as objectives, how much weight is given for each method, standards and criteria. Besides clearing up any vagueness around the system, informing teachers will help motivate them to reach the standards and work

hand in hand with their institutions. One plausible scenario for these issues could be the changing and updating of teacher evaluation systems without providing sufficient information for related parties. Hence, establishing a solid system can help in making it more stable and comprehensible for teachers. If the need for updates arises, then teachers need to be involved and informed. Another procedural challenge reported by the participants in this research is the lack of communication between teachers and evaluators. When asked in his interview about his views on the greatest challenge, Ayman replied:

*“To inform teachers **silently** through their sites about the results of evaluation to know the points of weakness and points of success.”*

Ayman draws attention to how EFL teacher evaluation is practised in his institution. What he experienced was the results of his evaluation coming through his online account of the university, without any opportunity to have a dialogue channel with the evaluators. Nevertheless, the analysis of the quantitative data did not seem to provide a strong support for the lack of dialogue issue. More than half of the participants, precisely 55.4%, agreed with the idea that the teacher evaluation system enhances dialogue between teacher and evaluators about effective teaching, while 22.9% remained neutral. These contradicting findings may be a result of involving different Saudi universities where varying teacher evaluation systems are implemented. It appears that, where a committee is assigned to carry out teacher evaluation duties, adequate communication usually takes place, as the selected committee members dedicate some time to communicating with the teachers they evaluate. In the following sub-category, ideas around the preparation of teachers will be presented.

#### **5. 3. 4. 4. Teachers' Preparation and Readiness**

As illustrated in the previous section, EFL teachers' awareness of the evaluation procedures is important to make the system successful. Preparing teachers is one way of ensuring they are aware about their evaluation and ready for it. Analysis of the data showed that preparing EFL teachers in any way can help build a strong evaluation scheme. When the informants were asked in the open-ended section of the questionnaire about the strengths of the current evaluation system at their university, some answers reinforced this notion, as, for instance, in the following quote:

*"Has plenty of **preparation** and **discussion** before the actual evaluation"*

Given the diversity of contexts and backgrounds of the participants, preparation was not always embraced by people in charge of teacher evaluation system. When preparation is present, it takes different forms, depending on how EFL teachers perceive those forms. The first and foremost available yet neglected source of information may be the faculty handbook or manual. As a way of preparing EFL teachers for teacher evaluation, the teachers' handbook was reported as a valuable source by only two participants during the one-to-one interviews and was completely overlooked in the focus group discussions and the questionnaire. For instance, the following extract from Omaira shows her perception of the faculty handbook as a source of valuable information:

*"We have a **faculty handbook** and there are some bulleted points in there regarding the points that we need to **be evaluated on**, and they are all those things, the general points that I outlined before."*

There were other forms of preparation reported by the participants, including pre-evaluation mentorship programmes, as illustrated in the following quote from one of the participants in the open-ended section of the questionnaire:

*“It is based on both peer-evaluation and committee evaluation. And it comes after a **mentorship programme**”*

Hatim referred to informal observation as a sort of preparation in the following extract:

*“There is a sort of preparation, as we have an **informal observation**, followed by a **feedback session** - the observer feedbacks the instructor on his performance-. But no, **I don't think it is sufficient.**”*

Majeda considered checklists as the only source for preparation:

*“Only that form, and the **form is like a checklist** (...) One-page checklist, and that's it.”*

Ferdaws mentioned criteria rubrics:

*“We just received a **rubric for the criteria** of the evaluation.”*

Mariam, however, took the issue further and explained how these modest forms of preparation are unlikely to be sufficient:

*“Actually, the teachers maybe have received something that you will be evaluated, but that's it. That's about it (...) **the email**. So just sending an email, I think it's still **not the ideal way** of doing it because like we said, this is their practice, this is their growth, so you should at least meet them, discuss things, tell them what's really going on, explain it to them one-on-one. Include them and tell them what's going on.”*



When asked about challenges to the current teacher evaluation system, Majeda raised the lack of appropriate preparation as an issue:

*“Second thing is that teachers are **not well-prepared** to be evaluated because most of the teachers, they don’t know the value of it.”*

In response to a question on whether he had received induction or written instructions on how evaluation works, Joseph stated:

*“No. No. Neither when I joined the department nor here. One **gleans it** from one’s colleagues.”*

Furthermore, in the open-ended section and in response to a question on the changes that should be made to their teacher evaluation system, one of the participants suggested the following:

*“It should take the staff **psychological and social readiness** before taking decisions based on his evaluation.”*

The quote highlights the importance of considering not only preparing teachers in terms of the official procedures but also in terms of the teachers’ psychological and social readiness. This point can be directly related to the lack of awareness presented in the previous sub-category. Raising EFL teachers’ awareness about teacher evaluation at their institutions can help in elevating the level of their readiness. The data also revealed that insufficient preparation of teachers may result in confusion

about the whole system of evaluation, as illustrated in the following extract from Hamdan in his interview:

*“At this institute, I was **unable to form a clear opinion** or idea regarding the teacher evaluation since I honestly thought initially that professional development was not the aim; it was rather an evaluation for retention or removal of the teacher. The first time formative teacher evaluation, I perceived it as just a process to make sure that I **was following certain instructional guidelines** rather than a **process for Professional Development**. I was **not given any instructions** regarding the evaluation process even the main points of what is it that teachers need to be conducting in their lessons in the classrooms.”*

The data suggests that EFL teachers may perceive teacher evaluation practices differently. In contexts where teachers' handbooks, mentorship programmes, and informal (formative) evaluations are available, these can be good opportunities for helping EFL teachers to be prepared in a way for teacher evaluation and to enhance their readiness. It appears that at some Saudi universities, where policymakers are concerned about implementing teacher evaluation systems for whatever goal they seek, the readiness and preparation of teachers beforehand may be overlooked; this is especially the case where there is a lack of specialized committees and well-qualified and certified trainers who work hand in hand with the evaluation committee in order to prepare teachers for evaluation beforehand and to help them achieve their evaluation goals afterwards.

#### **5. 3. 4. 5. Roles and Voice of Teachers**

The participants were asked about the role they play in the development of teacher evaluation practices and their voice in teacher evaluation at their institution. The data revealed that in seventeen out of twenty-four of the interviews and group discussions,

the participants confirmed in twenty different incidents that they do not have a role to play in the development of the practice. When asked in one of the interviews about the teacher's role, Majedah compared the teacher to a controlled robot:

*“She’s a **robot**. She has to have a course book and follow that course book page by page. Not only that, she has also to follow a pacing guide imposed on her (...) totally controlled.”*

However, a few participants reported certain roles EFL teachers do play in the development of teacher evaluation practices in their institutions. For instance, the following quote from Oaima shows that there may be a limited opportunity for teachers to reflect on strengths and weaknesses of the practice in a written formal observation form:

*“Well, the **lesson plan** that they make, for example, that has to have a **box** in which they [observed teachers] write their **strengths and their weaknesses** and things like that. So that is there. Otherwise, generally they have **never sent out a questionnaire** ever asking us what **we think of the observation** or their **evaluation**.”*

Furthermore, Fatin in her interview highlighted a further influential role of EFL teachers where some evaluation criteria have changed based on teachers' complaints:

*“It was done the year before, three years ago. They were having some complaints about certain **criteria** and it was discussed and things were **changed**.”*

Another participant, Farah, considered informal talks between teachers as a concealed role of EFL teachers that can help convey their opinions to the committee responsible:

*“Well, there is a role but it’s a hidden one. The committee starts to know of teachers’ practices from speeches, **from talks**, from chats and things like that, but not a real involvement into the case. So we don’t build the criteria ourselves. The teachers are not asked, “What criteria would you like to do?” The rules come or **the criteria come** to us from the **male section**.”*

She also contended that evaluation rules are imposed on them and usually come from the male section (given the gender-segregated educational system in Saudi, males and females have two separated campuses).

In the same vein, Ayman highlighted the idea that the evaluation process is launched by the institution without involving teachers:

*“In the evaluation process in my institute there is **no role** for teacher because the evaluation process is **launched from the institution**.”*

As the data suggests, the limited opportunities for EFL teachers’ involvement in developing the evaluation practices may be due to the fact that the system and criteria are developed by a specific group of people. This reflects the centralized nature of the evaluation system in Saudi higher education. Another critical stance that can be inferred from the two quotes above is that the segregated educational system appears to give extra power to its male members and also that sexism influences the practices of teacher evaluation in the Saudi context. Concerns of centralization, sexism and their effect on EFL teacher evaluation in educational institutions of Saudi higher education will be developed further in the forthcoming chapter.

Acknowledging the limited opportunities for EFL teachers' involvement in developing teacher evaluation systems, some participants claimed that teachers can have a positive role and suggested some channels to activate them. For example, Abdulhameed proposed distributing periodical surveys to collect teachers' feedback and reflection on the evaluation procedure:

*“According to reality, I think that **teacher does not have a worth mentioning role** in developing evaluation practices. Because evaluators do not think that evaluated subjects should have a role in such matters. However, in my opinion, teacher must have a significant role in this regard. Teacher can provide **constructive suggestions** to develop teacher evaluation practices. There should be a **periodical survey** based on which teacher evaluation practices developed.”*

When the participants were asked about channels open for them to air their voice, they seemed to have a common perspective, namely that they don't have enough opportunities to express themselves. For example, Joseph in his interview made the following comment:

*“It's only **if there's a problem** that I think you are given a chance to explain. Only then.”*

The data shows that in other universities there may be other channels, as illustrated by Karam:

*“We do have **teacher representatives** who discuss very few issues with the administration. By the end of the day, they do what they have in their closed and programmed minds.”*

As the quote suggests, this teacher's perception of teacher representatives is a negative one, in that they discuss very little and it is usually fruitless. In the same university, Hamdan reports on another channel allocated for EFL teachers to reflect:

*“Again, it has to be said that there is a **section** in the written report of the evaluation where the **teacher** being evaluated **can write a full statement of objection and clarifications** for such an objection. Additionally, I have heard of a couple of situations where the teacher has objected, to a minor degree, to a certain element in the evaluation and that objection was **looked at and positively reflected** in favour of the teacher and the grade was changed up slightly.”*

In this case the teacher's voice and comments were taken more seriously and the grade was changed accordingly. Also, from the same university, another participant referred to a further channel for communicating their concerns regarding teacher evaluation. In the following extract, Majeda comments on the right of teachers to disagree with the choice of evaluator:

*“And you have the luxury, if you feel uncomfortable by the evaluator, you have the luxury, for example, to **suggest another one or replace the evaluator** from the very beginning. You can just reject or say, ‘I would like a substitute, or to substitute that evaluator.’”*

Despite those opportunities, some of which were reported as ineffective, the majority of the participants believe that they need more chances to have a more effective role in the practices of teacher evaluation. For instance, in the open-ended section of the questionnaire, where participants were asked about the changes that should be made to the current evaluation system, a number of participants called for greater

involvement of teachers. In the following quote, the informant comments on the importance of involving teachers in preparing criteria and descriptors:

*“The issues included in the survey **should be discussed** with the teachers before preparing the questionnaire.”*

The data also shows that there can be obstacles preventing the teachers’ voices from reaching the people in charge, as illustrated by the following quote from Samar:

*“My point is that the girls have freedom to approach those authorities. And the **teachers don’t have the freedom to approach those authorities**. So doesn’t it look biased towards the students?”*

It appears, however, that when the results of the teacher evaluation are serious, EFL teachers can find a way to reach authorities and object. For example, in one of the focus group discussions, the participants told a story about a colleague who was given a letter informing him that he may not get his contract renewed. Brian and Daniel described the situation:

*“**Brian:** He went **straight** to the dean.*

***Daniel:** He did, and the dean said, ‘You will have to see what was **written about you**. You will have to see it.’ And that was a big **embarrassment** to the **head** of the department [who did the evaluation and wrote the report] at the time because he had to come down specially and open up what he wrote about him and give a reason as to why he wrote about it. But that was only done because **he went straight to the dean.**”*

The quote suggests that there can be pathways for teachers to express themselves and to object to evaluation reports that were written behind their backs by their direct

managers in contexts where teachers are not given a chance to reflect on the results, such as with a space in the report itself. The data also shows that issues of power relationship can contribute significantly to the lack of involvement of teachers and their voices in teacher evaluation. Teachers tend to see evaluators as representatives of higher authorities involved in teacher evaluation and hence avoid taking an active role in teacher evaluation. For instance, Hamdan articulated his opinion in this respect in the following extract:

*“It is worth noting that the evaluators might be seen as **representatives of the administration**, and when **elements of politics** are involved in professional development, I feel teachers will **not** take an **active and constructive role** in the process of teacher development.”*

Holding a similar standpoint, Joseph suggests that EFL teachers may intentionally avoid having a voice and the reasons behind that:

*“I worked for another educational institution (...) and when I went there, a senior British colleague said, ‘I’ll give you two pieces of advice, Joseph. Number one: **Keep out of the limelight**. And number two: If you do get in the limelight, **don’t dance**.’ And what event you know of course; just **keep your head down**. Don’t be conspicuous. But if you do get put on the spot, just **don’t say anything that is radical** or that is likely to **change** things. **Don’t suggest** anything that would change or **improve** things. Just **keep quiet** about things. And that’s the predominant philosophy in... I mean, it’s not just in the university or the society. It’s **in all top-down societies**.”*

It is worth noting that this idea of avoidance and keeping a low profile was mentioned by other participants too. Interestingly, this participant, Joseph, has linked this passive attitude to the top-down structure of the institutions and/or the structure of power on the whole society. This finding is also consistent with some researchers’ stance on



teachers' voice and role. For instance, Kang'd (2015) emphasises that when teachers have little or no voice, it can be a reaction to the top-down decisions made by administration and the hegemonic forces around them. This is significant and may lead to the isolated practices of the teachers and more centralization of decisions, an issue that will be discussed in the next chapter under "alienated teachers" section.

### 5. 3. 5. Current Teacher Evaluation Systems

In this category, issues reported by the participants and directly related to teacher evaluation systems in general will be presented. Analysis of the data revealed that there are a number of challenges related to the evaluation of EFL teachers at Saudi universities in different areas linked to the current practice, such as transparency, confidentiality, the criteria and, in some cases, the modular system. Before focusing attention on these areas, I will present some viewpoints on the system itself. For example, when the survey participants were asked in the open-ended section about the weaknesses of the system, one participant stated the following:

*“Due to **social and economical** circumstances, a teacher might not prove his real competence he already has. So, the **whole system** might be **unfair** to him.”*

This suggests that evaluation systems need to consider social and economic factors in order to be fair to those involved. The participant seems to link the unfairness of current EFL teacher evaluation system to the lack of flexible standards that allow for teachers' different economic and social conditions. For instance, if there are not sufficient teaching resources available for teachers, the evaluation of a teacher should not be affected accordingly. It seems wise for evaluators to consider that not all

teachers are equally capable of affording teaching resources and materials. The statistical findings also provided information regarding participants' level of satisfaction with the current evaluation system. Only 32.5% of the participants agreed that they are satisfied and 31.8% disagreed, while the majority, 35.7%, decided to remain neutral, suggesting that current teacher evaluation systems at Saudi universities are not seen as acceptable to EFL teachers. In the following sub-categories, more issues related to the current teacher evaluation practices will be highlighted. These sub-categories are displayed in figure 5.10.

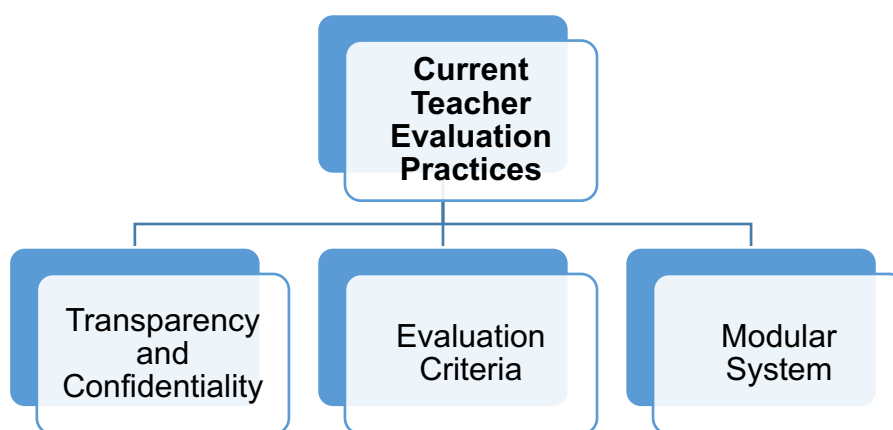


Figure 5. 10. Fifth category and sub-categories, theme 2

### **5. 3. 5. 1. Transparency and Confidentiality**

Transparency and confidentiality were reported as challenges to teacher evaluation systems in Saudi higher education. Transparency seems to be frequently used as a flexible sweeping term to convey different messages. As Michener and Bersch (2013) claim, a crucial issue with the term 'transparency' is its ambiguity: it offers a "nicely ambivalent concept, with a positive normative charge" (p. 233). The two conditions

they set to identify and evaluate transparency are information visibility (as in “light rendering an object entirely visible” p. 234) and inferability (i.e., ability to drawing precise assumptions from it “which can be inferred with some degree of accuracy” p. 234). It seems that visibility of information rather than inferability was associated with transparency of evaluation when it was mentioned as a challenge by the participants. In one of the focus group discussions, when asked about challenges to teacher evaluation, Karen and Fadwa considered transparency as one of the biggest challenges:

*“Karen: It is not convincing, but explaining to the teachers the purpose for the observations, any kind of evaluation. **Being transparent**, it is a big challenge.*

*Fadwa: Yeah. **The biggest challenge is transparency.** There should be only and only impartial observers, well-trained observers should come to observe.”*

As the quote suggests, with the absence of transparency the evaluation system may sound unconvincing to the teachers even if the purpose has been communicated. This may also be closely related to the lack of sufficient information and preparation discussed earlier (Sections 5.3.4.2. p. 209, and 5.3.4.3. p. 212). Given that people in charge tend to make decisions that are based on teacher evaluation in isolation, teachers may perceive the whole practice as lacking the required transparency. This may also leave EFL teachers wondering about the core reasons behind evaluators’ visits to the classroom. Farah, for example, articulated her confusion in the following quote:

*“We feel that the observers come to **terminate** or to write a report about the real **performance**. You feel that there is **something behind**. They came to observe me because there*

was a **complaint** (...). You don't know. There is not this kind of **transparency**."

In the same vein, Ferdaws raised not only the issue of transparency but also the issue of confidentiality with peer observation:

*"They did not encourage it [peer observation], plus there is **no transparency** (...). And it's not taken completely **anonymous** (...). So for example, if you go and observe one teacher, peer observation (...) normally the feedback and everything must stay between her and you. It shouldn't go outside. It should be **confidential**. So they **talked about each other's practices** inside the classroom. There is no transparency. There is **no Confidentiality**."*

Interestingly, the data shows some results which contradict previous research on transparency and peer evaluation of teachers. According to Goldstein (2007), peer assessment and review make teacher practices and decisions based on teacher evaluation more transparent than the administrators' evaluation of teachers. Statistical analysis of the survey provides partial support for the effectiveness of peer observation; when the participants were asked about the following statement "classroom observation done by peers is an effective method to evaluate teachers", 56.2% of the participants agreed, 21.7% disagreed and 22.1% remained neutral. Peer evaluation may help EFL teachers develop a better attitude towards teacher evaluation and perceive it as a valuable practice if it is carried out in a supportive way. In her one-to-one interview, Majeda articulated this idea as follows:

*"So just start with **the peer**. And then after encouraging teachers to that, teachers through a peer **will see the value** of evaluation and will get used to evaluation (...). I prefer it is on friendly basis. But yeah, I can't assure that, and some things could happen. It's our society. **We love gossips**."*

As the quote suggests, peer evaluation may encourage teachers to value the process but, on the other hand, it still may not ensure confidentiality. There is little debate in the literature on this issue, although Dodd (2015) contends that teacher evaluation written by academic leaders should be confidential and not available or approachable by outsiders. It appears that peer evaluation is more transparent but less confidential than the administrators' evaluation of teachers. A reasonable explanation may be that teachers get insufficient training in teacher evaluation and its related ethical considerations and hence pass on some confidential information about their colleagues.

### **5. 3. 5. 2. Evaluation Criteria**

In any teacher evaluation system, evaluation criteria seem to be the cornerstone and hence they should be clearly communicated to the teachers. As Seldin (2006) stresses, vague criteria can be seen as one of the most common deficiencies in programmes of teacher evaluation as they leave teachers in the dark about their expected performance. Complaints about teacher evaluation criteria were frequently repeated by the participants in many different ways. For instance, Ayman, when asked if he knows the evaluation criteria, contended that teachers were not informed about the criteria by which their performance will be assessed:

*“Yes, I do. The evaluation **criteria are a lot, but they aren't informed to me.**”*

From another Saudi university and in the same vein, Brian explained how he knew about the criteria of evaluation:

*“It’s just **by word of mouth**. They go on a 10-point scale. We get our modular evaluation results, and if the results are somewhere between 8 and 10, that’s good. If it’s below 8, it’s not so great. But if you’re getting 6’s, then you should be concerned. **That’s what I’ve been told.**”*

In other universities and where teacher criteria are communicated to EFL teachers, other factors related to those criteria were reported by the participants as challenges to teacher evaluation. For example in one of the focus groups, Karen highlighted how contradictions between criteria or descriptors within a specific criterion can be a challenge:

*“Contradicting between different **criteria**, between different **descriptors**. There are **contradictions**. And we very often find that even when the **observers** come to the classroom, they’re **not really aware** of what they should be looking for. They themselves are not aware of the criteria.”*

As the quote suggests, contradicting criteria may confuse teachers and evaluators alike, where evaluators become unsure what they should look for when they carry out teacher evaluation. Fatin also contended that some of the criteria used to evaluate EFL teachers at her university are unmeasurable:

*“The criteria itself. I think it should be something really **measurable**, something that doesn’t **depend on personal opinions** or things like that. It has to be like, 1, 2, 3 if you follow these things then you’ll get your best evaluation so it is not something personal (...). I think if they have like **specific criteria** it’s going to be **more accurate** and ends up with like better results.”*

As the data suggests, clear criteria that do not depend on personal opinion or the evaluator’s subjectivity can make teacher evaluation more accurate and help the

teacher to have better results. Danielson and McGreal (2000) claim that inappropriate or unclear criteria to evaluate teachers may turn teacher evaluation into a “meaningless exercise” (p. 7). Another challenge that was also reported by the participants is having random criteria. When asked about the importance of teacher evaluation, Majeda linked lack of importance to the aimless nature of teacher evaluation system at her university, which may result in random criteria:

*“It is not important because the aim is not clear. Because you don’t have an aim, so the **criteria**, I would say, they are very **random**. Random criteria, vague aim, and the system don’t help the evaluation itself.”*

It appears that identifying appropriate content for criteria by which EFL teachers can be evaluated is a complex task. Criteria for teacher evaluation may be linked not only to the general aim of evaluation, as the data indicates, but also to the emerging and traditional assessment standards (Dwyer, 2012). Hence, evaluation criteria should be based in general on teacher performance standards. Analysis of the quantitative data derived from the survey shows that the majority of participants (53%) agreed that the evaluation standards do a good job of defining effective teaching, and only 17.7% disagreed, while 29.3% remained neutral. Accordingly, these standards need to be used as a basis for teacher evaluation criteria.

### **5. 3. 5. 3. Modular System**

In relation to the previous sub-category, it seems that improving the standards and criteria for teacher evaluation has little effect on some Saudi universities. Another challenge to teacher evaluation as reported by the participants is the modular system employed in some language institutions at Saudi universities instead of the semester-

based system. Teachers teach one level of English over seven to eight weeks, officially known as a module, which is half of the term. The module usually includes mid and final exams along with other speaking and writing tests. In one of the open-ended questions where participants were asked to mention the weaknesses of the evaluation system, one participant commented as follows on the impact of the system on evaluation:

*“It’s **difficult** to achieve the required **standards** due to the pressure of the **modular system**.”*

In the same vein, Daniel in one of the focus group discussions asserted:

*“One of the challenges is the **modular system** is fixed in itself. I have students now (...) failed the first level, they passed it the next time, just **scrape** by, and then by accident they scrape by in second level. Okay? And I’ve spoken to at least two of their teachers, and they said, “They were terrible. I didn’t even think they will pass.” And I have them in a higher level, third level (...) and they’re absolutely **demotivated**. I don’t know what I can do with them.”*

As the quote suggests, the modular system may leave EFL teachers with groups of demotivated students who scrape by from one level to the next. In a different focus group discussion in another university Huda and Karen related modular system challenges to factors such as syllabus and time:

*“**Huda**: But the thing is the level of the students is not the main thing. It’s the **modular system** that necessitates that these students need to take this kind of syllabus (...) whether it is up to their level or not. So I have to teach it. I cannot change the pacing guide (...) I cannot modify it to—*

***Ghada**: Yes, modify it with my own content.*



**Karen:** *Taking students from A1 to B1 in such a short time is a real challenge as well, and especially because the institute has a **modular system** (...)*

In her interview Mariam contended that what she is doing inside the classroom cannot be considered teaching, as she explained:

*“I’m **not teaching**. I’m not evaluating the students or re-teaching them. I’m not giving them **a chance to learn**. I’m not nurturing them or making sure that they’re growing or learning or enjoying, anything. **I’m rushing** to finish because they have exams. Exam after exam after exam. So we all know that this is the situation with the **modular system**.”*

As the data suggests, EFL teacher evaluation can be negatively and indirectly affected by the modular system, where students can be found below the expected level and demotivated to learn the language, which may, in turn, discourage teachers. Another indirect factor is the amount of time allocated compared to the amount of materials to be covered in the syllabus in a semester. Given that there are committees in some universities that are responsible for setting the curriculum, syllabus and pacing guide, it may be important to bring these issues to their attention, especially if they are not involved in teaching or have limited teaching practice (as the data suggests). In universities where students’ evaluation is the primary method of evaluating EFL teachers, the participants reported a direct link between the modular system and teacher evaluation. For instance, Daniel in the following quote describes how in the first module students highly value their teachers and give them a high evaluation rate:

*“And another thing that has come in with that **modular system** is they give the **average of the two evaluations**. They put the two modules together for the semester, and they give you the average. Now when pupils come in September,*

*they are normally with a background that is rather dubious sometimes because the teacher always just spoke in Arabic a lot in the classroom. They often say that you're the first teacher that keeps speaking in English to us, and **they think you're the best teacher** very often. I find this anywhere. The only thing about the teacher evaluation is the fact that the modular system is **not really conducive to a good evaluation.**"*

The quote suggests that evaluating EFL teachers two times per semester, i.e. once per module, and calculating the average does not positively contribute to the evaluation of teachers. It appears that students in the preparatory year at Saudi universities may tend to generously evaluate EFL teachers when they first join the university and as a result produce unfair evaluations. Once again and as illustrated earlier (section 5.3.3.1. p. 200), students need to be educated and trained in how to take an effective part in teacher evaluation. They also need to know about its importance, practice and consequences so they can take it more seriously and do it more professionally.

#### **5. 4. Suggested Solutions for Better EFL Teacher Evaluation**

Having explored the participants' views and ideas on challenges related to EFL teacher evaluation at Saudi universities, I will now present their insights and suggestions regarding ways in which teacher evaluation can be improved. The solutions recommended by the participants mostly fall into two categories: solutions on a micro-level, to be attained with efforts from the institution itself, and solutions on a macro-level, where entities other than the institution need to be involved (figure 5.11. below). In the following section, solutions on the micro-level will be presented.

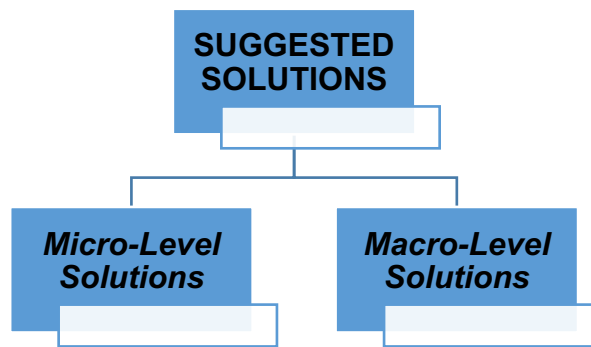


Figure 5. 11. Third theme and categories

### 5. 4. 1. Micro-Level Suggested Solutions

These suggestions include the following: introducing teacher evaluation, involving related parties, appointing qualified evaluators, advancing the practice, fixing existing problems, and introducing a different system. Each one of those suggested solutions will be presented in the following sub-categories (figure 5.12.).

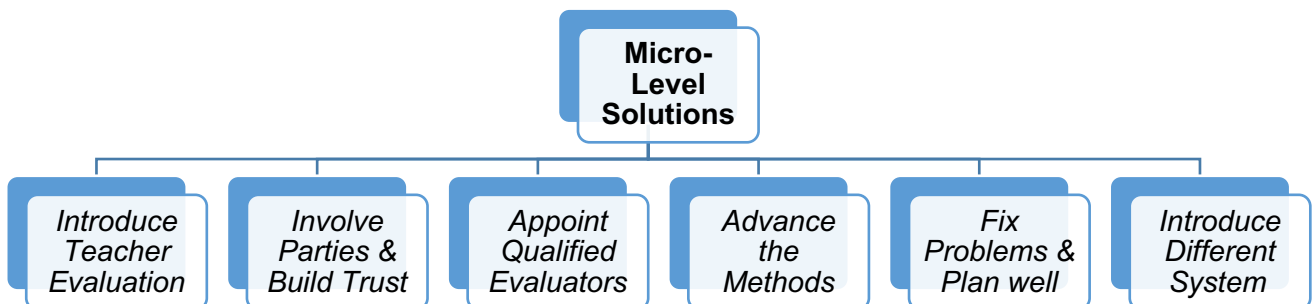


Figure 5. 12. First category and sub-categories, theme 3

#### 5. 4. 1. 1. Introduce Teacher Evaluation

In order for any evaluation system to succeed and achieve its goals, it should be made clear to those who are involved in it, especially the teachers. Introducing teacher

evaluation to EFL teachers will help them recognize its importance, comprehend practices around it and cooperate with the people in charge of it. The majority of the participants reported that they have little information about teacher evaluation and their views were presented earlier (Section 5.3.4.3. p., 212 and 5.3.4.4. p. 215), with some of them suggesting the introduction of teacher evaluation to the teachers. For example, Majeda in her interview stated:

*“I suggest if you want to do evaluation, you should at least give an **orientation** about the **value of evaluation** itself. So as a teacher, I can accept you in my classroom. I don’t look at you as a **threat** to me.”*

As the extract indicates, providing orientation to teacher evaluation may lessen the threat experienced by some teachers. The survey analysis also supports this finding, with many participants highlighting the importance of introducing the evaluation system to the teachers. In the open-ended question where participants were asked about the changes that should be made to their current evaluation system, one participant claimed:

*“There should be **sessions** given to teachers by their evaluators **before** any observations.”*

This idea was reinforced during the focus group discussions, where Brian suggested documenting the purpose and use of teacher evaluation as follows:

*“I think that there could be some additions to the **faculty handbook** about the **purpose and use** of student evaluations and self-evaluations. Or there could be a university **document** accessible to all the faculty members (...) so that we would have clear **understanding** of how these are used and what they’re used for.”*

The data suggests that introducing teacher evaluation would help teachers understand and appreciate the practice. This is, however, not an easy task, especially at Saudi universities, where teacher evaluation is not a stable fixed practice and was implemented to mainly satisfy administrative demands rather than to ensure the quality of the education. A relevant yet distinct issue will be presented in the following sub-category which relates to the involvement of other parties.

#### **5. 4. 1. 2. Involve Teachers, Students and Build Trust**

Involving people who take part in the practice of teacher evaluation received some attention from the participants. This involvement may take different shapes and forms. One of the participants, Ferdaws, in her interview highlighted using questionnaires and discussion groups to know more about teachers' ideas on teacher evaluation:

*“You can hand the **questionnaires** to everybody and then you collect the data (...). Take these which are said by the majority of them. And then when you finish, try another tool. Okay? Back up this questionnaire to a **discussion group**, for example.”*

In one of the focus group discussions, Sameyah argued that giving teachers an opportunity to present and/or to join training courses might enhance their involvement:

*“For example, the coordinator knows who are the best teachers, and she makes a meeting or interview with them and till them each one is going to make a **presentation** and come to a **training course** for teachers in one point and **develop it**, and follow these teachers, not one teacher, and makes a meeting for them once a week or once every two weeks”*

Findings from the questionnaire also support the idea of involving and empowering teachers. When asked about the changes they wish to be made to the current teacher evaluation practices, one of the participants reinforced the importance of focusing on empowering teachers:

*“Evaluation should be focused on teacher development and **empowerment** more than on promotion or tenure.”*

The involvement of students was also suggested by the participants, as illustrated in the following extract from Omaima:

*“I would also be very, very happy if they took the **students’ preferences** and their opinions in mind as well because we don’t do that. The professional development, the evaluation is **all about the teachers** (...). Give **students training**, help them to learn. Help them to study. Develop some strategies for them how to tackle this new phenomenon that is always arising.”*

The quote suggests that students’ involvement may also include training them besides taking their preferences into account. This supports earlier findings (Section 5.3.3.1. p. 200) focusing on the perception of students’ lack of training and ability to rate teachers as a challenge. Involving teachers and students in teacher evaluation may help in building trust in the system of evaluation itself. This idea was conveyed by Majeda in the following extract:

*“**Trust is important** for a teacher. If I don’t trust, I would just react negatively (...). If I had a hand in it, I will just focus on **building that trust** between the teacher and the evaluation. I would **involve** teachers”*

As the quote suggests, teachers need to trust their teacher evaluation system so that they will not react in a negative way. However, Samar saw it the other way round, where the teacher evaluation system might be seen as a result of lack of trust in teachers; hence, her following suggestion:

*“So this is my additional point that the teacher **should be trusted** anyway. She should be given some confidence that whatever she is doing is done in **good faith**. She is not dishonest.”*

Analysis of the survey provided support for this finding. When asked in one of the open ended questions about the changes to be made to the teacher evaluation system, some participants highlighted this idea, as illustrated in the following quote:

*“A need to **learn to trust** the hired foreign directors of the prep-year programs who are trained evaluators themselves to do a large part of the job.”*

It appears that building trust between the parties involved in teacher evaluation is important. This can be achieved by involving teachers, students and other parties involved in teacher evaluation such as the evaluators themselves. In the following sub-category, issues related to the evaluators will be presented.

#### **5. 4. 1. 3. Appoint Qualified Evaluators**

Evaluators are an essential element in any evaluation system as it is they who carry out and activate the procedures of teacher evaluation systems. Where evaluator training is offered, some evaluation systems require only qualified evaluators to carry out evaluation. According to Alexander (2016), a qualified evaluator is “an individual who has completed the prequalification process (...) and successfully passed the

state-developed assessments specific to evaluation of teachers. Each qualified evaluator maintains his or her qualification by completing the retraining as applicable” (p. 21). However, this appears not to be the case in some Saudi universities, where unqualified or poorly trained evaluators are frequently reported. In his one to one interview, Hamdan contended that evaluators need training:

*“I think teacher **evaluators** themselves have to undergo **training** so that they can be on the same page. Having said that, I do appreciate the fact the in the last two terms, highly **qualified professionals in the PDU** have been assigned to carry out teacher evaluations”*

The idea of trained evaluators and observers was also supported by data from the survey. When asked about the changes they would like to see made to their current teacher evaluation system, one participant said the following:

*“Those who observe teachers **should be trained** a lot.”*

Another participant replied to the same question suggesting specific qualifications for the evaluators:

*“**CELTA and DELTA trainers** should be hired to evaluate teachers. The current team of evaluators are **not properly qualified** for the job.”*

It seems that this participant considers holders of such certificates as sufficiently accredited to carry out the job of evaluating the performance of the teacher. As the data suggests, teacher evaluation can be better carried out in a professional manner



by qualified individuals who have been properly trained to evaluate EFL teachers' performance.

#### **5. 4. 1. 4. Advance the Methods, Improve the Criteria and Scoring Techniques**

The collected data shows that different universities use different methods to evaluate EFL teachers. Generally speaking, suggested solutions related to methods include introducing the least frequently used methods in that particular context or putting more weight on those that are not given much value. In other words, the participants suggested implementing peer evaluation/observation, committee evaluation, student rating, student achievement, teaching portfolios, academic research and self-evaluation. However, it seems that peer evaluation and student rating received more attention from the participants than the other teacher evaluation methods. In the following extract Hatim, in his one to one interview, explains the reasons behind calling for peer observation:

*“Peer observation could be very helpful as it creates a sort of team work spirit which is very lacking within the staff circle because of **the threat** of termination they feel on a daily basis.”*

In addition, in one of the group discussions, strong support for implementing students' ratings was expressed by the participants, as illustrated in the following extracts:

*“Ghada: Nevertheless, the **student survey** is a good way of evaluating teachers.*

*Fadwa: Yes, it's wonderful.*

*Ghada: It's the best thing because they see you every day. It's not just one visit a module or a year. But they **don't take it into consideration**. It's there, but I don't think it has a big value.*

*Huda: I think it's the most **unbiased way** of evaluating because it is repeated."*

Analysis of the survey showed that classroom observation carried out by administration members is the most frequently applied method in the universities investigated, as indicated by 61.8% of the participants. Student rating was the next most frequently used method, selected by 60.2% of the participants. Student achievement received mention from only 39%, teaching portfolio 35.7%, self-evaluation 25.7%, classroom observation by peers 24.9%, peer evaluation 16.5%, and other methods 7.6%. In the open-ended section, one of the participants suggested combining the evaluations:

*"Evaluation should be a combination of **peer, students and qualified, objective evaluators.**"*

In the same vein, in one of the open ended questions where changes were to be suggested, another participant proposed:

*"They should make it **more holistic**"*

Interestingly, some of the participants suggested special techniques for evaluation which could be tailored to suit teachers based on their experience or the different responsibilities they might have besides teaching, as illustrated in the following quote from Mariam:

*"So I would suggest two different criteria: for **novice teachers** and different criteria for **experienced teachers**. It will be more developing for me."*

According to King (2015), experienced EFL teachers in the Gulf area might find teacher evaluation a demotivating experience and show dissatisfaction with it as they may have experienced a “downturn in their feelings” (p. 176). Hamdan illustrates in the following extract his feelings when he received a special teacher evaluation based on what he was assigned to do as part of his position:

*“In my case for example, even though I was not evaluated on teaching, per say, I was evaluated **professionally** assessed by my superior on what exactly **have I achieved** since I took up the position of the head of research. That really felt brilliantly comforting.”*

It seems unfair to ignore other duties teachers might have and to evaluate them on teaching only.

The final suggestion put forward by the participants relates to removing numerical techniques of scoring. For instance, Ferdaws stated:

*“The **numerical scoring** or rating or these criteria should be **eliminated**. Instead of giving a score of like 100 or 55 etc., we should give **expressions**. Expressions go to the heart to be honest.”*

Analysis of the questionnaire provided support for this finding; when suggesting changes participants would like to see made, one participant wrote:

*“**Stop rating** teachers on a scale of 1-5. Recognise that all teachers have **strengths**.”*

As the data suggests, the participants seem to prefer receiving words over numbers as their performance evaluation results; they may perceive words as more expressive than numbers, where evaluators can acknowledge teachers' strengths and those in charge of teacher evaluation can consider their views as well. Nevertheless, a contextual factor that may support the application of numerical rating is the large number of EFL teachers in the institution and therefore the ease of application of this technique.

#### **5. 4. 1. 5. Fix the Existing Problems and Plan well**

As in any other academic *modus operandi*, teacher evaluation practices have different problems which vary in degree and effect on the success of the systems employed. In the following quote Mariam indicates how significant and divergent the issues can be:

*“We have a **big flaw in the system** (...). It’s a whole continuum of related elements in the whole process of evaluation (...). And the evaluation should be like if you have **problems**, you go **fix them**. There should be like an **analysis** of what’s happening, and we should be **voicing** it and **fixing** it. But that’s not happening”*

Acknowledging that there are flaws in the system, people in charge of teacher evaluation may need to analyse what has been done in practice in order to fix it. Given that teachers, and possibly other included parties, do not have much voice and/or role in developing the practices of teacher evaluation as presented earlier (Section 5.3.4.4. p. 215), it may be problematic for policymakers to become aware of the system flaws let alone fix them. This seems to reinforce the significance of EFL teachers' voice in the development of the system; this issue will be further discussed in the following chapter. In order to find solutions for current problems, different suggestions were

proposed by the participants: for instance, in the following extract from one of the focus group discussions, Karen and Fadwa suggest planning and consultation:

*“Karen: The first thing I would say is that before any kind of observation starts, there should be **careful planning**. The whole thing should be thought through carefully. What the **objectives** are, how those objectives might be **achieved**, and what kind of **training** people would need to be able to achieve those objectives (...)”*

As the quote suggests, in order to plan carefully, objectives need to be identified, ways to achieve them need to be implemented, and training people need to be encouraged. In addition, having a counselling group or body may help teachers to be more aware about the practices; similarly, a unit which offers advice and help for teachers could be beneficial or, as Arreola (2006) states, a “faculty advisory board” not only to help teachers overcome obstacles related to their evaluation but also to offer a mechanism for teachers to have a sort of involvement in developing the policies around teacher evaluation.

#### **5. 4. 1. 6. Introduce a Different Evaluation System**

Developing evaluation of teachers’ performance is a dynamic process where different methods can be introduced and new mechanisms for using them are a choice for policymakers and educational institutions. Some of the participants suggested making changes to current evaluation systems or even introducing different new ones. For instance, implementing the five-minute walk-through technique was recommended, by Mariam, in her one to one interview:

*“The **five-minute** walk-in and the three-minute walk-in, so as an instructional leader that you would visit teachers’ classes and you don’t just focus on the **teacher**. You focus on the*

*students, the whole instructional process, and the atmosphere of the class (...) and you have discussions, nothing written, no form, nothing.”*

As the quote suggests, walk-through visits might be seen as less formal than typical classroom observation; yet, and as Carraway and Young (2015) suggest, the five-minute walk-through strategy may be used to mainly enhance the instructional leadership skills of the people in charge of teacher evaluation but may not provide a reliable evaluative tool to ensure the quality of teaching. Analysis of the survey data produced similar findings. One of the participants said:

*“I'd like to see it **more regular but less formal**”*

However, the data shows that some participants may seem to hold contradicting views. For instance, the following extract from Albert represents his standpoint on the frequency of evaluation:

*“So I would have a mixture of teacher observation after the probation, **at the most once a year**, just to make sure things are okay”*

It appears that in Saudi universities, teacher evaluation is influenced by the way it is put into practice in that particular context; hence, the micro-level solutions suggested seem to vary accordingly.

#### **5. 4. 2. Macro-Level Suggested Solutions**

In this category, solutions were suggested for areas beyond the institution itself and requiring the involvement of other entities/bodies. This involvement may vary in the

type of input. For example, in the first sub-category, namely investing in teachers, funding from the university or the Ministry of Education will be needed and hence this is classified as a solution on the macro-level. The other sub-category is collaboration with other departments and/or universities. Both of these sub-categories will be presented in the following section and are shown in figure 5.13. below.

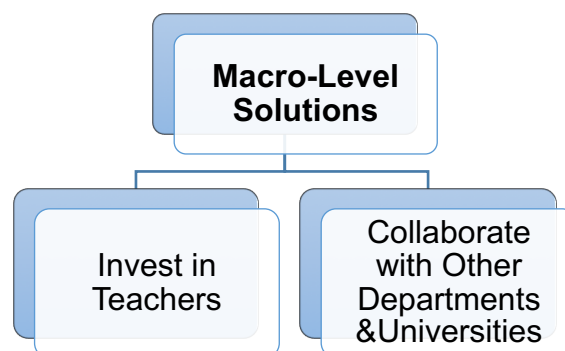


Figure 5. 13. Second category and sub-categories, theme 3

#### 5. 4. 2. 1. Invest in Teachers

Evaluation of teacher performance is the first step towards the professional development of teachers and investing in them. As reported by the participants, English language institutions can invest in EFL teachers in many different ways after evaluating them. In the open ended section where participants were asked to make any comments they would like to add, one teacher suggested:

*“**Workshops** should be held to discuss the common shortcomings of teachers’ performance. These workshops help teachers to **exchange** their **experiences**.”*

Exchanging experiences in workshops offered in their universities can encourage teachers to overcome common shortcomings. By doing so, they may become more

committed to their job and the academic institutions they work for. This idea was suggested by Fadwa, a participant in one of the focus group discussions:

*“So if I have these flaws in me, **help me to improve** myself. Then I may be more **committed**. I will never send my CVs to any other places. Even I will invite people to come because this is the best place to work. The first thing is my **sense of security**.”*

According to the participant, when institutions help teachers to flourish and improve, they feel more secure and accordingly committed. Moreover, some participants went further and suggested practical ways to invest in EFL teachers and activate their role in improving each other’s performance. For instance, the following quote from Brian shows how institutions can benefit from teachers who are good at some teaching skills:

*“so these kinds of teachers who are focused on these **specific skills** and get good results from their students, they could be observed or even **evaluated** or could themselves be the **evaluators** of other teachers in their skills in that area.”*

As the data shows, giving such teachers the chance to share their strengths with their colleagues through presenting training workshops can help language institutions to invest in their staff members. These workshops can be an enormous support for teachers and part of their follow-up plans, an area that is almost completely neglected in some Saudi universities (a challenge explored in Section 5.3.2.4. p. 196). Asher articulated this idea in the following quote from his one to one interview:

*“If teacher evaluation is to benefit the education system in Saudi Arabia the following should be done: **Professional development workshops** on evaluation approaches should be held for all educators (supervisors and teachers) with the help of **inside or outside experts** (...) until everyone*



*involved becomes familiar with the process. The benefits of such workshops are **immense** for supervisors, teachers and students.”*

As the quote suggests, workshops can be run by inside experts, which provides support for the earlier quote, or outside experts from training centres or other universities, which relates to the following sub-category. Whichever approach is taken, it can be considered as part of the teacher continuous professional development programmes, an area that may have received insufficient attention (Section 5.2.4.2. p. 170). Mariam in her interview highlighted the idea that educational institutions should invest in Saudi nationality teachers as they also need help:

*“And **observe the Saudis**, and give them a **development plan**, a real one. Give them a chance to learn and develop (...). Focus on them. Help them. They really need help. They should **invest** in them.”*

It appears that Saudi EFL teachers do not receive sufficient attention from evaluation systems in Saudi universities, as some seem to exclude them completely from teacher evaluation, as mentioned earlier (Section 5.3.4.2. p. 209). A possible reason for this exclusion is that Saudi teachers' contracts do not need to be renewed as they are on a long-term type of contracts and teacher evaluation is mainly carried out to make contractual decisions. Hence, those in charge of teacher evaluation make little effort to evaluate Saudis or to monitor their performance. Beside the fact that Saudi EFL teachers need evaluation to improve their instructional practices, having them on long-term contracts makes the need to invest in them as permanent teachers more urgent, and this should be an essential part of academic institutions' sustainability plans.

#### **5. 4. 2. 2. Collaborate with Other Departments and Universities**

Being aware of teaching practices and teacher evaluation in other departments of the same university or the same department in other universities can help to develop the quality of education in any educational institution. Such collaboration was frequently mentioned by participants in this research. For instance, the following quote from Fadwa in one of the focus groups highlighted the importance of having and collaborating with a counselling department within the same university:

*“First of all, there is need for counselling. There should be a **counselling department.**”*

The survey results also support the importance of consultation with people from outside the department. In one of the open ended questions where participants were asked about the changes that should be made to current evaluation systems, one participant claimed:

*“But most importantly there should be **highly-qualified outside consultants** vetting the managers at the PYs.”*

Apart from counselling, participants seem to prefer evaluators who are not their own colleagues, as illustrated in the following suggestion:

*“Professionals who are qualified who actually have some sort of related degree or certificate to evaluate should be brought in. **These people should not be our colleagues**”*

In the same vein, participants of one of the focus group discussions elaborated on the advantages of having a separate team of evaluators and added that they should be aware of the context and teaching at the same university:

***Karen:** You said there should be a **separate team** not from the actual teachers who are colleagues. It's very important. You'll find that in countries like the U.K., there are actually groups that come from the Ministry of Education and Ofsted, and they are specialized in observing and they know what they're looking for (...).*

***Fadwa:** still I feel there should be **one department**, only highly qualified, fully trained teachers who have special certifications, and teachers who have no interaction with other colleagues should be part of this.*

***Karen:** That's another thing. Yeah. The observers need to **understand the context**.*

***Ghada:** This is one positive point why the observers should be **teachers from the same university**.*

***Huda:** Yes. And they should teach for some time."*

The idea of having a separate body for teacher evaluation was also supported by findings from the open ended section of the survey, as in the following quote:

*"A **separate body** [need to] be made for Teacher Evaluation with specific and clear objectives charged with positive thinking to create a thorough educational atmosphere."*

Some participants widened the circle of involvement and suggested involving evaluators not only from other departments on campus but also from other campuses/branches of the same university or other universities. The following extract from one of the open ended questions is an example:

*“Evaluators **shouldn't be colleagues** teaching in the same workplace, they'd rather be from **other branches** (...) to avoid prejudice and underestimation.”*

Similarly, the following extract from one of the focus group discussion elaborates further:

*“**Basmah**: And one more point is that nobody gets the chance to go to **another university** to learn something from them. You are teaching in the same way, every day the same thing, like nothing changes. But in our countries, like in India and all, we people go somewhere else, some other people come to our places, and there is interaction between the panels of the universities to improve the chance of cooperation. So nobody feels **suffocated**. It's like an evaluation, a kind of evaluation to improve yourself. You **compare yourself to others**.*

***Sameyah**: And feel **self-confident**.”*

The idea of collaborating with other universities was also supported by results from the questionnaire. In one of the open ended questions, a participant contended:

*“Well, it isn't the sole thing to judge a teacher. More things come from (...), **active environment of development** through symposiums, seminars, workshops between different departments within a university and between same departments of **different universities**.”*

The quote suggests that departments and universities are working in isolation where they do not have any contact with the other local universities or the outside world. Thus, it can be argued that collaboration with other specialized entities that have individuals qualified to do teacher evaluation will help overcome some of the related challenges; issues such as bias, subjectivity and rancour between colleagues that were reported by the participants as challenges to EFL teacher evaluation will be

eliminated. Having said that, despite the fact that collaboration with colleagues from different departments and/or universities can help improve the practice of teacher evaluation, it may need more time and effort in order to make it successful and effective.

## **5. 5. Summary of the Chapter**

This chapter offered a detailed account of the analysis and interpretation of the data. Emerging themes, categories and sub-categories were presented. Three major themes that relate to the research questions were thoroughly explored and linked to related categories and sub-categories in order to draw a comprehensive picture of the participants' views and ideas about a number of relevant matters as described by EFL teachers in the five Saudi universities. The importance of teacher evaluation, challenges to EFL teacher evaluation, and suggested solutions were the three main themes. Based on the analysis of these themes, a number of social, cultural, and political related aspects were found and expected to lead into the findings. These aspects along with some others will be discussed in the following chapter.

## **Chapter Six**

### **Discussion of the Findings**

#### **6. 1. Introduction**

In this chapter, I provide some critical insights into the most significant findings of this research from both qualitative and quantitative strands in relation to the phenomenon of EFL teacher evaluation and its pertinent mechanisms within the higher education system in Saudi Arabia. Amongst all the themes, categories and sub-categories presented earlier in the former chapter, key findings on EFL teachers' perceptions of teacher evaluation are discussed in the following sections:

- EFL Teacher Evaluation: Values VS Demands
- What Makes a Qualified Evaluator?
- EFL Teachers' Aversion to Teacher Evaluation
- Students and the Intellectual Leap between General and Higher Saudi Education

The abovementioned four themes are highly related to the research findings and were arranged according to the same sequence of the findings' themes. In the first theme (related to research finding theme 1, section 5.2. p. 151), values and demands take the practice of teacher evaluation beyond findings on importance of teacher evaluation and were discussed in relation to different contextual factors. Second, third, and fourth themes (related to research finding theme 2, section 5.3. p. 180) discuss the main three challenges (section 5.3.2., 5.3.3., 5.3.4.) in light of the Saudi context.

#### **6. 2. EFL Teacher Evaluation: Values VS Demands**

As previously indicated in the analysis chapter, the majority of the participants acknowledged the importance of evaluating the performance of teachers regardless

of the mechanism that controls teacher evaluation practices at their university. When teacher evaluation is perceived to lose its significance, this is likely to be linked to flaws in the way it has been practised rather than the value of its general aim(s). This is in line with Palls's (2013, cited in Bridglall, Caines, and Chatterji (2014) stance that evaluation data can be used not simply to identify and eliminate those who are underperforming but rather to monitor and improve performance. It appears that, from a teacher's perspective, monitoring and improving performance are of more value than identifying and eliminating low performing teachers. Having said that, the old argument around the importance of teacher evaluation and why it is carried out in the first place emerges. The findings of the current study imply that teachers believe evaluation serves mainly two purposes: firstly, forming an authoritative body that holds teachers accountable to predetermined standards, and, secondly, improving teachers' performance. These two purposes will be discussed in the following sub-section.

### **6. 2. 1. The Demands of Accountability and Quality Assurance**

The findings reveal that teacher evaluation committees as authoritative bodies are in charge of setting standards in order to ensure the quality of teaching. With the rapid growth of higher education institutions in Saudi Arabia, a need for quality assurance which holds teachers accountable was created and led to the formation of an independent government agency for accreditation called The National Commission for Academic Accreditation and Assessment, abbreviated to NCAAA (Darandari et al., 2009). Under the auspices of the Ministry of Education, the NCAAA is held responsible for setting general standards, framework and procedures for quality assurance. These foundational standards are expected to be evident in accredited Saudi universities; however, the commission lays emphasis on "continuing quality improvement rather

than on satisfying required standards” (NCAAA, 2015a, p. 5). At Saudi universities, where evaluation standards are communicated to teachers, participants found these standards important, as they set benchmarks for teachers on the quality of teaching. Yet, the findings show that various standards and practices can be found in different universities in Saudi Arabia. This finding confirms the claim that Saudi higher educational institutions have to establish their own individualized quality assurance model (Albaqami, 2015), which ought to be specific and relevant to their particular context. As a result, models for and practices of EFL teacher evaluation should vary in different Saudi universities. This variation is in fact encouraged by the NCAAA, as it necessitates diversity and flexibility inside educational organizations. This organizational variation is important, given that some elements within teacher evaluation to meet quality standards may be more significant in some contexts than in others. From a personal perspective, this is an appropriate standpoint for two main reasons. First, it is not easy to set a model or framework that will fit all the various contexts within higher education institutions in Saudi Arabia. Second, it can help universities maintain and safeguard their academic identity; this identity is supposed to be unique in nature, appropriate to each university or department teaching practice and should serve its needs and demands.

Furthermore, the findings of this research indicate that, within each university, teacher evaluation standards were established by top management and policymakers in a top-down hierarchical structure and handed down to evaluators and teachers. These findings echo those of Albaqami (2015) on a local level and Shah, Wilson, and Nair (2010) on a wider international level and indicate that the systematic review of quality assurance is generated from top to bottom. The hierarchical interaction between those



on top who have power and those at the bottom who do not seems to be the essence of a widespread bureaucratic system found in most of Saudi universities. It reflects a hegemonic relationship between policymakers and the people affected by such policies within the system. Given that policymakers cannot ensure that teacher evaluation standards are checked by themselves, they tend to delegate others (evaluators and/or evaluation committees) to accomplish this mission. Thus, delegated entities are required to report the degree of success in achieving the given standards they have observed to the people in power on top of the hierarchy. This eventually results in a complex hierarchical system where the implementation of policymakers', who have the ultimate power, dogma is carried out by the evaluators and imposed on EFL teachers. According to Mathison and Ross (2002), such centralization of authority has become standard "by business and political alliances for standardization of processes ... under the guise of public good" (p. 97). This is true where systems of higher education need to prove their commitments to public good and serve the public interest (Eaton, 2015) in order to gain public trust in spending public funds, which is the case of Saudi public/state universities. This situation results in an increasing focus on performance, efficiency and accountability, where educational systems aim to generate individuals that are typically competitive and enterprising. As Ambrosio (2013) claims, through the state's power, neoliberalism in education aims at creating special forms of subjectivity that are arranged artificially, which enables both the society and the market culture to function in an efficient way. It seems that Saudi higher education is following that major trend, where quality is measured through state control mechanisms. In this respect, Jenlink's (2016) intriguing ideas on how accountability politics in an era of neoliberalism has distracted

democracy are of interest. Jenlink discusses how the neoliberal hegemony represents a threat to education both economically and/or politically. This is true given that the culture of accountability and performativity encroaches on the freedom and creativity of the teachers, the freedom that Freire (2000) alludes to as a function of education while arguing that educational process cannot be neutral. He states that education functions as the “practice of freedom, the means by which men and women deal critically and creatively with reality and discover how to participate in the transformation of their world” (p. 34).

As the findings suggest, the results of teacher evaluation are perceived by the participants to influence personnel decisions. In such cases, and as discussed above, policymakers expect evaluators to provide them with accounts that are value-free and reliable, which makes the evaluators’ key task one of assessing the extent to which a predetermined set of standards has been achieved. This type of evaluation has been theorized in the literature as the “*technocratic*” approach to evaluation (Hanberger, 2001). As Codd (1994) states within the context of accountability measures, this model endorses a set of educational technocratic evaluation practices that produce various managerial controlled structures rather than professional improvement in teaching, which was a finding reported by the participants of this study too. In his view, technocratic reductionism can be seen as the ultimate result of the new accountability measures that emerged as a display of the ‘economic rationalism’ that has been known to inform governmental policy-making.

In this respect, Green (2010) cautions that evaluation conducted in a “climate of control” (e.g., evidence-based decision-making, result-based management and

performance indicators) with technocratic ideas is the main challenge to the principles of an evaluation that is democratically-oriented. I believe that this is true, as a technocratic approach to evaluation typically promotes elitism and a culture where decisions are made for teachers rather than by teachers. In line with this, Codd (1994) contends that, with technocratic evaluation, judgement is taken to the people indirectly where evaluators mainly and directly serve bureaucrats and politicians and serve the people in an indirect way. This suggests that, 24 years after Codd's statement, the situation of teacher evaluation in Saudi Arabia has not changed to a more participatory or humanistic model; indeed, it may be even more restrictive and top-down. However, I believe that more involvement of teachers is needed in order to improve making decisions around policies and practices of teacher evaluation in the Saudi context.

Shifting our attention back to the NCAAA, its central mission is to ensure the consistency of evaluation standards with international benchmarks (NCAAA, 2015a). For courses developed outside Saudi Arabia, the NCAAA requires that they be accredited by a "proper authority in the country of origin" (NCAAA, 2015b, p. 11). Despite the fact that EFL programmes in the investigated universities are in general locally developed, where only the textbooks come from abroad, these universities usually seek academic accreditation from agencies that are overseas. This shows how a form of dependence on external academic agencies is built to ensure the quality of language teaching in Saudi universities.

In this regard, Pennycook (2017) maintains that ELT seems to be controlled mainly by two countries: Britain and the USA at the centre of an inner circle. These western countries exercise an additional level of authority (Llurda, 2016) over ELT in the Middle

East, which affects and influences EFL teachers. This is usually achieved by sending experts from the inner circle western countries to provide assessment and/or support to other countries, which has resulted in a sort of academic imperialism.

This dependence has greatly contributed to a consumerist culture of ready-made products and services to encourage the world-wide marketing of ELT. Pennycook (2017) points out that ELT can be seen as a business-driven practice that has spread all over the globe. Against this background, it seems that with the recent educational reforms taking place, not only in Saudi Arabia but also in other Arab countries, the quality assurance practices are heading towards a new direction. According to Eaton (2015), ministries of education and leaders of higher education in Arab countries are looking into establishing robust accreditation and quality assurance environments and considering the formation of new national quality assurance bodies in the region. Nevertheless, and as Furedi (2010) states, “marketization is a reality that academics have to live with” (p. 1) given that universities increasingly mimic private and public corporation managerial models. Various techniques of management and cost-oriented plans and strategies have gradually replaced academic practices and rituals around the world. There is a vast literature that can be found on the issue of the marketization of higher education and how students are globally perceived as consumers who can make demands of the university they join and highly contribute to the phenomenon of marketization in higher education (Naidoo & Williams, 2015; Nordensvard, 2010; Williams & Molesworth, 2010; Woodall, Hiller, & Resnick, 2014). However, it is worth mentioning that marketization as a phenomenon in higher education has not only economical aspects but also a political and/or ideological agenda attached to it.

Having said that and considering the various issues of power discussed earlier, I believe that EFL teacher evaluation should not be hegemonic. A form of a counter-hegemonic model that relies heavily on democratic assumptions derived from genuinely shared authority is needed in order to promote EFL teacher evaluation policies and practices in the country. As Mathison and Ross (2002) suggest, forms of evaluation that are participatory should encourage democratic principles and can be counter-hegemonic. They believe that counter-hegemonic practices are derived from an authentic partaking of authority amongst not only the few, powerful individuals so to speak, but also the many. This can be achieved when we create and empower a community that shares similar interests to “work toward greater clarity about and commitment to particular value positions” (p. 98). In my opinion the best way to approach a counter-hegemonic form of teacher evaluation is to engage a wide range of stakeholders in shaping the intended teacher evaluation scheme including its objectives, methods, standards and other aspects of the practice. This will help to achieve a participatory approach to teacher evaluation; however, the degree to which the involved parties are expected to take part and how to balance their contribution to the scheme is critical.

### **6. 2. 2. The Values of Improving Instructional Practices**

The findings indicate that EFL teachers believe that teacher evaluation is carried out essentially for summative purposes i.e., the results of the evaluation processes are directly linked to end of the academic year decisions on their contract renewal. Accordingly, teacher evaluation has hardly been proven to have a positive influence on continuous professional development, CPD, which aims to improve teaching practices. As Smits and Champagne (2008) write, “Evaluation approaches have long

been criticized because their results are often not used” (p. 427). This typically is the case where teacher evaluation does not have a formative lens or a developmental scope and where evaluation results and output are not considered to identify areas that need improvement. However, it is believed that the theory of Participatory Evaluation (PE) addresses this drawback through encouraging the members of educational organizations to observe and rethink their professional practices (Cousins & Earl, 1995).

According to Cousins and Earl (1992), PE refers to “applied social research that involves a partnership between trained evaluation personnel and practice-based decision makers, organizational members with program responsibility or people with a vital interest in the program” (p. 399-400). In his later work, Cousins develops the theory and differentiates between two different types of PE. In *Transformative Participatory Evaluation (T-PE)* and by involving the less powerful stakeholders in reflecting, creating knowledge and making decisions, evaluation is perceived as a developmental procedure that can be related to democratic and empowerment evaluation, participatory action research and other types of enquiry that are collaborative and driven ideologically. The other stream is *Practical Participatory Evaluation (P-PE,)* where the central interest is to generate knowledge that is useful for making decisions, solving problems and improving practice (Cousins, 2003). The main goal here is to increase the usefulness of the results of the evaluative participative processes and the empowerment of the so-to-speak less powerful stakeholders.

As indicated in the previous chapter, the results of evaluation procedures and feedback, where provided, were perceived by the majority of the participants to be of value. Constructive feedback is valuable as it helps EFL teachers improve their teaching practices and promote their professional growth. This is in line with Ahmed and Asghar's (2015) study at an English language centre in the Middle Eastern context, where they found the quality of post-observation feedback to be high and informative, which consequently leads to a positive impact on teaching practices. When teachers know about their weaknesses, it is possible for them to work independently on them and improve them by reading, discussing them with their colleagues or through other available resources.

Despite the reported quality of the feedback, the findings indicate that teacher evaluation results remain institutionally useless, as feedback, classroom observation notes and other evaluative outcomes are not very productive. If EFL teachers are not provided with an adequate follow-up developmental plan and well-designed professional workshops, they will not consider teacher evaluation beneficial to them. It has been found that university EFL teachers in the Saudi context support the common belief that the level of teachers' professional development can be enhanced by exploiting the results from teacher evaluation (Al Asmari, 2016; Hakim, 2015; Shah & AlHarthi, 2014). Similarly, Al-Harthy (2017) found that Omani senior teachers at college level hold a positive attitude towards the importance of teacher evaluation feedback to their professional development. However, the impact of teacher evaluation on their professional development is evidently weak and ultimately attributed to different factors. These factors include: conflict with motivational structures of teaching as an occupation, evaluation systems' technical shortcomings, implementation

problems, political complexities, and most importantly the misalignment between understanding teaching tasks and the design of evaluation systems (Darling-Hammond, 2013; Howard & Gullickson, 2010; Rowan, 1990). Smylie (2014) summarizes it all as one factor most consistently associated with the lack of teacher evaluation impact on professional development, namely the “troublesome relationship between evaluation and professional development” (p. 98).

Howard and Gullickson (2010) argue that this flimsy link between teacher evaluation and professional development is one of the major “threats” to teacher evaluation potential to improve practices of teaching. This notion, which is in line with findings from the current study, however, contradicts other findings in the literature. For instance, Ahmed and Asghar’s (2015) study reveals that EFL teachers believe that their professional development has been positively influenced by the annual faculty evaluation process despite the fact that they perceive their evaluative system as rigid and inconsiderate of cultural diversity. The quantitative data they collected from 45 participants suggests that teachers perceived the process of teacher evaluation as having a positive impact on professional growth. In the qualitative part, eight out of nine of the participants of that study were enrolled onto CELTA after they underwent a summative evaluation and hence saw it as a positive impact of EFL evaluation on their professional development.

As mentioned in the previous chapter, the participants indicated that they perceive the workshops which professional development bodies usually offer as more of a useless training routine than an instructional practice improvement opportunity. Therefore, ambitious EFL teachers tend to informally pursue their professional development



courses or workshops while coping with their everyday teaching and academic duties. Many researchers have emphasised the insignificance of the existing professional development workshops at Saudi universities (Al-Seghayer, 2013; Al Asmari, 2016; Hakim, 2015; Khan, 2011; Shah, 2016; Zohairy, 2012). Where there are opportunities for joining continuous professional development workshops and training programmes that are institutionally arranged, EFL teachers express their dissatisfaction with those offered courses. Regardless of Al Asmari's (2016) argument that professional development training provides EFL teachers in Saudi universities with a good chance to learn from colleagues and/or experts, a considerable number of EFL teachers see them as pointless and time-consuming practices. Such a standpoint can be attributed to the varied educational backgrounds and experiences of the teachers, where some of them have received high quality training before coming to Saudi Arabia.

Additionally, as the majority of EFL teachers at Saudi universities are expatriate teachers, the findings reveal that people in top management have some concerns that those teachers will leave the country before the end of the semester if they do not have any more classes to teach. It seems that there is a difference in views between EFL teachers and management about the purpose of these professional workshops. While EFL teachers expect the workshops to develop their teaching practices, they see them as a ploy by management to make them stay in the country and to keep them busy. As part of the faculty and staff employment processes standards set by the CNAAA, proof of "proportion of teaching staff participating in professional development activities" (NCAAA, 2015c, p. 60), is one of the required Key Performance Indicators (KPIs ). This proof should be evident for quality assurance purposes to hold higher education institutions accountable, as discussed earlier. A

conflict arises when those in management do not properly communicate the objectives of such workshops to EFL teachers beforehand. The management is known to have the upper hand over arrangements for and budgets allocated to such events and activities. Those in management assume teachers to be busy discharging their day-to-day duties. Accordingly, professional development activities are usually carried out during non-teaching dates of the year in order to avoid staff burnout.

Turning EFL teachers into partners and involving them in taking decisions about the type and schedules of workshops will help them perceive training and workshops as more effective. Being effective members within educational systems can help EFL teachers develop positive attitudes towards educational policies and practices that teachers feel are imposed on them in a top-down approach. In this regard, this study is in line with previous literature that globally supports teachers' involvement in planning professional development activities, such as the studies of Al Asmari (2016), Bayar (2014), Smith (2015), Starkey et al. (2009), and Whitworth and Chiu (2015). In the Saudi higher education system, Al Asmari (2016) emphasises that teacher educators and professional development organizers should conduct a careful needs analysis for their training activities with the help of EFL teachers to find out what they require to develop. I believe that language teachers, as practitioners and insiders, can be a useful source of information about the type of workshops they need in order to improve their teaching practices.

### **6. 3. What Makes a Qualified Evaluator?**

As the title suggests, this section discusses the required characteristics of a well-qualified evaluator as seen by the EFL teachers who participated in this research. This

category seems to be of utmost importance in the Saudi context, given the current lack of national EFL training centres/institutions that are adequately staffed to provide sufficient training for teacher evaluators and educators. Instead, an intensive one-day or one-week preparatory programme, which is insufficient to equip evaluators and to turn them into qualified practitioners in the field of teacher evaluation, is arranged by EFL institutions and given to a group of selected EFL teachers to train them to become evaluators of their colleagues. Accordingly, EFL teacher evaluators have never been through any proper professional training that enables them to be accredited evaluators, nor are they democratically chosen as ideal candidates for the position.

Shah and AlHarthi (2014) describe precisely the way in which teacher evaluators are chosen in Saudi higher education. Heads of units/committees responsible for teacher evaluation usually select the people they trust to be in charge of EFL teacher evaluation. The chosen evaluators very often have a close relationship with their bosses and they are selected to perform certain duties. Hence, it seems that the responsibility for selecting and training EFL teacher evaluators is down to the managerial bodies in Saudi universities, which tends to be the case in any typical centralized and top-down educational system. Nevertheless, the intricacies associated with the professional qualifications and personal skills as a complex network of evaluation capabilities, that is typically overlooked, must not be undervalued. As Zepeda (2014b) asserts, those who fulfil supervisory duties, such as evaluators and observers, are classified as instructional leaders whose vision is a combination of pedagogical beliefs, experience, reality and human values. This vision is expected to provide a critical “tangible representation of effective instructional planning and delivery” (DiPaola & Hoy, 2008, p. 10). As Green (2010) writes, the practice of

evaluation helps evaluators develop the ability to skillfully interlink supervision, evaluation, and professional development, which eventually lead to teacher growth.

The findings of this study echo findings from other studies in the literature in the way that classroom experience, context awareness, supervisory skills, qualifications and special training are found to be salient features of a qualified EFL teacher evaluator. For example, Zuheer's (2013) study on EFL teacher evaluation in Sana'a, the capital of Yemen, emphasized the importance for evaluators to be aware of the fact that classroom visits give merely a narrow sample of the teachers' performance. Furthermore, Albaiz's (2016) study on EFL teaching in Saudi universities arrived at a similar conclusion, namely that the evaluator must understand the content and objectives of the course taught by the evaluated teacher. This is significant, given that one of the main aims for evaluators is to collect objective data and evidence about the extent to which the teacher has achieved the course objectives.

Course objectives serve as a road map for teachers and evaluators and guide the evaluation practices. Thus, evaluators should familiarize themselves with the course and its objectives before deciding on the required data to fulfil their evaluative duties. Being aware of these aspects helps evaluators stay focused on the tasks of evaluation and links it to the educational institution's general objectives. It can also help evaluators, who are seen as educational leaders and an essential part of supervisory teams, to ensure that instructional practices taking place within the institution do not violate the institutional vision, aims and policies. And herein lie a number of concerns in the Saudi context, where the findings reveal that evaluators with less experience

and familiarity with EFL courses and their objectives than the teachers themselves have been assigned to such supervisory positions.

As mentioned in the previous chapter, having a less experienced evaluator than the teacher results in less confidence in the outcome of evaluation, and a lack of trust between EFL teachers and their evaluators. In this regard, this study contradicts some previous research. For instance, Shah and AlHarthi (2014) attribute lack of trust to the power relationship imbalance between EFL teachers at Saudi universities and their evaluators. In my opinion imbalance of relationship may put teachers under undesirable pressure or make teachers develop negative attitudes towards evaluation.

Furthermore, Rehman and Al-Bargi (2014) highlight different significantly related issues in the Saudi higher education context. They relate the perception of EFL teachers' trust to the non-verbal behaviours of evaluators in post-observation meetings. Evaluators' non-verbal behaviours are of considerable importance given that discussing evaluation results with a trusted evaluator helps teachers disclose their opinion about their teaching practices and feel that their comments will not be used against them. In other words, trust as a concept in EFL teacher evaluation has different dimensions and effects. I believe that, in order for evaluators to ensure they have an effective evaluation system and they can be perceived as qualified evaluators, they need to gain the trust of the teachers in the first place.

The findings also suggest that a crucial issue linked to the evaluator's qualifications, which can seriously compromise the integrity of the whole process of teacher evaluation, is favouritism. In this regard, Danielson and McGreal (2000) assert that in many educational institutions teachers believe that evaluators tend to reserve high

grades for their friends; that the practice of favouritism as such makes teachers have little confidence, or maybe no confidence at all, in the results of teacher evaluation.

In the TESOL context in Saudi Higher education, similar findings are found. Al-Jarf (2015) points out that evaluator's favouritism rather than the actual performance of the teacher determines the evaluation results: evaluators tend to be affected by their personal relationship with the teachers they evaluate. She argues that prior conflicts can lessen the chances of EFL teachers at Saudi universities getting fair evaluation results and induce low rating. Evaluators who act with subtle forms of bias by giving high ratings to their friends can be expected to show vindictive attitudes towards the teachers they do not like.

One of the major causes of favouritism and vindictiveness is the method by which evaluators themselves are selected. Evaluators are chosen by their superiors not according to qualification but rather based on the trust shown by their management system. This may be based on a sound rationale, in that firm requirements, standards and/or criteria for selecting evaluators are rarely found in Saudi universities and selection is commonly centred around personal relationship (Shah & AlHarthi, 2014). Favouritism also indicates that the system eliminates disliked teachers from occupying teacher evaluator positions. In my opinion, embracing discriminatory practices in such a way in educational institutions is one of the facets of professional injustice against those eliminated teachers. This practice will have a long-lasting effect on the feelings and attitudes of those teachers who were not fortunate enough to be selected as evaluators. For instance, they may become demotivated and less committed to their job. Indeed, work motivation for teachers is linked to how teachers feel and how they

are being treated at their work place (William, 2015, p. 13). As an ultimate result, the quality of their teaching will be negatively affected, which will subsequently affect the quality of students' learning.

It is possible that setting key job qualifications and adopting professional standards to control the selection of EFL teacher evaluators can help educational institutions have more qualified evaluators and better evaluation systems in Saudi universities. Equally significant, educational institutions need to implement a quality training course/programme for teacher evaluators as a compulsory requirement for occupying such a sensitive supervisory position. In this regard and within a similar centralized educational system in the Gulf, Sufean and Ahmad (2014) reinforce the significance of selecting, appointing and training teacher evaluators, placing it at the top in the execution of policy plans in their model for teacher evaluation policy implementation. This sheds light on both the importance of and the need for appointing well-trained teacher evaluators not only in Saudi Arabia but also in the Gulf area at large. As Murphy (2013) writes, the professional training of evaluators is necessitated by the fact that evaluation standards vary across evaluators when they put the standards into practice for different reasons, which results in evaluation outcomes being less effective and reliable.

In educational institutions where teacher evaluators are not properly trained, EFL teachers develop feelings of unease with what they perceive as incompetent evaluators and relinquish hope for teacher evaluation to be effective. In the context of Saudi higher education, this typically leads, as the findings of this study revealed, to concerns that evaluation outcomes will not support any learning and issues related to

job security will also arise (Rehman & Al-Bargi, 2014). Moreover, Thalib and Manda (2016) argue that incompetent teacher evaluators can negatively affect the performance and motivation of the teachers. Therefore, they deem that “it is time to reform orientation in training school supervisors and principals” in order to achieve the goal that “motivation and performance of teachers become a priority in managing education” (p. 7310). Similarly, other studies that were conducted in EFL contexts argue that lack of professionalism, part of which is professional training, on the side of teacher evaluators makes the whole procedure and/or system of evaluation ineffective (Borg, 2015; Copland, 2008; Hooton, 2008; Howard, 2010; Mallows, 2002; Murdoch, 2000; Shah & AlHarthi, 2014; Tennant, 2006). Based on all the previously discussed issues and in order to make teacher evaluation schemes successful and beneficial, I believe that teacher evaluators need to be well-qualified and professionally trained.

## **6. 4. EFL Teachers’ Aversion to Teacher Evaluation**

As pointed out in the previous chapter, there are significant aspects of teacher evaluation that are related to EFL teachers and have been influenced by problems in implementing policies and practices of teacher evaluation at Saudi universities. Most of the reported teacher-related issues are centred around two key concerns, namely threat caused by teacher evaluation and alienation of teachers within the system of teacher evaluation. In the following two sections, each of these is discussed in light of the existing literature on EFL teacher evaluation research and the context of the study.

### **6. 4. 1. Teacher Evaluation: Stress and Threat Aspects**

The findings of this research indicate that the teachers experience different kinds of work-related stress, one of which is teacher evaluation. In line with this, Aslrasouli and



Vahid (2014) claim that teaching as a profession has been widely acknowledged as full of stress and “teachers of English as a foreign language are not exceptions” (p. 305). The results also echo the findings of earlier researchers such as Borg (1990), who argues that throughout various studies, a third of the investigated teachers around the world perceived teaching as a stressful career and this has become a globally-shared concern. Teachers found that teacher evaluation procedures cause further stress and put them under undesirable pressure for different reasons. This idea has been confirmed by Merç (2015), who indicates that in the context of EFL teaching, instruments used for teacher evaluation such as classroom observation, peer-teachers, paperwork and university supervisors are among these stress-creating factors.

A contributing contextual factor that relates to EFL teacher evaluation in Saudi universities can help explain this particular finding. Teacher evaluation as practised in some Saudi universities tends to be judgmental in nature rather than developmental, whether it is carried out to fulfil accreditation requirements or to make decisions on contract renewal. Shah and AlHarthi (2014) and Shukri (2014) described how EFL teachers in preparatory year programmes in Saudi Arabia experienced stressful situations when they were observed as part of their evaluation. Shukri (2014) claimed that the stress caused by classroom observation is a result of having judgmental rather than developmental evaluation. The findings of this study provide support for this argument. Judgmental evaluation is almost always the case where evaluators attempt to assess the performance of the teacher according to a set of predetermined and fixed ranking criteria (e.g. satisfactory, good, distinguished), with the evaluator deciding what rank to give the teacher based on her/his judgement. As a result,

teacher evaluation is perceived as a threat by EFL teachers, causes considerable pressure and leads to protective and defensive reactions from the side of the teacher. Accordingly, evaluators may need to avoid being judgmental. In line with this, AlKhars's (2013) study investigating creativity in English language teaching in Kuwait indicates that the key to individuals in supervisory positions - such as teacher evaluators - is to remember how stressful it is to be an EFL teacher with all the duties and demands s/he needs to deal with every single day.

The findings have also revealed that when they feel threatened by teacher evaluation practices, teachers tend to fake a lesson when they are observed in order to get high evaluation results and to avoid contract termination or any other undesirable negative effect of low evaluation grades. The artificiality, disconnect from reality and unnatural performance while being observed make it an act of what Rennert-Ariev (2008) calls *bureaucratic ventriloquism*: an "inauthentic response so markedly detached from the individual's own beliefs, that the utterances themselves seem to be projected from elsewhere" (p. 111). In order to improve their evaluation grades, teachers get disconnected from their own beliefs and reality and embrace a deformed replica of their authentic performance. In such circumstances, chances for a genuine intellectual engagement of the teachers is undermined by bureaucratic considerations and instead they are "performing inauthentic gestures of compliance to mandates with which they disagreed" (Rennert-Ariev, 2008, p. 133).

Bradford (2016) arrived at the same conclusion: that teachers tend to shape what they do during classroom observation so that they can fit in with prescribed notions of outstanding or good teaching, which makes the observation become normalized. I

believe that this situation has a serious impact on the teachers, as their freedom and creativity is restricted by such a mechanism of teacher evaluation and observation. This shows how, in their quest for achieving a high mark from the observer, teachers may ignore their philosophy, beliefs and freedom and portray a false 'self' instead of their true nature and behaviour. Accordingly, such a forced change of behavior can harm teachers and cause different forms and levels of identity tensions. As Warner (2015) maintains, when teachers become the people they believe their evaluators want in order to fit in within the community of practice, they feel they are "turning their backs upon their deeply-seated beliefs in teaching as humanistic and personal work" (p. 156). Warner goes further and suggests that by doing so, teachers end up perceiving themselves as "bad teachers", where they measure their successes on achieving "mimetic rather than transformative outcomes" (p. 156).

TESOL researchers such as Al-Harthy (2017) have suggested that, for the purpose of teacher evaluation, a classroom observer should behave as a mentor and a guide rather than a threatening judge. In order to achieve this, the observer should bear in mind the teachers' experience, avoid being negative, and discuss any possible alternative approaches to teacher evaluation with the teachers themselves. As discussed in the section above, teachers perceive teacher evaluation as a threat when its purpose is mainly judgmental. In order to mitigate the psychological uncertainty of fear caused by judgmental classroom observation in particular and teacher evaluation in general, evaluators need to show more respect for the teacher and to acknowledge his/her effort and experience. Melibari (2016) stresses that teacher evaluation needs to be oriented as a constructive practice aimed at helping EFL teachers rather than a "destructive criticism" (p. 216). When its results are used for managerial decisions,

teacher evaluation becomes threatening and puts teachers under undesirable stress instead of helping them. The findings of this study suggest that teacher evaluation results which derive from peer-observation are perceived as less threatening than those from supervisors and/or administrators. As Ostovar-Nameghi and Sheikhahmadi (2016) contend, facilitating teacher collaboration where a friendly atmosphere is provided by collegial interaction allows teachers to exchange feedback, support and assistance in a non-threatening reciprocal approach. Amongst other techniques to encourage teacher collaboration, they suggest peer-observation. However, some issues related to peer observation, such as confidentiality and judgmental behaviours, have been reported. Also, and from a rather personal perspective, I believe that peer-observation requires all teacher-peers to be well prepared and trained to carry out this procedure. In fact, this can be challenging when put into practice, as language institutions in Saudi higher education need to spend a lot of time and money to get all EFL teachers ready to activate peer-observation as a legitimate exercise, whether to evaluate teachers' performance or to develop their teaching practices.

It is possible that adopting a well-structured and well-developed training programme for teacher evaluators can lead to a less threatening teacher evaluation. As discussed earlier, teacher evaluators do not go through sufficient training that makes them well-qualified for this position. Besides making their evaluation outcome questionable, this also contributes significantly to the threat aspects of teacher evaluation as perceived by EFL teachers, given the fact that decisions are made based on some of these evaluations. In this regard, Shah and AlHarthi (2014) indicate that in the context of TESOL in the Saudi universities' preparatory year, "teachers-observers' trust deficit

and the observers' lack of training contribute to the teachers' anxiety, stress and insecurity" (p. 1599). I believe that providing better training opportunities for the evaluators will allow them to develop their supervisory skills; they can learn how to evaluate with more objectivity, acknowledge teachers' strengths, and help teachers develop and improve their practices. They should also learn how to encourage a collaborative atmosphere between teachers which provides a less threatening work environment. Consequently, the evaluators become more reliable and qualified practitioners, which helps teachers not only feel less pressure but also develop faith in the teacher evaluation system.

#### **6. 4. 2. Alienated EFL Teachers**

The findings of this study reveal that EFL teachers' involvement in teacher evaluation at Saudi universities is very limited. In almost all the public universities investigated, EFL teachers are not offered proper induction sessions that introduce the teacher evaluation scheme at their institutions to them and instead they learn about it only when it happens. This unfortunate situation means that they are unaware of many of the aspects related to their own evaluation and hence the evaluation does not make much sense or have much value for many of them. The findings echo those of Bradford's (2016) study: that the authentic engagement of teachers in teacher evaluation is hindered. Bradford highlights that organizational routines do not usually create the space required for teachers to engage in sense-making about teacher evaluation practices, which is the basis for authentic engagement, and instead cause "inauthentic 'hoop jumping' during evaluation activities" (p. 6).

On a theoretical level, sense-making is tackled in two theories of evaluation, namely Value-Engaged Evaluation (VEE) and Emergent Realistic Evaluation (ERE), in terms of whose perspectives and to what degree they are to be included. A key feature distinguishing between the two theories is that, while sense-making is a fundamental concern of VEE, “values inquiry is dependent on the purpose of the evaluation in ERE” (Dillman, 2013, p. 58). Sense-making can be seen as an initial step in which findings are used to support further evaluative activities. Approaches to evaluation that involve sense-making to understand the phenomena around us are offered through emergent neo-realist theories that use the term ‘sense-making’ to “describe efforts to construct meaningful orders” (Julnes & Mark, 1998, p. 39). Given that knowledge is formally constructed through evidence-based approaches to evaluation at Saudi universities, which is totally rejected by neo-liberalists who believe in the innate capacity of humans to make sense, sense-making seems to contradict current evaluation that has a positivist epistemology orientation and hence is absent in the Saudi context.

These findings can be directly related to the way EFL teacher evaluation is practised at Saudi universities. As discussed earlier (section 6.2.2. p. 261), the evaluation of teachers tends to be summative in nature and does not lead to improved instructional practices, as described by the participants of this study. In this regard, Horn, Kane, and Wilson (2015) contend that at educational institutions where organizational routines are developed merely to monitor instructional practices instead of to improve them, the practice is fostered by an “*instructional management logic*” rather than “*instructional improvement logic*”. With this type of routine, educational organizations tend to oversimplify teachers’ sense-making and restrain teachers’ engagement in deliberation with administrators about practice problems (Bradford, 2016). In lieu of

this, they apply evaluation in a decision-making oriented approach and overlook the engagement of teachers in any discussion about the practice of teacher evaluation. Rennert-Ariev (2008) cautions that disregarding deliberation inhibits the authentic engagement of teachers and creates an atmosphere of false representation and performance, such as evaluation forms with boxes to check, which was reported by participants of this study. Furthermore, Osmond-Johnson (2015) warns that the constrained involvement of teachers in developing and improving educational practices has unfortunately “left a wealth of teacher knowledge untapped and undervalued” (p. 2).

As indicated in the previous chapter, EFL teachers at Saudi universities are given very limited opportunities to air their voice in the practice of teacher evaluation. These results echo findings from Aburizaizah’s (2016) study that in Saudi higher education system, policies and practices of EFL teacher evaluation are influenced by a specific group of people who occupy an equal authority level and work within the same circle. Other people who are situated outside that circle are not granted a voice in teacher evaluation and their influence is very limited and sometimes hindered. Aburizaizah attributes this phenomenon to the sociopolitical context and, given that Saudi universities are managed and run according to government policies, substantial improvement and change in practice do not occur. This issue of power structure in educational institutions was discussed earlier in light of the findings (section 6.2.1. p. 255). On the other hand, when educational institutions provide opportunities to empower teachers and give them a proper chance to air their voice in their evaluation and to improve the scheme, they will bring in positive consequences for the teacher’s “self-efficacy” (Reddy et al., 2016). This is likely to happen, given that teachers’ voices

provide genuine and firsthand engagement with educational practices within that particular context, which will lead to improving the policies and practices of both evaluation schemes and teaching. In this regard, Osmond-Johnson (2015) stresses that, when awareness of the breadth of issues related to teaching is created and means for a collective voice is exerted, teachers will be empowered and their professional development and growth also reinforced.

Some studies in the context of TESOL at Saudi universities, however, have produced contradicting findings. For example, Shah and AlHarthi (2014) argue that within the process of teacher evaluation and for the sake of carrying out classroom observation, EFL teachers are perceived as totally “passive participants” throughout the whole process. The reason behind their opinion was mainly the imposed choice of the lesson topic by the observers. The findings of the current study oppose that claim and reveal that EFL teachers have the privilege of not only choosing their lesson but also agreeing or disagreeing about who will observe them. It could be that the practice of classroom observation has improved over the course of time at Saudi universities, given that the above study was conducted and published prior to the data collection for this study, which gives more hope of having better teacher engagement in the future. However, my findings confirm Shah and Al Harthi’s argument that the lack of sufficient professional autonomy and teachers’ voice in the process of teacher evaluation results in teachers resenting the scheme of teacher evaluation and perceiving evaluators as exercising unjustifiable power over them and passing judgement on their teaching practices.



Personally, I believe that there is a desperate need for greater engagement of EFL teachers with teacher evaluation within the context of Saudi higher education. One possible way this may be achieved is by encouraging management to involve teachers in all the different stages of teacher evaluation. First, EFL teachers need induction sessions to introduce to them the teacher evaluation scheme, its purpose, methods and different practices; this will make the teacher evaluation meaningful and important to them, which in turn will help develop the sense-making discussed at the beginning of this section. Then, being at the heart of teacher evaluation, teachers need to reflect on the practice of evaluation in order to improve its policies and implementation. It is the responsibility of educational leaders at Saudi universities to put in place teacher evaluation systems and practices that encourage adequate engagement of EFL teachers.

I believe that this requires more positive approaches and liberal attitudes that help promote understanding and encourage criticality in order to create an inviting and supportive atmosphere for teachers' engagement. In this regard, Wenger (2010) describes what he calls a "*Regime of Accountability*", in which he theoretically defines two modes of accountability: vertical and horizontal. In the vertical type of accountability, matters such as decisional authority, conventional hierarchies, resources' management, bureaucracies, regulations and policies, prescriptions, accounting and audit inspections are associated with accountability. On the other hand, with horizontal accountability, engagement in shared activities, practice standards, negotiation of reciprocal relevance, peer recognition, reputation and identity, and obligation to collective learning are linked to accountability. Wenger warns that assuming the horizontal relationship lacks the required accountability is a mistake,

resulting in educational organizations tending to trust the vertical structure in order to create reliant accountability models. At Saudi universities, EFL teacher evaluation practices seem to be built typically upon vertical accountability. Nonetheless, I am greatly interested in the way in which teacher evaluation schemes could be better balanced with horizontal accountability so as to engage EFL teachers more with teacher evaluation and to give them more opportunities to air their voice.

With a more practical stance, Melibari (2016) suggests that, in order for EFL teachers in Saudi universities to be engaged with their peers to monitor their teaching practices in a non-hierarchical setting, forming “Critical Friends Groups” is highly recommended. Critical Friends Groups are defined as groups of professionals within the field of teaching who gather to discuss issues related to teaching practices (Curry, 2008). They are generally associated with the collegiality of professionals, which leads to a positive influence on learners’ achievement (Dunne, Nave, & Lewis, 2000). This could be particularly effective in the Saudi context, as the findings of this study reveal that, due to the hierarchical setting and threatening atmosphere of teacher evaluation and classroom observation, teachers tend to adjust their performance to please their evaluator and avoid showing their real everyday teaching routine and practices.

By utilizing the strategy of Critical Friends Groups, EFL teachers could be less apprehensive about discussing the difficulties they experience while teaching, without the fear that it will negatively reflect on their evaluation. There is a need to create a safe space for EFL teachers, where they can discuss their difficulties and problems in a supportive, non-prescriptive and constructive manner. It is unlikely that the majority of EFL teachers would want to become part of Critical Friends Groups; yet there should

be opportunities for EFL teachers in the Saudi context to partner with EFL teachers from different universities to establish these groups. In addition, it might be easier to create virtual Critical Friends Groups, given the advancement of online resources in the field of education in general and language teaching in particular.

## **6. 5. Students and the Intellectual Leap between General and Higher Education**

As indicated in the previous chapter, student rating of EFL teachers is questionable in the Saudi context, given that not all students have the required critical thinking and reasoning skills to evaluate their teachers. In all of the universities investigated, students are asked to rate their teachers; however, the effects of student ratings of teachers vary amongst them. When student ratings affect decisions including the retention of teachers, the capability of the students to evaluate the teachers and the students' reasoning skills should be given careful consideration. TESOL researchers have paid attention to student rating of teachers (Burden & Troudi, 2007; Davidson; Mazandarani, 2014; Sanif, 2015) and some have questioned the competency of students to evaluate the teachers. For example, Burden and Troudi (2007) suspect the aptitude of students who have just joined the university, where teaching differs from high school, to rate their teachers in their very first term. When EFL teachers' perception is taken into account, Davidson (2010) reaches the same conclusion as the current study, that many teachers believe students are generally incapable of evaluating "faculty in a fair and objective manner" (p. 187). He also argues that many EFL teachers believe that students generally do not take teacher evaluation seriously. In this respect, Burden and Troudi (2007) suggest that a considerable number of students have a rather cynical attitude towards the impact of teacher evaluation on the

performance of teachers. The Saudi context, however, is not an exceptional case and findings from the analysis chapter reveal relatively similar results.

There are a number of educational and socio-cultural contextual factors that contribute to such findings in Saudi Arabia. For instance, the educational system at the pre-university level and the everyday interaction and communication within Saudi society provide young learners with very limited opportunities to express their thoughts and stand up for their own ideas. Even inside one's family, Saudis do not have much chance to express their opinions. In this regard, Al-Seghayer (2016) highlights that expressing one's view at a Saudi home is subject to hierarchical structure. He argues that his "father and older siblings are entitled to give their opinions concerning, let's say, a family matter, but I am not" (p. 16). In Saudi education, schools tend to practice the very same hierarchical tradition but with the teacher as the authoritative figure.

Another key socio-cultural factor that affects education in Saudi Arabia is that the system of education embodies the spread of cultural norms of collectivist social values. These social values emphasize obedience to Saudi traditions and strict socio-cultural rules. In a way, the system of education in Saudi Arabia reproduces these collectivist values where learners must obtain knowledge from their teachers but refrain from revealing individual or personal insights and/or questioning their teachers' opinions (Moraya, 2012; Razek & Coyner, 2013). This reminds us of the role that educational systems and schools play in what Bourdieu (1977) identifies as the reproduction of cultural and social inequalities. It is then very significant that teachers, through the practice of teaching and other activities inside and outside the classroom, should act as agents of improvement and change, and encourage students to be critical instead

of encouraging the reproduction of social values. This might not sound feasible in a restrictive educational setting like that of Saudi Arabia. However, I believe that a move toward criticality of independent learning is now possible, especially within the wider context of political and social reforms introduced by the current leaders of the country and through Saudi Vision 2030. This Vision is defined as:

A plan to reduce Saudi Arabia's dependence on oil, diversify its economy, and develop public service sectors such as health, education, infrastructure, recreation, and tourism. Goals include reinforcing economic and investment activities, increasing non-oil industry trade between countries through goods and consumer products, and increasing government spending on the military, manufacturing equipment and ammunition. (Wikipedia, 2018)

It seems that young Saudi learners are made to follow the stream of restricted cultural norms where they should not deviate from the behavior of the majority of the population. Accordingly, they strive to fit into the accepted intellectual ready-made and highly valued model. Al-Seghayer (2016) warns that by doing so, a generation of young Saudi learners is being created who are either unwilling or unable to communicate their feelings and/or express their thoughts freely. If needs be, and they are asked about their opinion, these learners will immediately refer to what other people believe or think of the issues they are asked about instead of stating their own opinion. This is an inevitable consequence of the way in which they were raised and educated, where they did not develop the required intellectual skills to construct their own concepts.

In this respect and drawing on the transformative potential rather than the determinism of Bourdieu's theoretical constructs, Mills (2015) raises the need to ensure that the

classroom is linked to the world beyond, where teachers encourage learners to draw on their own experience in order to be more critical and to achieve academic success. Moreover, Alrajhi, Alkharusi, and Aldhafri (2016) in their study of 324 students from different school districts in the Sultanate of Oman assert that teachers need to train students to develop their critical thinking strategies to help them increase their academic achievements.

Turning our attention to the contributing educational factors to this phenomenon in the context of the study, there seems to be strong agreement in the literature that Saudi instructional practices are typically characterized by memorization and rote learning (e.g. Elyas & Al Grigri, 2014; Fageeh, 2003; Hall, 2013; Krieger, 2007; Rahman & Alhaisoni, 2013). Alghamdi (2014) quotes an English language expatriate teacher who taught in Saudi Arabia for many years and came to the conclusion that “Students are used to using their memory rather than to think, analyse, and critique” (p. 212). Additionally, when EFL teachers at middle and high schools explain an English text, they usually do not provide students with examples from everyday life nor do they encourage them to join discussions and draw upon their own opinions and ideas (Moraya, 2013). This break away from real world experience while teaching English to Saudi young learners suggests that students are indirectly encouraged to memorise passages and texts from their text books, normally issued by governmental centralised bodies, in order to precisely reproduce them in the national Saudi exams (Wong, 2016). Al-Seghayer (2014) argues that these national exams are typically known for requiring students to reproduce with perfect accuracy, usually from memory, an English passage selected from four or five passages in the textbook. The more accurate the student in reproducing the original text, the higher the grade s/he gets in

these national exams. Hence, memorisation practices continue to be reinforced in such learning environments.

Faruk and Rahman (2016) argue that in this respect Saudi higher education is no better than intermediate and secondary education, as learning is also confined to memorizing information. This is the unfortunate situation despite the obligation of higher education to follow the “National Qualifications Framework for Higher Education in the Kingdom of Saudi Arabia (...) which encourages critical thinking skills” (p. 202). It becomes apparent that in the Saudi context there is a contradiction between the national framework for higher education and current instructional practices. As Wong (2016) states, in order for learners to meet learning demands in higher education, they must be able to “interpret, synthesize, analyze, evaluate (...) identify patterns, compare and contrast ideas, categorize and differentiate between different pieces of information; and interpret information in a way that reflects sound reasoning” (p. 233).

Given the above description of the current Saudi learners’ situation, it is likely that helping them move away from merely memorising and encouraging them to develop some advanced thinking skills can help towards having a better student rating system at Saudi universities. We need to raise our young generation’s awareness of the different ways to communicate their thoughts and feelings, and to encourage them to independently express their opinions and ideas. By doing so we can help them to what Faruk and Rahman (2016) describe as “fly their ‘nest’ of remember and (...) to wean them off their dependence on memorization” (p. 209) in their conclusion about Saudi university students’ thinking skills. They reached the conclusion that students at Saudi universities adopt memorisation as the easiest learning strategy with which they are

familiar from the time when they start being educated informally at home. They warn that if Saudi university students continue with the same memorisation learning habit, they are going to miss their last and best chance to put an end to being like “lotus eaters” (ibid, p. 209) and to obtain higher thinking skills. In this regard Al-Seghayer (2016) suggests that, although evaluation, as a skill for Saudi students, is not generally a common practice in our schools, learners must be trained in how to critically think in order to be able to evaluate the information they are taught. This is likely to be true, given that the cognitive skills that are required to be able to evaluate and to have advanced thinking skills are not innate, and students cannot acquire these skills independently by themselves (Faruk & Rahman, 2016).

Faruk and Rahman also urge that memorization on its own does not enable learners to move even a step upward to “understanding”, let alone approach the rest of the higher learning domains such as applying, analysing, evaluating and finally creating. In sum, advanced thinking skills are essential learning skills that Saudi learners need to develop. Critical thinking is the needed ‘liberating force’ in educational systems that can produce “students who are well-informed, open-minded, flexible, fair-minded in evaluation, honest in dealing with personal bias, and prudent in judgement making” (Al-Seghayer, 2016, p. 24). However, it needs not only teachers or educational systems but also the whole society to support the development of these required thinking skills that enables students to evaluate teachers.

## **6. 6. Summary of the Chapter**

In this chapter, I have discussed the main points in the findings that were reported in the previous chapter with reference to the context of this study and the existing



literature on EFL teacher evaluation research. I have also highlighted different contextual factors that may have influenced the phenomenon of teacher evaluation in the Saudi context and discussed the findings considering global perspectives. I started the chapter by discussing the importance of EFL teacher evaluation as perceived by the teachers and linked it to other practices related to teacher evaluation in Saudi higher education. I then moved the focus to the evaluators and some critical issues related to their qualifications. In the section that follows, I drew attention to a number of aspects related to EFL teachers and associated with the system of their evaluation. The final point discussed in this chapter was the intellectual leap between general and higher education in Saudi Arabia. In the following chapter, the contribution of this study will be presented along with some recommendations for further research.

## **Chapter Seven**

### **Conclusion, Recommendations, and Contributions**

#### **7. 1. Introduction**

Drawing together the threads of this exploratory study, this chapter provides some implications and recommendations for the main stakeholders involved in EFL teacher evaluation in universities in the Kingdom of Saudi Arabia. The recommendations made are of both theoretical and pedagogical importance. The chapter proceeds with my proposed participatory model for EFL teacher evaluation that is informed by the findings of this study. Some suggestions for future research are subsequently discussed. I then offer some personal reflections on the journey of my PhD. Finally, the chapter is brought to an end with a number of concluding remarks.

#### **7. 2. Theoretical and Pedagogical Recommendations for Stakeholders**

The findings of this exploratory study raise a number of significant issues related to EFL teacher evaluation as practised in Saudi Arabia. Hence, they lead to some important theoretical and pedagogical recommendations. Whereas most of the ideas were generated through data analysis and are more or less in alignment with the existing literature, some may be of greater interest in reference to the context of Saudi Arabia. It is worth mentioning that the proposed recommendations are directed towards both macro- and micro-level policies on teacher evaluation in the Saudi system of higher education (Figure 7.1. below). In fact, most of the recommendations revolve around policymaker- and administrator-related factors, i.e. on administrative

and managerial level. This is to be expected, given that the participants are exclusively EFL teachers who perceive people with higher authority as the agents of change.

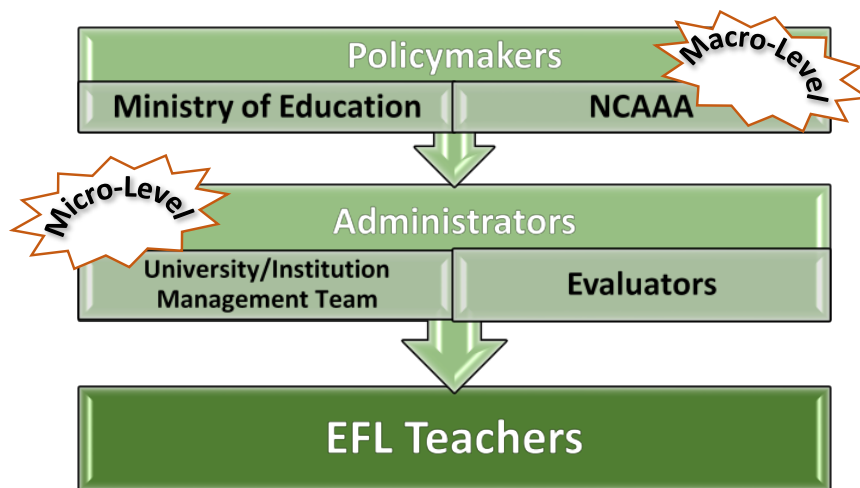


Figure 7. 1. Macro- and micro- level stakeholders

### 7. 2. 1. General Recommendations: Relationship between stakeholders

In the process of my interviews with the participants, it was brought to my attention that some teachers were not satisfied with the relationship between themselves and the people in charge of teacher evaluation within the current hierarchal structure. As discussed in previous chapters, it seems that there is a need to replace what has hitherto been perceived as the *instructional management logic*, where managing how teachers teach is the focus and is practised in hierarchal ways resulting in a vertical relationship between the involved parties, with an *instructional improvement logic*, where the focus is on improving, rather than managing, teaching and is practised in a same level approach resulting in a horizontal type of relationship.

In the recommended landscape mode, one can ensure that all parties involved, i.e. policymakers, administrators and teachers, are required to participate in improving practice. However, each party's contribution and/or participation is unique and cannot be subsumed by another party's participation. This is due to the fact that each party has a different perspective and a dissimilar role to play as a distinctive part within the practice of teacher evaluation. Hence they are expected to produce a different participation with diverse values to the advancement of the practice of teacher evaluation. Yet, each different participation should contribute by the same token to the overall improvement plan. In this way, stakeholders can achieve a horizontal relationship that leads to instructional *improvement* rather than *management*. Figure 7.2. below illustrates these ideas.

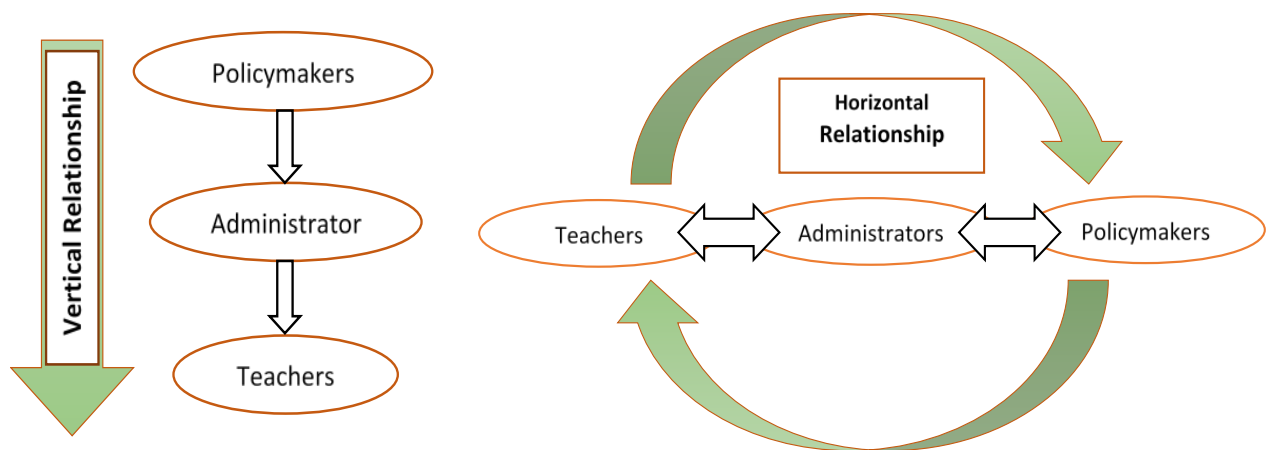


Figure 7. 2. Relationship amongst stakeholders

## **7. 2. 2. Recommendations for Policymakers**

In fact, any teacher evaluation scheme at Saudi universities can be seen as part of a wider system for educational evaluation that is linked to two main bodies in charge of education in the country, namely the Ministry of Education and the NCAAA (as an independent commission). Most policies and practices revolving around teacher evaluation are mainly encouraged, required, verified and examined by the NCAAA; however, there are some areas that the Ministry of Education can be held responsible for in relation to the evaluation of teachers. As has been argued constantly throughout the previous two chapters, there are a number of influential factors that are beyond the control of teachers in the Saudi context and have been reported by the participants. For that reason, it may be necessary for The Ministry Deputy of Planning and Development, The Ministry Deputy of Curricula and Educational Programs, and the NCAAA to revisit their current policies and realign these policies and practices with both the interests and needs of the involved stakeholders.

### **7. 2. 2. 1. The Ministry Deputy of Planning, and Development**

As indicated in the previous analysis and discussion chapters, teacher evaluators require better skills and more training in order to be better qualified for their position. It is well known that the *Ministry of Education* provides training opportunities for those in supervisory positions, such as evaluators of teachers for schools in public education, to ensure they are well prepared to carry out the mission effectively. Given that teacher evaluation is an administrative job, the people in charge of it need to develop and improve their administrative and professional skills. I found that the *Training Management* section provides training sessions on Strategic Performance

Measurement, which is relevant to evaluating the performance of university teachers given that both measure performance in educational institutions.

It could be claimed that the previously mentioned department does not have the capacity to provide proper training for EFL teacher evaluators at university level. This suggests the need to create local training centres and to adopt training as a core requirement to make teacher evaluators eligible for the work. As reported and discussed in the previous chapters, educational leaders and those who occupy supervisory positions in Saudi universities are in serious need of professional training programmes. Drawing on Al-Seghayer (2013) notion of partnerships with both local and overseas training centres, I would like to bring to the fore the idea that the adequacy of policies is contingent upon appropriate implementation. These policies received clear support required to put them into action as they are in alignment with the strategic objectives of the National Transformation Programme (NTP) as part of the Saudi Vision 2030. Given that the second objective of the NTP under the *Ministry of Education* section is to improve the recruitment, training and development of teachers, some critical highlights in relation to training programmes can be extracted.

I believe that a comprehensive framework needs to be developed for the professional development of teacher evaluators, including initial training and continuous training programmes. As a consequence, I would like to suggest forming a national policy and developing practical strategies that help upgrade training practices in particular and the teaching profession in general. I believe that raising the quality of services provided to teachers will raise the professional level of teachers and will result in better learning for students in our universities.

### **7. 2. 2. 2. The Ministry Deputy of Curricula and Educational Programs**

Although this study did not probe into issues related to the curriculum and pedagogical aspects of Saudi educational programmes, the findings show that the participants agree with the idea that, with current educational programmes, students do not develop the required critical thinking skills. This is important, given that university students are involved in teacher evaluation schemes, and it becomes even more important when student ratings feed into, if not form the main source for, personnel decisions. Thus, it is my belief that the Ministry of Education needs to update the national curriculum at a micro-level in order to develop educational programmes that encourage students to think more critically. Under the *Ministry Deputy of Curricula and Educational Programs* there is a department called *General Management of Curriculum Policies and Planning*. It seems that this department needs to launch different curricular policies that support and stimulate our students to gain better critical thinking skills.

As illustrated in the previous chapters, the current educational system in Saudi Arabia requires students to merely memorize information. In order to help students become independent critical thinkers, education programmes and curricula need to be revisited and reshaped according to the current national demands on Saudi individuals and the Saudi society as a whole. Drawing on the Saudi Vision 2030 and the transition of the economy from over-reliance on oil to a more balanced investment in human resources, education becomes a central building block of the Vision. I believe that in the coming few years, the expected progress in the *Ministry of Education*, and other Ministries alike, is likely to be scrutinized carefully. Therefore, policymakers need to consider critical issues that will help them achieve the desired change the country is looking for.

### **7. 2. 2. 3. The NCAAA**

Dependence on external academic agencies, as discussed previously in the findings and discussion chapters, seems to encourage educational institutions in Saudi higher education to remain as consumers of other people's goods. Besides the economic consequences of such practices, they also lead to more dependence and reliance on experts with different educational and cultural backgrounds. As an alternative, the NCAAA needs to take a more effective and independent leading role in accrediting higher education institutions. In order to achieve this goal, the NCAAA should join other accreditation commissions in the area (Gulf, Middle East and/or Arab countries) and collaborate with them in an attempt to establish a regional accreditation body that counts for teachers' voice and participation. It is not wise or acceptable, however, to claim that this will be an easy goal to achieve. As the literature on academic accreditation suggests, establishing, developing, piloting and improving mechanisms for quality assurance in higher education are both time-consuming and financially demanding. For instance, it took Europe more than ten years to establish the European Quality Assurance Register for Higher Education (EQAR), after the Bologna Declaration in 1999 (Cardoso, Rosa, & Stensaker, 2016). Accordingly, I would suggest that the NCAAA and other relevant bodies in the region take steps forward in order to play more effective and influential roles in promoting the establishment of an independent quality assurance regional body.

### **7. 2. 3. Recommendations for Administrators**

The findings of this exploratory study have a number of important implications for the future practice of EFL teacher evaluation in higher education in the Saudi context. The following recommendations relate to those in charge of teacher evaluation and



occupying supervisory positions. I make a distinction between two administrative categories, namely the management team and evaluators. The first group comprises people who are solely carrying out managerial duties, and their concerns are essentially carrying out administrative practices, monitoring regulations, setting committees and recruiting teachers. On the other hand, the second group fulfills a supervisory mission besides teaching and observing classes of their colleagues. Furthermore, the second group is in more direct contact with teachers than the first group. It is worth mentioning, however, that there is not such a clear distinction between the two groups in some Saudi universities and hence the following recommendations apply to both groups.

### **7. 2. 3. 1. University/Institution Management Teams**

As discussed in the previous chapters, institutions generally do not have a set of professional standards for selecting teacher evaluators; as a consequence, favouritism was frequently reported whenever the issue of selecting evaluators was brought to the table. For this particular reason, it is highly recommended that those in charge of teacher evaluation establish a set of required professional standards according to which institutions should recruit evaluators of teachers. By doing so, not only will favouritism be reduced, if not eliminated, but also evaluators will be encouraged to do their best to maintain high qualification standards to stay in their job. Evaluators will also be encouraged to improve their supervisory skills and strategies to match those sets of standards. I would also suggest the democratic election of evaluators, where teachers are provided with opportunities to vote for different candidates. This practice will support the engagement of teachers in teacher evaluation schemes, as discussed in the previous chapters. However, as with any

other democratic approach, this poses different challenges, such as the need for more time and effort.

Furthermore, one of the significant weaknesses reported by the participants and discussed in the previous two chapters is the insufficient training of evaluators. I recommended to policymakers in the first section (7.2.1.1.) that the Ministry of Education provide proper training opportunities for evaluators; however, higher education institutions can also provide quality training for evaluators via other different channels. There are various private local and international training centres that are certified and have the capacity to offer adequate training workshops for candidate evaluators. Not only do novice evaluators need training, but also competent and experienced evaluators need to stay updated by being enrolled into professional development programmes. It is highly recommended that evaluators get educated and certified using initial training courses and subsequently undergo frequent up-dating through special in-service continuous apprenticeship training. A licensing system, where evaluators need to go through compulsory training programmes whenever their practice license expires, would also be beneficial. This may be problematic in the Saudi context, but it would help university management ensure that evaluators are well-qualified, updated and on top of their practice to carry out their duties effectively and to make the system of teacher evaluation more trustworthy.

The findings also indicate that the participants perceive teacher evaluation as a threat that puts them under undesirable pressure and results in harmful feelings. In most cases this was frequently linked to the fact that teacher evaluation is typically summative in nature and evaluators are perceived as being judgmental in their

conclusions. Therefore, I recommend that institutions' management needs to add a formative lens to the scheme of teacher evaluation at their workplace. This can be achieved through improving the teachers' instructional practices, providing professional development sessions, and supporting teachers to help students learn efficiently as the ultimate goal of teacher evaluation. Universities' management needs to replace what has hitherto been practised as summative decision-making oriented evaluation with a more developmental approach to teacher evaluation.

Moreover, the study found that most of the participants expressed aversion to current teacher evaluation regulations and practices. Universities' and institutions' management alike need to listen more to the teachers and give them the chance to air their voice in an unthreatening atmosphere in order to reduce that feeling of aversion. The majority of EFL teachers at Saudi universities are expatriate teachers and, as reported in the analysis chapter, some of them may prefer to keep a very low profile about how they feel and might therefore be seen as passive individuals.

I believe it is the institution management's responsibility to encourage the engagement of EFL teachers in the development of current policies and practices of teacher evaluation. Hence, I recommend that those in charge of managing higher education institutions empower EFL teachers and engage them with the system of teacher evaluation within their institution. I therefore encourage them to introduce teacher evaluation to the teachers by arranging induction sessions and providing written documents that can be kept as a handy reference for the teachers; this will help make teacher evaluation become transparent. Also, and given that teachers often have something to say, as I witnessed in the data collection phase, I support the need to

involve teachers in improvement plans for teacher evaluation systems and to ask for their opinions and ideas. In this respect, I recommend that the management asks EFL teachers to reflect on the practices of teacher evaluation in a regular basis and consider their reflections when improving and updating these practices.

### **7. 2. 3. 2. Evaluators**

The findings of this exploratory study provide a number of insights into the outcomes of current EFL teacher evaluation as practised in Saudi universities. Having revealed a lack of follow-up improvement plans that are grounded in the results of teacher evaluation, this study calls for further attention to implementing a formal agreement, between EFL teachers and evaluators, on action plans for teachers in order to develop and improve their teaching practices. Planning is the cornerstone of any required performance development where both strengths and weaknesses are identified and ways of maintaining strengths and improving on weaknesses are explored. EFL teachers and evaluators need to reach an agreement on what the teacher is required to do, what workshops s/he may need to attend, and when these actions are required to take place. An adequate follow-up review should help both teacher and evaluator to remain informed about and alerted to the teacher's progress and achievement towards accomplishing the improvement plan.

Furthermore, the fruitless nature of workshops and the incongruity between EFL teachers' training needs and the professional development sessions and workshops on offer in the higher education institution have been extensively discussed in the previous chapters. The findings reveal that the participants expressed dissatisfaction with the professional development programmes currently provided. Accordingly, there is a need to replace what has been, up till now, a haphazard method of providing

developmental training opportunities for EFL teachers with better well-planned, structured and organized training programmes, for which needs analysis is a necessity. Additionally, this study highly recommends involving teachers in decisions about their professional development activities. This will greatly contribute to the empowerment of EFL teachers and give them an adequate chance to air their voice - an issue that the participants reported to be desired and needed in the context of Saudi Arabia. This will apparently encourage more engagement of EFL teachers and will also mitigate their aversion to the current system of evaluation of teachers in Saudi Arabia.

#### **7. 2. 4. Recommendations for Teachers**

The findings reveal that current practices of teacher evaluation *per se* put teachers under undesirable pressure in order for them to achieve evaluation standards. As discussed in previous chapters, the inauthentic performance detaches EFL teachers from their own beliefs and reality. As this was attributed to the summative nature of teacher evaluation practices, teachers should facilitate informal collaboration amongst themselves and employ some other informal strategies to enhance and monitor their performance. This requires EFL teachers to have a real intention and to take some courageous positive action to implement other techniques to achieve this purpose. For example, teachers can observe each other's classes or they may form Critical Friends Groups to help address weaknesses and improve their teaching practices in a friendly manner. The study found that some teachers believe working with peers should be encouraged by administrations and results should always remain confidential. Therefore, this study recommends that EFL teachers take the lead and encourage these informal approaches and also maintain confidentiality of the practice. Teachers

should collaborate with one another to mitigate the current reported pressure and help create a more relaxing environment. This also should be encouraged by administrative leaders, as it can help make teachers less passive in improving the practice of teacher evaluation in Saudi universities.

As one of the findings indicates, the professional development sessions educational institutions currently provide for EFL teachers do not satisfy the teaching practice needs of teachers, nor are the sessions informed by the outcomes of teacher evaluation. Accordingly, this study recommends that EFL teachers engage in independent self-directed professional development activities. Besides collaborating with colleagues from the same profession, EFL teachers should be encouraged to learn how to be reflective practitioners. The practice of reflection helps teachers to step back and critically observe and evaluate their teaching and instructional practices. Reading simple articles that focus on practice is an invaluable professional asset to language teachers. Remaining connected to a community of practice is one of the significant pathways to enhance teaching and to make teachers more aware of the most recent and effective teaching strategies and techniques that will help them make better learners. An additional way to help teachers update different aspects of their knowledge and practices is to join the professional development workshops, either traditional or on-line, that are offered by private training centres. In other words, EFL teachers need to feel more responsible about their own professional qualities and improvement. Accordingly, they should make an effort to provide themselves with other development opportunities and sources.

### **7. 3. Contextual, Methodological and Theoretical Contributions**

This exploratory study contributes to the body of knowledge in the field of EFL teacher evaluation in a number of ways. First, it has added to our understanding of English language teacher evaluation in the context of Saudi Higher Education; second, it has a number of methodological contributions; and, thirdly, it has contributions to make also with regard to theory. In the following, I shall elaborate on these contributions and present them in more detail.

#### **7. 3. 1. Contributions to the Saudi Context**

The current study contributes to the literature pertaining to higher education in the Kingdom of Saudi Arabia in many respects. As explained in earlier chapters in this thesis, this exploratory research fills a gap in the existing literature with regards to the challenges EFL teachers face under current teacher evaluation systems in the Saudi higher educational context. This is, to the best of my knowledge, the first study conducted in Saudi Arabia that has explored the importance of, challenges to, and suggested solutions for the evaluation of EFL teachers in Saudi universities. As a result, the current study can be beneficial to academic authorities and policymakers in Saudi Higher Education given that it sheds light on significant aspects related to the evaluation of language teachers in higher educational institutions. For instance, they can develop existing policies and practices of teacher evaluation to ensure it is perceived as fruitful and significant to the teachers. Alternatively, they can improve the current circumstances and workplace setting to mitigate the challenges and difficulties associated with teacher evaluation.

As the review of literature has revealed, previous research studies have investigated certain aspects related to EFL teacher evaluation in Saudi Higher Education. Yet, none of these studies, to my knowledge and up to the date of writing this thesis, has explored in depth the perception of EFL teachers of the importance of the scheme and the challenges related to the way it has been practised in different Saudi universities. The study also offers details of solutions proposed by EFL teachers for the current weaknesses and drawbacks of the teacher evaluation systems currently implemented. Consequently, it is hoped that this study has filled some of the existing gaps in knowledge and research in the area of TESOL about this particular phenomenon and some of its relevant aspects in this particular context.

### **7. 3. 2. Methodological Contributions**

On a methodological level, this study has attempted to show the value of mixed method sequential research that includes both quantitative and qualitative tools. The literature on EFL teacher evaluation has shown that previous studies have mainly used a single qualitative or quantitative approach. This study is one of the first exploratory studies that have provided both breadth and depth in examining the investigated phenomenon. The distribution of the questionnaire in three different regions in a large country like Saudi Arabia to cover five different universities gives the study the breadth that distinguishes it from other studies in the field within the Saudi context that tend to cover a single educational institution.

In addition, the insights from the individual and group interviews provide this study with the depth required to understand the explored phenomenon, which contributes to depicting a complete snapshot of EFL teacher evaluation in the Saudi Higher



Education context. A large number of the participants showed great enthusiasm and eagerness to answer the questions of all three tools utilized. For most of those who were interviewed, it was their first experience of being involved in a qualitative research study. The qualitative tools have also given me the chance to carefully listen to the participants and to pay attention to details which purely quantitative studies lack. In fact, using both quantitative and qualitative approaches jointly has enriched the findings of this exploratory study and helped it to distinctly stand out amongst other similar studies in the field of teacher evaluation in Saudi higher education.

### **7. 3. 3. Theoretical Contributions: Participatory Teacher Evaluation Model**

In the literature review chapter, I presented and examined a number of frameworks and models adopted in different contexts and utilized to evaluate EFL teachers. Building upon different issues raised by and explained throughout this exploratory study, I propose a participatory model for EFL teacher evaluation as illustrated in figure 7.3. in the following page. The proposed model is informed by the perceptions of Saudi and expatriate EFL teachers about teacher evaluation as currently practised in five Saudi universities. Given that the proposed participatory model is informed by EFL teachers' understanding and experience of the current system of teacher evaluation and concerns related to its policies and practices in the Saudi context, different components have been embedded in the proposed teacher evaluation model. These include the main stakeholders involved in teacher evaluation and participating in making decisions, the required actions that need to be addressed, and the four stages where the actions and stakeholders are to be involved.

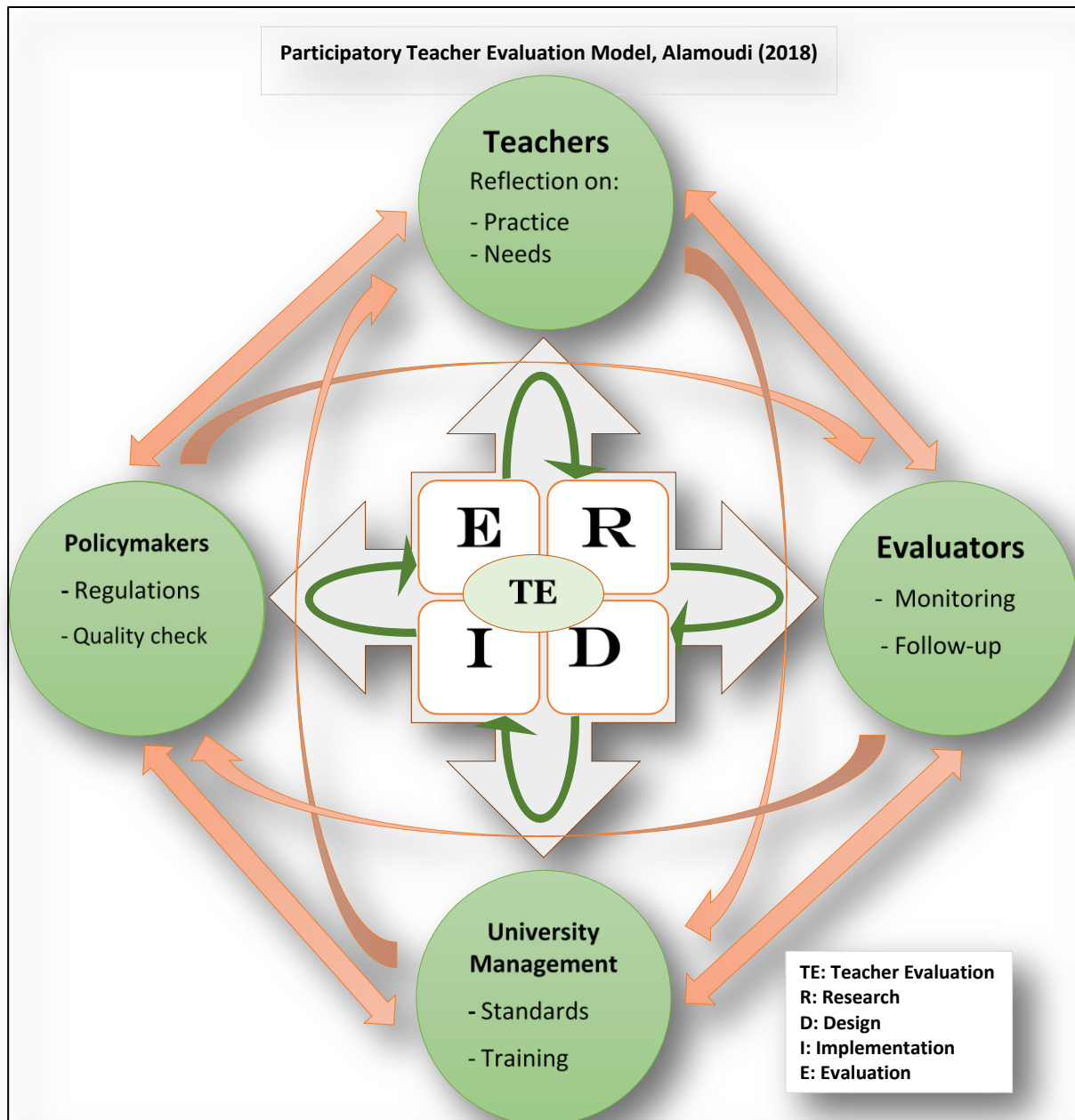


Figure 7. 3. Proposed participatory teacher evaluation model

As indicated in previous chapters and reinforced in the recommendations section above, vertical relationships in a hierarchical structure with the reported power imbalance within current teacher evaluation programmes do not help overcome the obstacles or face the challenges to current practices of teacher evaluation. In the

current reported model of EFL teacher evaluation, administrators and higher authorities are more empowered than those who are being evaluated, i.e. the 'marginalized' teachers who have little or no voice to participate in the development of the current evaluation schemes.

Instead, the participatory approach to teacher evaluation I propose can foster teacher's empowerment and promote the genuine investment of the parties involved. As Newman and McNamara (2016) state, social work values such as anti-oppression and social justice can be highlighted through a participatory approach. This involves sharing information amongst different stakeholders, especially those who are marginalized and vulnerable. Moreover, the participatory approach includes exploring and generating new knowledge that aims at addressing situations that need improvement. In the following section, I give a general description of the proposed participatory model.

### **7. 3. 3. 1. The Model: A General Idea**

Informed by findings from the current exploratory study, key stakeholders (a total of four) are identified, arranged according to the authority and power given to them within the current teacher evaluation system, as reported by the participants. The least empowered and most marginalized stakeholders, i.e. teachers, come first, followed by those who are more empowered, in a sequential order. Stakeholders are placed on the four points of the diamond-shaped model. Major actions related to the evaluation of teachers are summarized and listed under the relevant stakeholders (more details about stakeholders and actions will follow in Section 7.3.3.2. p. 308). Relationships between stakeholders, formed in order to make decisions on the actions, are represented in the model with two-headed arrows or two reversed one-way arrows to

reinforce the notion of a reciprocal participatory relationship, where the power of decision-making is being shared rather than just flowing from the more powerful authority at one end to another passive party at the other end, which would be represented with a normal one-headed arrow.

Informed by the works of Tufte and Mefalopulos (2009), this proposed participatory model of teacher evaluation adapts and visualizes the four stages of the Participatory Development Project framework they identify, namely research, design, implementation and evaluation. These four stages are represented by their initials and placed at the centre of the proposed model and connected in a dynamic cyclical mode with single-headed arrows to indicate a sequence between them. In this sequence, every stage leads to the one next to it. It is anticipated that each participatory development project cycle would begin with research as its first stage and end with evaluation as the fourth and final stage, where that end leads to the beginning of a new participatory cycle (more details about the four stages will follow in Section 7.3.3.3. p. 311). Participation in decision-making is intended to take place throughout these four stages amongst those key stakeholders who are expected to be affected by the intended changes (Tufte & Mefalopulos, 2009).

### ***7. 3. 3. 2. The Model: Who and What? Stakeholders Involved and Actions Required***

Informed by the results of this exploratory study, the four key stakeholders as perceived by the participants are teachers, evaluators, university/institution management and policymakers. Each of these stakeholders is allocated a space at one of the ends of the semi-diamond shaped model. The action(s) that is/are required by each party will be explained below; however, and as indicated earlier, these are

interactive communicative actions where relevant stakeholders are expected to participate in making the suggested decision(s).

- **Teachers**

By being evaluated, teachers are at the very heart of the practices of teacher evaluation and the stakeholder most affected by its consequences. Hence, teachers should be encouraged to reflect on the current practices and how they affect, whether positively or negatively, their teaching practices. It is the teachers who are typically responsible for the progress of their students and achieving the instructional objectives of educational programmes and, hence, they can critically relate all these to the process of teacher evaluation. Teacher evaluation is supposed to support the pedagogical elements of any programme where teachers can reflect on them (Brookfield, 2017) as they are the link between almost all of the elements in education. Teachers are also required to report on what they feel is needed in order to improve the current practices of teacher evaluation. These actions should be carried out in a participatory approach with other stakeholders through planned and structured activities in groups such as one-day retreat workshops or individual participation in short surveys or one-to-one meetings.

- **Evaluators**

Teacher evaluators are typically in charge of monitoring the performance of the teachers. They are also expected to set follow-up plans to ensure that the performance of EFL teachers meets the desired standards. However, evaluators are not recommended to work in isolation and instead they should engage with other involved

stakeholders. In this way, educational institutions can achieve better teaching and accordingly better learning and student outcomes.

**- *Universities/Institutions Management***

Management of universities/institutions, with insights from other stakeholders, are expected to set standards to select teachers' evaluators and to introduce proper procedures for arranging continuous training for evaluators. As these areas of standards and training may vary in different universities depending on their needs, managements can tailor standards and training requirements to these needs. In the proposed model, university/institution management is encouraged to involve other relevant stakeholders such as teachers and evaluators.

**- *Policymakers***

Policymakers are still in charge of making decisions in relation to teacher evaluation in the proposed model. However, these decisions should be informed by ideas from other relevant stakeholders. Policymakers will still make regulations for teacher evaluation and be held responsible for quality checks; this will ensure that clear, updated and suitable regulations are available for other parties to decide on and to implement in their educational institutions and programmes. Furthermore, related stakeholders are expected to be involved in the quality check procedures and their implementation, which is mainly part of the policymakers' duties. This reinforces the idea of the two-way communication this study calls for.

It is worth adding that, on a regular basis, all the involved stakeholders need to go through a kind of 'participatory communication' for a new cycle of teacher evaluation

project development, i.e. to begin a cycle of the four stages. The following section explains each of these stages.

### **7. 3. 3. 3. The Model: Stages of a Participatory Development Project**

Participation is essential for the development of any project to bring in different perspectives. As suggested by Tufte and Mefalopulos (2009), this participation can take place in four different stages as follows:

#### **- Research Stage**

In this stage, all relevant stakeholders are involved in accurately defining problems/challenges related to the effectiveness of the teacher evaluation programme. The required research at this stage is expected to revolve around the development problem and might include the study of previous experience, current policies, the knowledge and attitudes of individuals and communities, and other contextual information that is in direct relation to culture, socio-economic circumstances, gender, religion and other factors related to teacher evaluation.

#### **- Design Stage**

In this stage, the actual activities are defined. The active participation by the involved stakeholders not only improves the quality of the outcome but also determines the relevance of the recommended interventions. Moreover, the commitment and ownership of the communities involved can be further secured by utilizing the participatory approach.

#### **- Implementation Stage**

The planned intervention is implemented at this stage. Furthermore, relevance, commitment and sustainability can be increased by participation during this phase.

### **- Evaluation Stage**

The participatory approach here ensures that the required changes are achieved, brought to the attention of relevant parties and assessed. In order to achieve meaningful evaluation, measurements should be defined at the beginning of the initiative in participatory processes that allow the involvement of all relevant stakeholders. It is finally worth mentioning that a consistent application of two/multi-way communication in each stage is of a vital importance to make the whole cycle of programme development genuinely participatory.

### **7. 3. 3. 4. The Model: Concerns and Complexities**

Before concluding this section on the proposed model, it is worth emphasising that the current study does not claim that the proposed participatory EFL teacher evaluation model could be generalized to all contexts or that it is entirely new. The proposed model is informed by the findings reported in this thesis, which are a reflection of the perceptions and understanding of the participants in the Saudi context. Informed by the Saudi university participants' points of view, the proposed model is intended to help in dealing with issues and concerns which revolve around the practices and policies of EFL teacher evaluation in higher education in the kingdom of Saudi Arabia.

Given the exploratory nature of the current study, it cannot be assumed that the proposed model is entirely applicable without any need for modification and/or addition. The model may remain theoretical until its practicality is examined, revisited and modified according to the needs that may arise as soon as it is implemented. As Chouinard and Cousins (2015) write, the use of a participatory approach to evaluation



has been advocated, as it is perceived to afford better ways that help capture different realities that are genuinely local. However, in contexts such as Saudi Arabia, where most of the systems are top-down oriented, the notion of participation *per se*, let alone its implementation, might turn out to be much more problematic and complicated. In such a context, not all people have an equal voice; as a result, it is quite common to discover different types of struggle with the programme community in terms of its cultural norms (Bradley, Mayfield, Mehta, & Rukonge, 2002; Ofir, Kumar, & Kuzmin, 2013), given that evaluators often come across complex processes of conflict resolution and/or negotiation (Chouinard & Cousins, 2015).

As a result, concerns, compromises and constraints may emerge in the procedures of adopting a participatory approach to evaluation in the context of the study. For instance, Newman (2008) reinforces the idea that the term “participation” is subjective and she raises different issues related to participatory evaluation in terms of the dynamic and balancing of power. Before reaching this step, the selection of stakeholders may be a challenging task in itself (Cullen, Coryn, & Rugh, 2011). McGee and Gaventa (2011) conclude that one of the issues of the participatory approach, besides power-related issues, is mainly to do with who defines the problem. Thus, and as Newman (2008) claims, balancing the needs of diverse stakeholders is essential. There may also be issues that are related to other contextual factors depending on the university where the participatory approach is practiced.

#### **7. 4. Suggestions for Further Research**

To the best of my knowledge, this study is one of very few attempts to expound different dimensions of the phenomenon of EFL teacher evaluation from the teachers’

perspective in Saudi Arabia. The findings of this study have produced a number of perspectives that stimulate careful consideration in order to attest to the aforementioned recommendations for the three main stakeholders, namely policymakers, administrators and teachers. In addition, the findings raise many significant yet unforeseen concerns and questions, for which more research is necessary. Due to the limitations of funding and time allocated for this PhD research, I have focused my attention on EFL teachers as the main and central stakeholder around whom the procedures and effects of teacher evaluation occur. Accordingly, it might be a good idea to explore other stakeholders' perceptions in order to reach a more comprehensive understanding of the very nature of EFL teacher evaluation at Saudi universities. Separate studies can be carried out for each of the stakeholders, where the research gives focused attention to each party exclusively.

This study has attempted to explore EFL teacher evaluation systems in Saudi higher education as experienced by the teachers in five different universities. All the participants in this study worked at public universities and hence have experienced teacher evaluation as it is practised in that specific sector only. Including universities from the private sector might have enriched the outcome of this study, as it would illustrate how teacher evaluation is being practised in that sector and would introduce some different dimensions of the areas of EFL teacher evaluation explored. Although I made an effort to approach the gatekeepers of some private universities in Saudi Arabia, as I strongly believed that data from that sector would bring to the table different issues, I was unable to get a positive response from any of the three private universities I contacted. Thus, further research should be conducted to explore how EFL teachers in private universities perceive teacher evaluation. This is of high

importance in Saudi Arabia, since there is seemingly a growing interest in private sector education in Saudi Arabia. It is worth emphasising that one of the Saudi Vision 2030 Objectives in Education as reported in the National Transformation Plan is to increase private sector participation in education and to expand the privatization of governmental services.

As my last suggestion, I would like to reinforce the importance of embarking on research that might be seen as a fresh approach within the context of Saudi Arabia. As most research conducted in the Saudi context is mainly exploratory and/or experimental in nature, more research from other paradigms is needed. One possibility would be for action research to be conducted, utilizing the suggested participatory EFL teacher evaluation model, or some of its aspects, to operate an intervention and seek a change. In order for this change to happen, some trustworthy research from other paradigms, e.g. the critical or postmodernist paradigm, needs to be conducted. Such research did catch my early attention and thoughts; however, given that very little research had been done at that stage, I decided to lead the initiative and to make my exploratory study a base for further research under the umbrella of these other paradigms. Introducing research from other paradigms will enrich the research on EFL teacher evaluation in the Saudi context and will similarly verify outcomes from recent studies.

## **7. 5. My PhD Journey: Reflections on an Academic and Socio-Cultural Adventure**

As I am reaching this final stage of my doctoral journey, I would like to conclude with some personal reflections on the whole experience. When I reached this very last chapter, I had a flashback of how this doctoral journey started. Indeed, finding the right

school and having good advisors can turn the doctoral journey into a magical ride, and I feel I was very fortunate to have them both. Pursuing a postgraduate degree from the School of Education at the University of Exeter, which is one of the top one hundred universities around the world, has made my PhD journey a noble academic venture. Living in the UK with my loving family has also added extra socio-cultural lenses and dimensions to the journey I have been through and enriches it.

On an academic level, the MSc programme in the first year helped polish my research skills. Bearing in mind my previous research background, which was highly influenced by the interpretive approach towards academic research, the MSc programme was a great opportunity for me to learn more about other philosophical and theoretical stances. This was a significant step I needed to take before gaining the academic and research confidence to tackle a phenomenon at doctoral level. At that early stage of my doctoral journey, I had the privilege to know about and to learn how to utilize different computer programmes and software such as SPSS, NVivo, Mendeley and EndNote, all of which are technological advancements that assist researchers in this era of technological revolution. Amongst different academic qualities I have attained during my PhD journey I would like to highlight the skills I have had to acquire in order to become an independent critical researcher; these will help me to further my knowledge, lead future projects and advance my career.

Living in a historical and cultural country like the United Kingdom can be a delightful and daunting experience at the same time. Having lived in Saudi Arabia my entire life, I found living in Britain to be completely different from what I was used to. I have enjoyed the recurring natural beauty of the four different seasons over the years in

Exeter, which is less apparent in my home-town of Jeddah. However, I also have experienced some rough cold winter nights, which were a shock to the system. Different dress-codes and different interaction styles and communication were some issues I came across at the beginning but I eventually learned how to adjust in order to be well settled in the new society that I became part of. Given the availability of different means of public transport, I was able to explore different historical and cultural destinations in the country. Besides the amount of reading I had done about living in the United Kingdom, some while preparing to come to Britain, these visits helped me to better understand British history and culture. Drawing upon what I mentioned in this final section, it is my contention that the PhD journey is a major turning point in my life and marks a new chapter in which I hope to be able to contribute to education, academic research and society.

## **7. 6. Final Thoughts**

The main purpose of this thesis was to gain a deep understanding of EFL teachers' perceptions about teacher evaluation as it is currently practiced in Saudi universities. The participants of this exploratory study, EFL teachers from five different public universities in the Kingdom of Saudi Arabia, provided many insights into the importance of teacher evaluation, the challenges in making teacher evaluation meaningful to EFL teachers, and the suggested solutions to overcome current challenges. The study has also provided some practical and theoretical recommendations that might help develop current practices of EFL teacher evaluation in Saudi higher education. Furthermore, the study has suggested potential areas for further academic investigation and research.

In the final chapter, this exploratory study has been brought to its end and conclusions drawn. However, this concluding chapter has also opened several doors for further investigation in the future. In view of this, I hope my thesis will constitute a point of departure for other researchers in the field of TESOL to further investigate this significant phenomenon of teacher evaluation in order to improve the teaching and learning of English in the context of this study as well as in other similar contexts.

## Appendix 1: The Questionnaire of the Study



### Survey: EFL Teacher Evaluation at Saudi Universities – Study

#### EFL Teacher Evaluation at Saudi Universities

#### Help Make Our "Evaluation of EFL Teachers" Work Better

Dear colleagues,

My name is Khadija Alamoudi and I am a lecturer at the English Language Institute, King Abdulaziz University, Jeddah. Currently, I am a doctoral student at the University of Exeter and my research focuses on EFL (English as a Foreign Language) teachers' evaluation in Saudi universities.

You are kindly invited to participate in this survey. As a participant, you are expected to be:

- 1- EFL teacher,
- 2- Teaching for preparatory stage students in Saudi higher education,
- 3- Should have been evaluated at least ONCE in your current university.

Your participation is voluntary and your responses are highly valued. Your decision to participate will provide you, as teachers experiencing the situation, with a chance to voice out your opinions in order to evolve the current practices. You might also like to request a synopsis of the study findings at a later stage.

There is no right or wrong answer and I value each individual opinion. Make sure that your personal details will not be disclosed to any party and all your responses will be treated with strict confidentiality and anonymity.

The questionnaire takes an average of 8 to 15 minutes and it consists of four sections:

- [Section 1:](#) Methods used for Teacher Evaluation,
- [Section 2:](#) Purposes for Teacher Evaluation,
- [Section 3:](#) Teacher Evaluation as Experienced by EFL Teachers in Saudi Universities,
- [Section 4:](#) Demographic Information.

The software used for this questionnaire is user-friendly and all questions can be easily answered using laptops, tablets, smartphones or any other devices that have an active access to the internet.

In case you are interested in receiving whatsapp, line, or viber link feel free to contact me on +447477337475. Or you can get in touch through Twitter @KhadijaAlamoudi.

Should you have further queries or comments, please do not hesitate to contact me at kaaa201@exeter.ac.uk

I Agree

#### Section 1: Methods used for Teacher Evaluation

1- At my university, the following methods are used to evaluate EFL teachers: (please select all that apply) \*

- Students rating
- Students achievements

Classroom observation (by Peers)

Classroom observation (by Admins)

Peer-evaluation

Self-evaluation

Teaching Portfolio

Others, please specify

**2- Do you think there can be other methods for EFL teacher evaluation than the ones mentioned above? \***

No

Yes, please specify

**3- Student rating is an effective method to evaluate EFL teachers. \***

Strongly Agree       Agree       Neutral       Disagree       Strongly Disagree

**4- Collecting students rating of teachers in mid-course would eliminate the grading bias effect (i.e. students rating of teacher will not be affected by their course grades).**

**\***

Strongly Agree       Agree       Neutral       Disagree       Strongly Disagree

**5- Peer-evaluation is an effective method to evaluate teachers. \***

Strongly Agree       Agree       Neutral       Disagree       Strongly Disagree

**6- Classroom observation done by peers is an effective method to evaluate teachers. \***

Strongly Agree       Agree       Neutral       Disagree       Strongly Disagree

**7- Self-evaluation is an effective method to evaluate EFL teachers. \***

Strongly Agree       Agree       Neutral       Disagree       Strongly Disagree

**8- When they do self-evaluation, teachers tend to rate themselves higher than they deserve. \***

Strongly Agree       Agree       Neutral       Disagree       Strongly Disagree

**9- Teaching-Portfolio (a documented statement of a faculty member's teaching responsibilities, philosophy, goals and accomplishments as a teacher) is an effective method to evaluate EFL teachers. \***

Strongly Agree       Agree       Neutral       Disagree       Strongly Disagree

**10- Constructing one's own portfolio can boost desire to improve weak areas. \***

Strongly Agree       Agree       Neutral       Disagree       Strongly Disagree

**Section 2: Purposes for Teacher Evaluation**

**11- Teacher evaluation should mostly focus on formative purposes, i.e. professional development. \***

Strongly Agree       Agree       Neutral       Disagree       Strongly Disagree

**12- Teacher evaluation should mostly focus on summative purposes, i.e. promotion, tenure, etc. \***

Strongly Agree       Agree       Neutral       Disagree       Strongly Disagree

**13- At my university, EFL teachers are evaluated for:**

**\***

Developmental reasons, i.e. training and workshops

Decision-making reasons, i.e. promotion, tenure, or contract renewal

Both developmental and decision-making reasons



Others, please specify

**14- Student ratings should be used for formative purposes only. \***  
 Strongly Agree     Agree     Neutral     Disagree     Strongly Disagree

**15- Peer evaluation should be used for formative purposes only. \***  
 Strongly Agree     Agree     Neutral     Disagree     Strongly Disagree

**16- Self-evaluation should be used for formative purposes only. \***  
 Strongly Agree     Agree     Neutral     Disagree     Strongly Disagree

**17- Teaching portfolios should be used for formative purposes only. \***  
 Strongly Agree     Agree     Neutral     Disagree     Strongly Disagree

**Section 3: Teacher Evaluation as Experienced by EFL Teachers in Saudi Universities**

**18- Students have the critical thinking skills that enables them to participate in staff evaluation. \***  
 Strongly Agree     Agree     Neutral     Disagree     Strongly Disagree

**19- Students tend to give high ranks for teachers who give students high marks. \***  
 Strongly Agree     Agree     Neutral     Disagree     Strongly Disagree

**20- The process used under the Teacher Evaluation System fosters a climate for instructional improvement. \***  
 Strongly Agree     Agree     Neutral     Disagree     Strongly Disagree

**21- The teacher evaluation system provides me with objective information about my teaching. \***  
 Strongly Agree     Agree     Neutral     Disagree     Strongly Disagree

**22- The teacher evaluation system enhances dialogue between teachers and evaluators about effective teaching. \***  
 Strongly Agree     Agree     Neutral     Disagree     Strongly Disagree

**23- The teacher evaluation system recognizes teachers' contributions to the academic institution as a whole (e.g. relationships with co-workers, professional development, and document completion) \***  
 Strongly Agree     Agree     Neutral     Disagree     Strongly Disagree

**24- The teacher evaluation system increases my reflection on choice of teaching strategies. \***  
 Strongly Agree     Agree     Neutral     Disagree     Strongly Disagree

**25- I regularly receive focused follow-up based on my evaluations. \***  
 Strongly Agree     Agree     Neutral     Disagree     Strongly Disagree

**26- I regularly receive instructional support based on my evaluations. \***  
 Strongly Agree     Agree     Neutral     Disagree     Strongly Disagree

**27- I focus my professional development efforts on activities that directly help me achieve the evaluation standards. \***  
 Strongly Agree     Agree     Neutral     Disagree     Strongly Disagree

**28- The feedback I receive from my evaluators is valuable. \***  
 Strongly Agree     Agree     Neutral     Disagree     Strongly Disagree

**29- The teacher evaluation process is helpful to my professional growth. \***  
 Strongly Agree     Agree     Neutral     Disagree     Strongly Disagree

**30- I am satisfied with the current evaluation system. \***  
 Strongly Agree     Agree     Neutral     Disagree     Strongly Disagree

**31- My evaluators have been properly trained to consistently evaluate my teaching. \***  
 Strongly Agree     Agree     Neutral     Disagree     Strongly Disagree

**32- The standards used in the teacher evaluation system are fair. \***  
 Strongly Agree     Agree     Neutral     Disagree     Strongly Disagree

33- The evaluation standards do a good job of defining effective teaching. \*

Strongly Agree       Agree       Neutral       Disagree       Strongly Disagree

34- Working towards improving my performance on the evaluation standards will help me to improve the quality of my teaching. \*

Strongly Agree       Agree       Neutral       Disagree       Strongly Disagree

35- The Teacher Evaluation Instrument includes clear explanations for each performance descriptor. \*

Strongly Agree       Agree       Neutral       Disagree       Strongly Disagree

36- How often do you prefer evaluation to be done (once, twice, four times a year, every semester, etc.)? \*

37- What are the strengths of the current Teacher Evaluation System? \*

38- What are the weaknesses of the current Teacher Evaluation System? \*

39- What are the changes that should be made to the Teacher Evaluation System? Explain how or why. \*

40- In two words, describe what Teacher Evaluation means to you. \*

41- Are there any other comments you wish to make about Teacher Evaluation? \*

**Section 4: Participant's Demographic Information**

42- Please specify the range of your age. \*

- 22-30 years
- 31-40 years
- 41-50 years
- 51-60 years
- 61-70 years
- 71 years or over

**43- Please specify your gender. \***

- Male
- Female

**44- Please specify your highest qualification. \***

- Bachelor
- Master
- Doctorate
- Others, please specify

**45- Please specify your job title. \***

- Teaching Assistant
- Language Instructor
- Lecturer
- Assistant Professor
- Associate Professor
- Professor
- Others, please specify

**46- Please specify for how long is your contract. \***

- One year
- Two years
- Three years
- Four years
- Five years
- Six years or more
- Others, please specify

**47- Please specify for how long you have been teaching EFL in higher education. \***

- Less than a year
- 1-3 years
- 3-6 years
- 7-10 years
- 11-15 years
- More than 15 years

**48- Please specify how many hours per week are you assigned to teach. \***

- 3 hours
- 6 hours

- 9 hours
- 12 hours
- 15 hours
- 18 hours
- Others, please specify

**49- Please specify which Saudi University you are teaching at. \***

- King Abdulaziz University
- King Fahd University of Petroleum and Minerals
- King Saud University
- Taif University
- Umm Al-Qura University
- Other, please specify.

**50- Please specify your nationality. \***

**51- I would like to conduct both focus group discussion and one-to-one interview to discuss teacher evaluation in more details, would you be interested in being a participant? \***

- yes, for focus group discussion
- yes, for one-to-one interview
- yes, for **both** of them
- no, **none** of them

**52- If you select any yes for the previous question, please provide the following details:**

**1- Your first and surname**

**2- Your email address**

**3- Your phone/mobile number**

for more information or support, please contact me at [kaaa201@exeter.ac.uk](mailto:kaaa201@exeter.ac.uk)

## Appendix 2: One-to-One Interview Guide

Title: EFL Teacher Evaluation at Saudi Universities  
Semi Structured Interview

Interview Number: ..... Day and Date: ....., .../.../2016. Venue: .....

Interviewee (Code): ..... Gender: M / F Highest Qualification: .....

How long have been working at this university: .....

I would like to thank you for your willingness to take part in this interview. Can I again confirm that the interview will remain completely anonymous and no record will be kept with your name on it?

I would like to confirm that teacher evaluation in general (and not classroom observation) is what I am interested in.

Major Themes to Explore	Main Question	Follow up Qs and prompts
Introduction	1- Please tell me briefly about Teacher Evaluation in your university?	Why and how teachers are being evaluated? Does it include: teaching, research, community service?
	2- How important is teacher evaluation process to you?	Why do you think so? Does it affect your salary or contract?
Teachers Readiness	3- Have you received any induction or written instructions on how evaluation works?	Have you ever asked for them?
	4- Do you know the evaluation criteria? Are they informed to you?	What are they? May I ask for a copy?
	5- How EFL teachers are prepared for teacher evaluation at your institute?	Do you think this could be sufficient? Why?
Challenges	6- What do you believe is the biggest challenge in making performance evaluation effective and meaningful to EFL teachers?	Why do you think so? <ul style="list-style-type: none"> <li>- Evaluator/observer</li> <li>- Timing</li> <li>- Pre/post-observation meetings</li> <li>- Feedback</li> </ul>
	7- Do you think there are other challenges?	What are they?
Evaluator	8- At your institute, who is the evaluator?	Is he/she the observer too?
	9- What do you think about the evaluator?	What kind of: <ul style="list-style-type: none"> <li>- Knowledge,</li> <li>- Qualifications,</li> <li>- Skills</li> </ul>

<b>Feedback</b>	10- Tell me about the quality of the feedback you receive after evaluation. (Timely, constructive, sufficient, useful, etc., )	
	11- Do you have the right to defend yourself against the evaluator comments?	Will the evaluation results will be modified accordingly?
<b>Consequences of evaluation</b>	12- Do the evaluation results motivate your professional development?	Why? Or why not?
	13- In what way do the evaluation results affect your instructional practices or the quality of your work experience?	
	14- Can evaluation affect your job security?	Why do you think so?
	15- Do you believe that the evaluation process affects your self-esteem?	If yes, in what way?
	16- What do "rewards and recognition" mean in your teacher evaluation system?	
	17- Do you believe that the evaluation process affects your sense of belonging to the institution?	If yes, how does it affect the sense of belonging? If no, why it does not?
<b>Teacher's role</b>	18- What role does the teacher play in the development of teacher evaluation practices?	Is it an effective role? Why? Or why not?
	19- Tell me about the teacher's voice in Teacher Evaluation at your institute.	Is it well heard by people in charge of Teacher Evaluation?
<b>Conclusion</b>	20- Thank you very much for helping me and giving me some of your time. Can I finally ask if you think there is any aspect of your experience with teacher evaluation that has not been covered in this interview?	

### Appendix 3: Focus Group Discussion Schedule

Title: EFL Teacher Evaluation at Saudi Universities

Focus Group Discussion

FG Number: ..... Day and Date: ....., .../.../2016. Venue: .....

Number of Participants: ..... Gender: M.... / F .....

Participants' Code: .....


I would like to thank you for your willingness to take part in this Focus Group Discussion. Can I again confirm that the interview will remain completely anonymous and no record will be kept with your name on it?

Major Themes to Explore	Main Question	Follow up Qs and prompts
Introduction	1- How often is the performance review at the University?	
	2- What do you think is the university key purpose of EFL teacher evaluation process?	Do you think that evaluation could be done for a better purpose?
Methods	3- What are the methods of evaluation used at your university?	Is research and community service included in teacher evaluation?
	4- What do you think of the methods that should be used to evaluate EFL teachers at a university level?	Anything more?
Challenges	5- What do you think about the discourse around teacher evaluation at your institute?	Written or spoken discourse? How about the other type of discourse?
	6- What do you believe is the biggest challenge in making performance evaluation meaningful to EFL teachers?	Why do you think so? - Evaluator/observer - Timing - Pre/post-observation meetings - Feedback
	7- Do you think there are other challenges?	What are they?
Suggestions for Solutions	8- Can you suggest some ways to overcome these challenges?	
	9- Have you ever offered or communicated these suggestions with people in charge?	Why? Or why not?

	10- How will these suggestions help making teacher evaluation meaningful?	
<b>Conclusion</b>	11- Thank you very much for helping me and giving me some of your time. Can I finally ask if you think there is any aspect of your experience with teacher evaluation that has not been covered in this interview?	



## Appendix 4: Ethical Approval Form

	Ref (for office use only) <div style="border: 1px solid black; width: 150px; height: 20px; margin: 5px auto;"></div>	
<b>COLLEGE OF SOCIAL SCIENCES AND INTERNATIONAL STUDIES</b>		
<p>When completing this form please remember that the purpose of the document is to clearly explain the ethical considerations of the research being undertaken. As a generic form it has been constructed to cover a wide-range of different projects so some sections may not seem relevant to you. Please include the information which addresses any ethical considerations for your particular project which will be needed by the SSIS Ethics Committee to approve your proposal.</p>		
<p>Guidance on all aspects of the SSIS Ethics application process can be found on the SSIS intranet:  <a href="https://intranet.exeter.ac.uk/socialsciences/staff/research/researchenvironmentandpolicies/ethics/">https://intranet.exeter.ac.uk/socialsciences/staff/research/researchenvironmentandpolicies/ethics/</a></p>		
<p>All staff and postdoctoral students within SSIS should use this form to apply for ethical approval and then send it to one of the following email addresses:</p>		
<p><a href="mailto:ssis-ethics@exeter.ac.uk">ssis-ethics@exeter.ac.uk</a> This email should be used by staff and postdoctoral students in Egenis, the Institute for Arab and Islamic Studies, Law, Politics, the Strategy &amp; Security Institute, and Sociology, Philosophy, Anthropology.</p>		
<p><a href="mailto:ssis-gseethics@exeter.ac.uk">ssis-gseethics@exeter.ac.uk</a> This email should be used by staff and postdoctoral students in the Graduate School of Education.</p>		
<b>Applicant details</b>		
Name	Khadija Abdullah Alamoudi	
Department	Education	
UoE email address	Kaaa201@exeter.ac.uk	
<b>Duration for which permission is required</b>		
<p>You should request approval for the entire period of your research activity. The start date should be at least one month from the date that you submit this form. Students should use the anticipated date of completion of their course as the end date of their work. Please note that <u>retrospective ethical approval will never be given.</u></p>		
Start date:10/10/2015	End date:30/12/2018	Date submitted:09/10/2015
<b>Students only</b>		
<p>All students must discuss their research intentions with their supervisor/tutor prior to submitting an application for ethical approval. The discussion may be face to face or via email.</p>		
<p>Prior to submitting your application in its final form to the SSIS Ethics Committee it should be approved by your first and second supervisor / dissertation supervisor/tutor. You should submit evidence of their approval with your application, e.g. a copy of their email approval.</p>		
Student number	630057281	
Programme of study	Doctor of Philosophy (PhD)	
Name of Supervisor(s)/tutors or Dissertation Tutor	Prof Troudi, Salah Dr Abdollahzadeh, Esmaeel	
Have you attended any ethics training that is available to students?	Yes, I have taken part in ethics training at the University of Exeter Entitled: Understanding Research Ethics <a href="http://as.exeter.ac.uk/rdp/postgraduateresearchers">http://as.exeter.ac.uk/rdp/postgraduateresearchers</a>	
<p>c:\Users\kaaa201\AppData\Local\Microsoft\Windows\Temporary Internet Files\Content.IE5\UGEHLDXJ\SSIS_Ethics_Application_Khadija_Alamoudi.docx</p>		
Page 1 of 4		

	If yes, please give the date of the training:02/02/2014
<b>Certification for all submissions</b>	
<p>I hereby certify that I will abide by the details given in this application and that I undertake in my research to respect the dignity and privacy of those participating in this research. I confirm that if my research should change radically I will complete a further ethics proposal form.</p> <p>Khadija Abdullah Alamoudi</p> <p>Double click this box to confirm certification <input checked="" type="checkbox"/></p> <p><i>Submission of this ethics proposal form confirms your acceptance of the above.</i></p>	
<b>TITLE OF YOUR PROJECT</b>	
EFL Teachers Perceptions on Teacher Evaluation at Saudi Universities	
<b>ETHICAL REVIEW BY AN EXTERNAL COMMITTEE</b>	
No, my research is not funded by, or doesn't use data from, either the NHS or Ministry of Defence.	
<b>MENTAL CAPACITY ACT 2005</b>	
No, my project does not involve participants aged 16 or over who are unable to give informed consent (e.g. people with learning disabilities)	
<b>SYNOPSIS OF THE RESEARCH PROJECT</b>	
<i>As a guide - 750 words.</i>	
<p>Teacher evaluation has been the focus of much research recently; yet, due attention has not been given to EFL instructor (Tovar, 2011). In an attempt to fill in this gap in the literature, my research aims at exploring the challenges EFL teachers might face along with their needs under the current teacher evaluation system in some Saudi Universities. Taking the International Standards and Models for Teacher Evaluation into account (e.g., Value-Added Model, Peer Assistance and Review Program, and Standard-Based Evaluation), special attention will be directed to the methods applied to evaluate EFL teachers and the purposes behind their evaluation in Saudi higher education system. An exploratory mixed methods design of quantitative followed by qualitative methods of data collection and analysis will be applied in order to investigate EFL teacher evaluation. Questionnaires, focus groups and semi-structured interviews following the purposive sampling approach will be used to form an informational corpus of data that will be analysed. Implications will be eventually made as an attempt to improve the current system.</p> <p>Research Questions are:</p> <ol style="list-style-type: none"> <li>1. What are the methods used for EFL teacher evaluation in Saudi universities?</li> <li>2. What are the purposes for EFL teacher evaluation in Saudi universities?</li> <li>3. What are the challenges of the current evaluation system in Saudi universities as experienced by EFL teachers in Saudi universities?</li> <li>4. How do EFL teachers cope with these challenges?</li> </ol>	
<b>INTERNATIONAL RESEARCH</b>	
<p>My research will take place in the Kingdom of Saudi Arabia. Research protocols there include seeking permission to collect data from gatekeepers, which may include people like deans, vice-deans, heads of research or postgraduate units, etc. A written permission is required before going to the field and seeking participants' responses. At King Abdulaziz University, for instance, a form entitled "Data Collection Permission Form" is needed to be issued prior to commencing any fieldwork. I am totally aware about the required procedures and I will adhere to them.</p>	
SSIS Ethics Application form_template_v10	Page 2 of 4

The following sections require an assessment of possible ethical consideration in your research project. If particular sections do not seem relevant to your project please indicate this and clarify why.

#### RESEARCH METHODS

Methods to collect the data includes:

- 1- Questionnaire: which will be sent online to the entire population where all EFL teachers at the selected Saudi Universities will have equal opportunity to participate. Mainly close-ended questions will be used and few open-ended questions will be included. The questionnaire link will be sent to gatekeepers who will in return send it to the population. Data will be analysed using proper software (SPSS in mind). Results will help in constructing schedules and guidelines for interviews.
- 2- Focus Group Discussion: will be conducted with those who show interest in the questionnaire and who fulfil the research requirements. Audi or videotape record will be used depend upon the participant consent. Collected materials will be analysed using proper software (Invivo in mind) and results will help in guiding on-to-one semi-structured interviews.
- 3- Semi-structured Interviews: will be carried out with participants who are willing to participate and show interest in the topic. After signing consent forms, interviews will be scheduled to the participants' convenience. Interviews will be audio recorded and transcribed later on for analysis using proper software (Invivo in mind).

#### PARTICIPANTS

Participant would mainly those teachers or staff who are involved in teaching English as a foreign language at Saudi universities. The sample may involve: teaching assistants, lecturers, language instructors, assistant professors, associate professors, or professors. All participants should have been evaluated at least once. For the quantitative phase the number is expected to be 200-300 or a bit more depends on the access I might get. As for the qualitative the number is expected to be 20 for the FGD (4 groups, 5 people each) and around 6-8 for the one-to-one interview. I might give a small souvenir (e.g. bookmark) to FGD and on-to-one interview participants for their timing.

#### THE VOLUNTARY NATURE OF PARTICIPATION

For the questionnaire, participants will be recruited through an email that will be sent to them by admins of their institutions. There is an information page at the beginning of the questionnaire that gives brief about the research for those who are interested in participating voluntarily, freely and willingly. They will be told they have the right to withdraw from the research at any time. Anonymity and confidentiality will be also maintained by utilising a consent form that will be signed by all participants to assures the following: a) confidentiality; b) anonymity; c) information about the project and d) the right to withdraw at any time without disadvantage to the participant. An example of the consent form is submitted with this proposal.

#### SPECIAL ARRANGEMENTS

As far as I am aware, no special arrangement is needed.

#### THE INFORMED NATURE OF PARTICIPATION

All participants will be given sufficient information about the research project and about the nature of their participation. This can be part of the introductory phase along with the consent

form.

#### ASSESSMENT OF POSSIBLE HARM

There would be no potential harm to the participant or the researcher. The experience of the participants and their opinions will be recorded and analysed anonymously. The researcher will try to avoid any disclosure of distinguishing features of participants or incidents. If any harm comes to the attention of the researcher, serious cautions will take place then, adjustment or cancellation of particular part would be also considered.

#### DATA PROTECTION AND STORAGE

The information participants provide will be used for research purposes and their personal data will be processed in accordance with current data protection legislation and the University's notification lodged at the Information Commissioner's Office. Participants' personal data will be treated in the strictest confidence (by keeping a separate list of names and addresses linked to raw data by a numerical key) and will not be disclosed to any unauthorised third parties. For secure storage of data, all files will be password protected stored on University U- Drive then all files will be destroyed after five years. The results of the research will be published in anonymised form.

#### DECLARATION OF INTERESTS

The Saudi Ministry of Higher Education funds the research for educational purposes. At this point, I am not considering any publication; however, the results might be published by King Abdulaziz University or University of Exeter.

#### USER ENGAGEMENT AND FEEDBACK

Some institutes, and rarely participants, may ask for feedback on their input. If this is the case, general feedback will be provided.

#### INFORMATION SHEET

Attached

#### CONSENT FORM

Attached

#### SUBMISSION PROCEDURE

Staff and students should follow the procedure below.

In particular, students should discuss their application with their supervisor(s) / dissertation tutor / tutor and gain their approval prior to submission. Students should submit evidence of approval with their application, e.g. a copy of the supervisors email approval.

This application form and examples of your consent form, information sheet and translations of any documents which are not written in English should be submitted by email to the SSIS Ethics Secretary via one of the following email addresses:

[ssis-ethics@exeter.ac.uk](mailto:ssis-ethics@exeter.ac.uk) This email should be used by staff and postdoctoral students in Egenis, the Institute for Arab and Islamic Studies, Law, Politics, the Strategy & Security Institute, and Sociology, Philosophy, Anthropology.

## Appendix 5: Research Information and Consent Sheet



### INFORMATION SHEET AND CONSENT FORM FOR RESEARCH

#### Title of Research Project

*EFL Teachers Perceptions on Teacher Evaluation at Saudi Universities*

#### Details of Project

*Dear colleagues,*

*My name is Khadija Alamoudi and I am a lecturer at the English Language Institute, King Abdulaziz University, Jeddah. Currently, I am a doctorate student at the University of Exeter and my research focuses on EFL (English as a Foreign Language) teacher evaluation in Saudi universities.*

*Participants are expected to be EFL teachers in higher Saudi education who should have been evaluated at least once in their current university. Your personal details will not be disclosed to any party and all your responses will be treated with strict confidentiality and anonymity. Participation is voluntarily and you can withdraw at any stage for a good reason. Results of this project will be used for educational purposes only. Thank you in advance for your contribution.*

#### Contact Details

*Should you have any queries or comments, please do not hesitate to contact me,*

*Name: Khadija Alamoudi*

*Postal address: Staff House, University of Exeter St Luke's Campus, Heavitree Road, Exeter, EX1 2LU, UK*

*Telephone: 00 44 7477337475*

*Email: kaaa201@exeter.ac.uk*

*If you have concerns/questions about the research you would like to discuss with someone else at the University, please contact: Prof. Salah Troudi*

*Email: S.Troudi@exeter.ac.uk*

#### Confidentiality

*Questionnaires and Interview tapes with transcripts will be held in confidence. They will not be used other than for the purposes described above and third parties will not be allowed access to them (except as may be required by the law). However, if you request it, you will be supplied with a copy of your interview transcript so that you can comment on and edit it as you see fit (please give your email below so that I am able to contact you at a later date). Your data will be held in accordance with the Data Protection Act.*

#### Data Protection Notice

*The information you provide will be used for research purposes and your personal data will be processed in accordance with current data protection legislation and the University's notification lodged at the Information Commissioner's Office. Your personal data will be treated in the strictest confidence (by keeping a separate list of names and addresses linked to raw data by a numerical key) and will not be disclosed to any unauthorised third parties. For secure storage of data, all files will be password protected stored on University U- Drive, then all files will be destroyed after five years. The research is funded by the Ministry of Education and results of the research will be published in anonymised form.*



**INFORMATION SHEET AND CONSENT  
FORM FOR RESEARCH**

**Anonymity**

*Questionnaire and Interview data will be held and used on an anonymous basis, with no mention of your name.*

**Consent**

I have been fully informed about the aims and purposes of the project.

I understand that:

- there is no compulsion for me to participate in this research project and, if I do choose to participate, I may withdraw at any stage for a good reason;
- I have the right to refuse permission for the publication of any information about me;
- any information which I give will be used solely for the purposes of this research project, which may include publications or academic conference or seminar presentations;
- If applicable, the information, which I give, may be shared between any of the other researcher(s) participating in this project in an anonymised form;
- all information I give will be treated as confidential;
- the researcher(s) will make every effort to preserve my anonymity.

.....  
(Signature of participant)

.....  
(Date)

.....  
(Printed name of participant)

.....  
(Email address of participant if they have requested to view a copy of the interview transcript)

.....  
(Signature of researcher)

**Khadija Abdullah Alamoudi**  
(Printed name of researcher)

One copy of this form will be kept by the participant; a second copy will be kept by the researcher(s). Your contact details are kept separately from your interview data.

## Appendix 6: Data Analysis / Questionnaire: Categorical Variables

FREQUENCIES VARIABLES=Q49 Q43 Q50 Q45 Q51

/STATISTICS=MINIMUM MAXIMUM

/BARCHART FREQ

/ORDER=ANALYSIS.

### Frequencies

Notes		
Output Created		22-MAR-2016 12:24:16
Comments		
Input	Data	(Completed_Checked).sav
	Active Dataset	DataSet1
	File Label	Generated by Java SPSS Writer v1.83 <a href="http://spss.pmStation.com">http://spss.pmStation.com</a>
	Filter	<none>
	Weight	<none>
	Split File	<none>
	N of Rows in Working Data File	249
Missing Value Handling	Definition of Missing	User-defined missing values are treated as missing.
	Cases Used	Statistics are based on all cases with valid data.
Syntax		FREQUENCIES VARIABLES=Q49 Q43 Q50 Q45 Q51 /STATISTICS=MINIMUM MAXIMUM /BARCHART FREQ /ORDER=ANALYSIS.
Resources	Processor Time	00:00:00.73
	Elapsed Time	00:00:00.70

**Statistics**

		At Which Saudi University	Gender	Nationality	Job Title	Interested in Interviews and FGD
N	Valid	248	249	249	249	249
	Missing	1	0	0	0	0
	Minimum	1	1	1	1	1
	Maximum	5	2	21	7	4

**Frequency Table**

**At Which Saudi University**

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	King Abdulaziz University	99	39.8	39.9	39.9
	King Fahd University of Petroleum and Minerals	11	4.4	4.4	44.4
	King Saud University	35	14.1	14.1	58.5
	Taif University	53	21.3	21.4	79.8
	Umm AlQura University	50	20.1	20.2	100.0
	Total	248	99.6	100.0	
Missing	0	1	.4		
Total		249	100.0		

**Gender**

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Male	111	44.6	44.6	44.6
	Female	138	55.4	55.4	100.0
	Total	249	100.0	100.0	



		<b>Nationality</b>			
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	American	16	6.4	6.4	6.4
	Algerian	1	.4	.4	6.8
	British	21	8.4	8.4	15.3
	Canadian	5	2.0	2.0	17.3
	Caucasian	1	.4	.4	17.7
	Egyptian	17	6.8	6.8	24.5
	Filipino	1	.4	.4	24.9
	Indian	14	5.6	5.6	30.5
	Iranian	1	.4	.4	30.9
	Irish	1	.4	.4	31.3
	Jordanian	22	8.8	8.8	40.2
	New Zealand	1	.4	.4	40.6
	Pakistani	38	15.3	15.3	55.8
	Saudi	65	26.1	26.1	81.9
	South African	5	2.0	2.0	83.9
	Sudanese	11	4.4	4.4	88.4
	Syrian	4	1.6	1.6	90.0
	Tunisan	10	4.0	4.0	94.0
	Yemeni	2	.8	.8	94.8
	Zimbabwean	1	.4	.4	95.2
	Not Specified	12	4.8	4.8	100.0
	Total	249	100.0	100.0	

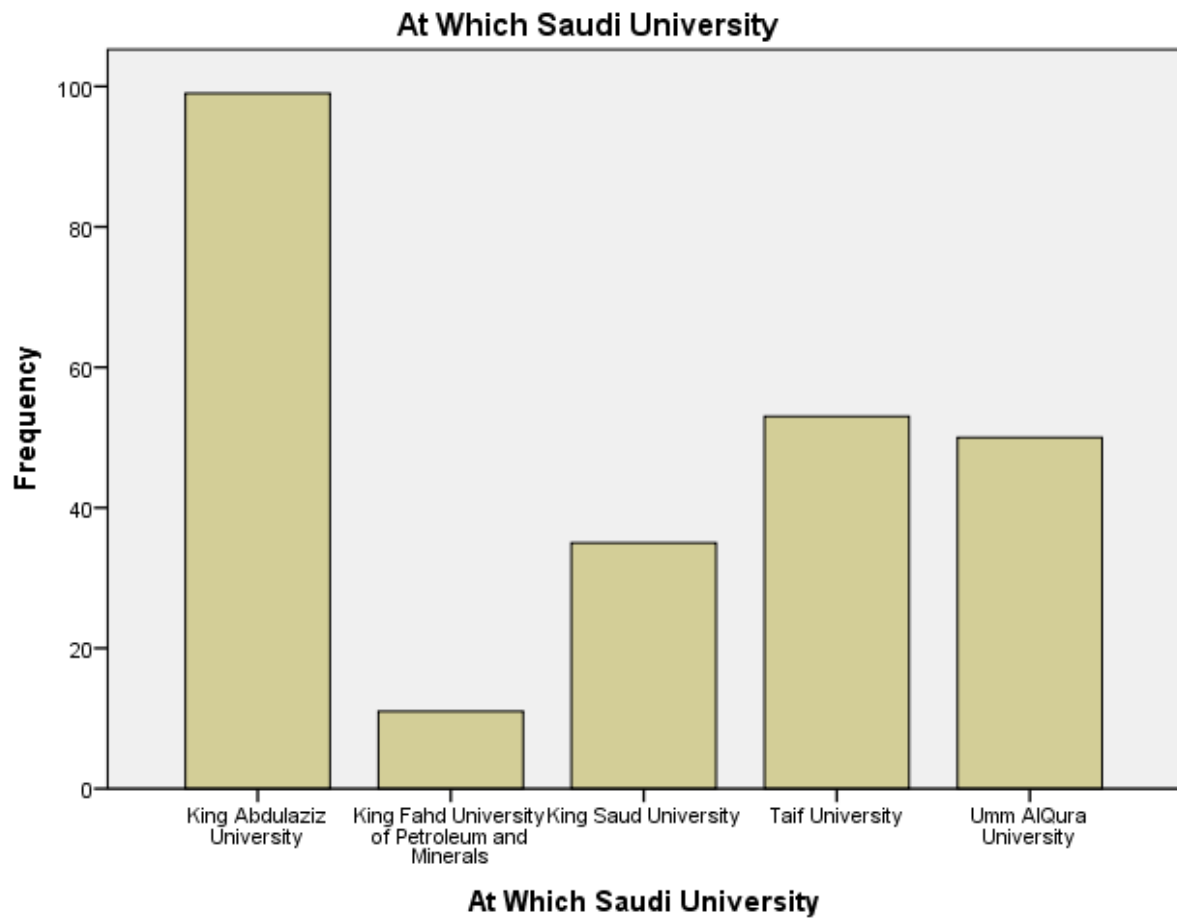
**Job Title**

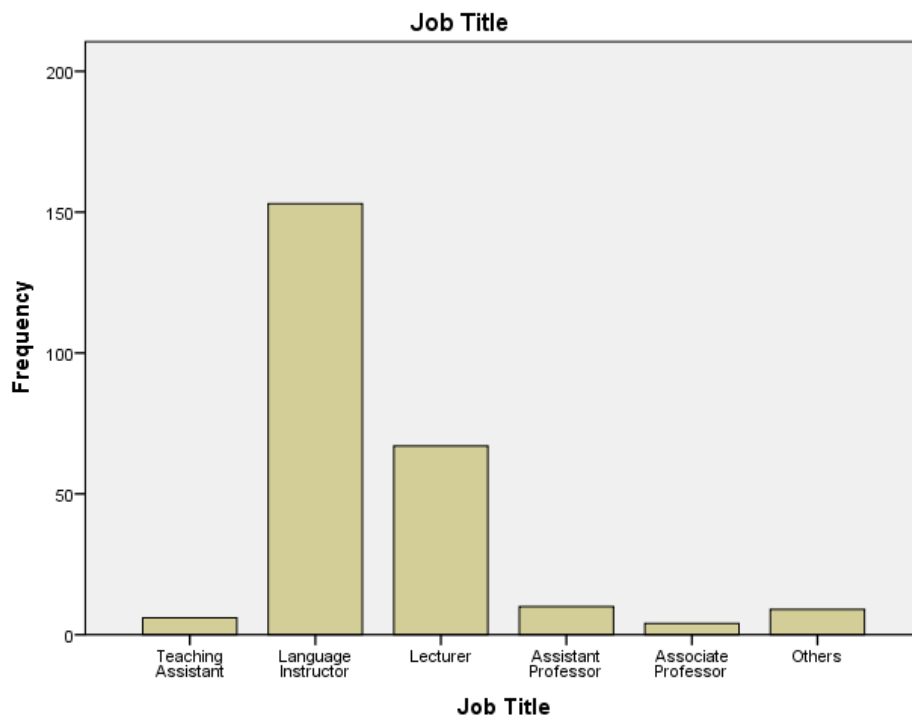
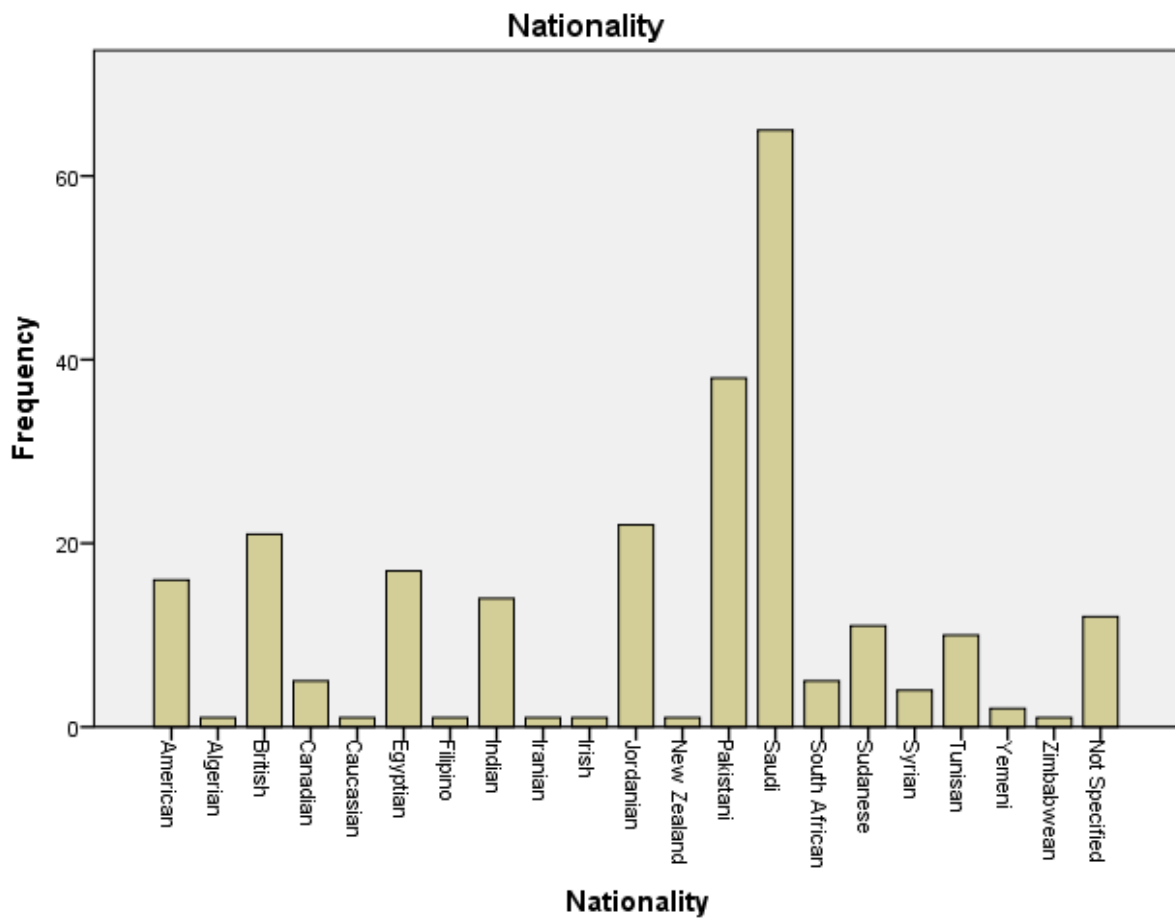
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Teaching Assistant	6	2.4	2.4	2.4
	Language Instructor	153	61.4	61.4	63.9
	Lecturer	67	26.9	26.9	90.8
	Assistant Professor	10	4.0	4.0	94.8
	Associate Professor	4	1.6	1.6	96.4
	Others	9	3.6	3.6	100.0
	Total	249	100.0	100.0	

**Interested in Interviews and FGD**

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	yes for focus group discussion	14	5.6	5.6	5.6
	yes for onetoone interview	25	10.0	10.0	15.7
	yes for both of them	43	17.3	17.3	32.9
	no none of them	167	67.1	67.1	100.0
	Total	249	100.0	100.0	

### Bar Chart





## Appendix 7: Data Analysis / Questionnaire: Continuous Variables

FREQUENCIES VARIABLES=Q42 Q44 Q46 Q47 Q48

/STATISTICS=MINIMUM MAXIMUM

/BARCHART FREQ

/ORDER=ANALYSIS.

### Frequencies

#### Notes

Output Created	22-MAR-2016 12:35:01	
Comments		
Input	Data	(Completed_Checked).sav
	Active Dataset	DataSet1
	File Label	Generated by Java SPSS Writer v1.83 <a href="http://spss.pmStation.com">http://spss.pmStation.com</a>
	Filter	<none>
	Weight	<none>
	Split File	<none>
	N of Rows in Working Data File	249
Missing Value Handling	Definition of Missing	User-defined missing values are treated as missing.
	Cases Used	Statistics are based on all cases with valid data.

Syntax		FREQUENCIES VARIABLES=Q42 Q44 Q46 Q47 Q48  /STATISTICS=MINIMUM MAXIMUM  /BARCHART FREQ  /ORDER=ANALYSIS.	
Resources	Processor Time		00:00:00.69
	Elapsed Time		00:00:00.70

**Statistics**

		Range of Age	Highest Qualificatio n	How Long Is the Contract	For How Long Teaching EFL in Higher Education	Teaching Hours per Week
N	Valid	249	249	249	249	249
	Missing	0	0	0	0	0
	Minimum	1	1	1	1	1
	Maximum	5	4	7	6	7

## Frequency Table

### Range of Age

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	22-30 years	33	13.3	13.3	13.3
	31-40 years	101	40.6	40.6	53.8
	41-50 years	82	32.9	32.9	86.7
	51-60 years	29	11.6	11.6	98.4
	61-70 years	4	1.6	1.6	100.0
	Total	249	100.0	100.0	

### Highest Qualification

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Bachelor	55	22.1	22.1	22.1
	Master	157	63.1	63.1	85.1
	Doctorate	20	8.0	8.0	93.2
	Others	17	6.8	6.8	100.0
	Total	249	100.0	100.0	

**How Long Is the Contract**

	Frequency	Percent	Valid Percent	Cumulative Percent
Valid One year	117	47.0	47.0	47.0
Two years	13	5.2	5.2	52.2
Three years	8	3.2	3.2	55.4
Four years	10	4.0	4.0	59.4
Five years	8	3.2	3.2	62.7
Six years or more	39	15.7	15.7	78.3
Others	54	21.7	21.7	100.0
Total	249	100.0	100.0	

**For How Long Teaching EFL in Higher Education**

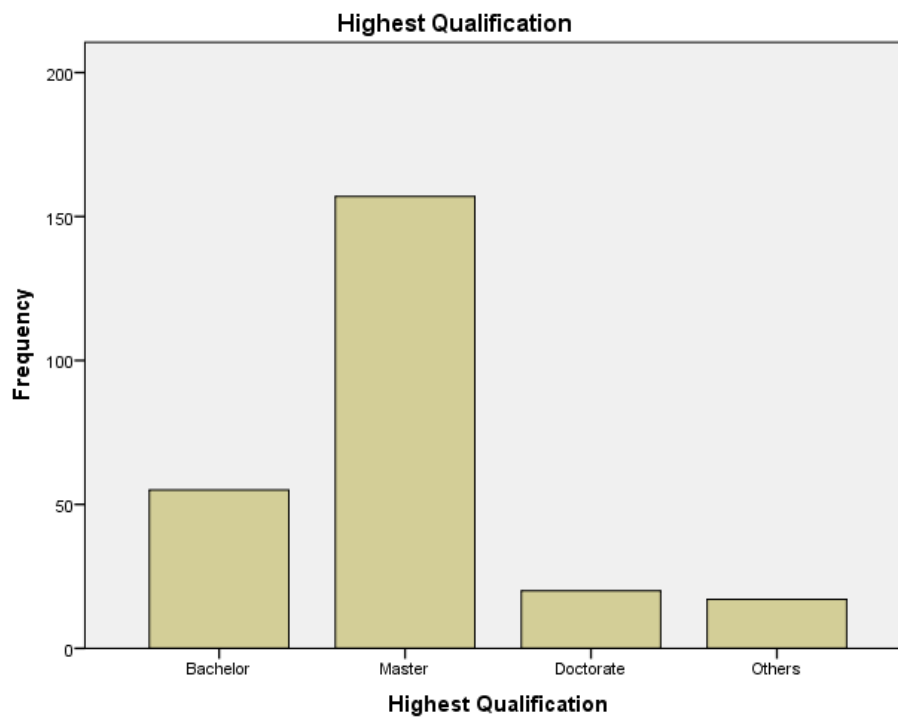
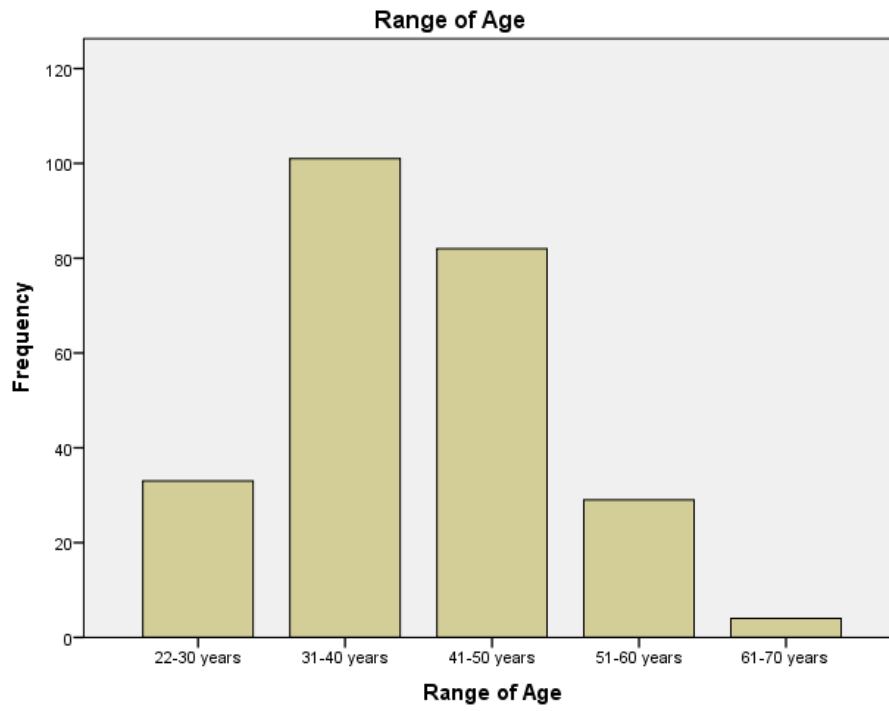
	Frequency	Percent	Valid Percent	Cumulative Percent
Valid Less than a year	6	2.4	2.4	2.4
1-3 years	34	13.7	13.7	16.1
3-6 years	66	26.5	26.5	42.6
7-10 years	66	26.5	26.5	69.1
11-15 years	34	13.7	13.7	82.7
More than 15 years	43	17.3	17.3	100.0
Total	249	100.0	100.0	

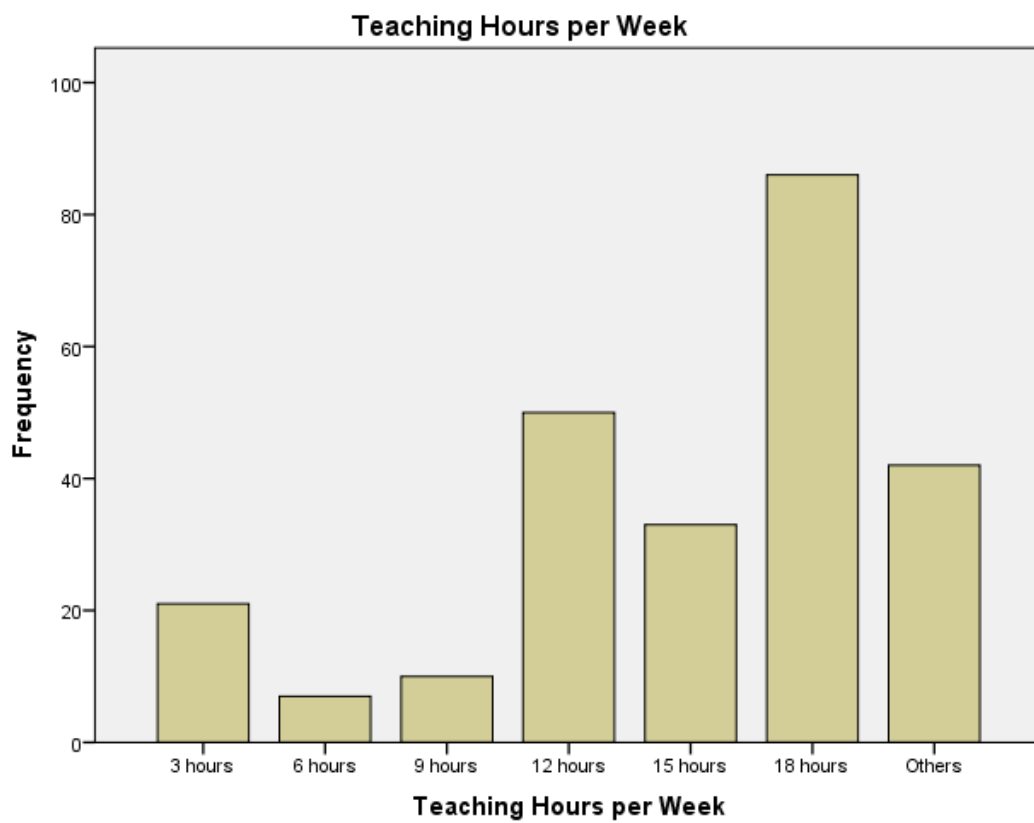
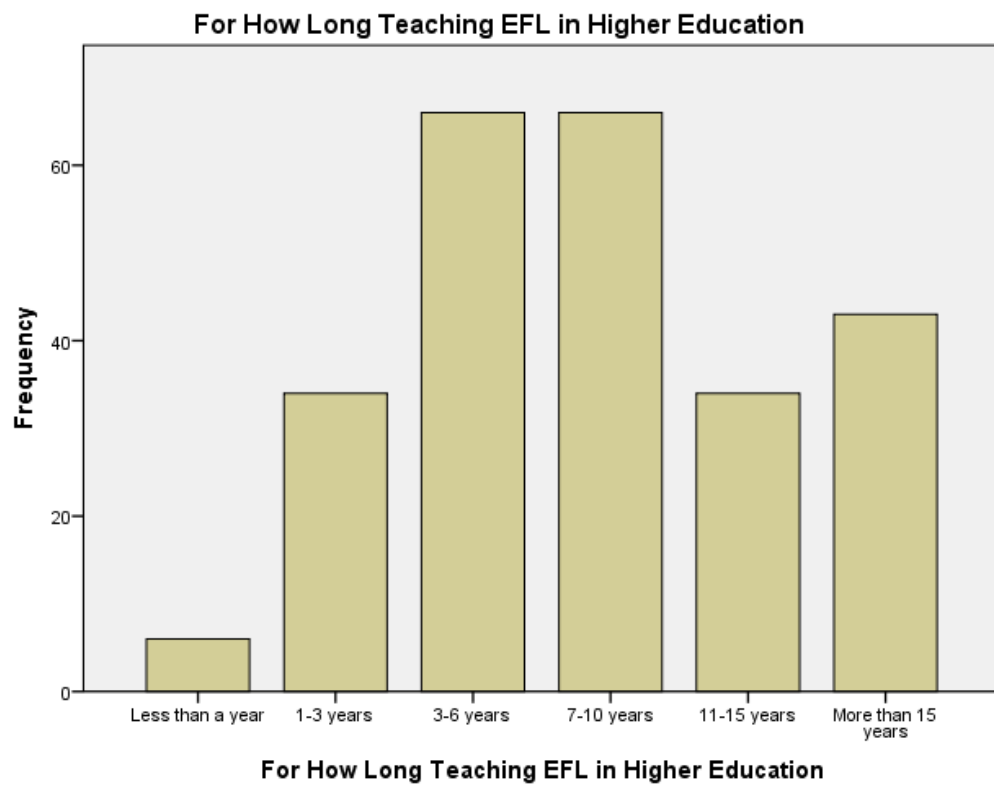


**Teaching Hours per Week**

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	3 hours	21	8.4	8.4	8.4
	6 hours	7	2.8	2.8	11.2
	9 hours	10	4.0	4.0	15.3
	12 hours	50	20.1	20.1	35.3
	15 hours	33	13.3	13.3	48.6
	18 hours	86	34.5	34.5	83.1
	Others	42	16.9	16.9	100.0
	Total	249	100.0	100.0	

### Bar Chart





## Appendix 8: Data Analysis / Questionnaire: Frequencies and Frequency Tables

FREQUENCIES VARIABLES=Q3 Q4 Q5 Q6 Q7 Q8 Q9 Q10 Q11 Q12 Q13 Q14  
Q15 Q16 Q17 Q18 Q19 Q20 Q21 Q22 Q23

Q24 Q25 Q26 Q27 Q28 Q29 Q30 Q31 Q32 Q33 Q34 Q35

/ORDER=ANALYSIS.

### Frequencies

#### Notes

Output Created		05-DEC-2017 15:39:28
Comments		
Input	Data	(Completed_Checked).sav
	Active Dataset	DataSet1
	File Label	Generated by Java SPSS Writer v1.83 <a href="http://spss.pmStation.com">http://spss.pmStation.com</a>
	Filter	<none>
	Weight	<none>
	Split File	<none>
	N of Rows in Working Data File	249

Missing Value Handling	Definition of Missing	User-defined missing values are treated as missing.
	Cases Used	Statistics are based on all cases with valid data.
Syntax		<p>FREQUENCIES</p> <p>VARIABLES=Q3 Q4 Q5 Q6 Q7 Q8 Q9 Q10 Q11 Q12 Q13 Q14 Q15 Q16 Q17 Q18 Q19 Q20 Q21 Q22 Q23</p> <p>Q24 Q25 Q26 Q27 Q28 Q29 Q30 Q31 Q32 Q33 Q34 Q35</p> <p>/ORDER=ANALYSIS.</p>
Resources	Processor Time	00:00:00.02
	Elapsed Time	00:00:00.01

### Statistics

		3-Student rating is an effective method to evaluate EFL teachers.	4-Collecting students rating of teachers in mid-course would eliminate the grading bias effect.	5-Peer-evaluation is an effective method to evaluate teachers.	6-Classroom observation done by peers is an effective method to evaluate teachers.	7-Self-evaluation is an effective method to evaluate EFL teachers.
N	Valid	249	249	249	249	249
	Missing	0	0	0	0	0

**Statistics**

		8-When they do self-evaluation, teachers tend to rate themselves higher than they deserve .	9-Teaching-Portfolio is an effective method to evaluate EFL teachers.	10-Constructing one's own portfolio can boost desire to improve weak areas.	11-Teacher evaluation should mostly focus on formative purposes, i.e. professional development.	12-Teacher evaluation should mostly focus on summative purposes, i.e. promotion, tenure, etc.
N	Valid	249	249	249	249	249
	Missing	0	0	0	0	0

**Statistics**

		13-At my university, EFL teachers are evaluated for:	14-Student ratings should be used for formative purposes only.	15-Peer evaluation should be used for formative purposes only.	16-Self-evaluation should be used for formative purposes only.	17-Teaching portfolios should be used for formative purposes only.
N	Valid	249	249	249	249	249
	Missing	0	0	0	0	0

**Statistics**

		18-Students have the critical thinking skills that enables them to participate in staff evaluation.	19-Students tend to give high ranks for teachers who give students high marks.	20-The process used under the Teacher Evaluation System fosters a climate for instructional improvement.	21-The teacher evaluation system provides me with objective information about my teaching.	22-The teacher evaluation system enhances dialogue between teachers and evaluators about effective teaching.
N	Valid	249	249	249	249	249
	Missing	0	0	0	0	0

**Statistics**

		23-The teacher evaluation system recognizes teachers' contributions to the academic institution as a whole.	24-The teacher evaluation system increases my reflection on choice of teaching strategies.	25-I regularly receive focused follow-up based on my evaluations.	26-I regularly receive instructional support based on my evaluations.	27-I focus my professional development efforts on activities that directly help me achieve the evaluation standards.
N	Valid	249	249	249	249	249
	Missing	0	0	0	0	0

**Statistics**

		28-The feedback I receive from my evaluators is valuable.	29-The teacher evaluation process is helpful to my professional growth.	30-I am satisfied with the current evaluation system.	31-My evaluators have been properly trained to consistently evaluate my teaching.	32-The standards used in the teacher evaluation system are fair.
N	Valid	249	249	249	249	249
	Missing	0	0	0	0	0

**Statistics**

		33-The evaluation standards do a good job of defining effective teaching.	34-Working towards improving my performance on the evaluation standards will help me to improve the quality of my teaching.	35-The Teacher Evaluation Instrument includes clear explanations for each performance descriptor.
N	Valid	249	249	249
	Missing	0	0	0



## Frequency Table

### 3-Student rating is an effective method to evaluate EFL teachers.

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Strongly Agree	33	13.3	13.3	13.3
	Agree	87	34.9	34.9	48.2
	Neutral	65	26.1	26.1	74.3
	Disagree	43	17.3	17.3	91.6
	Strongly Disagree	21	8.4	8.4	100.0
	Total	249	100.0	100.0	

### 4-Collecting students rating of teachers in mid-course would eliminate the grading bias effect.

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Strongly Agree	53	21.3	21.3	21.3
	Agree	97	39.0	39.0	60.2
	Neutral	63	25.3	25.3	85.5
	Disagree	29	11.6	11.6	97.2
	Strongly Disagree	7	2.8	2.8	100.0
	Total	249	100.0	100.0	

**5-Peer-evaluation is an effective method to evaluate teachers.**

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Strongly Agree	28	11.2	11.2	11.2
	Agree	115	46.2	46.2	57.4
	Neutral	57	22.9	22.9	80.3
	Disagree	41	16.5	16.5	96.8
	Strongly Disagree	8	3.2	3.2	100.0
	Total	249	100.0	100.0	

**6-Classroom observation done by peers is an effective method to evaluate teachers.**

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Strongly Agree	27	10.8	10.8	10.8
	Agree	113	45.4	45.4	56.2
	Neutral	55	22.1	22.1	78.3
	Disagree	43	17.3	17.3	95.6
	Strongly Disagree	11	4.4	4.4	100.0
	Total	249	100.0	100.0	

**7-Self-evaluation is an effective method to evaluate EFL teachers.**

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Strongly Agree	57	22.9	22.9	22.9
	Agree	120	48.2	48.2	71.1
	Neutral	44	17.7	17.7	88.8
	Disagree	24	9.6	9.6	98.4
	Strongly Disagree	4	1.6	1.6	100.0
	Total	249	100.0	100.0	

**8-When they do self-evaluation, teachers tend to rate themselves higher than they deserve .**

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Strongly Agree	17	6.8	6.8	6.8
	Agree	102	41.0	41.0	47.8
	Neutral	72	28.9	28.9	76.7
	Disagree	54	21.7	21.7	98.4
	Strongly Disagree	4	1.6	1.6	100.0
	Total	249	100.0	100.0	

**9-Teaching-Portfolio is an effective method to evaluate EFL teachers.**

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Strongly Agree	36	14.5	14.5	14.5
	Agree	110	44.2	44.2	58.6
	Neutral	48	19.3	19.3	77.9
	Disagree	39	15.7	15.7	93.6
	Strongly Disagree	16	6.4	6.4	100.0
	Total	249	100.0	100.0	

**10-Constructing one's own portfolio can boost desire to improve weak areas.**

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Strongly Agree	44	17.7	17.7	17.7
	Agree	143	57.4	57.4	75.1
	Neutral	39	15.7	15.7	90.8
	Disagree	15	6.0	6.0	96.8
	Strongly Disagree	8	3.2	3.2	100.0
	Total	249	100.0	100.0	

**11-Teacher evaluation should mostly focus on formative purposes, i.e. professional development.**

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Strongly Agree	68	27.3	27.3	27.3
	Agree	132	53.0	53.0	80.3
	Neutral	22	8.8	8.8	89.2
	Disagree	25	10.0	10.0	99.2
	Strongly Disagree	2	.8	.8	100.0
	Total	249	100.0	100.0	

**12-Teacher evaluation should mostly focus on summative purposes, i.e. promotion, tenure, etc.**

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Strongly Agree	20	8.0	8.0	8.0
	Agree	81	32.5	32.5	40.6
	Neutral	64	25.7	25.7	66.3
	Disagree	79	31.7	31.7	98.0
	Strongly Disagree	5	2.0	2.0	100.0
	Total	249	100.0	100.0	

**13-At my university, EFL teachers are evaluated for:**

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Developmental reasons ie training and workshops	55	22.1	22.1	22.1
	Decisionmaking reasons ie promotion tenure or contract renewal	46	18.5	18.5	40.6
	Both developmental and decisionmaking reasons	137	55.0	55.0	95.6
	Others please specify	11	4.4	4.4	100.0
	Total	249	100.0	100.0	

**14-Student ratings should be used for formative purposes only.**

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Strongly Agree	37	14.9	14.9	14.9
	Agree	114	45.8	45.8	60.6
	Neutral	45	18.1	18.1	78.7
	Disagree	45	18.1	18.1	96.8
	Strongly Disagree	8	3.2	3.2	100.0
	Total	249	100.0	100.0	

**15-Peer evaluation should be used for formative purposes only.**

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Strongly Agree	39	15.7	15.7	15.7
	Agree	123	49.4	49.4	65.1
	Neutral	52	20.9	20.9	85.9
	Disagree	30	12.0	12.0	98.0
	Strongly Disagree	5	2.0	2.0	100.0
	Total	249	100.0	100.0	

**16-Self-evaluation should be used for formative purposes only.**

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Strongly Agree	38	15.3	15.3	15.3
	Agree	115	46.2	46.2	61.4
	Neutral	56	22.5	22.5	83.9
	Disagree	35	14.1	14.1	98.0
	Strongly Disagree	5	2.0	2.0	100.0
	Total	249	100.0	100.0	

**17-Teaching portfolios should be used for formative purposes only.**

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Strongly Agree	32	12.9	12.9	12.9
	Agree	117	47.0	47.0	59.8
	Neutral	60	24.1	24.1	83.9
	Disagree	37	14.9	14.9	98.8
	Strongly Disagree	3	1.2	1.2	100.0
	Total	249	100.0	100.0	

**18-Students have the critical thinking skills that enables them to participate in staff evaluation.**

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Strongly Agree	13	5.2	5.2	5.2
	Agree	74	29.7	29.7	34.9
	Neutral	59	23.7	23.7	58.6
	Disagree	68	27.3	27.3	85.9
	Strongly Disagree	35	14.1	14.1	100.0
	Total	249	100.0	100.0	



**19-Students tend to give high ranks for teachers who give students high marks.**

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Strongly Agree	64	25.7	25.7	25.7
	Agree	98	39.4	39.4	65.1
	Neutral	48	19.3	19.3	84.3
	Disagree	37	14.9	14.9	99.2
	Strongly Disagree	2	.8	.8	100.0
	Total	249	100.0	100.0	

**20-The process used under the Teacher Evaluation System fosters a climate for instructional improvement.**

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Strongly Agree	21	8.4	8.4	8.4
	Agree	122	49.0	49.0	57.4
	Neutral	67	26.9	26.9	84.3
	Disagree	31	12.4	12.4	96.8
	Strongly Disagree	8	3.2	3.2	100.0
	Total	249	100.0	100.0	

**21-The teacher evaluation system provides me with objective information about my teaching.**

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Strongly Agree	22	8.8	8.8	8.8
	Agree	131	52.6	52.6	61.4
	Neutral	57	22.9	22.9	84.3
	Disagree	30	12.0	12.0	96.4
	Strongly Disagree	9	3.6	3.6	100.0
	Total	249	100.0	100.0	

**22-The teacher evaluation system enhances dialogue between teachers and evaluators about effective teaching.**

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Strongly Agree	30	12.0	12.0	12.0
	Agree	108	43.4	43.4	55.4
	Neutral	57	22.9	22.9	78.3
	Disagree	42	16.9	16.9	95.2
	Strongly Disagree	12	4.8	4.8	100.0
	Total	249	100.0	100.0	

**23-The teacher evaluation system recognizes teachers' contributions to the academic institution as a whole.**

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Strongly Agree	30	12.0	12.0	12.0
	Agree	114	45.8	45.8	57.8
	Neutral	47	18.9	18.9	76.7
	Disagree	42	16.9	16.9	93.6
	Strongly Disagree	16	6.4	6.4	100.0
	Total	249	100.0	100.0	

**24-The teacher evaluation system increases my reflection on choice of teaching strategies.**

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Strongly Agree	31	12.4	12.4	12.4
	Agree	139	55.8	55.8	68.3
	Neutral	39	15.7	15.7	83.9
	Disagree	36	14.5	14.5	98.4
	Strongly Disagree	4	1.6	1.6	100.0
	Total	249	100.0	100.0	

**25-I regularly receive focused follow-up based on my evaluations.**

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Strongly Agree	14	5.6	5.6	5.6
	Agree	70	28.1	28.1	33.7
	Neutral	85	34.1	34.1	67.9
	Disagree	60	24.1	24.1	92.0
	Strongly Disagree	20	8.0	8.0	100.0
	Total	249	100.0	100.0	

**26-I regularly receive instructional support based on my evaluations.**

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Strongly Agree	8	3.2	3.2	3.2
	Agree	73	29.3	29.3	32.5
	Neutral	88	35.3	35.3	67.9
	Disagree	58	23.3	23.3	91.2
	Strongly Disagree	22	8.8	8.8	100.0
	Total	249	100.0	100.0	

**27-I focus my professional development efforts on activities that directly help me achieve the evaluation standards.**

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Strongly Agree	40	16.1	16.1	16.1
	Agree	126	50.6	50.6	66.7
	Neutral	46	18.5	18.5	85.1
	Disagree	30	12.0	12.0	97.2
	Strongly Disagree	7	2.8	2.8	100.0
	Total	249	100.0	100.0	

**28-The feedback I receive from my evaluators is valuable.**

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Strongly Agree	43	17.3	17.3	17.3
	Agree	121	48.6	48.6	65.9
	Neutral	57	22.9	22.9	88.8
	Disagree	25	10.0	10.0	98.8
	Strongly Disagree	3	1.2	1.2	100.0
	Total	249	100.0	100.0	

**29-The teacher evaluation process is helpful to my professional growth.**

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Strongly Agree	45	18.1	18.1	18.1
	Agree	127	51.0	51.0	69.1
	Neutral	52	20.9	20.9	90.0
	Disagree	18	7.2	7.2	97.2
	Strongly Disagree	7	2.8	2.8	100.0
	Total	249	100.0	100.0	

**30-I am satisfied with the current evaluation system.**

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Strongly Agree	11	4.4	4.4	4.4
	Agree	70	28.1	28.1	32.5
	Neutral	89	35.7	35.7	68.3
	Disagree	55	22.1	22.1	90.4
	Strongly Disagree	24	9.6	9.6	100.0
	Total	249	100.0	100.0	

**31-My evaluators have been properly trained to consistently evaluate my teaching.**

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Strongly Agree	8	3.2	3.2	3.2
	Agree	62	24.9	24.9	28.1
	Neutral	88	35.3	35.3	63.5
	Disagree	63	25.3	25.3	88.8
	Strongly Disagree	28	11.2	11.2	100.0
	Total	249	100.0	100.0	

**32-The standards used in the teacher evaluation system are fair.**

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Strongly Agree	7	2.8	2.8	2.8
	Agree	100	40.2	40.2	43.0
	Neutral	84	33.7	33.7	76.7
	Disagree	41	16.5	16.5	93.2
	Strongly Disagree	17	6.8	6.8	100.0
	Total	249	100.0	100.0	

**33-The evaluation standards do a good job of defining effective teaching.**

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Strongly Agree	11	4.4	4.4	4.4
	Agree	121	48.6	48.6	53.0
	Neutral	73	29.3	29.3	82.3
	Disagree	31	12.4	12.4	94.8
	Strongly Disagree	13	5.2	5.2	100.0
	Total	249	100.0	100.0	

**34-Working towards improving my performance on the evaluation standards will help me to improve the quality of my teaching.**

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Strongly Agree	37	14.9	14.9	14.9
	Agree	155	62.2	62.2	77.1
	Neutral	39	15.7	15.7	92.8
	Disagree	12	4.8	4.8	97.6
	Strongly Disagree	6	2.4	2.4	100.0
	Total	249	100.0	100.0	



**35-The Teacher Evaluation Instrument includes clear explanations for each performance descriptor.**

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Strongly Agree	15	6.0	6.0	6.0
	Agree	129	51.8	51.8	57.8
	Neutral	65	26.1	26.1	83.9
	Disagree	32	12.9	12.9	96.8
	Strongly Disagree	8	3.2	3.2	100.0
	Total	249	100.0	100.0	



**Respondent:** Student achievement. Yes. And then the third one is student evaluation. I'm not aware of any other... Oh, yes. Teachers are, I believe, quite rightly given some kind of credit if they participate in other committee work or examination preparation, other additional work. But as far as I know, there are no other factors, no other criteria for evaluating teachers.

**Interviewer:** Not research? Community service?

**Respondent:** Those would be appreciated along with the committee service and things like this, but you're not expected as a matter of course to do them.

**Interviewer:** So it's not part of your official evaluation?

**Respondent:** No.

## **Q2, Interviewer: And how important is teacher evaluation to you?**

**Respondent:** I'm at the end of my career. I've evaluated so many teachers myself. I have a kind of perhaps a deeper understanding of it. So I like to know that I'm popular, but I prefer to think that popularity is because I keep the students interested in what I'm doing rather than just because I get good grades or give good grades. Personally, as I said when Dennis was here, one thing I do get in the prep year that I didn't get in the Jebel is job satisfaction. There are many reasons for that, but that's important. So that is one. Fortunately, I've had good evaluation, so it's nice. I think it's because the students enjoy my lessons, and because they enjoy my lessons, they're attentive.

**Interviewer:** Yeah. And does it affect your salary or contract? I mean the results of evaluation.

**Respondent:** It doesn't affect mine because any salary increments stop once you reach the age of 60 at this university. So it's the pressure of ageism. You could be the best teacher in the world, but after you reach the age of 60, you don't get any increments. Before that, yes, I understand. I don't know the precise details, but I understand that awards are given on the basis of the things I've talked about, including students' evaluation.

## **Teachers Readiness**

### **Q3, Interviewer: Right. Okay. Now, we move to teachers' readiness. Have you received any induction or written instructions on how evaluation works?**

**Respondent:** No.

**Interviewer:** Not at all?

**Respondent:** No. Neither when I joined the Jebel nor here. One gleans it from one's colleagues.

**Interviewer:** You heard about it?

**Respondent:** Yeah. I was never given any kind of orientation about it.

**Interviewer:** And have you ever asked for them?

**Respondent:** On the website, they're available. We one learn about that, and then you go and look, and it's self-evident. There are comments. Some of the comments are in Arabic, but I think I translated it. The grades are--

**Q4, Interviewer: Do you mean the criteria for--**

**Respondent:** Yeah. No. I don't think we see the criteria. I think we see the overall grade, if I'm right. I don't generally look it up, to be honest. I kind of know what they are.

**Interviewer:** So even the students' rating, the criteria for students' rating, you don't have any access to them?

**Respondent:** I want to be careful of what I say because I want to be accurate. I'm not sure. I haven't seen it. But that made it my mission. I don't go searching for these things, so--

**Interviewer:** So you just receive the final grade?

**Respondent:** Yeah. And even that, it's for me to fish on the website. It's not given to me as a sort of certificate and kind of like that. And I know because I'm invited to do extra work, which I wouldn't be if the grade wasn't at a certain level, so that's how I kind of--

**Interviewer:** Yeah. Like indications, let's call it that. And for the way you're being rated by students, is it like marks, or like satisfactory or excellent or...?

**Respondent:** What I understand is there is a checklist, a tick box sort of thing, and then I think some of them have a rating of 1 to 10, like a Likert scale, and the summary of it generates... I don't know what the summary of it generates, but I'm told. In fact, I was discussing this with a colleague today who has more experience than me, and he said it then depends on the dean. And apparently, this was what I was told. I have no evidence to this. Our dean is very strict. This colleague today said to me, "If you're in some departments, like mechanical engineering, if your students give you sort of 80 out of 100, the dean will give you an A." He said, "But here, you will only get a B for that." So apparently, there is some kind of grade normalization takes place at the dean level. That's what I understand.

**Interviewer:** But they are not clear to you?

**Respondent:** That's not clear. No.

**Interviewer:** Right.

**Respondent:** Not precisely clear.

**Q5, Interviewer:** Then how are EFL teachers prepared for teacher evaluation?

**Respondent:** As far as I know, again, they just gather the information from colleagues and from the website where the evaluations appear.

**Interviewer:** So you try to fish for them?

**Respondent:** Yeah.

## Challenges

**Q6, Interviewer:** Okay. Now, we will move to challenges. What do you believe is the biggest challenge in making performance evaluation effective and meaningful to EFL teachers?

**Respondent:** Here, I think it would be if you could find a way of convincing them that the students were led to or guided in or had an honest approach to it, and did not simply give a low grade to a teacher because they themselves didn't get a good grade.

**Interviewer:** You said convincing them. Who?

**Respondent:** I don't know. It's a good question. I don't know what kind of orientation, if any, the students have. You see? That's another question.

**Interviewer:** Okay. So you're questioning the qualification of students to evaluate?

**Respondent:** Well, I think no. I think teachers are suspicious that students will give them a good grade not necessarily because they're good teachers but because they're getting them good grades, which isn't necessarily the same thing.

**Interviewer:** Right.

**Respondent:** I'll give you a classic case in point. This university persists in using the TOEFL, in spite of the fact that all the American universities have rejected it in favor of the IELTS, which is a direct test, not an indirect test. All the key universities in the States now prefer IELTS to TOEFL. Now, I'll give you a very good example of why. And this was my doctorate, and I've spent the whole of my career really arguing in favor of direct testing rather than indirect. You think about the processes—the processes you use to answer a multiple-choice question and compare those with the processes that you have to use to generate grammar to write an essay. Huge difference.

**Interviewer:** Huge. Yes.

**Respondent:** Now, a classic case. The last course I taught, I had a student. He was a very nice student, but he was the weakest writing student in the class and he passed the TOEFL requirement to go to the Jebel, which is primarily a writing program. The other thing, of course, about TOEFL is that it's easy to teach, and Dennis loves it, because you just stand there and you spout grammar, and they can learn. It's actually quite artificial as well. The reading test, there is no time to read so you have to teach them. There are books. The actually book that we've got teaches them tricks. "Okay, what is this paragraph about?" You don't read the paragraph. You read the topic sentence. So it teaches you to read the topic sentence. So it's not actually a test of reading.

**Interviewer:** Test of skills.

**Respondent:** Yeah. So I think therefore what you've got is students will give a good rating to a teacher who is enabling them to get good grades. Whether that teacher is a good language facilitator is often not necessarily the case. I believe you can be an excellent teacher of TOEFL and a pretty awful teacher of language.

**Q7, Interviewer: Yes. True. I totally agree with you, because teaching for exam or for proficiency exams is totally different from teaching the language, with all its cultural aspects as well. Yeah. Nice. Do you think there are other challenges?**

**Respondent:** Yeah. I think apart from the ones we discussed, like the timescales of packing everything in, there's always too much to do, where often you have to move on where you'd like to dedicate another lesson to a particular aspect of language. You can't do that. You've just got to move on. That's a challenge. The standard of English that the students bring to us is a persistent challenge because with language, it's a skill. It's different from any other subject. It's acquisition, not learning. An acquisition means habits. And by the time they come to us, because this is primarily an oral culture, not just in Arabic too, but in English, they have used the English that they've got, which is totally... Well, it's like Pidgin English. It's an interlanguage, but they've used it. It's become fossilized and it's far, far difficult to move them on. I often wish they had never been taught English at school than mistaught the way they are. That's a huge challenge because--

**Interviewer:** So you prefer to start from scratch?

**Respondent:** I think I would. I think we can do a better job if they haven't mistaught before. But of course, it's an idealized world. Because English is the language of the world, they're going to have some anyway. But I explained it by Arabic. I worked for the British Government in Yemen for four years, and I lived in Sana'a and travelled all over the country because I was the director of the project. And of course, in Yemen, very few people speak English, unlike Saudi Arabia. So in Yemen, I had to pick up a bit of Arabic. And that was about 30 years ago. It was before the First Gulf War. It was a long time ago. And I tell my students this, and then I speak to them in Arabic. And of course, I made lots of mistakes, but I said, "Did you understand what I said?" and they said, "Yes, of course." So that's exactly... See, what has happened

is because I don't use Arabic, it's fossilized. I said, "It would be extremely difficult for me to change it."

**Interviewer:** True.

**Respondent:** And I warned them. I said, "Look, you're in a great position. You've got far more English than I've got Arabic, but if you're not careful, it will stick." And that's what's happening a lot. Unfortunately, with many it's already stuck because they've been using it for so long, and to move it on, you've got to have so many different techniques. They've got awareness raising, but they've got to believe you.

**Interviewer:** Yes.

**Respondent:** And if you come with the belief that just by learning some more words and some grammatical rules, my English is going to get better, that doesn't work. And a lot of them bring only instrumental motivation. They don't actually want to study English. They know they have to because theoretically it's an English medium university. But if they did it from love, I think probably, at the maximum, a third of them would be doing it because they're always the better ones as well.

## Evaluator

**Q8, Interviewer:** Yeah. It's always the case. True. Okay. We will talk now about the evaluator. At your institute, who is the evaluator?

**Respondent:** I suppose... Well, we have a director of the whole school that is the prep year, and then each course has its coordinator. But I think the evaluator is the director.

**Interviewer:** Not the coordinator?

**Respondent:** That's what I think.

**Interviewer:** Okay. Do you have classroom observation?

**Respondent:** I did recently, but it's unusual. The reason I did was because I asked for a transfer from down from the Jebel to here, which is unusual. I mean, people usually see it as promotion going up.

**Interviewer:** Exactly.

**Respondent:** It isn't promotion, but for reasons I've explained, I wanted to come down. So the only way to do that was Paul, who is the director, had to put me through some hoops, and one of them was being observed. But it only happens if there's a problem with the teacher normally.

**Interviewer:** Okay. And usually, is the observer also the evaluator?

**Respondent:** Yes.

**Q9, Interviewer:** Yes. When it happens. Okay. What do you think about the evaluator?

**Respondent:** Do you mean in his capacity as an evaluator?

**Interviewer:** Exactly. His knowledge, qualifications, skills, everything?

**Respondent:** Well, I'm afraid I don't know. I don't know what his qualifications... I'm not quite happy with him being there. He handled my transition very well. I have a good working relationship with him. I don't know what attention is paid to qualifications and experience when people are put in that position. You'd have to ask the deans that.

**Interviewer:** Right. And students are also part of the whole scheme of evaluation. Am I right?

**Respondent:** Just as we've talked about before. Every course, the students, towards the end of the course, but before the final results come out, but towards the end of the course, I think they do a questionnaire and submit it online.

**Interviewer:** Right. Are they well-qualified to do this? I mean, not thinking about their training, if they are trained to do that. But are they qualified to evaluate teachers?

**Respondent:** Absolutely not. No, except insofar as they are the learners. They are the clients, the end users.

**Interviewer:** Exactly.

**Respondent:** So there is an extent to which that qualifies them. But over and above that, nothing additional.

**Interviewer:** Their thinking skills and their abilities?

**Respondent:** Yeah.

## Feedback

**Q10, Interviewer:** Okay. Tell me about the quality of the feedback you receive after evaluation.

**Respondent:** The only thing I get is a grade, or if there are comments, I can see the comments. And the comments are either negative or positive. They're never constructive, if you know what I mean, because of the reason we've talked about before. The students don't know anything about language teaching. There's no reason why they should, other than their experience as learners.



**Interviewer:** True.

**Respondent:** I find I'm quite surprised sometimes when I do things. If I ask them to work in a group in a particular way, they often don't see the reason for that until they've done it. But whether they think when they're doing the evaluation, "Oh, this teacher did something interesting," I don't know. Do you know what I think? I think they really just vote onto things, which is the grades they've been getting and basically how much fun you are as a teacher. I mean, if you laugh and have a joke with them and engage them all, then I think they'll give you a good evaluation. And I think that's important. I don't denigrate that.

**Interviewer:** So they usually judge the personality of the teacher?

**Respondent:** I think that's got a big part to play in it. Yes.

**Q11, Interviewer: Okay. Do you have the right to defend yourself against the evaluation?**

**Respondent:** I believe so. It hasn't happened in my case where I've had to. Well, I did, but it wasn't for academic reasons. When I was on the Jebel, I was accused of saying something about religion that I didn't say.

**Interviewer:** Maybe misunderstood?

**Respondent:** Yes. But I was able to defend that. But it was very much like I was accused and had to defend myself. It wasn't a sort of... And finally, I mean, it was dismissed. But these things can happen.

**Interviewer:** Yeah. I know what you mean. And will the evaluation results be modified accordingly? According to your defense?

**Respondent:** I don't think the evaluation results are modified, but I think the consequences might be.

**Interviewer:** Okay. Which is, I think, good.

**Respondent:** Yeah. I mean, if I've done what I was accused, if I had said what I was accused of saying, well, I don't think it's right, but I was accused of saying I didn't believe in God. Now, I didn't say that or anything like it. In fact, I'd have great difficulty answering that question. But if I had said, "Yes, I said that in class," I think that would be a dismissible offense in this university.

## Consequences of evaluation

**Q12, Interviewer: True. Yes. Anyhow, now we move to the consequences of evaluation. Do the evaluation results motivate your professional development?**

**Respondent:** I think I can speak on behalf of my colleagues when I say no. Myself as well. I don't need any motivation, but that's me. I can't be in a classroom and not be engaged with the learners, which I think is the most important thing. If there was no syllabus, if there was no pacing schedule, I would still go in there and I would still be engaging them all, and we would be doing something to develop their language. But I don't know. That's just me.

**Interviewer:** Yeah. And how about workshops? The evaluation results, do they feed into the workshops that you're offered?

**Respondent:** No. And workshops are used rather cynically here. We often finish the semester before the holiday begins, and I think we have workshops then to ensure that people aren't leaving before the end of the semester.

**Interviewer:** Just to keep them busy?

**Respondent:** Absolutely.

**Interviewer:** But not to improve their practices as EFL teachers?

**Respondent:** I think that is secondary.

**Interviewer:** Okay. Got it.

**Respondent:** It's very cynical of me, but I think it's true. That's how it is.

**Q13, Interviewer: Okay. In what way do the evaluation results affect your instructional practices or the quality of your work experience?**

**Respondent:** Personally, they don't.

**Q14, Interviewer: They don't. Okay. Can evaluation affect your job security?**

**Respondent:** Yes.

**Interviewer:** In what way?

**Respondent:** If I got a bad evaluation below, I think it's 7, like 70%--

**Interviewer:** 70% of...?

**Respondent:** Especially because of my age, and I think there's a pressure from outside the university to... We've got an awful lot of people in the... I mean, I'll say that. I feel like a young guy, and my son is 33. But yeah, I think a lot of people feel insecure because it happened actually this semester as well. Somebody who didn't get a 7, and he's already on the age of it doesn't matter what, but he was in his 70s, has been told, "That's it."

**Interviewer:** So it's a matter of hiring or firing?

**Respondent:** Yeah. Absolutely. Yeah.

**Interviewer:** Okay. Interesting.

**Respondent:** Given certain parameters. I don't think it happens to the young guys. And I'll be very honest with you. I don't think it happens to Muslim converts either.

**Interviewer:** Okay. There are special cases.

**Respondent:** There are special cases.

**Q15, Interviewer: Okay. Do you believe that the evaluation process or processes affect your self-esteem?**

**Respondent:** No.

**Interviewer:** No? Why no?

**Respondent:** I've been doing it too long.

**Interviewer:** Okay.

**Respondent:** As I said, I'm a qualified teacher-trainer myself. I know when I'm doing a good job or not. I mean, I don't want that to sound arrogant. No. I mean, I don't like to walk out of the classroom and feel that I wasn't very good. It happens sometimes still. Of course, it does, occasionally. I know I could have done that better. I could have done something in a different way. But I'm not affected by the evaluation.

**Q16, Interviewer: Results. Okay. Now, what do rewards and recognition mean in your teacher evaluation system?**

**Respondent:** In my...?

**Interviewer:** Teacher evaluation system. In the evaluation system in here, I mean.

**Respondent:** For me?

**Interviewer:** Yes, for you.

**Respondent:** As I said before, they don't really mean anything because it doesn't affect my salary. It doesn't affect my self-esteem because I'm kind of an expert in what I do anyway. So I think it has a negligible effect. I don't think I would be, in any way, a different teacher if it wasn't there.

**Interviewer:** Right. But how about rewards and recognition?

**Respondent:** There aren't any.

**Interviewer:** Okay.

**Respondent:** The only reward, if it is a reward, was when I've been asked to teach on an evening class, which is overtime, which most teachers want but I'm doing this because I enjoy teaching. I'm not really doing it because I desperately need the money. And in fact, I said to Paul this time, "Please, don't include me," and he came knocking on the door and he said, "Look, we've only got a certain number of teachers who have got..." because they insist on having teachers teach on those programs who got at least a B plus, so there aren't so many. And he said, "Look, I'm short on teachers. Will you please do it again?" My wife is loving it. So I'm teaching morning, afternoon and evening, five days a week.

**Interviewer:** Right. And for the highly performing teachers, are they recognized in a way or another?

**Respondent:** That's the only recognition, as far as I know.

**Interviewer:** You get more work?

**Respondent:** Yeah. They get offered more work. It's extra paid work, so I supposed it's seen as a... And summer school. It's like these are used as rewards. Yeah. I suppose they are rewards.

**Interviewer:** Okay. So good performance is recognized this way.

**Respondent:** In that way. Yeah.

**Interviewer:** Okay. But you don't usually have any certificate of appreciation or something?

**Respondent:** Absolutely not. No. I think at the end of your service, you'll probably get a letter from the university saying, "Thank you for your..." That's all.

**Q17, Interviewer: Yeah. Okay. Do you believe that the evaluation process affects your sense of belonging to the institution?**

**Respondent:** That's a good question. Well, I mean, sense of belonging to the institution? I do have a better sense of belonging to the prep year than I did in the Jebel, but that's because we're all in one building and I see my colleagues, whereas when I was in the Jebel, we all used to go off and teach in different buildings, so you'd very seldom see your colleagues. And if you're a different shift from your officemate, because we share offices, then you might only see him once or twice a week. But is that to do with the institution? I'm not sure. It's more to do with the social environment.

**Interviewer:** Exactly.

**Respondent:** I mean, would love to feel proud of working at this university, but I would have to tell you that coming here, my primary impression has been one of disappointment at the level. I expected it to be higher. With the English language standard, it's not very high.

**Interviewer:** I see what you mean.

**Respondent:** Really. I don't know what happens in the schools, but very little it seems.

**Interviewer:** Do you think that being physically close to your colleagues does affect your sense of belonging?

**Respondent:** Yeah. Definitely.

**Interviewer:** So you didn't feel it there?

**Respondent:** I didn't feel this in the Jebel because you're just stopping and chatting in corridors. It makes a big difference. Yeah.

### Teacher's role

**Q18, Interviewer:** Sure. Now, we move to teacher's role. What role does the teacher play in the development of teacher evaluation practices?

**Respondent:** As far as I know, none.

**Q19, Interviewer:** None. Tell me about the teacher's voice in teacher evaluation at your institute.

**Respondent:** As we said before, it's only if there's a problem that I think you are given a chance to explain. Only then.

**Interviewer:** But not to develop the practice?

**Respondent:** No. I haven't seen anything that is... In fact, I don't think there's time for positive development. It's only put in as an afterthought to stop teachers going home early at the end. So it's very cynical. It's all very negative.

**Interviewer:** Yeah.

**Respondent:** I mean, it's like the society as a whole. It's top-down. There isn't much... One learns to... Even when you have ideas. And that I'm not just talking about this university. I'm talking about all the jobs I've had in the Middle East, including many of them in management and many of these awful jobs where you end up as the intermediary between, for example, the Saudi deans and your colleagues who are Westerners. And you've got to work between two completely different worldviews vis-à-vis learning. I'm not talking about anything else, just learning. I mean, for example, our dean is a dean of math. He has a PhD in math. So his English isn't wonderful. I mean, it's okay for conversation, but it's not that good. I'm an IELTS examiner. He's not that high, yet he's in charge of the institution. So he's not a language expert. He's not a language specialist. So I can see faults in the exams that we give, but I wouldn't go and try to explain to him. Even one idea of

mathematics. Because of the PhD I did, I know quite a lot about statistics. We have some statistical procedures which are ludicrous, but I wouldn't dream of telling him because what might happen is he might give me something to do, like a responsibility to fix it in addition to the work that I already have, which is quite enough. And then if I did it and it didn't work, I would get blamed for it. So there's no inducement for people giving ideas, and you're thought of by your colleagues as being pretty stupid if you do.

**Interviewer:** Yeah. I see what you mean.

**Respondent:** And lots of people have ideas. We all know more about language learning and teaching than the dean does, but you can't say that.

**Interviewer:** Yeah. And with the top-down hierarchy we have, I think we all believe that things are imposed on us also as staff and as teachers.

**Respondent:** Yeah. Absolutely. You do as you're told, and you don't really question it. You can have a little grumble about it sometimes, but--

**Interviewer:** When you feel too bad about it.

**Respondent:** Yeah, but you can't do anything to change it, and it's not very wise to... I'll give you a very nice analogy.

**Interviewer:** Yes. Go on, please.

**Respondent:** I worked for another educational institution not very far from here, the military one. And when I went there, a senior British colleague, he said, "I'll give you two pieces of advice, Joseph. Number one: Keep out of the limelight. And number two: If you do get in the limelight, don't dance." And what event you know of course; just keep your head down. Don't be conspicuous. But if you do get put on the spot, just don't say anything that is radical or that is likely to change things. Don't suggest anything that would change or improve things. Just keep quiet about things.

**Interviewer:** Okay.

**Respondent:** And that's the predominant philosophy in... I mean, it's not just in the university or the society. It's in all top-down societies.

**Interviewer:** Yeah. It's the structure of the whole society.

**Respondent:** I mean on the extreme point, with Stalin in communist Russia, he would give a lecture in a great big lecture theater, and the people who applauded last were taken out and shot.

**Interviewer:** Yeah. I see what you mean. Okay. Thank you very much for helping me and giving me some of your valuable time.

**Respondent:** I hope I was of some help.

## Conclusion

**Q20, Interviewer:** Can I finally ask if you think there are any aspects of your experience with teacher evaluation that has not been covered in this interview?

**Respondent:** Well, all the things that we've said don't happen here, of course, are things that would normally happen. What one would like to think would happen in a more open environment where people could express their feelings without feeling threatened... I mean, I had observed teachers. When I was doing my teacher training, one of the things I was taught was the first thing, you always consider yourself to be an honored guest if you go into a classroom to observe, and you do not go in to find faults.

**Interviewer:** Exactly.

**Respondent:** You go in with the cooperation of the teacher and you look for the good things.

**Interviewer:** Exactly.

**Respondent:** And of course, you may see something where you think, "Well, I think that could be improved." But rather than saying, "You could have done that better," you say something like, "What would you think if you had done this instead of that? What would have been the result?" I mean, he may have a perfectly good reason for doing or not doing it. Sadly, it's not what people come here for. The whole setup is difficult to attract young teachers, first of all, because there's no life here in our terms. And I'm not talking about discos and everything. I'm just talking about normal feeling of freedom. Middle-aged people generally are more interested in their careers, and if you come work in Saudi Arabia for any length of time, it can be seen as a slur in your career because the best teachers don't come here. They've got a lot of older teachers because they've had their career. They're good, and the good thing about them is they'll often stay for a longer period of time, so they learn how to teach, although of course, they tend to be less radical, but that's exactly what the system wants.

**Interviewer:** Yes.

**Respondent:** I mean, I don't know how you could change it. I don't think it could be changed, frankly, and I'm surprised sometimes how well it works because it does sometimes. You get fantastic students, and then you can sometimes see students visibly making progress, which is very rewarding.

**Interviewer:** And you should feel very happy.

**Respondent:** Oh, God, yeah. You do. Absolutely. I mean, I'm struggling with a group at the moment, but the group I taught before, even though they were all repeaters, it

was a small group of about 14. Right from the start, we just clicked like that. I felt an empathy with all of them. I felt that there was nobody who disliked English and was there just because they had to be, and I felt very good at the end of that. So yeah, I mean, I'm sure they learned something from me. I'm sorry. I've forgotten the question.

**Interviewer:** It's okay. I felt very good about what you said that you were happy though they were lower achievers.

**Respondent:** Yeah. I don't know why they were repeaters. To me, they seemed to have the ability to make progress, and now they did. It was reflected in the results as well, so that was good, though that was the group where the one who passed the TOEFL was the worst writer. But he's very good at speaking English.

**Interviewer:** Cool. Thank you very much.

**Respondent:** Okay.



## Appendix 10: Data Analysis / Qualitative Thematic Analysis: NVivo Sample Nodes

EF1\_TeacherEvaluation7 (NVivo 11) (3)jmp - NVivo Pro

Nodes	Name	References	Created On	Created By	Modified On	Modified By
Challenges	Challenges	24	365 30/05/2016 11:12	KA	30/06/2016 09:59	KA
Classroom Related Challenges	Classroom Related Challenges	12	28 08/06/2016 09:20	KA	19/06/2016 11:57	KA
Evaluator Related Challenges	Evaluator Related Challenges	23	124 08/06/2016 09:20	KA	19/06/2016 12:00	KA
Other Types of Challenges	Other Types of Challenges	21	61 08/06/2016 09:24	KA	14/07/2016 09:37	KA
Students Related Challenges	Students Related Challenges	15	74 08/06/2016 09:20	KA	13/07/2016 20:30	KA
Teacher Related Challenges	Teacher Related Challenges	19	80 08/06/2016 09:21	KA	19/06/2016 12:00	KA
Discourse	Discourse	2	3 11/07/2016 11:58	KA	19/09/2016 17:19	KA
Distinctive Phrases	Distinctive Phrases	0	0 24/06/2016 09:49	KA	19/09/2016 17:19	KA
Feedback	Feedback	5	6 20/06/2016 10:38	KA	19/09/2016 17:19	KA
Importance	Importance	23	149 17/06/2016 14:36	KA	19/09/2016 17:19	KA
Consequences	Consequences	23	135 19/06/2016 12:59	KA	19/09/2016 17:19	KA
Important	Important	8	9 21/06/2016 14:20	KA	21/06/2016 14:21	KA
Not Important	Not Important	4	4 21/06/2016 09:40	KA	12/07/2016 12:41	KA
Involvement of Teachers	Involvement of Teachers	23	141 19/06/2016 12:59	KA	19/09/2016 17:19	KA
Methods	Methods	22	78 30/05/2016 11:11	KA	19/06/2016 12:01	KA
Purposes	Purposes	5	5 30/05/2016 11:11	KA	15/06/2016 13:59	KA
Suggestions for Solutions	Suggestions for Solutions	16	116 30/05/2016 11:13	KA	13/07/2016 20:31	KA
Macro-Level	Macro-Level	5	18 16/11/2016 11:18	KA	23/11/2016 10:27	KA
Micro-Level	Micro-Level	15	98 16/11/2016 11:18	KA	23/11/2016 10:27	KA

## Appendix 11: Data Analysis / Qualitative Thematic Analysis: NVivo Interview Coding Sample

The screenshot displays the NVivo software interface. On the left, a tree view shows the following node structure:

- Nodes
  - Sources
  - Classifications
  - Collections
  - Queries
  - Reports
  - Maps
  - Folders

The main workspace shows a node named "Classroom Related Challenges" selected. Below it, a table lists its sub-nodes and their reference counts:

Node Name	Sources	References
Challenges	24	355
Classroom Related Challenges	12	268
Classroom Observation	5	8
Observers	4	4
Perceived Notion	1	1
Large Classes	3	5
Give Attention to Each St in L	1	1
Time Can Be a Challenge	10	13
Long Teaching Hours	2	2
Time Consumed	1	2
Time Scale	8	9
Evaluator-Related Challenges	23	124
Fairness and Inequality	2	4
Feedback Issues	14	31
Delayed Feedback	3	3
Insufficient Feedback	3	3
No Feedback	7	8
Questionable Feedback	5	7
Users Feedback	5	9
Issues of Power	3	14
Personality and Exposures to Diff	11	19

The main text area shows a snippet of an interview transcript with several nodes applied as codes:

**Internal Focus Group Discussions IV 2 - 118 - 5 3 references coded [2.67% Coverage]**

**Reference 1 - 0.75% Coverage**

**Respondent 1:** Taking students from A1 to B1 in such a short time is a real challenge as well, and especially because the EU has a modular system, so there are like six or seven weeks for each level, whereas the rest of the university have a semester system, so even like practically the exams for other departments which clash with teaching weeks for EU, so we lose time there as well. So there are time constraints. There are curriculum constraints. The teachers are expected to cover a certain amount of materials in a very short time. Again, you're going to deal with students which are really not in the level they should be, for whatever reasons. So these are all big challenges that teachers face. And I feel that... Okay, I'm going to talk about the people who are preparing the curriculum. They have no classroom experience. They sit in offices somewhere, have never taught at EU, and say, "Do this, do that," without really knowing what the teachers are going through in the classroom. Let's do that, "without really knowing what the teachers are going through in the classroom."

**Moderator:** So it's unrealistic plans?

**Respondent 1:** Yes.

**Respondent 4:** There is need to engage teachers.

**Respondent 1:** I agree.

**Reference 1 - 0.75% Coverage**

**Respondent 2:** Sometimes, she tried her best, but I think the time for leading her, it's not enough. Sometimes I can't sit with her, so I have to talk to her at home, I have to send her some links in Whatsapp. Then I point out some figures for her and arouse them.

## **Appendix 12: Faculty Recruitment Standards**

### **1. Teaching Assistants**

It is clarified in Article 4 that anyone should be appointed as a teaching assistant must have the following:

- 1- A university degree from a Saudi university or other recognized university
- 2- A level of “very good” at least for the undergraduate degree
- 3- Whatever issued by the Council of the League of other conditions

Furthermore, different levels of teaching assistant are stated in Article 9.

### **2. Lecturers and Language Teachers**

In article 5, the requirements for the appointment of a lecturer and a language teacher are as follows:

- 1- A master degree or equivalent from a Saudi university or other recognized universities
- 2- A level of “very good” at least for the undergraduate degree
- 3- Whatever issued by the Council of the League of other conditions

### **3. Researcher Assistant**

Article 6 states that the following criteria are required for the appointment of the Assistant of a Researcher:

- 1- For those holding Master degree (Researcher Assistant “A”):
  - A. To obtain a master degree from a Saudi university or other recognized universities with a level of at least “very good”
  - B. Any other conditions deemed appropriate by the university
- 2- For those holding Bachelor degree (Researcher Assistant “B”):

- A. To obtain a bachelor degree from a Saudi university or other recognized universities with a level of at least “good”
- B. Any other conditions deemed appropriate by the university

It is also stated in Article 7 that teaching assistants, lecturers and language teachers are appointed based on the recommendation of the department in which they will work with the College Board and Standing Committee for the affair of teaching assistants, lecturers, language teachers and research assistants. The University Council eventually makes the appointment decision.

As for appointment decision of research assistants, Article 8 indicates that they are appointed by the President of the university decision upon the recommendation of the board of the department and the College Board followed by the Standing Committee for the affair of teaching assistants, lecturers, language teachers and research assistants. In Article 10, it is stated that: applied to language teachers and research assistants whatsoever listed in the approved educational employments document by the Civil Service Council resolution No. 590 and the dated 11/10/1401 AH and its amendments.

#### **4. Assistant Professors**

A doctoral degree, or its equivalent from a Saudi university or other recognized universities, is required for the appointment of an Assistant Professor. The University Council may add any other conditions, as stated in Article 11.

Article 12 indicates that in case of necessity and upon the recommendation of the board of the department and the College Board or Scientific Council, The Board of the University can appoint Assistant Professor, without the requirement to obtain a

doctorate in the disciplines to those who do not grant doctoral degrees according to the following criteria:

- 1- Having a master degree or equivalent from a Saudi university or another recognized university
- 2- Having spent at least three years as a lecturer
- 3- Providing scientific production of not less than three units published after earning a master degree, including at least one unit conducted individually, and that the provided scientific production should be consistent with what is stated in Article 29 of this document

## **5. Associate Professors**

Article 13 declares that subject to the provisions of Article 12, it is required for the appointment of the Associate Professor:

- 1- Obtaining a doctoral degree from a Saudi university or other recognized universities
- 2- Having spent at least four years of experience as a faculty member in the university or other recognized universities after being appointed as an Assistant Professor
- 3- Being promoted as an Associate Professor by one of the Saudi universities or other recognized universities

## **6. Professors**

In Article 14, it is stated that subject to the provisions of Article 12, it is required for appointment of the professor:

- 1- Obtaining a doctoral degree from a Saudi university or other recognized universities

- 2- Having spent at least four years of experience as a faculty member in the university or other recognized universities after being appointed as an Associate Professor
- 3- Being promoted as a Professor by one of the Saudi universities or other recognized universities

Faculty members are appointed based on the recommendation of the board of the department and the college specialists and recommendation of the Scientific Council; and the University Council shall make the appointment decision, as clarified in Article 15.

It is explained in Article 16; however, that faculty members who have previous experience, and for the purpose of salary, they are appointed according to the following criteria:

- 1- Experience must be gained after the stipulated qualification for appointment.
- 2- The experience should be in the field of specialization and has been gained while working in a Saudi university, other recognized universities, at one of the government departments, or at international organizations.
- 3- Experience is counted for the purposes of salary according to the following:
  - A. experience in the membership of the faculty of the university on a year-for-year basis.
  - B. experience that is not in the field of teaching (if homogeneous with the field of specialization) on the basis of a year-for-half year basis.

## Appendix 13: Teacher Evaluation Form Sample/ Other Qualifications Holders

**KINGDOM OF SAUDI  
ARABIA**

Ministry of Higher Education  
King Abdulaziz University

بِسْمِ اللَّهِ الرَّحْمَنِ الرَّحِيمِ



المملكة العربية السعودية  
وزارة التعليم العالي  
جامعة الملك عبدالعزيز  
إدارة شؤون المجلس العلمي

Dean of \_\_\_\_\_  
Vice-Dean of the Women's Campus

The following decree was issued by the Academic Council in its fourth meeting, dated 7/8/1422 H:

Decree N°. 3

The Academic Council hereby approves the Academic File Form for Lecturers and Teaching Assistants (males and females) holding full-time academic posts and not enrolled in any academic programs. The Council recommends the updated form as a substitute for the previous form.

You are kindly requested to take the necessary steps to put the decree in action. The form will be available as soon as it is printed.

The University Dean of Graduate Studies and Scientific Research  
And Head of the Academic Council

Dr. Mustafa Al-Idrisi

## The Academic Council

### *The Academic File for Lecturers and Teaching Assistants.*

#### Instructions

1. The file is filled out once a year, two months before the end of the academic year.
2. The lecturer/teaching assistant fills out the information in the first three sections
3. The department chair or department supervisor fills out Section 4.
4. The vice-dean of the college adds their input in Section 5-A.
5. The dean/dean of the Women's Campus adds their input in Section 5-B.
6. The file is archived in the dean's office and the department is given a copy.
7. The lecturer/teaching assistant is entitled to view the data in Section Four of their academic file and sign their acknowledgement, not their approval.
8. An asterisk (\*) is added on the left-side margin of the page opposite items with enclosed documents.
9. This academic file is filled out by lecturers and teaching assistants with full-time academic posts who are not enrolled in any academic programs.



### Section 1: Personal Information

Name:			
College:	<b>Department:</b>		
Position:	<input type="checkbox"/> Lecturer <input type="checkbox"/> Teaching Assistant <input type="checkbox"/> Language <b>Insturctor</b>		
BA Major:	<b>University:</b>		
	<b>Year:</b>		
MA MAjor:	<b>University:</b>		
	<b>Year:</b>		
Number and Date of Current Job Assignment:			
Current Position and Place of Work:			

### Section 2: Courses Taught in the Academic Year 1434/1435

Semester	Course Name	Course Number	Number of Sections	Teaching Units		
				Lectures	Practical	Clinical
First						
Second						

Section 3: Activities (Including student activities, academic activities, and administrative activities):

<input type="checkbox"/> No Activities	<input type="checkbox"/> Participated in the Following Activities:
.....	
.....	
.....	
.....	
.....	
.....	
.....	

Department Chair/Supervisor:  
Assistant/Language Instructor  
Name:  
Signature:  
Date:

Lecturer/Teacher  
Name:  
Signature:  
Date:

Section 4: Academic Performance Evaluation:

	Evaluation Topic*	/100
1	Commitment to teaching load	
2	Commitment to office hours	
3	Extracurricular activities	
4	Respecting academic traditions	
5	Commitment to department decisions	
6	Responsibility	
7	Cooperation and communication with students	
8	Cooperation and communication with colleagues	
Total		
Percentage ( /100)		

\*In case of inapplicability, please write N/A. The points will be disregarded and the percentage will be calculated according to the remaining topics.

Department Chair/Supervisor:  
 Assistant/Language Instructor  
 Name:  
 Signature:  
 Date:

Lecturer/Teacher  
 Name:  
 Signature:  
 Date:

Section 5: Approval of the Academic Performance of the Lecturer/Teaching Assistant/Language Instructor:

(a) Opinion of the Vice Dean of the College

.....

.....

.....

.....

.....

.....

Name:  
Signature:

(b) Opinion and Approval of Dean/Dean of the Women's Campus

.....

.....

.....

.....

.....

.....

Name:  
Signature:

## Appendix 14: Teacher Evaluation Form Sample/ Doctorate Holders



المملكة العربية السعودية  
وزارة التعليم العالي  
جامعة الملك عبد العزيز

### الملف الأكاديمي عضو هيئة التدريس بالجامعة Faculty Academic Portfolio

Name : الاسم :  
College : الكلية :  
Dept. القسم :

## **Preparing Instructions**

- The Faculty member fills out pages (1,2,3 and 4) and sign on page(4).
- The Chairman fills out pages (5,6,7, and 8) , signs and approves page(4).
- The concerned Vice-Dean verifies the upper part of page(9).
- The Dean approves the lower part of page (9).
- This file should be prepared periodically, once a year, two months before the end of the academic year.
- The file should be kept at the Dean's office and a copy of it should be given to Department.
- The faculty member has the right to look at his file and sign for knowing the contents and not for the approval (He may attach his comments separately).
- A(\*) should be used to the left of each item with attached documents.

## الجزء الأول

### نشاط عضو هيئة التدريس

الفترة من / / 14 هـ إلى / / 14 هـ

### ACTIVITIES OF A FACULTY MEMBER

Period: FROM / / TO / /

Name: الاسم :  
General Specialization: التخصص العام  
Specialization: التخصص الدقيق  
Last Qualification: المؤهل العمي الأخير :  
University: الجامعة المتخرج منها :  
Date of Last Degree: تاريخ التخرج :  
Appointment Date: تاريخ التعيين في عضوية هيئة التدريس :  
المرتبة العلمية : تاريخ الحصول عليها : / /

• يعبأ من قبل عضو هيئة التدريس .

• ويعتمد من رئيس القسم (المشرفة على القسم).

الوحدات التدريسية *			رقمها Course No.	اسم المادة Course title	الفصل
الطبي Lab	السريرية Clinical	المحاضرات Lecture			
					الأول
					الثاني

\* تشمل الإشراف على الرسائل العلمية ومشاريع التخرج .

الندوات والمؤتمرات والحلقات الدراسية		التاريخ Date	الجهة والمكان Place	عنوان المؤتمر Conf. Tit.
بدون مشاركة	بمشاركة *			

\* **المشاركات:**

- 1- بحث أو ورقة عمل
- 2- حضور ورشة عمل
- 3- رئاسة جلسة علمية
- 4- حضور مجلس إدارة المؤتمر
- 5 - عضو منظم للمؤتمر
- 6- للاستزادة العلمية
- 7 - أسباب أخرى .....

Published and Accepted  
Papers for Publication



## أوراق الأبحاث المنشورة والمقبولة للتشر :

السنة	الصفحات	المجلد	عنوان الدورية	عنوان البحث

## Research Projects Participation

## المشاركة في مشاريع الأبحاث

.....

.....

.....

## Books and Translation Works

## الكتب وأعمال الترجمة

.....

.....

.....

## هل أستدت إليك أنشطة داخل القسم ؟ Did the dept. assign to you any activities?

Yes  No  نعم  لا

If the answer is Yes, Pls. list the activities.

في حالة الإجابة بنعم أذكرها

.....

.....

.....

## Did the college assign to you any activities?

## هل أستدت إليك أنشطة بالكلية ؟

Yes  No  نعم  لا

If the answer is Yes, Pls. list the activities.

في حالة الإجابة بنعم اذكرها

Any assignment in committees of students activities?

هل أستدت إليك أنشطة في لجان الطلاب؟

Yes  No

لا  نعم

If the answer is Yes, Pls. list the activities.

في حالة الإجابة بنعم اذكرها

.....

.....

.....

Did the University assign to you any activities?

هل أستدت إليك أنشطة بالجامعة؟

Yes  No

لا  نعم

If the answer is Yes, Pls. list the activities.

في حالة الإجابة بنعم اذكرها

.....

.....

.....

Any assignments given to you outside the University?

هل أستدت إليك أنشطة خارج الجامعة؟

Yes  No

لا  نعم

If the answer is Yes, Pls. list the activities.

في حالة الإجابة بنعم اذكرها

.....

.....

.....

توقيع واعتماد  
رئيس القسم ( المشرقة على القسم )  
Sign. of Chairman

توقيع  
عضو هيئة التدريس  
Sign. of Faculty Member

## الجزء الثاني

### تقويم عضو هيئة التدريس

يعبأ من قبل رئيس القسم (المشرفة على القسم).

يوقع عليه عضو هيئة التدريس بالاطلاع .

لعضو هيئة التدريس إرفاق وجهة نظره في خطاب مستقل إن أراد ذلك.

## تقويم الأداء الأكاديمي لعضو هيئة التدريس ( يعده رئيس القسم العلمي )

## (أ) الأداء التعليمي :

درجة التقويم		عناصر التقويم*
لا ينطبق	من 100	
		1- الالتزام بالأعباء التدريسية
		2- الالتزام بالساعات المكتبية
		3- استخدام وسائل وتقنيات التعليم
		4- تحديث المواد التعليمية التي يقوم بتدريسها
		5- المشاركة في وضع المناهج ومقررات المقررات.
		6- احترام الأعراف والتقاليد في التدريس.
		7- المشاركة في تقويم الدارسين.
		8- الإشراف على الرسائل والمشاركة في مناقشتها.
		9- الالتزام بقرارات القسم
المعدل الإجمالي ( من 100 )		
المعدل المحتسب للترقية العلمية 25% من الإجمالي		
* في حالة عدم انطباق العنصر يكتب "لا ينطبق" وتحذف علامته ويؤخذ معدل العناصر المنطبقة.		
ملاحظات رئيس القسم العلمي (المشرفة على القسم).		
.....		
.....		
.....		
.....		
رئيس القسم العلمي (المشرفة على القسم) :		
.....		
التوقيع :		
.....		
التاريخ : / / 14 هـ		

## (أ) خدمة الجامعة والمجتمع :

درجة التقويم		عناصر التقويم*
لا ينطبق	من 100	
		1- حضور جلسات مجلس القسم العلمي.
		2- تنفيذ ما يكلف به من مهام.
		3- الإسهام في نشاطات القسم العلمي.
		4- الإسهام في نشاطات الكلية.
		5- الإسهام في نشاطات الجامعة.
		6- المشاركة فيما يقدم من برامج خدمة المجتمع.
		7- التفرغ الكلي لما يحقق أهداف القسم والكلية والجامعة .
		8- التعاون مع الآخرين في أداء الواجبات العلمية الإدارية والفنية للقسم العلمي.
		9- الالتزام بتقديم الخدمات العلاجية (طب بشري – أسنان)
المعدل الإجمالي ( من 100 )		
المعدل المحتسب للترقية العلمية 15% من الإجمالي		
* في حالة عدم انطباق العنصر يكتب "لا ينطبق" وتحذف علامته ويؤخذ معدل العناصر المنطبقة.		
ملاحظات رئيس القسم العلمي (المشرفة على القسم).		
.....		
.....		
.....		
.....		
.....		
رئيس القسم العلمي (المشرفة على القسم) :		
.....		
التوقيع :		
.....		
التاريخ : / / 14 هـ		

## (ج) السلوكيات :

## عناصر التقويم

1- هل له مخالفات في احترام القيم المهنية التربوية والأعراف الجامعية

2- متعاون مع رؤسائه

3- متعاون مع زملائه

الإجابات ذات الدلالة السلبية لا بد أن تكون مبررة أو موثقة .

رئيس القسم العلمي (المشرفة على القسم) :

التوقيع :

التاريخ : / / 14 هـ

اسم عضو هيئة التدريس :

توقيع عضو هيئة التدريس بالاطلاع

 توجد ملاحظات مرفقة لا توجد ملاحظات

توثيق واعتماد تقويم الأداء الأكاديمي لعضو هيئة التدريس :

رأي وتوثيق وكيل (وكيلة) الكلية المختص

اسم وكيل (وكيلة) الكلية المختص :

التوقيع:

التاريخ: / / 14هـ

## رأي واعتماد عميد الكلية / عميدة قسم الطالبات

عميد الكلية (عميدة قسم الطالبات):

التوقيع:

التاريخ: / / 14هـ

- ترسل صورة من الملف ومرفقاته بعد اعتماده إلى القسم
- ترفق صورة من آخر ملف أكاديمي مع كل طلب يشتمل على عرض من أغراض التقويم.



## Evaluation of the Academic Performance of a Faculty Member of the University

<p>The University Regulations, which prescribe a faculty member's duties and obligations, describe him/her as a well-qualified individual, capable of fulfilling his/her responsibilities towards the following :</p> <ul style="list-style-type: none"> <li>• Pursuit, enhancement, development and proper application of knowledge.</li> <li>• Practice of effective teaching which aims to prepare efficient manpower.</li> <li>• Undertaking of scientific research in order to make valuable contribution in the field of specialization.</li> <li>• Preservation of the religious and moral values of the society through exemplary conduct.</li> </ul> <p>The academic performance is evaluated in light of the following :</p> <ul style="list-style-type: none"> <li>• Full time commitment towards duties and obligations.</li> <li>• Distinguished execution and scholarly achievement in field of specialization.</li> <li>• Strict adherence to the laws and rules that govern the relationship between the University and its faculty members.</li> </ul> <p>The Regulations also detail the duties and responsibilities of the faculty member towards the following :</p> <ul style="list-style-type: none"> <li>• Teaching</li> <li>• Scientific research</li> <li>• Duties required by University.</li> </ul> <p><b>1. Objectives of Academic Evaluation</b></p> <p>By evaluating its faculty members, the University aims at :</p> <ol style="list-style-type: none"> <li>i) Reviewing the performance of faculty members to ensure conformity with conditions dictated by the University regulations.</li> <li>ii) Revealing aspects of strength and weakness in that performance.</li> <li>iii) Enabling members to improve their academic performance.</li> <li>iv) Considering distinguished performance as a basis for granting faculty members the privileges of taking sabbatical leaves, attending conferences, seminars, workshops, and training programs, or of any other academic incentives.</li> <li>v) Linking promotion with the comprehensive academic performance.</li> <li>vi) Renewing and terminating contracts for expatriates on the basis of academic performance.</li> <li>vii) Establishing the principle of appreciating and encouraging exceptional academic performance in the University at large.</li> </ol>	<p><b>2. Purpose of Academic Evaluation</b></p> <p>The evaluation of academic performance will be utilized as a determining factor for the following purposes :</p> <ol style="list-style-type: none"> <li>i) Academic promotion.</li> <li>ii) Honorarium.</li> <li>iii) Sabbatical leaves.</li> <li>iv) Conferences, seminars, workshops, and training programs.</li> <li>v) Scholarships.</li> <li>vi) Nomination for senior administrative posts.</li> <li>vii) Extension of service beyond retirement.</li> <li>viii) Secondment and consultations.</li> </ol> <p><b>3. Components of the Evaluation</b></p> <p>The evaluation of the academic performance for faculty members includes the following components :</p> <ol style="list-style-type: none"> <li>i) Teaching performance.</li> <li>ii) Academic duties.</li> <li>iii) General conduct.</li> </ol> <p><b>4. Frequency of Evaluation</b></p> <p>The academic evaluation of faculty members should be conducted annually, two months before the end of the academic year.</p> <p><b>5. Design of the Evaluation Forms</b></p> <p>The following factors were taken into consideration for the design of the evaluation form :</p> <ol style="list-style-type: none"> <li>i) Conformity with the requirements furnished by the University's criteria for the staff academic evaluation and promotion.</li> <li>ii) Clarity and simplicity.</li> <li>iii) Presenting the evaluator with various options (five numbered options).</li> <li>iv) Assuring a high degree of review by the Vice-Dean and the Dean of the College.</li> <li>v) Reducing the confidentiality restrictions to allow the evaluator more objective assessment by working with Vice Dean, and Dean of College.</li> <li>vi) In case that certain criteria do not apply to certain member, such as supervising theses where no post-graduate programs exists, the non-applicable criteria should not be considered and the average of the applicable criteria should be considered.</li> </ol>
--	---

## List of References

- Abdelgadir, M. A. (2012). *Learning English in large classes at tertiary level in Sudan: Influence of socio-cultural factors, learning strategies utilised by students, and students' perceptions*. (Doctor of Education in TESOL), University of Exeter, Exeter.
- Abu Arrab, E., & Gadadah, I. (2008). Evaluation of Faculty Members' Performance in Higher Education Institutions. *The Arab Journal for Quality Assurance of Higher Education*, 1(1), 69-107.
- Aburizaizah, S. J. (2014). Developing a framework for classroom lesson delivery to improve English teachers' performance in the foundation year programme at a Saudi university. *English Language Teaching*, 7(1), 1-8.
- Aburizaizah, S. J. (2016). Teacher development in higher education in an EFL context: Why evaluate and how to evaluate? *US-China Foreign Language*, 14(4), 269-283.
- Ahmed, A., & Asghar, J. (2015). An investigation into annual faculty evaluation process of EFL teachers in the middle east: Challenges for leadership & management. *Theory and Practice in Language Studies*, 5(6), 1140-1150.
- Akbari, R., & Tavassoli, K. (2014). Developing an ELT context-specific teacher efficacy instrument. *RELC Journal*, 45(1), 27-50.
- Akbari, R., & Yazdanmehr, E. (2012). EFL teachers' recruitment and dynamic assessment in private language institutes of Iran. *Journal of English Language Teaching and Learning*, 3(8), 29-51.
- Al-Abdulkader, A. A. (1979). *A survey of the contribution of higher education to the development of human resources in the kingdom of Saudi Arabia*. University of Kansas.
- Al-Anzi, B. S. (2015). *The situation of the preparatory year programs from the perspective of newspapers authors*. Paper presented at the 1st National Conference for Prep Year in Saudi Universities, University of Dammam, Saudi Arabia.
- Al-Hammad, Y. (2011). *Teaching performance assessment for the intermediate school English teachers in the city of Hail as per the Teaching Quality Standards*. (Mater's Thesis), Imam Muhammad Bin Saud Islamic University, Saudi Arabia.
- Al-Harthy, W. S. A. (2017). The perception of the role of post-lesson discussion in the professional development of Omani English teachers. *Journal of Basic and Environmental Sciences*, 4, 1-17.
- Al-Hazmi, S. (2003). EFL teacher preparation programs in Saudi Arabia: Trends and challenges. *Tesol Quarterly*, 37(2), 341-344.
- Al-Jarf, R. (2015). Assessing EFL college instructors' performance with digital rubrics *Teaching and Learning in Saudi Arabia: Perspectives from Higher Education* (pp. 1-30). The Netherlands: Sense Publishers.
- Al-Juhani, S. H. (2012). A study of failure and withdraw phenomena at College of Preparatory Year, Princess Norah bint Abdulrahman University. *The Future of Arabic Education*, 19(79), 245-334.

- Al-Mahrooqi, R., Denman, C., Al-Siyabi, J., & Al-Maamari, F. (2015). Characteristics of a good EFL teacher: Omani EFL teacher and student perspectives. *SAGE Open*, 5(2), 1-15.
- Al-Maliki, T. A. (2013). The preparatory year programs at universities: A superior goal .. but! *Al-Madina Online*. Retrieved from <http://www.al-madina.com/node/458408>
- Al-Otaibi, M. N. (2015). *A proposal for the management improvement of preparatory year programs at Saudi universities guided by current international experiences*. (Doctor Dissertation), Al-Imam Muhammad ibn Saud Islamic University, Riyadh.
- Al-Qahtani, N. (2010). Critical thinking of king Saud university and Emirates University students, school of education: A Comparative study. *Journal of Modern Education Association*, 3(7), 379-406.
- Al-Saadat, A. I. (2004). A close look at Saudi EFL teachers preparation in language testing. *Scientific Journal of King Faisal University (Humanities and Management Sciences)*, 5(2), 227-249.
- Al-Sameti, R., & Al-Seraji, H. (2016). The preparatory year: A direction or tertiary business? *Al-Madina Online*. Retrieved from <http://www.al-madina.com/node/665261>
- Al-Seghayer, K. (2005). Teaching English in the Kingdom of Saudi Arabia: Slowly but steadily changing. *Teaching English to the world: History, curriculum, and practice*, 125-134.
- Al-Seghayer, K. (2013). The actuality, inefficiency, and needs of EFL teacher-preparation programs in Saudi Arabia. *International Journal of Applied Linguistics & English Literature*, 3(1), 143-151.
- Al-Seghayer, K. (2014). The four most common constraints affecting English teaching in Saudi Arabia. *International Journal of English Linguistics*, 4(5), 17-26.
- Al-Seghayer, K. (2016). *Real face of Saudi Arabia: critical insider perspective on educational, lifestyle, and social issues in the Kingdom*. Riyadh: Hala Print Co.
- Al-Shehri, S. (2017). A developmental paradigm for English language instruction at preparatory year programs. *Arab World English Journal*, 8(3), 432-447.
- Al-Shumaimeri, Y. A. (2003). *A study of classroom exposure to oral pedagogic tasks in relation to the motivation and performance of Saudi secondary school learners of English in a context of potential curriculum reform*. (Doctor of Philosophy ), University of Leeds, Leeds.
- Al Asmari, A. (2016). Continuous professional development of English language teachers: Perception and practices. *Advances in Language and Literary Studies*, 7(3), 117-124.
- Al Boqami, S. (2014). *The Impact of job security on promoting creativity for administrative employees at King Abdulaziz University*. (Master ), King Abdulaziz University Jeddah, Saudi Arabia.
- Alaidarous, A. (2011). *Proposed perspective for development of comprehensive system for Saudi university staff performance evaluation from the university staff points of view in Holy Makkah area*. (Doctor Dissertation), Um Al Qura University, Makkah.

- Alamoudi, K., & Troudi, S. (2017). EFL teacher evaluation: A theoretical perspective. In S. Hidri & C. Coombe (Eds.), *Evaluation in foreign language education in the Middle East and North Africa* (pp. 29-41). London: Springer.
- Alamri, A. A. M. (2008). *An evaluation of the sixth grade English language textbook for Saudi boys' schools*. (MA thesis), King Saud University.
- Albaiz, T. A. (2016). Enhancement of higher education teaching of English in Saudi Arabia. *US-China Education Review A*, 6(6), 327-344.
- Albaqami, S. (2015). Implementing quality assurance in Saudi Arabia: a comparison between the MESO and the MICRO levels at PSU. *Higher Education Studies*, 5(3), 66-81.
- Alexander, E. (2016). *Teacher evaluation: The relationship between performance evaluation ratings and student achievement*. (Doctor of Education), Liberty University, Lynchburg.
- Alghamdi, A. K. H. (2014). The road to culturally relevant pedagogy: Expatriate teachers' pedagogical practices in the cultural context of Saudi Arabian higher education. *McGill Journal of Education/Revue des sciences de l'éducation de McGill*, 49(1), 201-226.
- Alharbi, J. S. (2015). Standards of the instructional performances in higher-education institutions; conformity to quality requirements: Arrass College of Science and Arts—mode. *Universal Journal of Educational Research*, 3(7), 427-434.
- AlKhars, D. A. M. A. (2013). *Creativity in English language teaching in Kuwait: A TESOL study*. (Doctor of Philosophy), University of Exeter, Exeter.
- Alrajhi, M. N., Alkharusi, H. A., & Aldhafri, S. S. (2016). Learning processes and academic achievement among Omani school students. *Review of European Studies*, 8(4), 62-71.
- Arora, S. (2014). *Defossilising the errors of ESL learners through feedback*. Paper presented at the Innovation in English Language Teacher Education, Hyderabad, India.
- Arreola, R. A. (2006). *Developing a comprehensive faculty evaluation system: A guide to designing, building, and operating large-scale faculty evaluation systems* (Vol. 86). San Francisco: Jossey-Bass.
- Asher, H. (2016). *Polling and the public: What every citizen should know*. London: SAGE Publications, Inc.
- Ashour, I. (2011). *A rhetorical critique of teacher supervision and the potential for pedagogy to serve the purposes of liberation in Syria*. (Doctor of Philosophy), Indiana University of Pennsylvania.
- Aslrasouli, M., & Vahid, M. S. P. (2014). An investigation of teaching anxiety among novice and experienced Iranian EFL teachers across gender. *Procedia-Social and Behavioral Sciences*, 98, 304-313.
- Au, W. (2011). Teaching under the new Taylorism: high-stakes testing and the standardization of the 21st century curriculum. *Journal of Curriculum Studies*, 43(1), 25-45.
- Austin, E. J., Saklofske, D. H., & Egan, V. (2005). Personality, well-being and health correlates of trait emotional intelligence. *Personality and Individual Differences*, 38(3), 547-558.
- Azike, A. A. (2014). Transformation Agenda: Using the R and E Model for Educational Reforms. *Educational Research International*, 3(6), 63-71.

- Bailey, K. M. (2006). *Language teacher supervision: A case-based approach*. Cambridge: Cambridge University Press.
- Bailey, M. R. (2016). Invasion of the Body Snatchers: Pedagogicide and Teacher Evaluation in the Age of Accountability. *Curriculum and Teaching Dialogue*, 18(1/2), 139-154.
- Barqawi, B., & Abuelrub, E. (2016). *Ittar Namouthaj Litaqweem Jawdat Ada' A'dha' Hai'at Attadrees (A framework for faculty performance evaluation)*. Paper presented at the The Sixth International Arab Conference on Quality Assurance in Higher Education.
- Barr, A. S. (1948). The Measurement and Prediction of Teaching Efficiency. *The Journal of Experimental Education*, 16(4), 203-283.
- Bayar, A. (2014). The components of effective professional development activities in terms of teachers' perspective. *International Online Journal of Educational Sciences*, 6(2), 319-327.
- Bell, T. R. (2005). Behaviors and attitudes of effective foreign language teachers: Results of a questionnaire study. *Foreign Language Annals*, 38(2), 259-270.
- Benton, S. L., & Cashin, W. E. (2014). Student ratings of instruction in college and university courses. In P. M. (Ed.), *Higher education: Handbook of theory and research* (Vol. 29, pp. 279-326). Dordrecht: Springer.
- BERA. (2011). Ethical Guidelines for Educational Research. Retrieved from <https://www.bera.ac.uk/researchers-resources/publications/ethical-guidelines-for-educational-research-2011>
- Borg, M. G. (1990). Occupational stress in British educational settings: A review. *Educational Psychology*, 10(2), 103-126.
- Borg, S. (2015). *Teacher cognition and language education: Research and practice*. London: Bloomsbury Publishing Plc.
- Bourdieu, P. (1977). *Outline of a Theory of Practice* (Vol. 16). Cambridge: Cambridge university press.
- Boyd, D., Grossman, P., Ing, M., Lankford, H., Loeb, S., & Wyckoff, J. (2011). The influence of school administrators on teacher retention decisions. *American Educational Research Journal*, 48(2), 303-333.
- Bradford, C. (2016). *What counts as quality teaching? The importance of sensemaking in teacher evaluation activities*. (Doctor of Philosophy), University of Wisconsin-Madison, Madison.
- Bradley, J. E., Mayfield, M. V., Mehta, M. P., & Rukonge, A. (2002). Participatory evaluation of reproductive health care quality in developing countries. *Social Science & Medicine*, 55(2), 269-282.
- Braskamp, L. A., & Ory, J. C. (1994). *Assessing faculty work: Enhancing individual and institutional performance*. San Francisco: Jossey-Bass
- Bridglall, B. L., Caines, J., & Chatterji, M. (2014). Understanding validity issues in test-based models of school and teacher evaluation. *Quality Assurance in Education*, 22(1), 19-30.
- Brinkman, S., & Kvale, S. (2015). Interviews: Learning the craft of qualitative research interviewing. *Aalborg*. Accessed January, 24, 2017.
- Brookfield, S. (2017). *Becoming a critically reflective teacher*. San Francisco: Jossey-Bass.

- Brophy, J., & Good, T. (1986). Teacher behaviour and student achievement. In M. Wittrock (Ed.), *Handbook of research on teaching* (3rd ed., pp. 328–375). New York: Macmillan.
- Brown, A. V. (2006). *Students' and teachers' perceptions of effective teaching in the foreign language classroom: A comparison of ideals and ratings*. (Doctor of Philosophy), The University of Arizona.
- Brown, I. I., & Crumpler, T. (2013). Assessment of foreign language teachers: A model for shifting evaluation toward growth and learning. *The High School Journal*, 96(2), 138-151.
- Brown, J. (2014). *Mixed methods research for TESOL*: Edinburgh University Press.
- Brown, J. D. (2008). 19 Research Methods for Applied Linguistics: Scope, Characteristics, and Standards. *The handbook of applied linguistics*, 476.
- Bryman, A. (2015). *Social research methods*: Oxford university press.
- Burden, P. (2010). Creating confusion or creative evaluation? The use of student evaluation of teaching surveys in Japanese tertiary education. *Educational Assessment, Evaluation and Accountability*, 22(2), 97-117.
- Burden, P., & Troudi, S. (2007). An evaluation of student ratings of teaching in a Japanese university context. In M. A.-H. C. Coombe, P. Davidson, & S. Troudi (Ed.), *Evaluating teacher effectiveness in ESL/EFL contexts* (4 ed., pp. 152-166). Michigan: The University of Michigan Press.
- Burr, D., & Hooser, M. (1995). Alterations to the en bloc basic fuchsin staining protocol for the demonstration of microdamage produced in vivo. *Bone*, 17(4), 431-433.
- Carbone, A., Ross, B., Phelan, L., Lindsay, K., Drew, S., Stoney, S., & Cottman, C. (2015). Course evaluation matters: improving students' learning experiences with a peer-assisted teaching programme. *Assessment & Evaluation in Higher Education*, 40(2), 165-180.
- Cardoso, S., Rosa, M. J., & Stensaker, B. (2016). Why is quality in higher education not achieved? The view of academics. *Assessment & Evaluation in Higher Education*, 41(6), 950-965.
- Carraway, J. H., & Young, T. (2015). Implementation of a districtwide policy to improve principals' instructional leadership: Principals' sensemaking of the skillful observation and coaching laboratory. *Educational Policy*, 29(1), 230-256.
- Castleberry, A. (2014). NVivo 10 [software program]. Version 10. QSR International; 2012. *American Journal of Pharmaceutical Education*, 78(1).
- Centra, J. (1977). The how and why of evaluating teaching. *h'mt Directions for Htghtr Educatan*, 17, 93-106.
- Chism, N. V. N. (1999). *Peer Review of Teaching: A Sourcebook*. Bolton, MA Anker Publishing CO
- Chouinard, J. A., & Cousins, J. B. (2015). The journey from rhetoric to reality: Participatory evaluation in a development context. *Educational Assessment, Evaluation and Accountability*, 27(1), 5-39.
- Christie, C. A., & Alkin, M. C. (2013). An evaluation theory tree. In M. C. Alkin (Ed.), *Evaluation roots: A wider perspective of theorists' views and influences* (2nd ed., pp. 12-65). London: SAGE Publications, Inc.
- Clark, C., & Peterson, P. (1986). Teachers' thought process. In M. Wittrock (Ed.), *Handbook of research on teaching* (pp. 255–296). New York: Macmillan.

- Codd, J. A. (1994). Educational reform and the contradictory discourses of evaluation. *Evaluation & Research in Education*, 8(1-2), 41-54.
- Cohen, L., Manion, L., & Morrison, K. (2013). *Research methods in education*. London: Routledge.
- Colby, S. A., Bradshaw, L. K., & Joyner, R. L. (2002). *Teacher evaluation: A review of the literature*. Paper presented at the Annual Meeting of the American Educational Research Association, New Orleans, LA.
- Copland, F. (2008). Deconstructing the discourse: Understanding the feedback event. In G. S. & R. K. (Eds.), *Professional encounters in TESOL* (pp. 5-23). London: Springer.
- Cousins, J. B. (2003). Utilization effects of participatory evaluation. In T. Kellaghan & D. L. Stufflebeam (Eds.), *International Handbook of Educational Evaluation* (pp. 245-265). Dordrecht: Springer Netherlands.
- Cousins, J. B. (2004). Commentary: Minimizing evaluation misuse as principled practice. *American Journal of Evaluation*, 25(3), 391-397.
- Cousins, J. B., & Earl, L. M. (1992). The case for participatory evaluation. *Educational Evaluation and Policy Analysis*, 14(4), 397-418.
- Cousins, J. B., & Earl, L. M. (1995). The case for participatory evaluation: Theory, research, practice. *Participatory evaluation in education: Studies in evaluation use and organizational learning*, 3-18.
- Cowan, J., & Goldhaber, D. (2015). *National Board certification and teacher effectiveness: Evidence from Washington*. Retrieved from Washington Bothell: <http://www.CEDR.us/publications.html>
- Creswell, J. (2017). *Research design: Qualitative, quantitative, and mixed methods approaches*. London: SAGE Publications, Inc.
- Creswell, J. W. (2012). *Educational research: planning, conducting, and evaluating quantitative and qualitative research*. Boston: Pearson Education.
- Creswell, J. W. (2013). *Research design: Qualitative, quantitative, and mixed methods approaches*. London: SAGE Publications, Inc.
- Creswell, J. W., & Clark, V. L. P. (2016). *Designing and Conducting Mixed Methods Research*. SAGE.
- Creswell, J. W., Gutmann, M. L., & Hanson, W. E. (2010). Advanced mixed methods research designs. In A. Tashakkori & C. Teddlie (Eds.), *Handbook of mixed methods in social and behavioral research* (Vol. 209, pp. 240).
- Creswell, J. W., & Poth, C. N. (2017). *Qualitative inquiry and research design: Choosing among five approaches*. London: SAGE Publications, Inc.
- Crotty, M. (1998). *The foundations of social research: Meaning and perspective in the research process*. London: SAGE Publications, Inc.
- Cullen, A. E., Coryn, C. L., & Rugh, J. (2011). The politics and consequences of including stakeholders in international development evaluation. *American Journal of Evaluation*, 32(3), 345-361.
- Curry, M. (2008). Critical friends groups: The possibilities and limitations embedded in teacher professional communities aimed at instructional improvement and school reform. *Teachers College Record*, 110(4), 733-774.
- Daniel, J. (2011). *Sampling essentials: Practical guidelines for making sampling choices*. London: SAGE Publications, Inc.
- Danielson, C. (2011). *Enhancing professional practice: A framework for teaching*. Alexandria: ASCD.

- Danielson, C., & McGreal, T. L. (2000). *Teacher evaluation to enhance professional practice*. Alexandria: ASCD.
- Darandari, E. Z., Al-Qahtani, S. A., Allen, I. D., Al-Yafi, W. A., Al-Sudairi, A. A., & Catapang, J. (2009). The quality assurance system for post-secondary education in Saudi Arabia: A comprehensive, developmental and unified approach. *Quality in Higher Education*, 15(1), 39-50.
- Daresh, J. C. (2007). *Supervision as proactive leadership*. Prospect Heights, IL: Waveland Press Inc.
- Darling-Hammond, L. (2000). Teacher quality and student achievement. *Education policy analysis archives*, 8(1), 1-44.
- Darling-Hammond, L. (2013). *Getting teacher evaluation right: What really matters for effectiveness and improvement*. New York: Teachers College Press.
- Davidson, P. (2010). Faculty attitude towards three methods of evaluating teacher effectiveness. In C. Coombe, M. Al-Hamly, P. Davidson, & S. Troudi (Eds.), *Evaluating teacher effectiveness in ESL/EFL contexts* (4 ed., pp. 183-195). Michigan: The University of Michigan Press.
- Davison, C., & Leung, C. (2009). Current issues in English language teacher-based assessment. *Tesol Quarterly*, 43(3), 393-415.
- Day, C., Sammons, P., & Stobart, G. (2007). *Teachers matter: Connecting work, lives and effectiveness*. Berkshire: McGraw-Hill Education.
- De Boer, H. (2013). Trust, the essence of governance. *Governing higher education: National perspectives on institutional governance*, 43-61.
- Delvaux, E., Vanhoof, J., Tuytens, M., Vekeman, E., Devos, G., & Van Petegem, P. (2013). How may teacher evaluation have an impact on professional development? A multilevel analysis. *Teaching and Teacher Education*, 36, 1-11.
- Denzin, N. K., & Giardina, M. D. (2016). *Ethical futures in qualitative research: Decolonizing the politics of knowledge*. London: Routledge.
- Dillman, L. M. (2013). Comparing evaluation activities across multiple theories of practice. *Evaluation and program planning*, 38, 53-60.
- DiPaola, M. F., & Hoy, W. K. (2008). *Principals improving instruction: Supervision, evaluation, and professional development*. Charlotte, NC: Information Age Publishing, Inc.
- Disterheft, A., Caeiro, S., Azeiteiro, U. M., & Filho, W. L. (2013). Sustainability Science and Education for Sustainable
- Development in Universities: A Way for Transition. In S. Caeiro, W. Leal Filho, C. Jabbour, & U. M. Azeiteiro (Eds.), *Sustainability assessment tools in higher education institutions* (pp. 3-28). London: Springer.
- Dodd, T. J. (2015). *Influences on principal evaluations of teacher performance*. (Doctor of Philosophy), The Pennsylvania State University, Pennsylvania.
- Doherty, J. F. (2009). *Perceptions of Teachers and Administrators in a Massachusetts Suburban School District Regarding the Implementation of a Standards-Based Teacher Evaluation System*. (Doctor of Education), Seton Hall University.
- Domas, S. J., & Tiedeman, D. V. (1950). Teacher competence: An annotated bibliography. *The Journal of Experimental Education*, 19(2), 101-218.
- Dörnyei, Z. (2007). Research methods in applied linguistics.



- Dörnyei, Z., & Taguchi, T. (2009). *Questionnaires in second language research: Construction, administration, and processing*. London: Routledge Ltd.
- Downing, P. R. (2016). *Effects of Teacher Evaluation on Teacher Job Satisfaction in Ohio*. (Doctor of Education), The University of Findlay.
- Drever, E. (2006). *Using Semi-structured Interviews in Small-scale Research: A Teacher's Guide*. Glasgow: SCRE Publication129.
- Dunne, F., Nave, B., & Lewis, A. (2000). Critical friends groups: Teachers helping teachers to improve student learning. *Phi Delta Kappan*, 28(4), 31-37.
- Dwyer, C. A. (2012). Criteria for performance-based teacher assessments: validity, standards and issues In A. J. Shinkfield & D. L. Stufflebeam (Eds.), *Teacher evaluation: Guide to effective practice* (Vol. 41, pp. 62-80). Boston: Springer Science & Business Media.
- Eaton, J. (2015). Quality and an international higher education sapce. *International higher education*, 40, 3-4.
- Elyas, T., & Al Grigri, W. H. (2014). Obstacles to teaching English in Saudi Arabia public schools: Teachers' and supervisors' perceptions. *International Journal of Language Teaching*, 2(3), 74-89.
- Evans, C. M., Lee, J. C., & Thompson, W. C. (2016). "First, do no harm?": A framework for ethical decision-making in teacher evaluation *Student Growth Measures in Policy and Practice* (pp. 169-188). London: Springer.
- Fageeh, A. I. (2003). *Saudi college students' beliefs regarding their English writing difficulties*. (Doctor of Philosophy), Indiana University of Pennsylvania, Pennsylvania.
- Fareh, S. (2010). Challenges of teaching English in the Arab world: Why can't EFL programs deliver as expected? *Procedia-Social and Behavioral Sciences*, 2(2), 3600-3604.
- Farrell, T. S. (2014). *Promoting teacher reflection in second language education: A framework for TESOL professionals*. London: Routledge.
- Faruk, S. M. G., & Rahman, M. R. (2016). "Excellent results" of teaching listening at a Saudi university: Appearance and reality. *International Journal of Humanities and Cultural Studies*, 2(1), 201-212.
- Field, A. (2013). *Discovering statistics using IBM SPSS statistics*. London: SAGE Publications, Inc.
- Flick, U. (2014). *An introduction to qualitative research*. London: SAGE Publications, Inc.
- Fominaya, C. F. (2014). *Social movements and globalization: How protests, occupations and uprisings are changing the world*. Hampshire, UK: Palgrave Macmillan.
- Freire, P. (2000). *Pedagogy of the oppressed*. New York: Bloomsbury Publishing.
- Freire, P. (2014). *Pedagogy of hope: Reliving pedagogy of the oppressed*. New York: Bloomsbury Publishing.
- Fulton, O. (2013). Higher education governance in the UK: Change and continuity *Governing higher education: National perspectives on institutional governance* (pp. 187-211). London: Springer.
- Furedi, F. (2010). Introduction to the marketisation of higher education and the student as consumer. In M. Molesworth, R. Scullion, & E. Nixon (Eds.), *The marketisation of higher education and the student as consumer* (pp. 1-8). London: Routledge.

- Galbraith, C. S., & Merrill, G. B. (2012). Predicting student achievement in university-level business and economics classes: Peer observation of classroom instruction and student ratings of teaching effectiveness. *College Teaching*, 60(2), 48-55.
- Garrick, J. (1999). Doubting the philosophical assumptions of interpretive research. *International journal of qualitative studies in education*, 12(2), 147-156.
- Genesee, F. (2001). Evaluation. In D. N. Ronald Carter (Ed.), *The Cambridge Guide to Teaching English to Speakers of Other Languages* (pp. 144-150). Cambridge: Cambridge University Press.
- Getzels, J. W., & Jackson, P. W. (1963). The teacher's personality and characteristics. In N. Gage (Ed.), *Handbook of research on teaching* (pp. 506-582). Chicago: Rand McNally.
- Glanz, J. (1998). Histories, antecedents, and legacies of school supervision. *Handbook of research on school supervision*, 39-79.
- Glasman, N. S. (1986). *Evaluation-based leadership: School administration in contemporary perspective*. New York: SUNY Press.
- Glisan, E. W. (2005). Leaving no foreign language teachers behind: A grassroots response. *The Modern Language Journal*, 89(2), 268-272.
- Goldhaber, D. (2015). Exploring the potential of value-added performance measures to affect the quality of the teacher workforce. *Educational Researcher*, 44(2), 87-95.
- Goldstein, J. (2007). Easy to dance to: Solving the problems of teacher evaluation with peer assistance and review. *American Journal of Education*, 113(3), 479-508.
- Green, E. (2010). Building a better teacher. *New York Times Magazine*, 1-9.
- Greene, J. C. (2007). Evaluation, democracy, and social change. In I. G. Shaw, J. C. Greene, & M. M. Mark (Eds.), *The Sage handbook of evaluation* (pp. 118-140). London: SAGE Publications, Inc.
- Groen, J. F. (2017). Engaging in enhancement: Implications of participatory approaches in higher education quality assurance. *Collected Essays on Learning and Teaching*, 10, 89-100.
- Guba, E. G., & Lincoln, Y. S. (1981). *Effective evaluation: Improving the usefulness of evaluation results through responsive and naturalistic approaches*. San Francisco: Jossey-Bass.
- Gullatt, D. E., & Ballard, L. M. (1998). Choosing the right process for teacher evaluation. *American Secondary Education*, 26(3), 13-17.
- Gutmann, L. (2014). *How teacher evaluation shapes conceptions of good practice: Policy intention, implementation & interpretation*. (Doctor of Education), The University of North Carolina at Chapel Hill, North Carolina.
- Habbash, M. M. R. a. (2011). *Status change of English and its role in shaping public education language policy and practice in Saudi Arabia: A postmodernist critical perspective*. (Doctor of Philosophy), University of Exeter, Exeter.
- Habermas, J. (1994). Three normative models of democracy. *Constellations*, 1(1), 1-10.
- Hakim, B. (2015). Teacher evaluation as a tool for professional development: A case of Saudi Arabia. *Advances in Language and Literary Studies*, 6(5), 97-103.

- Hall, T. R. (2013). *Saudi male perceptions of study in the United States: An analysis of King Abdullah Scholarship program participants*. (Doctor of Education), Western Kentucky University, Kentucky.
- Hanberger, A. (2001). Policy and program evaluation, civil society, and democracy. *American Journal of Evaluation*, 22(2), 211-228.
- Hanberger, A. (2004). *Democratic governance and evaluation*. Paper presented at the Uppsats presenterad vid sjätte EES (European Evaluation Society) konferensen, 30/9-2/10 2004, Berlin., Berlin, Germany.
- Hawkins, M., & Norton, B. (2009). Critical language teacher education. In A. Burns & J. Richards (Eds.), *Cambridge guide to second language teacher education* (pp. 30-39). Cambridge: Cambridge University Press.
- Henderson, B. B., & Buchanan, H. E. (2007). The scholarship of teaching and learning: A Special Niche for Faculty at Comprehensive Universities? *Research in Higher Education*, 48(5), 523.
- Hooton, N. K. (2008). The design of post-observation feedback and its impact on student teachers *Professional encounters in TESOL* (pp. 24-41). London: Springer.
- Horn, I. S., Kane, B. D., & Wilson, J. (2015). Making sense of student performance data: Data use logics and mathematics teachers' learning opportunities. *American Educational Research Journal*, 52(2), 208-242.
- House, E., & Howe, K. R. (1999). *Values in evaluation and social research*. London: SAGE Publications, Inc.
- Howard, A. J. (2010). *Teacher appraisal: The impact of observation on teachers' classroom behaviour*. (Doctor of Philosophy), University of Warwick, Coventry.
- Howard, B. B., & Gullickson, A. R. (2010). Setting standards for teacher evaluation. In K. M. M. (Ed.), *Teacher assessment and the quest for teacher quality: A handbook* (pp. 337-353). San Francisco: Jossey-Bass.
- Howells, G., & Meiring, L. (2015). Towards greater objectivity in classroom observation. *Revista de Linguas para Fines Específicos*, 6, 239-256.
- Hughes, S. P. (2007). *The identification of quality indicators in English language teaching: A study in compulsory and non-compulsory secondary level language education in the province of Granada*. (Doctor Thesis Doctoral Thesis), University of Granada, Granada.
- Jenlink, P. M. (2016). Democracy Distracted in an Era of Accountability Teacher Education Against Neoliberalism. *Cultural Studies ↔ Critical Methodologies*, 1-10.
- Johnson, R. B., & Onwuegbuzie, A. J. (2004). Mixed methods research: A research paradigm whose time has come. *Educational Researcher*, 33(7), 14-26.
- Julnes, G., & Mark, M. M. (1998). Evaluation as sensemaking: Knowledge construction in a realist world. *New directions for evaluation*(78), 33-52.
- Kang, G. (2015). *The power of teachers: Exploring identity, agency, and collaboration in the context of literacy*. (Doctor of Philosophy), University of Illinois at Urbana-Champaign, Illinois.
- Karanxha, Z., Agosto, V., & Bellara, A. (2014). Modeling social justice educational leadership: Self-assessment for equity (SAFE) *International Handbook of Educational Leadership and Social (In) Justice* (pp. 1187-1206). Dordrecht: Springer.

- Khan, I. A. (2011). Professional development of English teachers: The Saudi Arabian context. *Interdisciplinary Journal of Contemporary Research in Business*, 3(2), 1583-1591.
- Kim, D., Wolf-Wendel, L., & Twombly, S. (2011). International Faculty: Experiences of Academic Life and Productivity in U.S. Universities. *The Journal of Higher Education*, 82(6), 720-747.
- King, F. (2014). Evaluating the impact of teacher professional development: an evidence-based framework. *Professional Development in Education*, 40(1), 89-111.
- King, M. (2015). Evaluating experienced teachers. In A. Howard & H. Donaghue (Eds.), *Teacher Evaluation in Second Language Education* (pp. 167-180). London: Bloomsbury Publishing Plc.
- Klem, A. M., & Connell, J. P. (2004). Relationships matter: Linking teacher support to student engagement and achievement. *Journal of school health*, 74(7), 262-273.
- Klinger, D. A., DeLuca, C., & Merchant, S. (2016). The intersection of international achievement testing and educational policy development. In L. Volante (Ed.), *The Intersection of International Achievement Testing and Educational Policy: Global Perspectives on Large-Scale Reform* (pp. 140–159). New York: Routledge.
- Korenev, A., Westbrook, C., Merry, Y., & Ershova, T. (2016). Language Teachers' Target Language project: Language for specific purposes of language teaching. *Language Teaching*, 49(1), 146-148.
- Krieger, Z. (2007). Saudi Arabia puts its billions behind western-style higher education. *Chronicle of Higher Education*, 54(3), 1-6.
- Krishna, V. (2004). *School Psychology*. New Delhi: Discovery Publishing House.
- Krueger, R. A., & Casey, M. A. (2014). *Focus groups: A practical guide for applied research*. London: SAGE Publications, Inc.
- Kushner, S. (2016). Living Life at the Edge: Democracy, Equity and Pragmatics in Evaluation. *Evaluation for an Equitable Society*.
- Kyriakides, L., & Demetriou, D. (2007). Introducing a teacher evaluation system based on teacher effectiveness research: an investigation of stakeholders' perceptions. *Journal of Personnel Evaluation in Education*, 20, 43-64.
- Lee, T. (2016). Outcome mapping for planning evaluations in American K-12 urban education: Potent possibilities. *Education*. Retrieved from <https://www.slideshare.net/drtlee/outcome-mapping-for-planning-evaluations-in-american-k12-urban-education-potent-possibilities-57017594>
- Lichtman, M. (2012). *Qualitative research in education: A User's Guide: A user's guide*. London: SAGE Publications, Inc.
- Litosseliti, L. (2017). *Research methods in linguistics*: Bloomsbury Publishing.
- Llurda, E. (2016). 'Native speakers', English and ELT. In G. Hall (Ed.), *The Routledge Handbook of English Language Teaching* (pp. 51-63). London: Routledge.
- Mackenzie, N., & Knipe, S. (2006). Research dilemmas: Paradigms, methods and methodology. *Issues in educational research*, 16(2), 193-205.
- Mahamadou, G. N. (2010). *Teacher professional development: The needs of TESOL teachers in the Republic of Niger*. (Doctor of Philosophy), Indiana University of Pennsylvania.

- Mallows, D. (2002). Non-linearity and the observed lesson. *ELT journal*, 56(1), 3-10.
- Mann, S. J. (2016). *The research interview*. Dordrecht: Springer.
- Martin, B., Breunig, M., Wagstaff, M., & Goldenberg, M. (2017). *Outdoor leadership: Theory and practice*. US: Human Kinetics.
- Mathison, S., & Ross, W. (2002). The hegemony of accountability in schools and universities. *Workplace*, 9, 88-102.
- Mazandarani, O. (2014). *EFL lecturers' perceptions of teacher effectiveness and teacher evaluation in Iranian universities*. University of Exeter.
- McCaffrey, D. F., Lockwood, J., Koretz, D. M., & Hamilton, L. S. (2003). *Evaluating value-added models for teacher accountability*. Santa Monica, CA: RAND Corporation.
- McGee, R., & Gaventa, J. (2011). *Shifting power? Assessing the impact of transparency and accountability initiatives (2040-0209)*. Retrieved from London:
- McGreal, T. L. (1983). *Successful Teacher Evaluation*. Alexandria: ASCD Publications.
- McMullen, M. G. (2014). The value and attributes of an effective preparatory English program: Perceptions of Saudi university students. *English Language Teaching*, 7(7), 131-140.
- Medley, D. M. (1979). The effectiveness of teachers. In P. L. Peterson & H. Walberg (Eds.), *Research on teaching: Concepts, findings, and implications* (pp. 11-26). Berkeley: McCutchan.
- Medley, D. M., Coker, H., & Soar, R. S. (1984). *Measurement-based evaluation of teacher performance: An empirical approach*. London: Longman Publishing Group.
- Melibari, G. (2016). *ELT teaching quality and practice in Saudi Arabia: A case study of the perspectives of ESP and EGP students, teachers and managers at the ELC in Umm al-Qura University*. (Doctor of Philosophy), University of Hertfordshire, Hertfordshire.
- Merç, A. (2015). Assessing the Performance in EFL teaching practicum: Student teachers' views. *International Journal of Higher Education*, 4(2), 44-56.
- Michener, G., & Bersch, K. (2013). Identifying transparency. *Information Polity*, 18(3), 233-242.
- Miles, M. B., Huberman, A. M., & Saldaña, J. (2014). *Qualitative data analysis: A methods sourcebook*. London: SAGE Publications, Inc.
- Mills, C. (2015). Implications of the My School website for disadvantaged communities: A Bourdieuan analysis. *Educational Philosophy and Theory*, 47(2), 146-158.
- Ministry of Education, K. o. S. A. (2013). General goals for teaching english in saudi arabia. Retrieved from <https://www.slideshare.net/tamma07/general-goals-for-teaching-english-in-saudi-arabia-pdf>
- Ministry of Education, K. o. S. A. (2015). Dr Al Sabti: Budget reflects leadership valuation of HE crucial role at building capital and knowledge society. Retrieved from <http://www.mohe.gov.sa/en/news/Pages/29-12-2014.aspx>
- Ministry of Education, K. o. S. A. (2017). Higher Education Statistics Retrieved from <https://www.moe.gov.sa/en/HigherEducation/governmenthighereducation/Pages/AdmissionStatistics.aspx>

- Ministry of Higher Education, K. o. S. A. (1999). System of higher education. Retrieved from <http://www.mohe.gov.sa/en/news/Pages/11-12-2014.aspx>
- Mohan, N. (2011). On the use of non tenure track faculty and the potential effect on classroom content and student evaluation of teaching. *Journal of Financial Education*, 37(1), 29-42.
- Monin, P., Noorderhaven, N., Vaara, E., & Kroon, D. (2013). Giving sense to and making sense of justice in postmerger integration. *Academy of Management Journal*, 56(1), 256-284.
- Moraya, A. (2013). *The influence of culture and cultural background knowledge on Saudi students' English language learning*. (Master of Art), Southern Illinois University Carbondale, Carbondale.
- Morgan, D. L. (1996). Focus groups. *Annual review of sociology*, 22(1), 129-152.
- Muijs, D. (2010). *Doing quantitative research in education with SPSS*. London: SAGE Publications, Inc.
- Murdoch, G. (2000). Introducing a teacher-supportive evaluation system. *ELT journal*, 54(1), 54-64.
- Murphy, R. (2013). *Testing teachers: What works best for teacher evaluation and appraisal*. Paper presented at the Improving social mobility through education.
- Nagel, C. I. (2012). *Teacher perceptions regarding portfolio-based components of teacher evaluations*. (Doctor of Education), Illinois State University.
- Naidoo, R., & Williams, J. (2015). The neoliberal regime in English higher education: Charters, consumers and the erosion of the public good. *Critical Studies in Education*, 56(2), 208-223.
- NCAAA. (2015a). *Handbook for Quality Assurance and Accreditation in Saudi Arabia, Part1, The System for Quality Assurance and Accreditation*. Saudi Arabia.
- NCAAA. (2015b). *Handbook for Quality Assurance and Accreditation in Saudi Arabia, Part3, External Reviews for Accreditation and Quality Assurance*. Saudi Arabia.
- NCAAA. (2015c). *Handbook for Quality Assurance and Accreditation in Saudi Arabia Part2, Internal Quality Assurance Arrangements*. Saudi Arabia.
- Neuman, L. W. (2002). *Social research methods: Qualitative and quantitative approaches*.
- Newman, A., & McNamara, Y. (2016). Teaching qualitative research and participatory practices in neoliberal times. *Qualitative Social Work*, 15(3), 428-443.
- Newman, K. (2008). Whose view matters? Using participatory processes to evaluate Reflect in Nigeria. *Community Development Journal*, 43(3), 382-394.
- Nordensvard, J. (2010). The consumer metaphor versus the citizen metaphor: different sets of roles for students. In M. Molesworth, R. Scullion, & E. Nixon (Eds.), *The marketization of higher education and the student as consumer* (pp. 157-169). London: Routledge.
- Norris, J., & Mills, N. (2014). *Innovation and accountability in foreign language program evaluation*. Paper presented at the Language Program Evaluation Symposium, Harvard University.
- Norris, N. (1990). *Understanding educational evaluation*. London: Kogan Page Ltd.
- Ofir, Z., Kumar, S., & Kuzmin, A. (2013). Evaluation in developing countries: What makes it different. In S. I. Donaldson, T. Azzam, & R. F. Conner (Eds.),

- Emerging practices in international development evaluation* (pp. 11-24).  
Charlotte, NC: Information Age Publishers.
- Ohio Department of Education, O. (2015, 12/23/15). OTES Model Packet.  
Retrieved from  
<http://education.ohio.gov/getattachment/Topics/Teaching/Educator-Evaluation-System/Ohio-s-Teacher-Evaluation-System/Teacher-Performance-Ratings/OTES-Model-122315.pdf.aspx>
- Ohio Department of Education, O. (2016, 9/16/2016). Teacher Evaluation.  
Retrieved from <http://education.ohio.gov/Topics/Teaching/Educator-Evaluation-System/Ohio-s-Teacher-Evaluation-System/Teacher-Evaluation-FAQs#FAQ1032>
- Ohio Department of Education, O. (2017, 7/25/2017 ). Educator evaluation overview.  
Retrieved from <http://education.ohio.gov/Topics/Teaching/Educator-Evaluation-System/Educator-Evaluation-Overview>
- Oppenheim, A. N. (2000). *Questionnaire design, interviewing and attitude measurement*: Bloomsbury Publishing.
- Osmond-Johnson, P. (2015). *Alternative discourses of teacher professionalism: A study of union active teachers and teacher unions*. (Doctor of Philosophy), University of Toronto, Toronto.
- Ostovar-Nameghi, S. A., & Sheikahmadi, M. (2016). From teacher isolation to teacher collaboration: Theoretical perspectives and empirical findings. *English Language Teaching*, 9(5), 197-205.
- Pallant, J. (2013). *SPSS survival manual*: McGraw-Hill Education (UK).
- Papay, J. (2012). Refocusing the debate: Assessing the purposes and tools of teacher evaluation. *Harvard Educational Review*, 82(1), 123-141.
- Patton, M. Q. (2008). *Utilization-focused evaluation*. London: SAGE Publications, Inc.
- Peeters, A., & Robinson, V. (2015). A Teacher Educator Learns How to Learn from Mistakes: Single and Double-loop Learning for Facilitators of In-service Teacher Education. *Studying Teacher Education*, 11(3), 213-227.
- Pennycook, A. (2004). Critical moments in a TESOL praxicum. In B. Norton & K. Toohey (Eds.), *Critical pedagogies and language learning* (pp. 327-345). Cambridge: Cambridge University Press.
- Pennycook, A. (2017). *The cultural politics of English as an international language*. London: Routledge.
- Percy, A. J. (2015). A critical turn in higher education research: Turning the critical lens on the academic language and learning educator. *Discourse: Studies in the Cultural Politics of Education*, 36(6), 881-893.
- Peters, M. A. (2017). *Education for ecological democracy*: Taylor & Francis.
- Peterson, K. D. (2000). *Teacher evaluation: A comprehensive guide to new directions and practices*. Thousand Oaks, California: Corwin Press.
- Petrides, K. V., & Furnham, A. (2001). Trait emotional intelligence: Psychometric investigation with reference to established trait taxonomies. *European journal of personality*, 15(6), 425-448.
- Philip, T. M., Martinez, D. C., Lopez, E., & Garcia, A. (2016). Toward a teacher solidarity lens: former teachers of color (re) envisioning educational research. *Race ethnicity and education*, 19(1), 182-199.
- Podems, D. (2017). *DEMOCRATIC EVALUATION AND A NEW DEMOCRACY*.

- Popham, W. J. (2000). BIG CHANGE QUESTIONS." SHOULD LARGE-SCALE ASSESSMENT BE USED FOR ACCOUNTABILITY?"--Answer: Depends on the Assessment, Silly! *Journal of Educational Change*, 1(3), 283-289.
- Poster, C., Poster, D., & Benington, M. (1991). *Teacher appraisal: A guide to training*. London: Routledge.
- Powell, M., & Beard, J. W. (1984). Teacher effectiveness. *An Annotated Bibliography and Guide to Research*. Garland.
- Power, M. (1997). *The audit society: Rituals of verification*: OUP Oxford.
- Price, J. (2004). Research Limitations and the Necessity of Reporting Them. *American Journal of Health Education*, 35(2), 66-67.
- Pring, R. (2017). Importance of philosophy in the conduct of educational research. *Journal of International and Comparative Education (JICE)*, 23-30.
- Punch, K. F. (2013). *Introduction to social research: Quantitative and qualitative approaches*. London: SAGE Publications, Inc.
- Quirke, P. (2010). A coherent approach to faculty appraisal. In C. Coombe, M. Al-Hamly, P. Davidson, & S. Troudi (Eds.), *Evaluating teaching effectiveness in EFL/ESL contexts* (4 ed., pp. 89-105): Ann Arbor: University of Michigan Press
- Rahman, M. M. u. (2011). *ELT in Saudi Arabia: A study of learners' needs analysis*. Germany: LAP LAMBERT Academic Publishing.
- Rahman, M. M. u., & Alhaisoni, E. (2013). Teaching English in Saudi Arabia: prospects and challenges. *Academic Research International*, 4(1), 112-118.
- Ravitch, D. (2016). *The death and life of the great American school system: How testing and choice are undermining education*: Basic Books.
- Reddy, L. A., Dudek, C. M., Kettler, R. J., Kurz, A., & Peters, S. (2016). Measuring educators' attitudes and beliefs about evaluation: Construct validity and reliability of the teacher evaluation experience scale. *Educational Assessment*, 21(2), 120-134.
- Rehman, A. A., & Al-Bargi, A. (2014). Teachers' perspectives on post observation conferences: A study at a Saudi Arabian university. *Theory and Practice in Language Studies*, 4(8), 1558-1568.
- Rennert-Ariev, P. (2008). The hidden curriculum of performance-based teacher education. *Teachers College Record*, 110(1), 105-138.
- Richards, L. (2014). *Handling qualitative data: A practical guide*. London: SAGE Publications, Inc.
- Robinson, G. W. (2009). *Principals' perceptions regarding the role of the professional development and appraisal system in teacher supervision in Texas*. (Doctor of Education), University of Houston.
- Robles-Gómez, A., Ros, S., Hernández, R., Tobarra, L., Caminero, A. C., & Agudo, J. M. (2015). User acceptance of a proposed self-evaluation and continuous assessment system. *Journal of Educational Technology & Society*, 18(2), 97-109.
- Robson, C. (1993). Real world research: A resource for social scientists and practitioners-researchers. *Massachusetts: Blackwell Publishers*.
- Rossi, P. H., Lipsey, M. W., & Freeman, H. E. (2003). *Evaluation: A systematic approach*. London: SAGE Publications, Inc.



- Rowan, B. (1990). Commitment and control: Alternative strategies for the organizational design of schools. *Review of research in education*, 16, 353-389.
- Ryan, K. E. (2004). Serving public interests in educational accountability: Alternative approaches to democratic evaluation. *American Journal of Evaluation*, 25(4), 443-460.
- Safta, C. G. (2016). The ethical dimension of school evaluations. *Jus et Civitas*, 67(1), 47-54.
- Sandberg, J. (2005). How do we justify knowledge produced within interpretive approaches? *Organizational research methods*, 8(1), 41-68.
- Sanif, S. (2015). *Implementation of peer observation of teaching in elt tertiary education system in malaysia: a social-constructionism view*. (Doctor of Philosophy), University of Exeter, Exeter.
- Santiago, P., Roseveare, D., Van Amelsvoort, G., Manzi, J., & Matthews, P. (Writers). (2009). Teacher evaluation in Portugal: OECD review.
- Sapsford, R. (2006). Methodology. In V. Jupp (Ed.), *The Sage dictionary of social research methods* (pp. 176-178). London: SAGE Publications, Inc.
- Schumpeter, J. A. (2013). *Capitalism, socialism and democracy*. London: Routledge.
- Seldin, P. (2006). *Evaluating faculty performance: A practical guide to assessing teaching, research, and service*. San Francisco: Anker Publishing Company, Inc.
- Seldin, P., Miller, J. E., & Seldin, C. A. (2010). *The teaching portfolio: A practical guide to improved performance and promotion/tenure decisions*: John Wiley & Sons.
- Shah, M., Wilson, M., & Nair, C. S. (2010). *The Australian higher education quality assurance framework its success*. Paper presented at the The Australian Associations for Institutional Research conference, Mel-bourne, Australia
- Shah, S. R. (2016). *Teacher leadership: A case study of teacher leaders' professional development in an EFL institute of a Saudi Arabian university*. (Doctor of Education), University of Exeter, Exeter.
- Shah, S. R., & AlHarthi, K. (2014). TESOL classroom observations: a boon or a bane? An exploratory study at a Saudi Arabian University. *Theory and Practice in Language Studies*, 4(8), 1593-1602.
- Shannon-Baker, P. (2016). Making paradigms meaningful in mixed methods research. *Journal of Mixed Methods Research*, 10(4), 319-334.
- Shenton, A. K. (2004). Strategies for ensuring trustworthiness in qualitative research projects. *Education for information*, 22(2), 63-75.
- Shough, L. L. (2010). *Standards-based teacher evaluation: A professional development tool for improving instruction*: Arizona State University.
- Shukri, N. (2014). Exploring female teachers' perceptions towards teacher observation: Issues and challenges in the Arab context. *Arab World English Journal*, 5(2), 153 -166.
- Skaalvik, E. M., & Skaalvik, S. (2015). Job satisfaction, stress and coping strategies in the teaching profession - what do teachers say? . *International Education Studies*, 8(3), 181-192.
- Smith, E. (2015). Professional development and the impact of teacher relationships. *Journal of Initial Teacher Inquiry*, 1, 43-44.

- Smits, P. A., & Champagne, F. (2008). An assessment of the theoretical underpinnings of practical participatory evaluation. *American Journal of Evaluation, 29*(4), 427-442.
- Smylie, M. (2014). Teacher evaluation and the problem of professional development. *Mid-Western Educational Researcher, 26*(2), 97-111.
- Spillane, J. (2017). Leadership and learning: Conceptualizing relations between school administrative practice and instructional practice *How School Leaders Contribute to Student Success* (pp. 49-67). London: Springer.
- Starkey, L., Yates, A., Meyer, L. H., Hall, C., Taylor, M., Stevens, S., & Toia, R. (2009). Professional development design: Embedding educational reform in New Zealand. *Teaching and Teacher Education, 25*, 181-189.
- Stebbins, R. A. (2001). *Exploratory research in the social sciences* (Vol. 48). London: SAGE Publications, Inc.
- Stronge, J. H., & Xu, X. (2017). *What Makes a World-Class School and How We Can Get There*. Alexandria: ASCD.
- Sudman, S., & Bradburn, N. M. (1983). Asking questions: a practical guide to questionnaire design.
- Sufean, H., & Ahmad, S. (2014). Spiral dynamics in the implementation of performance-based teacher evaluation policy in a centralized education system. *Malaysian Online Journal of Educational Management, 2*(2), 22-37.
- Tashakkori, A., & Teddlie, C. (1998). *Mixed methodology: Combining qualitative and quantitative approaches* (Vol. 46). London: SAGE Publications, Inc.
- Tennant, A. (2006). Key issues in classroom observation. *IATEFL VOICES, 191*(8), 8-9.
- Thalib, S. B., & Manda, D. (2016). The effect of school supervisors competence and school principals competence on work motivation and performance of junior high school teachers in Maros Regency, Indonesia. *International Journal of Environmental and Science Education, 11*(15), 7309-7317.
- Tovar, D. (2011). *Teacher Evaluation: Foreign-Language Teachers' and Administrators' Perspectives on Effective Evaluation Strategies and Processes*. (Doctor of Education), University of San Francisco.
- Towe, P. B. (2012). *An investigation of the role of a teacher evaluation system and its influence on teacher practice and professional growth in four urban high schools*. (Doctor of Education), Seton Hall University, New Jersey.
- Troudi, S. (2005). Critical content and cultural knowledge for teachers of English to speakers of other languages. *Teacher Development, 9*(1), 115-129.
- Troudi, S., & Rich, S. (2012). Teaching Portfolios: An Opportunity for Reigniting and Retooling. In C. A. Coombe, L. England, & J. Schmidt (Eds.), *Reigniting, Retooling, and Retiring in English Language Teaching* (pp. 55-65). Michigan University of Michigan Press.
- Trujillo, T. (2014). The modern cult of efficiency: Intermediary organizations and the new scientific management. *Educational Policy, 28*(2), 207-232.
- Tufte, T., & Mefalopulos, P. (2009). *Participatory communication: A practical guide* (Vol. 170). Washington, D.C.: The World Bank.
- Tyler, R. W. (1949). *Basic principles of curriculum and instruction*. Chicago: University of Chicago press.

- Vaismoradi, M., Jones, J., Turunen, H., & Snelgrove, S. (2016). Theme development in qualitative content analysis and thematic analysis. *Journal of Nursing Education and Practice*, 6(5), 100.
- Van Aaken, A., & List, C. (2017). *Deliberation and decision: Economics, constitutional theory and deliberative democracy*. London: Routledge.
- Van Uden, J. M., Ritzen, H., & Pieters, J. M. (2014). Engaging students: The role of teacher beliefs and interpersonal teacher behavior in fostering student engagement in vocational education. *Teaching and Teacher Education*, 37, 21-32.
- Vesely, A. K., Saklofske, D. H., & Leschied, A. D. (2013). Teachers - The vital resource: The contribution of emotional intelligence to teacher efficacy and well-being. *Canadian Journal of School Psychology*, 28(1), 71-89.
- Waring, H. Z. (2013). Two mentor practices that generate teacher reflection without explicit solicitations: Some preliminary considerations. *RELC Journal*, 44(1), 103-119.
- Warner, C. K. (2015). *Formative impacts of high-stakes portfolio assessment on pre-service English teachers: A qualitative study of belief, attitude, and identity*. (Doctor of Philosophy), University of Kansas, Lawrence, KS.
- Weaver, L., & Cousins, J. B. (2007). Unpacking the participatory process. *Journal of Multidisciplinary evaluation*, 1(1), 19-40.
- Weimer, M., Parrett, J., & Kerns, M.-M. (1988). *How Am I Teaching? Forms and Activities for Acquiring Instructional Input*: ERIC.
- Wellington, J. (2015). *Educational research: Contemporary issues and practical approaches*: Bloomsbury Publishing.
- Wenger, E. (2010). Communities of practice and social learning systems: the career of a concept *Social learning systems and communities of practice* (pp. 179-198). London: Springer.
- Whitworth, B. A., & Chiu, J. L. (2015). Professional development and teacher change: The missing leadership link. *Journal of Science Teacher Education*, 26(2), 121-137.
- Wikipedia. (2018). Saudi Vision 2030. Retrieved from [https://en.wikipedia.org/wiki/Saudi\\_Vision\\_2030](https://en.wikipedia.org/wiki/Saudi_Vision_2030)
- William, A. (2015). *Use of non financial motivation as a strategy to improve teacher performance in community secondary schools in Tanzania: a case of Ilala municipality*. (Master of Education), The Open University of Tanzania, Dar es Salaam.
- Williams, J., & Molesworth, M. (2010). Constructing consumption: What media representations reveal about today's students. In M. Nixon, R. Scullion, & M. Molesworth (Eds.), *The marketisation of higher education and the student as consumer* (pp. 170).
- Williams, N. V., & Crates, K. (2015). Administrator perspectives of Ohio's teacher evaluation system: Implications for educational administration programs in higher education. *Leadership and Research in Education*, 2(1), 25-38.
- Wong, B. L. (2016). *Using critical-thinking strategies to develop academic reading skills among Saudi IEP students*. (Master of Arts), Hamline University, Minnesota.

- Woodall, T., Hiller, A., & Resnick, S. (2014). Making sense of higher education: Students as consumers and the value of the university experience. *Studies in Higher Education*, 39(1), 48-67.
- Wragg, E. C., Wikeley, F. J., Wragg, C. M., & Haynes, G. S. (1996). *Teacher appraisal observed*. London: Routledge.
- Youn, T. I., & Price, T. M. (2009). Learning from the experience of others: The evolution of faculty tenure and promotion rules in comprehensive institutions. *The Journal of Higher Education*, 80(2), 204-237.
- Yuan, K. (2015). A value-added study of teacher spillover effects across four core subjects in middle schools. *Education Policy Analysis Archives/Archivos Analíticos de Políticas Educativas*(23), 1-21.
- Zaid, M. A. (1994). *Comprehensive analysis of the current system of teaching English as a Foreign Language in the Saudi Arabian intermediate schools*. (Doctorate of Philosophy), University of Colorado at Boulder.
- Zepeda, S. J. (2014a). *The instructional leader's guide to informal classroom observations*. London: Routledge.
- Zepeda, S. J. (2014b). *The principal as instructional leader: A handbook for supervisors*. London: Routledge.
- Zohairy, S. (2012). Professional development challenges and possible solutions. In W. Al-Amri, H. Noor, and , & I. McGee (Eds.), *Saudi Preparatory Year English Program: The Future and Beyond: student, teacher, pedagogy and curricular issues* (pp. 162-183). Madinah: Taibah University.
- Zuheer, K. M. M. (2013). *Developing EFL teachers' performance at Sana'a secondary schools in the light of their professional and specialist needs*. Cairo University, Giza.
- Zwozdiak-Myers, P. (2012). *The teacher's reflective practice handbook: Becoming an extended professional through capturing evidence-informed practice*. London: Routledge.