Graduate School of Education

EFL Lecturers’ Perceptions of Teacher Effectiveness and Teacher Evaluation in Iranian universities

Submitted by Omid Mazandarani to the University of Exeter

as a thesis for the degree of

Doctor of Philosophy in Education

In May 2014

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

Teacher effectiveness research (TER), as a multifaceted phenomenon, is a seminal part of most educational agendas upon which a successful teacher appraisal system tends to be contingent. Whereas there is a wealth of research on teacher effectiveness in mainstream (general) education, there is a dearth of studies on it in second/foreign language education, thereby sowing seeds of doubts apropos of the extent to which findings in mainstream education can be applied to L2 education. A paucity of cutting-edge research in the Middle-eastern context is another missing piece of the jigsaw testifying to a need for further research on teacher effectiveness. Taking such lacunae into consideration, this study endeavours to inquire into EFL teacher effectiveness in the Iranian higher education context as its main objective. With this end in view, a number of research questions are formulated whereby the main constructs are identified. This study is aimed at investigating lecturers’ understanding of teacher effectiveness and its pertinent appraisal model, and more specifically, delving into their perceptions of teacher appraisal in Iran. Measures of evaluation, opportunities of which lecturers can avail themselves to improve their effectiveness, and lecturers’ ideal appraisal system are other areas which are examined in this research. In this study, a mixed methods exploratory sequential design is adopted to address the proposed research questions. Close-ended and open-ended questionnaires and semi-structured interview are the instruments utilised for data collection. The collected quantitative and qualitative data are analysed with the help of SPSS and NVivo, respectively. The analysis of both sets of data culminated in the emergence of six major themes, i.e. lecturers’ understanding of teacher effectiveness and the qualities of an effective teacher, measures of evaluation, opportunities and strategies conducive to improving teacher effectiveness, the Iranian appraisal system, non-teacher-controlled factors impacting on teacher effectiveness, as well as lecturers’ perceptions of an ideal appraisal system. Following a myriad of ideas garnered through data analysis, a differentiated appraisal model informed by lecturers’ voices is proposed. Based on the findings which provided evidence for some imperfections in the nexus between policy and implementation, this study concludes that there is still some room for improvement in teacher appraisal in
Iran. Important amongst others are better alignment between teacher appraisal and teachers' professional development needs, transparency of the appraisal, and use of all types and forms of teacher evaluation. The study brings to the fore further implications, conclusions and suggestions for future research which are presented in the final chapter of this thesis.
Acknowledgments

This study could not have been completed without the support of those who have continuously helped me throughout my studies. I would like to thank God for helping me complete such a phenomenal journey.

I would like to extend my deepest gratitude to my first supervisor Dr Salah Troudi for his encouragement and support. He has always been of invaluable academic and psychological support throughout my studies. His feedback and comments have always been critical, insightful and constructive. I would also like to express my sincere appreciation to my second supervisor Professor Wendy Robinson for her support, guidance and insightful feedback.

My wholehearted gratitude and deepest appreciation are extended to my wife and son for their love, patience, companionship and encouragement. It was their endless love and support which gave me the strength to withstand and overcome the challenges and complete my studies. I am also indebted to my parents who have always been a precious source of encouragement and support throughout my life. This study would not have been completed without the endless support and encouragement of my family.
# Table of contents

Abstract .......................................................................................................................... 2  
Acknowledgments .......................................................................................................... 4  
Table of contents .......................................................................................................... 5  
List of Tables .................................................................................................................. 12  
List of Figures ................................................................................................................. 13  
List of Abbreviations and Acronyms ............................................................................ 14  
Chapter One: Introduction ............................................................................................ 16  
1.1. Introduction ............................................................................................................. 17  
1.2. English Language Teaching .................................................................................. 17  
1.3. Rationale of the study ............................................................................................ 18  
1.4. Statement of the problem ....................................................................................... 22  
1.5. Significance of the study ....................................................................................... 24  
1.6. Research aims and objectives ............................................................................... 26  
1.7. Research questions ............................................................................................... 27  
1.8. Definition of key terms ......................................................................................... 28  
1.8.1. Appraisal ............................................................................................................ 28  
1.8.2. Teacher evaluation ............................................................................................. 29  
1.8.3. Teacher effectiveness .......................................................................................... 29  
1.9. An overview of the thesis ....................................................................................... 29  
1.10. Summary .............................................................................................................. 30  
Chapter Two: Context of the Study ............................................................................. 31  
2.1. Introduction ............................................................................................................ 32  
2.2. Country profile ..................................................................................................... 32  
2.3. A brief history of higher education in Iran ......................................................... 33  
2.4. Universities and higher education institutions .................................................... 35  
2.5. ELT in the Iranian higher education ................................................................... 36  
2.6. Policy making in the Iranian higher education system ....................................... 36  
2.7. Supervision and Evaluation of Higher Education .............................................. 37  
2.8. Summary .............................................................................................................. 39  
Chapter Three: Literature Review ................................................................................ 40  
3.1. Introduction ............................................................................................................ 41
### 3.2. Teacher effectiveness: an introduction ................................................................. 41

### 3.3. Teacher effectiveness, a multidimensional phenomenon ......................................... 44

#### 3.3.1. Subject area ........................................................................................................... 45

#### 3.3.2. Level of education .............................................................................................. 45

#### 3.3.3. Measures of evaluation (source of information) .................................................... 45

#### 3.3.4. Western vs. non-Western (Middle-eastern) education .......................................... 46

#### 3.3.5. Other less well-investigated areas ........................................................................ 46

### 3.4. Teacher effectiveness in mainstream education ....................................................... 47

### 3.5. Teacher effectiveness research in higher education ................................................ 51

### 3.6. Teacher effectiveness in ESL/EFL education .......................................................... 54

### 3.7. Teacher effectiveness research in Iran ...................................................................... 57

### 3.8. Teacher education, recruitment, development, performance and appraisal .......... 60

### 3.9. Teacher Evaluation: the challenges of standards ..................................................... 61

#### 3.9.1. What counts as teacher effectiveness ................................................................. 63

#### 3.9.2. Teacher effectiveness and quality in education .................................................... 65

### 3.10. Forms of teacher evaluation (Measures of evaluation) ........................................ 65

#### 3.10.1. Students’ evaluations of teaching (SETs)/Students’ ratings ............................... 66

#### 3.10.2. Students’ learning outcome/student achievement ............................................... 69

#### 3.10.3. Peer evaluation ................................................................................................... 70

#### 3.10.4. Teachers’ Self-evaluation .................................................................................... 71

#### 3.10.5. Appraiser’s observations ..................................................................................... 72

### 3.11. Types of evaluation models .................................................................................... 73

#### 3.11.1. Formative and summative approaches ............................................................... 73

#### 3.11.2. Generic vs. differentiated models ....................................................................... 75

#### 3.11.3. Qualitative vs. quantitative models .................................................................... 76

#### 3.11.4. Subjective vs. objective models ......................................................................... 77

### 3.12. An overview of some schemes and models developed for teacher appraisal around the world ............................................................................................................ 78

### 3.13. Reflections on the existing lacunae (Support for a differentiated model) .......... 90

### 3.14. Summary ................................................................................................................. 92

Chapter Four: Methodology ................................................................................................. 94

#### 4.1. Introduction .............................................................................................................. 95

#### 4.2. Theoretical and philosophical assumptions ........................................................... 96

#### 4.2.1. Ontological assumption ....................................................................................... 98

#### 4.2.2. Epistemological assumption ................................................................................ 99
4.2.3. Research, theory and practice ........................................................................... 100
4.3. Methodological assumptions ............................................................................. 101
4.4. Research design ................................................................................................. 102
  4.4.1. Mixed methods research ............................................................................... 103
    4.4.1.1. Implementation ....................................................................................... 105
    4.4.1.2. Priority .................................................................................................. 106
    4.4.1.3. Integration .............................................................................................. 106
    4.4.1.4. Theoretical perspectives ........................................................................ 107
  4.4.2. Sequential Exploratory Design ...................................................................... 107
    4.4.2.1. A few remarks on the design differentiated for this study ....................... 108
  4.4.3. Participants (Sample of the study) ............................................................... 109
  4.4.4. Methods and instruments ............................................................................ 112
    4.4.4.1. Questionnaire ....................................................................................... 113
    4.4.4.2. Interview ................................................................................................. 114
  4.4.5. Procedure ..................................................................................................... 114
4.5. Data Collection .................................................................................................. 115
  4.5.1. Quantitative data collection .......................................................................... 116
    4.5.1.1. Questionnaire ....................................................................................... 117
    4.5.1.2. Designing a researcher-developed questionnaire ................................. 118
    4.5.1.3. Piloting the questionnaire ...................................................................... 120
    4.5.1.4. Constructing the final questionnaire ...................................................... 122
    4.5.1.5. Administering the questionnaire ............................................................ 122
  4.5.2. Qualitative data collection ............................................................................. 123
    4.5.2.1. Semi-structured Interview .................................................................... 124
      4.5.2.1.1. Interview guide/schedule ................................................................. 126
      4.5.2.1.2. Piloting the interview guide ............................................................. 127
      4.5.2.1.3. Constructing the final interview guide ............................................. 127
      4.5.2.1.4. Conducting the interviews ............................................................... 128
    4.5.2.2. Open-ended questionnaire ..................................................................... 129
      4.5.2.2.1. Piloting the Open-ended questionnaire ........................................... 130
  4.5.3. Trustworthiness ............................................................................................ 131
4.6. Data Analysis ..................................................................................................... 132
  4.6.1. Quantitative data analysis ............................................................................ 133
  4.6.2. Qualitative data analysis .............................................................................. 134
<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.6.2.1</td>
<td>Interviews data analysis</td>
<td>135</td>
</tr>
<tr>
<td>4.6.2.1.1</td>
<td>Data Management: preparing and organising the data</td>
<td>136</td>
</tr>
<tr>
<td>4.6.2.1.2</td>
<td>Data exploration: reading and memoing</td>
<td>137</td>
</tr>
<tr>
<td>4.6.2.1.3</td>
<td>Data reduction: describing, classifying and Coding the data</td>
<td>138</td>
</tr>
<tr>
<td>4.6.2.1.4</td>
<td>Data interpretation: representing and visualizing the data</td>
<td>139</td>
</tr>
<tr>
<td>4.6.2.2</td>
<td>Open-ended questionnaire data analysis</td>
<td>140</td>
</tr>
<tr>
<td>4.7</td>
<td>The qualitative-quantitative data nexus</td>
<td>140</td>
</tr>
<tr>
<td>4.8</td>
<td>Ethical considerations</td>
<td>142</td>
</tr>
<tr>
<td>4.8.1</td>
<td>Informed consent and withdrawal right</td>
<td>142</td>
</tr>
<tr>
<td>4.8.2</td>
<td>Institutional ethical approval</td>
<td>143</td>
</tr>
<tr>
<td>4.8.3</td>
<td>Privacy, confidentiality, anonymity, data storage and disclosure</td>
<td>143</td>
</tr>
<tr>
<td>4.8.4</td>
<td>Accuracy</td>
<td>144</td>
</tr>
<tr>
<td>4.9</td>
<td>Limitations and Challenges of the study</td>
<td>144</td>
</tr>
<tr>
<td>4.10</td>
<td>Summary</td>
<td>146</td>
</tr>
<tr>
<td>Chapter 5: Data Analysis &amp; Findings</td>
<td>147</td>
<td></td>
</tr>
<tr>
<td>5.1</td>
<td>Introduction</td>
<td>148</td>
</tr>
<tr>
<td>5.2</td>
<td>Six Major Themes</td>
<td>149</td>
</tr>
<tr>
<td>5.3</td>
<td>General perceptions of TE &amp; Qualities of an effective teacher</td>
<td>151</td>
</tr>
<tr>
<td>5.3.1</td>
<td>Personal qualities</td>
<td>152</td>
</tr>
<tr>
<td>5.3.1.1</td>
<td>Friendly and Approachable</td>
<td>154</td>
</tr>
<tr>
<td>5.3.1.2</td>
<td>Informative &amp; Dedicated</td>
<td>155</td>
</tr>
<tr>
<td>5.3.1.3</td>
<td>Respectfulness</td>
<td>155</td>
</tr>
<tr>
<td>5.3.1.4</td>
<td>Inspirational and Self-motivated</td>
<td>156</td>
</tr>
<tr>
<td>5.3.1.5</td>
<td>Confidence</td>
<td>157</td>
</tr>
<tr>
<td>5.3.1.6</td>
<td>Patience</td>
<td>158</td>
</tr>
<tr>
<td>5.3.1.7</td>
<td>Openness &amp; Adaptability</td>
<td>158</td>
</tr>
<tr>
<td>5.3.2</td>
<td>Cognitive qualities</td>
<td>159</td>
</tr>
<tr>
<td>5.3.2.1</td>
<td>Language proficiency</td>
<td>160</td>
</tr>
<tr>
<td>5.3.2.2</td>
<td>Subject Matter knowledge</td>
<td>162</td>
</tr>
<tr>
<td>5.3.2.3</td>
<td>Knowledge of curriculum &amp; syllabus</td>
<td>162</td>
</tr>
<tr>
<td>5.3.2.4</td>
<td>ICT Literacy</td>
<td>164</td>
</tr>
<tr>
<td>5.3.3</td>
<td>Pedagogical skills</td>
<td>165</td>
</tr>
<tr>
<td>5.3.3.1</td>
<td>Instructional planning and alternative instruction</td>
<td>166</td>
</tr>
<tr>
<td>5.3.3.2</td>
<td>Simplicity/Tailoring materials to students’ needs</td>
<td>168</td>
</tr>
</tbody>
</table>
5.3.3.3. Needs analysis & student engagement .................................................. 170
5.3.3.4. Attentive to affective filter ................................................................. 171
5.3.3.5. Transferring knowledge ..................................................................... 172
5.3.3.6. Assessment ......................................................................................... 172
5.3.4. Professional skills ................................................................................ 174
5.3.4.1. Sense of responsibility & accountability ........................................... 175
5.3.4.2. Creativity & Innovation ........................................................................ 176
5.3.4.3. Authority & Management Skills ....................................................... 177
5.3.4.4. Building Rapport & Trust .................................................................. 178
5.3.4.5. Fulfilling course objectives/Intended learning outcomes .................. 178
5.3.4.6. Improving Students' Performance ..................................................... 179
5.3.4.7. Teachers' Networking ......................................................................... 180
5.3.4.8. Overcoming Problems ........................................................................ 181
5.3.4.9. Fair evaluation of Students ............................................................... 181
5.3.4.10. Reflective practice ............................................................................ 182
5.4. Measures of Evaluation .......................................................................... 184
5.4.1. Student Evaluation of Teachers (SETs)/Student Ratings ..................... 185
5.4.2. Students’ learning outcomes/achievement .......................................... 187
5.4.3. Peer evaluation/observation .................................................................. 188
5.4.4. Self-evaluation ....................................................................................... 190
5.4.5. Observation (external/internal) ............................................................. 190
5.4.6. Some further points to consider .......................................................... 192
5.5. Ways to improve Teacher Effectiveness in HE ........................................ 194
5.5.1. Experience ............................................................................................. 196
5.5.2. Technology Enhanced Teaching and Learning Strategies/ ICT Literacy ...... 196
5.5.3. Knowledge of Assessment ..................................................................... 197
5.5.4. Self-Regulation & Reflective skills ....................................................... 198
5.5.5. Keeping updated .................................................................................... 200
5.5.6. Use of Students’ feedback ..................................................................... 201
5.5.7. Universities’ accountability for promoting TE ...................................... 201
5.6. Lectures’ perceptions of teacher evaluation in Iran .................................. 203
5.6.1. Some perceived challenges .................................................................... 204
5.6.1.1. One-size-fits-all approach .................................................................. 205
5.6.1.2. Power issue: Conflicts among stakeholders’ interests/Subjectivity ...... 205
<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>6.3.4. Observation (External or Internal)</td>
<td>246</td>
</tr>
<tr>
<td>6.4. Ways to improve teacher effectiveness in HE</td>
<td>248</td>
</tr>
<tr>
<td>6.4.1. Teachers as front-line pathfinders</td>
<td>249</td>
</tr>
<tr>
<td>6.4.2. Administrators as logisticians</td>
<td>249</td>
</tr>
<tr>
<td>6.5. The Iranian appraisal system</td>
<td>250</td>
</tr>
<tr>
<td>6.5.1. Teacher appraisal: an enigma</td>
<td>251</td>
</tr>
<tr>
<td>6.5.2. Some major challenges</td>
<td>252</td>
</tr>
<tr>
<td>6.5.3. The missing pieces of the appraisal jigsaw</td>
<td>256</td>
</tr>
<tr>
<td>6.6. Some further issues to consider</td>
<td>258</td>
</tr>
<tr>
<td>6.6.1. Non-teacher-controlled factors affecting TE in HE</td>
<td>258</td>
</tr>
<tr>
<td>6.6.2. Lecturers’ ideal appraisal model</td>
<td>261</td>
</tr>
<tr>
<td>6.7. Summary</td>
<td>265</td>
</tr>
<tr>
<td>Chapter Seven: Implications &amp; Conclusion</td>
<td>266</td>
</tr>
<tr>
<td>7.1. Introduction</td>
<td>267</td>
</tr>
<tr>
<td>7.2. Theoretical and pedagogical implications for main stakeholders</td>
<td>267</td>
</tr>
<tr>
<td>7.2.1. Implications for policymakers</td>
<td>268</td>
</tr>
<tr>
<td>7.2.2. Implications for administrators</td>
<td>271</td>
</tr>
<tr>
<td>7.2.3. Implications for teachers</td>
<td>275</td>
</tr>
<tr>
<td>7.3. Theoretical contributions: A suggested model</td>
<td>277</td>
</tr>
<tr>
<td>7.3.1. Contributions to the existing knowledge</td>
<td>280</td>
</tr>
<tr>
<td>7.4. Policy-implementation nexus</td>
<td>284</td>
</tr>
<tr>
<td>7.5. Suggestions for further research</td>
<td>285</td>
</tr>
<tr>
<td>7.6. Conclusion</td>
<td>287</td>
</tr>
<tr>
<td>7.7. Reflection on my PhD journey</td>
<td>289</td>
</tr>
<tr>
<td>Appendices</td>
<td>291</td>
</tr>
<tr>
<td>Appendix 1: Researcher-developed questionnaire</td>
<td>292</td>
</tr>
<tr>
<td>Appendix 2: Interview Protocol</td>
<td>300</td>
</tr>
<tr>
<td>Appendix 3: Questionnaire respondents’ background data</td>
<td>301</td>
</tr>
<tr>
<td>Appendix 4: Informed Consent Form</td>
<td>304</td>
</tr>
<tr>
<td>Appendix 5: An excerpt from a sample interview manuscript</td>
<td>305</td>
</tr>
<tr>
<td>Appendix 6: A sample of coding process using NVivo</td>
<td>308</td>
</tr>
<tr>
<td>Appendix 7: Ethical Approval Form</td>
<td>309</td>
</tr>
<tr>
<td>References</td>
<td>311</td>
</tr>
</tbody>
</table>
## List of Tables

<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.1</td>
<td>Interviewees background information</td>
<td>112</td>
</tr>
<tr>
<td>5.1</td>
<td>Frequency of sources and references (Themes, Categories and Sub-categories)</td>
<td>151</td>
</tr>
<tr>
<td>7.1</td>
<td>Some rudiments of teacher evaluation</td>
<td>278</td>
</tr>
</tbody>
</table>
## List of Figures

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Figure 2.1</td>
<td>Iran Geographical Location (adapted from Islamic Republic of Iran Mission to the UN in New York)</td>
<td>32</td>
</tr>
<tr>
<td>Figure 3.1</td>
<td>The structure of total teacher effectiveness model (Cheng &amp; Tsui, 1996, p. 13; 1998, p. 41)</td>
<td>81</td>
</tr>
<tr>
<td>Figure 3.2</td>
<td>“The Onion: a model of levels of change” (Korthagen, 2004, p. 80)</td>
<td>84</td>
</tr>
<tr>
<td>Figure 3.3</td>
<td>“Elements of appraisal effectiveness” (Piggot-Irvine, 2003, p. 173)</td>
<td>86</td>
</tr>
<tr>
<td>Figure 3.4</td>
<td>“The measures of teacher effectiveness” (McBer, 2000, p. 6)</td>
<td>87</td>
</tr>
<tr>
<td>Figure 3.5</td>
<td>“The teaching skills” (McBer, 2000, p. 10)</td>
<td>88</td>
</tr>
<tr>
<td>Figure 3.6</td>
<td>“The model for professional characteristics” (McBer, 2000, p. 19)</td>
<td>89</td>
</tr>
<tr>
<td>Figure 4.1</td>
<td>The research design based upon Creswell et al.’s (2003, p. 224) four elements</td>
<td>108</td>
</tr>
<tr>
<td>Figure 4.2</td>
<td>The Sequential Exploratory Design</td>
<td>109</td>
</tr>
<tr>
<td>Figure 4.3</td>
<td>The procedure for Sequential Exploratory Design of the study</td>
<td>115</td>
</tr>
<tr>
<td>Figure 5.1</td>
<td>Major themes generated for Teacher effectiveness in HE</td>
<td>150</td>
</tr>
<tr>
<td>Figure 5.2</td>
<td>Themes I. (Perception of TE &amp; Qualities of an effective teacher)</td>
<td>152</td>
</tr>
<tr>
<td>Figure 5.3</td>
<td>Theme I. Category I (Personal qualities)</td>
<td>153</td>
</tr>
<tr>
<td>Figure 5.4</td>
<td>Theme I. Category II (Cognitive qualities)</td>
<td>160</td>
</tr>
<tr>
<td>Figure 5.5</td>
<td>Theme I. Category III. (Pedagogical skills)</td>
<td>165</td>
</tr>
<tr>
<td>Figure 5.6</td>
<td>Theme I. Category IV. (Professional skills)</td>
<td>175</td>
</tr>
<tr>
<td>Figure 5.7</td>
<td>Theme II. (Measures of Evaluation)</td>
<td>185</td>
</tr>
<tr>
<td>Figure 5.8</td>
<td>Theme III (Ways on increase TE)</td>
<td>195</td>
</tr>
<tr>
<td>Figure 5.9</td>
<td>Theme IV (Iranian appraisal system)</td>
<td>204</td>
</tr>
<tr>
<td>Figure 5.10</td>
<td>Theme V. Non-teacher factors affecting TE</td>
<td>216</td>
</tr>
<tr>
<td>Figure 5.11</td>
<td>Theme VI. Ideal/critical appraisal model</td>
<td>223</td>
</tr>
<tr>
<td>Figure 7.1</td>
<td>Implications for main stakeholders</td>
<td>268</td>
</tr>
<tr>
<td>Figure 7.2</td>
<td>Proposed Appraisal Model</td>
<td>279</td>
</tr>
<tr>
<td>Figure 7.3</td>
<td>Teacher development and Appraisal Cycle in HE</td>
<td>282</td>
</tr>
<tr>
<td>Figure 7.4</td>
<td>Formative-summative continuum of evaluation</td>
<td>284</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Full Form</td>
<td></td>
</tr>
<tr>
<td>--------------</td>
<td>-----------</td>
<td></td>
</tr>
<tr>
<td>BERA</td>
<td>British Educational Research Association</td>
<td></td>
</tr>
<tr>
<td>CALx</td>
<td>Critical Applied Linguistics</td>
<td></td>
</tr>
<tr>
<td>CEQ</td>
<td>Course Experience Questionnaire</td>
<td></td>
</tr>
<tr>
<td>CLT</td>
<td>Communicative Language Teaching</td>
<td></td>
</tr>
<tr>
<td>EFL</td>
<td>English as a Foreign Language</td>
<td></td>
</tr>
<tr>
<td>ELT</td>
<td>English Language Teaching</td>
<td></td>
</tr>
<tr>
<td>ESEA</td>
<td>Elementary and Secondary Education Act</td>
<td></td>
</tr>
<tr>
<td>ESL</td>
<td>English as a Second Language</td>
<td></td>
</tr>
<tr>
<td>HEA</td>
<td>Higher Education Academy</td>
<td></td>
</tr>
<tr>
<td>IAU</td>
<td>Islamic Azad University</td>
<td></td>
</tr>
<tr>
<td>ICT</td>
<td>Information and Computer Technology</td>
<td></td>
</tr>
<tr>
<td>ITEP</td>
<td>Initial Teacher Education Programme</td>
<td></td>
</tr>
<tr>
<td>L1</td>
<td>First Language</td>
<td></td>
</tr>
<tr>
<td>L2</td>
<td>Second/Foreign Language</td>
<td></td>
</tr>
<tr>
<td>MHTME</td>
<td>Ministry of Health, Treatment and Medical Education</td>
<td></td>
</tr>
<tr>
<td>MOH</td>
<td>Ministry of Education</td>
<td></td>
</tr>
<tr>
<td>MSRT</td>
<td>Ministry of Science, Research and Technology</td>
<td></td>
</tr>
<tr>
<td>NBC</td>
<td>National Board Certification</td>
<td></td>
</tr>
<tr>
<td>NBPTS</td>
<td>The National Board of Professional Teaching Standards</td>
<td></td>
</tr>
<tr>
<td>NCATE</td>
<td>The National Council for the Accreditation of Teacher Education</td>
<td></td>
</tr>
<tr>
<td>NCLB</td>
<td>The No Child left Behind Act</td>
<td></td>
</tr>
<tr>
<td>NGDIR</td>
<td>National Geoscience Database of Iran</td>
<td></td>
</tr>
<tr>
<td>NRHERT</td>
<td>National Report of Higher Education, Research and Technology</td>
<td></td>
</tr>
<tr>
<td>OECD</td>
<td>Organisation for Economic Cooperation and Development</td>
<td></td>
</tr>
<tr>
<td>OFSTEAD</td>
<td>Office for Standards in Education</td>
<td></td>
</tr>
<tr>
<td>PI</td>
<td>Performance Indicator</td>
<td></td>
</tr>
<tr>
<td>PNU</td>
<td>Payame Noor University, Iran</td>
<td></td>
</tr>
<tr>
<td>SCI</td>
<td>Statistical Centre of Iran</td>
<td></td>
</tr>
<tr>
<td>SLA</td>
<td>Second Language Acquisition</td>
<td></td>
</tr>
<tr>
<td>SLTE</td>
<td>Second Language Teacher Education</td>
<td></td>
</tr>
<tr>
<td>TALIS</td>
<td>Teaching and Learning International Survey</td>
<td></td>
</tr>
<tr>
<td>Acronym</td>
<td>Full Form</td>
<td></td>
</tr>
<tr>
<td>---------</td>
<td>-----------</td>
<td></td>
</tr>
<tr>
<td>TPDP</td>
<td>Teacher Professional Development Programme</td>
<td></td>
</tr>
<tr>
<td>TE</td>
<td>Teacher Effectiveness</td>
<td></td>
</tr>
<tr>
<td>TER</td>
<td>Teacher Effectiveness Research</td>
<td></td>
</tr>
<tr>
<td>TEFL</td>
<td>Teaching English as A Foreign Language</td>
<td></td>
</tr>
<tr>
<td>TEL</td>
<td>Technology Enhanced Learning</td>
<td></td>
</tr>
<tr>
<td>TER</td>
<td>Teacher Effectiveness Research</td>
<td></td>
</tr>
<tr>
<td>TEQ</td>
<td>Teacher Effectiveness Questionnaire</td>
<td></td>
</tr>
<tr>
<td>TESOL</td>
<td>Teaching English to Speakers of other Languages</td>
<td></td>
</tr>
<tr>
<td>TPAI</td>
<td>Teaching Performance Appraisal Instrument</td>
<td></td>
</tr>
<tr>
<td>TL</td>
<td>Target Language</td>
<td></td>
</tr>
<tr>
<td>UT</td>
<td>University of Tehran</td>
<td></td>
</tr>
</tbody>
</table>
Chapter One: Introduction
1.1. Introduction

This chapter presents an overall picture of the history and the state-of-the-art status of Teaching English as a Foreign Language (TEFL) and the challenges and opportunities associated with teacher effectiveness and its pertinent appraisal model in Iran. The chapter commences with a brief introduction of English Language Teaching (ELT) and then proceeds to the rationale of the study. Explicating the existing problems and challenges and then discussing the main constructs of this research, the chapter touches on the significance of the study. The research aims and objectives are presented thereafter. As a subsequent section, research questions will be introduced after which some key terms and concepts will be explained. The chapter concludes with an outline of the overall structure of the thesis.

1.2. English Language Teaching

Language teaching in a sense that is perceived as a profession came into its own in the twentieth century (Richards & Rodgers, 2001, p. 1). Throughout its history, language teaching has been of seminal importance to practitioners and researchers, given the fact that it is estimated that approximately 60 percent of the world population was multilingual by 2001 raising the need for foreign language teachers (p. 3). ELT nowadays is approached as a career in education, thereby necessitating a specialised requisite knowledge base, and to that end is imbued with a high level of professionalism (Burns & Richards, 2009, p. 2). Perhaps, one difference between mainstream education teachers and language teachers lies with ‘language’ per se; an unknown object (Cook, 1999, p. 190) which functions as a social and spatial activity (Pennycook, 2010, p. 3). Teaching in a language other than students’ mother tongue has its own challenges. Further research is needed to bring into light the less well-understood and controversial dimensions such as learning and teaching a foreign/second language (L2). It seems unwise to approach L2 teacher education and evaluation without pondering the underpinning theoretical and conceptual dimensions of language teaching and learning. Similar to several other disciplines in educational sciences, language teaching as a microcosm of mainstream general education has witnessed an evolutionary
development throughout its history. As Richards and Rodgers contend, different teaching methods and approaches have appeared in the last 60 years (2001, p. 15). Such development has often been discussed from at least three interrelated perspectives, namely philosophical, linguistic and psychological aspects. In a nutshell, language teaching evolution has commenced from the so-called pre-method to method to post-method era. Whereas there is a general consensus on concepts such approach, method, and methodology in mainstream education, their applications and implications in language education context have been given rather scant attention. Today, in some academic texts, concepts such as method, approach, methodology tend to be used interchangeably due to a paucity of consensus on the constituents of the aforementioned concepts in language education domain. Although TEFL is highly informed by theories and practices in mainstream education, understanding what counts as teacher effectiveness in language teaching appears to have its roots in a comprehensive appreciation of the particularities of foreign language teaching.

1.3. Rationale of the study

In line with that of mainstream education, teacher effectiveness in TEFL serves as an important constituent of a wider domain of educational effectiveness thereby resonating with several other interwoven areas such as teacher education and teacher development. As a prominent element in educational reform and school effectiveness (Cheng & Tsui, 1996, p. 7), teacher effectiveness has become high on the agenda in most educational contexts. Teacher effectiveness has turned out to be a concern in some educational contexts on the grounds of high expectations lecturers in higher education tend to face on the part of other minor and major stakeholders such as students, parents, administrators and policymakers. While teacher effectiveness in the domain of mainstream education appears to be more or less controversial in its generic sense, it has been a source of much contention in in the domain of (foreign) language teacher education. Despite the burgeoning literature on effective teaching methods and approaches in TEFL during the last decades, there is a dearth of studies operationalising the concepts of effective teaching and teacher effectiveness. This partly emanates from researchers'
discrepant views on the very nature of some notions and concepts in language teaching. For instance, Kumaravadivelu maintains that approach and method have been used interchangeably in some L2 teaching literature (2006, p. 85). Critiquing Antony (1963) and Richards and Rodgers’ (1982) frameworks, Kumaravadivelu alludes to the difference between method and methodology (pp. 84-86). Whereas Kumaravadivelu (2006) sees methods as what he calls ‘established methods conceptualised and constructed by experts in the field’, methodology, as he asserts, refers to what teachers do in the classroom to gain their teaching objectives (p. 84). As a consequence, prior to designing any appraisal model for assessing teacher effectiveness, we need to consider the very dynamic nature of language teaching. This is why the answer to the question posed in this study as to ‘what counts as an effective teacher?’ remains yet less well-understood. What I have argued in the above lines is merely a snapshot of some central lines of ideas pertaining to the challenging nature of teacher effectiveness. It is worth highlighting that teacher effectiveness research is a complex phenomenon. The evaluation of teacher effectiveness is the other side of the coin for which policymakers need to turn their attention to good account, e.g. how the appraisal is implemented and the data are interpreted.

Learners’ unbounded enthusiasm for comparing their teachers has existed throughout history. As a natural human instinct for the best, students have always been in the habit of comparing their teachers. The emergence of large online databases such as ‘ratemyprofessors.com’ and ‘myedu.com’ (Clayson, 2013) provides evidence for to the enthusiasm students are often fired with for evaluating their professors and teachers. Irrespective of students’ intentions, teachers do matter to a great deal to most students. In recent years, in line with mass education movement around the world, there has been a growing demand on the part of main stakeholders for effective teaching. Nevertheless, teacher effectiveness or effective teaching cannot be addressed unless the two core distinct yet interrelated concepts of effective and teaching are explained. This is of an utmost importance as different stakeholders may have a different understanding of what makes an effective teacher. Indeed, the literature on teacher effectiveness alludes to a continuum of effective teaching. Whereas some researchers maintain that
“teaching without learning is just talking” (Angelo, 1990, p. 75; Angelo & Cross, 1993, p. 3), other scholars prevail upon teachers to adopt duties and responsibilities other than what they call the traditional conception of teachers’ role, i.e. transferring knowledge. Repudiating the old aphorism “Those who can, do; those who can’t, teach”, Scriven (1994) argues how “Those who can do a hundred difficult duties, can teach well and can change the world” (p. 151). To many researchers and practitioners, teacher effectiveness is equivalent to teacher performance which is observable. Considering the quality and effectiveness of a teacher as a determining factor of teacher performance and student learning outcome, these researchers have endeavoured to develop new strategies and techniques in order to improve it (Cheng & Tsui, 1996, p. 7). However, it is unfortunate that little attention has been given to teacher centrality in education as compared to a preponderance of studies on students. According to Freeman and Johnson (1998), only 9% of the featured articles of TESOL Quarterly from 1990 to 1997 were germane to teacher education which from their perspective implies the scant attention paid to teacher continuing professional education (pp. 397-398). The challenging nature of teacher effectiveness has led researchers and practitioners to define and redefine these concepts and their practical ramifications. As stated earlier, prior to dealing with the issue of teacher effectiveness, we need to address some prerequisite questions, e.g. ‘what is meant by effective teaching’ or ‘effective in terms of what’. Indeed, the evaluation of teacher effectiveness appears to be meaningless unless there is a set of transparent criteria/standards. Furthermore, the mechanism of evaluation, i.e. how it is implemented and then interpreted, is another important aspect of teacher effectiveness appraisal which will be addressed, amongst other aspects, in this study.

Given the convoluted nature of teaching and learning in higher education, it seems important to investigate teacher effectiveness in relation to other constituents operating in a wider context of educational effectiveness. As a multidimensional phenomenon, teacher effectiveness appraisal extends to other interwoven areas such as institutional/organisational effectiveness, curriculum effectiveness, etc. Today, the above-mentioned concepts and notions appear to be as buzzwords in education studies with an old history which indubitably testifies to their significance.
Teacher education and teacher development programmes designed for pre-service and in-service teachers respectively are other areas which can potentially wield influence on teacher effectiveness.

Obviously, any appraisal model designed for evaluating teacher effectiveness should meet the dynamic and multifaceted nature of language teaching as far as possible. That’s why there appears a rise in the need for a so-called discipline-specific appraisal model, e.g. TEFL/TESL-specific models which can accommodate the peculiarities of teaching English as a foreign (EFL) or second (ESL) language. Nevertheless, the existing literature on TE gives evidence to the contrary as the majority of studies have addressed TE in its generic sense bringing up a question as to the extent to which such findings could be generalised to language and in particular, EFL teacher effectiveness. Addressing this issue, Crandall contends that ‘language teacher education is a microcosm of teacher education’ (2000, p. 34) rooted in similar theories and practices. As she further continues, the direction of language teacher education has been much influenced by general educational theories and practices (2000, p. 34). Despite such a judicious point of view, I think there exist a number of concerns pertaining to the nature of teaching in a language other than students’ mother tongue. Second/foreign language teachers have often been struggling with the question of best teaching method(s) whereby they have been expected to select effective teaching methods from among different language teaching methods and approaches. Given the multitudinous nature of language teachers and learners’ needs, wants and situations, no idealised method can be conceived of as the one which can provide teachers with situation-specific suggestions they may need in order to overcome the challenges they tend to face in their practices (Kumaravadivelu, 2012, p. 18). Therefore, despite the few studies in the last two decades (e.g. Brown, 2009; Fradd & Lee, 1998), further research is needed to understand the dynamics of teacher effectiveness and its pertinent appraisal in L2 context.

It is also worth referring to my personal interest in researching teacher effectiveness and teacher evaluation which is rooted in my previous experience as
a member of academic staff. As an EFL lecturer, I can clearly recollect the way I used to be evaluated each semester. Moreover, I remember that, similar to most of my colleagues, I had often a relatively vague idea of the nature and dynamics of teacher evaluation. It is my contention that in order for an appraisal model to be effective, it needs to be clear, fair and convincing.

1.4. Statement of the problem

As it was stated earlier, the contentious concept of ‘effectiveness’ and the question as to ‘what counts as an effective teacher?’ have provoked controversy among researcher and practitioners. As Campbell, Kyriakides, Muijs, and Robinson (2004a, p. 2) maintain, school effectiveness, teacher effectiveness and educational effectiveness have been inconsistent in the literature, albeit interconnected. Any mismatches among different stakeholders’ understanding and hence expectations of teacher effectiveness could exert a detrimental influence on the achievement of educational objectives. This appears to be one of the problems in Iran. In other words, policymakers in the Iranian higher education need to make their goals and expectations of effective teaching explicit so that all main stakeholders such as teachers, students, administrators, etc., have a similar impression of teacher effectiveness, though it might seem less tenable at first glance. Such an approach would contribute to establishing trust whose importance, to my understanding, has not been well-appreciated by researchers. The contradiction between teachers’ expectations and those of students could yield negative impacts on students’ study (Horwitz, 1990; Kern, 1995; Schulz, 1996, cited in Brown, 2009, p. 46). Therefore, as Brown asserts, foreign language teachers are required to explore any gaps between their beliefs and those of their students (p. 46). He further continues that L2 teachers and students might be of congruent or different perceptions of ‘effective teaching’ (p. 46). Hence, by virtue of a suitable and transparent appraisal model for teacher effectiveness, teachers, learners and administrators who are all perceived as the main stakeholders would have a mutual understanding of the qualities of an effective teacher. And herein lies a challenge in that many educational contexts suffer from a lack of a transparent and reliable teacher evaluation system or a faulty implementation of the policies as in Iran.
Since the history of teacher evaluation before 1970 is not clear (Shinkfield & Stufflebeam, 1995, p. 37), the current understanding of teacher evaluation has evolved from the 1970s onwards. Whereas teacher effectiveness and teacher evaluation have kept the attention in most Western contexts for which several appraisal models/schemes have been developed accordingly, their history in most Middle-eastern countries such as Iran seems to be narrowly defined and confined to merely accountability approaches. The emergence of some governmental and non-governmental organisations and institutions as well as special legislation and schemes/models on high quality teaching in some countries indicate the importance of teacher evaluation and the prestige it has thus far gained. The Office for Standards in Education, Children’s Services and Skills (Ofstead) in England, The National Board of Professional Teaching Standards in the United States (NBPTS), The Organisation for Economic Co-operation and Development (OECD) with 34 member countries from North and South America, Europe and the Asia-Pacific region (OECD) are examples of organisations dealing with excellence of education. Such sensitivities to effective teaching have led governments and organisations to develop and propose different schemes, frameworks or models for teacher effectiveness. The Hay McBer model of teacher effectiveness (McBer, 2000), the differential model of teacher effectiveness (Campbell, Kyriakides, Muijs, & Robinson, 2003; Campbell et al., 2004a; Campbell, Kyriakides, Muijs, & Robinson, 2004b; Cheng, Mok, & Tsui, 2001; Cheng & Tsui, 1996, 1998, 1999) are instances of researchers’ endeavour to propose models for teacher effectiveness.

Nevertheless, teacher effectiveness research and its appraisal in Iran vis-à-vis the above-mentioned educational contexts seem less satisfactory. As stated earlier, a review of the literature reveals that non-Western contexts especially the Middle-eastern ones such as Iran seem to be rather left out in teacher effectiveness research. This gives rise to speculation about the extent to which the appraisal models developed and used for teacher evaluation in Western countries can fit those of non-Western contexts such as Iran, given the cultural and socio-economic underpinnings of such models. The need for further investigation into teacher evaluation system in the Iranian higher education seems inevitable as the current understanding of teacher effectiveness and its corresponding appraisal model is a
bit confusing and less well-understood and, as suggested by this study, fails to embrace different characteristics and qualities of an effective teacher transparently. Several lines of concerns arise as to the dynamics and mechanism of teacher appraisal in the Iranian context, especially the ones germane to data collection and interpretation. Deciding on the measures of evaluation (sources of information) and the weight each measure should carry in the overall appraisal are, amongst other issues, pieces of the jigsaw which need to be considered. The idiosyncratic challenges associated with EFL teaching, I surmise, add to such concerns. As Campbell et al. (2004a) eloquently remind us, we need to move beyond generic aspects of teacher effectiveness in that teacher effectiveness of the same teacher might vary with respect to the students, subjects, context, etc. (p. 4).

Another drawback to teacher effectiveness evaluation which has been seemingly less well-appreciated is the current understanding of the objectives of teacher appraisal, i.e. *formative* and *summative* teacher evaluation. This aspect of appraisal is of an utmost importance as it has led researchers throughout history to develop different models to accommodate different expectations of teacher evaluation, i.e. professional growth and accountability purposes. As to the universities and higher education institutions in Iran, and further to the findings of this study, it appears that the developmental (formative) dimension of teacher evaluation has been rather left out by administrators. As a consequence, further research is needed to cast new light onto the strengths and weaknesses associated with the current understanding of teacher effectiveness and evaluation in higher education in Iran. This can help policymakers, administrators, lecturers and students move from the generic yet often simplistic approach towards teacher effectiveness to a more specific perception of teacher effectiveness and hence a more specific appraisal model for EFL lecturers.

### 1.5. Significance of the study

The need to answer the above-mentioned problems and challenges with EFL teacher effectiveness appraisal, I think, will suffice for explaining the significance of this study. It is unfortunate that there is a dearth of research studies on EFL lecturers’ effectiveness in the Iranian higher education context. Such paucity of
research echoes Campbell et al. (2004b) notion that much of the research on educational effectiveness has been confined to Western contexts, viz. the United States, the Netherlands and the UK (p. 451). I hope that this study can cast light on the status of teacher effectiveness and the merits and demerits of teacher evaluation system in universities and higher education institutions in Iran by virtue of EFL lecturers’ views and opinions. Elucidating the rather implicit but nonetheless important dimensions of L2 teacher evaluation as well as identifying the challenges and opportunities associated with a successful teacher appraisal system, this study calls for a context-specific and discipline-specific appraisal model for appraising EFL lecturers in Iran.

The study is expected to contribute to a better appreciation of the key concepts of ‘effectiveness’, ‘appraisal model’, ‘evaluation system’, etc. and to raise stakeholders’ awareness of ‘what makes an effective teacher?’ and ‘what counts as an effective teacher appraisal?’ Adopting an exploratory stance, this study has tried to address this phenomenon from lecturers’ perspectives, and to that end it is relatively new with regard to the context of the study, i.e. Iran. The findings of this research are expected to yield promising contributions to the existing knowledge of teacher effectiveness. This will in turn help us move beyond the traditional conception of teacher roles and duties to a more holistic understanding of teacher effectiveness which can thereby inform current policies. As a core element in teacher appraisal policy, the conception of teacher effectiveness needs to embrace a wider understanding of teachers’ activities (Campbell et al., 2004a, p. 105). This study will have several seminal implications for policymakers, administrators, and teachers such as the influence of different constituents of an educational system such as curriculum, syllabi, teacher recruitment, teacher development programme, etc., exerted on teacher effectiveness.

Given Postareff, Lindblom-Ylänne, and Nevgi (2008, p. 30) notion that teachers’ conceptions of teaching impact their stances towards teaching, e.g. teacher-centered and student-centered approaches, the findings of this research bring to light such conceptions by virtue of exploring lecturers’ perceptions of teacher effectiveness. Another aspect of significance of this study is directed towards initial
teacher education and teacher professional development. As Fradd and Lee (1998, p. 763) contend, teacher change in terms of knowledge and skill is a lingering process. Building on lecturers’ experiences and stories, this study deepens insights into the nature of challenges EFL lecturers tend to face in bringing their ideas into action. This is of paramount importance, on the grounds that teachers’ beliefs are rarely translated into action (Fradd & Lee, 1998, p. 763). This simply testifies to the need of devising a teacher evaluation system which is capable of not only rendering summative evaluations but also yielding formative data on teachers’ professional needs. Teacher educators then may avail themselves of a plethora of ideas emerging through the analysis of teachers’ likes and dislikes, needs and ideals based on which they can enrich Initial Teacher Education (ITEP) and Teacher Professional Development (TPDP) programmes.

As stated earlier, one of the merits of this research project is the level of education on which it focuses. It is unfortunate that much of the research conducted on TE has mostly focused on primary or secondary education. Indeed, research on teacher effectiveness in tertiary and higher education is sparse not only in the context of this study but also in most educational systems around the world. The dearth of research in inquiring into TE in Iran as well as a lack of understanding of the extent to which research on TE in mainstream education relates to that of TESOL/TEFL are the ostensible reasons for the significance of this study. Notwithstanding the appraisal model proposed in this study is informed by lecturers’ perceptions as the main stakeholder, I think, it can help bring about a deeper understanding of teacher appraisal. It is hoped that the findings of this research help fill the existing gap to some extent even though further research is needed to address other dimensions of teacher effectiveness and evaluation.

1.6. Research aims and objectives

The present study is aimed at investigating teacher effectiveness from a newly critical perspective which helps the researcher address the Iranian EFL lecturers’ concerns and expectations about teacher effectiveness and its pertinent appraisal. Problematising the current awareness of teacher effectiveness and teacher evaluation, this study endeavours to respond to several concerns and questions
revolving around the mechanism and dynamics of the evaluation of teacher effectiveness in Iran. Indeed, the aforementioned ideas in Sections 1.3, 1.4 and 1.5 are the areas which this study attempts to address. To do so, this study investigates into teacher effectiveness appraisal from six distinct yet interrelated perspectives. As the core objective which serves as a platform for other research goals, this study investigates lecturers’ understanding of teacher effectiveness and the qualities and characteristics of an effective teacher. Afterwards, lecturers’ perceptions of measures of evaluation (sources of information) are explored. Collecting lecturers’ views about teacher effectiveness and teacher appraisal, this study then proceeds to inquire into the strategies through which lecturers can improve their effectiveness. As an important part of this research, lecturers’ stances towards the merits and demerits of the Iranian teacher evaluation system are addressed subsequently. The so-called non-teacher-controlled factors within a wider domain of educational effectiveness affecting teacher effectiveness are another important aspect this study endeavours to elucidate. Finally, drawing upon the information obtained from the above-mentioned five areas, this study pursues the matter further by virtue of lecturers’ views about an ideal teacher evaluation system based on which a new context-specific appraisal model will be proposed.

1.7. Research questions

As discussed earlier, this study aims to explore the EFL lecturers’ understanding of teacher effectiveness in the Iranian higher education context. Given the fact that investigating teacher effectiveness is inevitably imbued with teacher evaluation, the present study endeavours to delve into teacher appraisal system from the standpoint of EFL lecturers. Adopting an exploratory stance towards the phenomenon, this study tries to give voice to lecturers as the main stakeholders who are perceived to be rather marginalised in the current evaluation system. Given the extensive and broad nature of research on teacher effectiveness and teacher appraisal, this study has been narrowed down to higher education context. Although the present study is rather exploratory by its very nature, the proposed research questions resonate with not only exploration but also description and
Six major research questions as well as two sub-questions have been formulated for this study as follows:

1. What is EFL lecturers’ understanding of teacher effectiveness and the qualities of an effective EFL teacher in higher education?
2. What are EFL lecturers’ perceptions of teacher appraisal?
   2.1. What are the measures of evaluation?
3. What are EFL lecturers’ perceptions of teacher effectiveness appraisal in the Iranian higher education context?
   3.1. What are the merits and demerits associated with the current teacher evaluation system in Iranian higher education?
4. How can EFL lecturers improve their teaching effectiveness?
5. What are the non-teacher-controlled factors affecting lecturers’ effectiveness?
6. What is lecturers’ ideal appraisal model?

1.8. Definition of key terms

In view of the interchangeable use of some terminologies in the literature, I will briefly define the most important ones used throughout this thesis in this section as follows:

1.8.1. Appraisal


“In language teaching, procedures that an institution, school or organization has in place to provide for regular review and assessment of teachers’ performance. Appraisal may include appraisal by a supervisor, by a colleague, by students, or self-appraisal (p. 29).”

Poster and Poster (1993) propose the following definition of appraisal:

“Appraisal is a means of promoting, through the use of certain techniques and procedures, the organisation’s ability to accomplish its mission of maintaining or improving what it provides while at the same time seeking to maintain or enhance staff satisfaction and development (p. 1).”
1.8.2. Teacher evaluation

The overall understanding of teacher evaluation/assessment is rather consistent in the literature. However, in order to minimise the potential misinterpretation of the concept, I am inclined to refer to the definition proposed by Longman Dictionary of Language Teaching and Applied Linguistics as follows:

“Procedures used to gather information about how and how well a teacher teaches. Teacher evaluation may be based on observation, learner evaluations, student results, self-evaluation, interviews, portfolios, etc., and serves a number of different purposes, including identifying strengths and weaknesses, contract renewal and promotion, and as part of the process of staff development” (Richards & Schmidt, 2002, p. 453).

1.8.3. Teacher effectiveness

The core concept of teacher effectiveness has received undivided attention in the literature for which several definitions have been thus far proposed. From among different explanations of teacher effectiveness, I refer to the following definition:

“The power to realise socially valued objectives agreed for teachers’ work, especially, but not exclusively, the work concerned with enabling students to learn” (Campbell et al., 2003, p. 354).

Given the popularity of teacher effectiveness in the literature and also for the ease of discussion, it is worth highlighting that I have used teacher effectiveness and lecturer effectiveness with respect to the findings of this study interchangeably. In other words, when referring to teacher effectiveness in reporting and discussing the findings of this study, I mean university lecturers’ effectiveness in the Iranian higher education context.

1.9. An overview of the thesis

This thesis embraces seven chapters, viz. Introduction, Context of the Study, Literature Review, Methodology, Data Analysis and Findings, Discussion, and finally Conclusions and Implications. The first chapter provides some introductory insights into the very nature of this research, inasmuch as it introduces the aims of the study as well as research questions. The second chapter offers some rudimentary information germane to the context of the study, i.e. Iran. The
evolution of teacher effectiveness research throughout history will be addressed in the chapter of literature review. Identifying the gap in the literature, this chapter touches on the state-of-the-art knowledge of teacher evaluation and will briefly review some teacher appraisal models and schemes adopted in some countries around the world. Afterwards, the thesis proceeds to chapter four which is aimed at casting light on the methodological issues. Referring to the theoretical and philosophical assumptions deployed in this study in order to approach the phenomenon under investigation, the chapter thereafter examines research design, data collection and analysis strategies along with ethical considerations. Chapter five is allocated to data analysis in which the obtained findings will be reported. Drawing on the findings, chapter six will discuss the most important dimensions of the ideas garnered and generated through the analysis of the collected data. The thesis will eventually end by drawing conclusions and examining the subsequent implications of this study.

1.10. Summary

Introducing the crux of this research, this chapter took effort to bring to attention the rationale behind this study, thereby explaining the current problems and challenges in the context of this study. Justifying the significance of study, the chapter discussed the aims and objectives set for this research project. The research questions and sub-questions were also examined in this section of the thesis. Definition of some key terms and presenting the overall structure of the thesis were the final parts of this chapter.
Chapter Two: Context of the Study
2.1. Introduction

This chapter provides a general overview of the context of this study, i.e. Iran, from historical, educational and socio-cultural perspectives. Delving into the contextual dimensions of the site of the study will hopefully yield up some precise information thereby contributing to a better understanding of the present research. To do so, some general but nonetheless requisite information pertaining to the country is presented in the country profile section. Then, the history of the Iranian higher education system will be explained. A brief account of higher education system in the last century with a special focus on a 35-year time span after the Iranian Islamic Revolution will be presented afterwards. The status of ELT/TEFL in Iran and some facts and figures about the Iranian universities and higher education institutions will be highlighted, thereupon.

2.2. Country profile

Iran is a country in the Southwest Asia which borders Turkmenistan, Azerbaijan and Armenia on the northern border, Afghanistan and Pakistan on the east, and Iraq and Turkey on the west as illustrated in Figure 2.1. ("About Iran: Facts & Figures,").

![Figure 2.1. Iran Geographical Location (adapted from Islamic Republic of Iran Mission to the UN in New York)](image)

The history of Iran, as one of the oldest ancient civilisations, dates back to around 10,000 BC when humans lived in the southern region of the Caspian Sea ("About Iran: Facts & Figures,."). The official language of the country is Persian (Farsi) even though other languages such as Azeri, Kurdish, Arabic and Armenian are used in some publications and broadcasts ("About Iran: Facts & Figures,"). According to
National Geoscience Database of Iran, the country is about 1,648,195 square kilometers in area and is located in Iran Plateau district (NGDIR). According to Statistical Center of Iran (SCI), the total population of the country was 75,149,669 by 2011 (*Selected Findings of the 2011 National Population and Housing Census*). While 71.4% of the total population of the country live in Urban areas, some 28.5% live in Rural areas and the rest (1%) tend to be considered as Unsettled Households (*Selected Findings of the 2011 National Population and Housing Census*, p. 3). The SCI report (2011) introduces Tehran province as the most populated province with the total population of 12,183,391 (16.2%). With regard to the religion of the country, the census shows that the total population embraces Muslim (99.4%), non-Muslim (0.3%) and not stated (0.3%). Moreover, the literacy rate of the country population aged 10 to 49 was 92.4% by 2011 (*Selected Findings of the 2011 National Population and Housing Census*). As the report adds, 18.2% of male and 18.4% of female population had higher education by 2011 either as graduates or students (p. 35). With regard to the mean age of population, Iran is a relatively young country with the mean age of 29.86 by 2011 (*Selected Findings of the 2011 National Population and Housing Census*, p. 12).

### 2.3. A brief history of higher education in Iran

Higher learning has a 25 century history in Iran (Bazargan, 2007, p. 781). According to the Iranian National Commission for UNESCO (1977), the first higher learning center was established in the 6th century B.C. by King Darius of Persia (cited in Bazargan, 2007, p. 781). While The University of Gondishapoor (UG) is considered as the first formal university which was established in the third century A.D. (Hekmat, 1972, cited in Bazargan, 2007, p. 781), the first higher education institution similar to European ones was *Dar al-Funun* [House of Arts] which was established in 1851 by the country’s chief minister, *Amir Kabir* (1807-52) (Ekhtiar, 2001). Engineering, medical and industrial sciences were among the courses taught at Dar ul-Funun College ("History. University of Tehran,"). The evolutionary development of higher education centers finally led to the establishment of Iran’s most important university, *University of Tehran* in 1934. Being regarded as the symbol of higher education in Iran, the University of Tehran is the largest and
oldest university in Iran with some 42,486 registered students ("History. University of Tehran,"). As of the establishment of the University of Tehran, other universities gradually were established in other major cities, e.g. University of Tabriz (1946) and Shiraz University (1949) (MSRT). The country witnessed an exceeding expansion of higher education since 1985 to such an extent that the enrolment rate germane to the 18-24 age group increased threefold between 1979 and 1995 (Bazargan, 2007, p. 784). Since the present research is aimed at investigating the current status of teacher effectiveness and its appraisal in higher education in Iran, the following sections tend to focus on the Post-Islamic Revolution era (1979 onwards). As a starting point, it is worth starting with the position of education in the country’s constitution. Article 30 of the Constitution of the Islamic Republic of Iran (I.R. IRAN) urges the government upon providing free education and higher education facilities (The Constitution of the Islamic Republic of Iran):

“Article 30: The government must provide all citizens with free education up to secondary school, and must expand free higher education to the extent required by the country for attaining self-sufficiency.” ("The Constitution of the Islamic Republic of Iran. Last amended in 1989.")

The current Iranian Ministry of Science, Research and Technology (MSRT) has gone through different stages of development. After the Islamic revolution in 1979, the previously known Ministry of Culture and Arts and Ministry of Science and Higher Education united under the title of ‘Ministry of Culture and Higher Education’ (MCHE) (Bazargan, 2007, p. 782). Nevertheless, MSRT was the only organisation responsible for higher education affairs until 1985. Ministry of Culture and Higher Education entrusted all its duties and responsibilities pertaining to medical education to the newly established Ministry of Health, Treatment and Medical Education (MHTME) (Bazargan, 2007, p. 783). In addition to the universities and higher education institutions run by these two major ministries, there are a number of higher education institutions which are affiliated to some other ministries. For instance, The School for International Relations (SIR) is a higher education institution which is affiliated to the Ministry of Foreign Affairs (SIR). Fifteen years later, Ministry of Culture and Higher Education was renamed ‘Ministry of Science, Research and Technology’ (MSRT). As of 2000, MSRT and MHTME have been responsible for all higher education affairs (Bazargan, 2007, p. 783) even though
some educational policies are made at a higher national level the mechanism of which will be discussed shortly. In the next section, I will provide some detailed information apropos of types of universities, academic qualifications and degrees, etc., in Iran.

2.4. Universities and higher education institutions

The Iranian higher education system offers different types of education. Similar to most countries, the Iranian universities and higher education institutions are of two major types, i.e. public which are state-owned and operate on state budgets or private. Public universities embrace all universities run by MSRT and MOHME. However, as stated earlier, there are some other higher education institutions affiliated to other ministries which are accredited by MSRT. Both public and private universities offer full-time and part-time education. Whereas most state universities are located in one city or town, there are few universities which are geographically located nationwide. In other words, headquartered in Tehran, some universities have branches all over the country, e.g. Islamic Azad University (IAU), Payame Noor University (PNU) and University of Applied Science and Technology (UAST). Payame Noor University (PNU), established in 1988, is the only state distance education university which is the host of 3,500 academic staff and about 1,100,000 students ("Payame Noor University (PNU),"). The most prominent and the largest private university in Iran is Islamic Azad University (hereafter IAU) which is said to be one of the largest universities in the world. Founded in 1982, IAU hosts 30,000 academic staff and 35,000 administrative staff and has more than 4,000,000 graduates with a total number of 1,700,000 students currently studying in some 750 academic fields (IAU; "IAU, Vice Presidency for International Affairs,"). The university has more than 440 branches and education centres as well as some foreign branches in the UK, Dubai, Lebanon, etc. ("IAU, Vice Presidency for International Affairs,"). The extensive geographical distribution of public and private higher education institutions across the country emanates from the policy of higher education expansion in the last two decades (Bazargan, 2007, p. 785). The burgeoning rate of enrollments in public and private universities from 180,000 in 1979 to 1,321,752 in 1997 (Bazargan, 2007, p. 785) and to 4,367,901 in 2013-
2014 (IRPHE) apparently assents to the policy of expanding higher education in the country.

2.5. ELT in the Iranian higher education

Having provided a general overview of education system in the previous sections, the chapter briefly reviews the status of ELT education in the Iranian higher education in terms of the academic programmes and faculty recruitment. The programmes and courses offered by ELT departments in Iran are currently of two different types. The first and foremost type revolves around the programmes and courses which directly relate to English Language Teaching. The academic disciplines in this group include Teaching English as a Foreign Language (TEFL), English Language and Literature, English Language Translation, and Linguistics. These programmes are offered at all academic levels of education including Associate’s, BA, MA and PhD degrees. However, ELT departments are also deemed to support universities with some limited English Language courses which are offered to students of other academic disciplines, e.g. students of Science, Engineering, Arts, etc. All students regardless of their academic disciplines need to pass at least two and sometimes three English Courses. These modules are usually Pre-university English (depending on students’ performances on English test), General English and English for Specific Purposes (ESP). Given the different nature of the above-mentioned programmes and courses, ELT departments tend to employ academic staff with different academic backgrounds. In a nutshell, ELT departments currently tend to recruit academic staff who major in TEFL/TESOL/Applied Linguistics, English Language and Literature, English Translation and finally Linguistics. Depending on what programme they teach, e.g. TEFL, Literature, etc., lecturers including the participants of this study may teach different modules such as Research Methods, English Language Teaching Methodology, Language Testing and Assessment, Practicum, etc.

2.6. Policy making in the Iranian higher education system

The policy making and planning institutions in the Iranian higher education system operate at three levels, namely national, sector and university levels ("A National
According to the report, the policy making bodies at three levels are as follows:

**National level**

- “The Islamic parliament
- The Supreme Council for Cultural Revolution (SCCR)
- The Government
- The Supreme Council of Science, Research and Technology

**Sector (ministerial) level**

- The Ministry of Science, Research and Technology (MSRT) and the Ministry of Health, Treatment and Medical Education (MHTME)
- The Council for Higher Education Development
- The Supreme Council for Planning
- The Council of Medical Education
- The Central Council of Scholarship
- The Council of Talented and Gifted Education
- The Council of Supervision and Evaluation of Higher Education
- The Council of Centers of Excellence

**University level**

- Board of Trustees
- The Chancellor
- The Governing Council of University
- The University Council
- The University Specialized Councils
- The Councils for Learning and Research

Despite such a diverse policymaking bodies in the Iranian higher education system, the majority of the country’s educational affairs rest with MSRT and MHTME.

**2.7. Supervision and Evaluation of Higher Education**

As stated previously, all universities and higher education institutions in Iran including both state and non-state affiliated ones tend to be monitored, supervised and accredited by MSRT and MHTME. In recent years, these two ministries have started some measures to assess the quality of universities one of which was the MHTME pilot study on self-evaluation in medical education departments in 1997 (Bazargan, 1999, cited in Bazargan, 2007, p. 786). Similarly, MSRT has been
monitoring the quality of education through different mechanisms the most important of which relates to the Council of Supervision and Evaluation of Higher Education (hereafter CSEHE). CSEHE is one of the key departments within MSRT which is deemed to supervise and evaluate the overall quality of higher education. The council embodies different offices and committees and is responsible for, amongst other duties, proposing policies germane to evaluation and supervision of universities and higher education institutions and improving the quality of education and research thereupon (Mohammadnejad, Roshan, & Motahari, 2010, p. 22; "A National Report of Higher Education, Research and Technology (2009-2010)," 2010, p. 12). MSRT and MHTME have a broad range of activities apropos of the supervision and evaluation of higher education for which they adopt versatile policies. A major step towards the realisation of the council policies was made in 2007 when all universities and higher education institutions were required under the new regulations to open offices for supervision and evaluation which are deemed to send regular reports to the secretariat of CSEHE at the end of each semester (MSRT, 2007). The duties and responsibilities of CSEHE are beyond the scope of this study. Nevertheless, teacher evaluation has always been an indispensable part of a wider evaluation of higher education which is indeed the focus of this research and hence will be briefly explained in this section. Lecturers are normally evaluated as part of their annual applications for promotion under MSRT professional promotion scheme which per se has been developed by The Supreme Council for Cultural Revolution (SCCR). Teacher professional promotion scheme is rather complex in that it embraces a variety of teachers’ activities in different areas, e.g. socio-cultural, educational, research-technological and scientific-executive (managerial) activities (MSRT, 2011, pp. 349-354). The data on the quality of lecturers’ educational effectiveness which is the focus of this study are usually collected from students, top graduates (with distinction), Head of the department and the Deputy Dean (Education).

It is worth highlighting that this section was merely aimed at giving a vivid picture of the current legislation and rules pertaining to teacher evaluation in the Iranian universities and higher education centres. The strengths and weakness associated with such scheme and the extent to which it is implemented in practice (policy-
implementation) are the major concerns and questions to which I will return in the next chapters in which I will report some important findings.

2.8. Summary

This chapter was aimed at providing a vivid picture of the context of the study. The chapter started with introducing Iran and reviewing the history of higher education, in brief. Afterwards, the Iranian higher education system with regard to types of universities and higher education institutions were explained. A brief review of the status of ELT in Iran was the next section which was presented in 2.5. The chapter then proceeded to introduce the processes of policymaking in higher education in Iran. The mechanism of supervision and evaluation of higher education was the final section of this chapter.
Chapter Three: Literature Review
3.1. Introduction

In previous chapters, the notions of teacher effectiveness and teacher evaluation system with special reference to the Iranian higher education context were introduced. Drawing on the proposed research questions, this chapter is aimed at reviewing the literature from different perspectives thereby identifying the lacunae in teacher effectiveness research in the world as well as those of the Iranian context. To so do, the chapter commences with revisiting the major concepts and constructs. Reviewing the literature on teacher effectiveness in mainstream (general) education and thereafter in second/foreign language education, the chapter then proceeds to review the current understanding of TER and the pertinent studies in the Iranian context. This will provide a platform for formulating the conceptual framework underpinning this study. A brief account of some legislation and Acts, schemes and models which are currently in use in some educational systems will be discussed afterwards. Reflecting on the existing gaps in the literature, the chapter ends with some concluding remarks.

3.2. Teacher effectiveness: an introduction

Perhaps, the history of students’ practice of drawing comparisons among their teachers and professors is as old as the history of education. Such comparisons appear to be a natural phenomenon born of humans’ quest for the better and/or the best. It is argued that research on language teaching has been informed by advancements in mainstream education, e.g. the paradigmatic shift from information processing to sociocultural theories of learning in general education in the 1990s (Tsui, 2011, p. 278). Moir (2009) maintains that teachers could be considered as the most seminal dimension of children’s education (p. 15). Teachers can exert differential influences on students’ academic life. Despite the importance of other influential components of an educational system, teachers’ centrality is more or less agreed on by researchers and practitioners. In order to reach a comprehensive understanding of TE, it appears inevitably important to address teachers’ learning to teach and teaching practices and quandaries from different angles, e.g. level of education, context of the study, etc. I will argue how
the related literature tends to reflect contentious gaps with respect to each of these dimensions whereby I will review some current approaches towards TE.

It is not easy to apportion priority among different aspects of research on teachers such as teacher education, teacher development, teacher evaluation, etc., given the fact that they are closely entwined with one another. By the outset of the third millennium along with its technological advancements in education, teachers' awareness of teacher appraisal started to increase to such an extent that administrators nowadays are increasingly required to convince teachers of the appropriateness and fairness of their appraisal systems which are often deprecated by teachers. At the heart of teacher effectiveness research rests the issue of quality. As Darling-Hammond (2000b, p. 33) suggests, well-preparedness of teachers exerts more influence on students’ success and achievements than any other factors such as students’ backgrounds or teachers’ education level. Teachers are exceedingly facing up expectations on the part of different stakeholders, viz. students, administrators, policymakers leading to their increased responsibilities and/or accountability. Most importantly, these stakeholders each have their own convictions that their voices need to be included in teacher appraisal. In other words, whereas teachers tend to possess deep conviction that they are the predominant stakeholder in the phenomenon whose voices need to be included as the basis for making decisions on their effectiveness, others are under the impression that students’ voices can best provide administrators with a platform for teacher evaluation. This is why the trade-off among different measures of evaluation has always been open to discrepancy. However, as stated earlier, the ultimate goal for teacher evaluation is mostly the issue of ‘quality’. The tendency to explore and establish methods for assessing teacher effectiveness seems to emanate from the current movement to improve the quality of teaching (Darling-Hammond, 2009, p. 1). High quality teachers are of prominent role in any educational system and as Darling-Hammond maintains are at the heart of any school reform (p. 1). In a similar vein, Korthagen (2004, p. 77) argues that the essential qualities of good teachers and how people can become good teachers are two important questions resting at the heart of the pedagogy of teacher education.
An important aspect of TER which is yet the mainspring of debate in the literature is the very concept of ‘effectiveness’. As Campbell et al. (2004a, p. 2) remind us, in spite of their interrelatedness, teacher effectiveness, school effectiveness and educational effectiveness tend to be used inconsistently in the literature. In addition, as Arthur, Tubre, Paul, and Edens (2003, p. 275) eloquently put it forward, deciding on effectiveness needs to take into account the fundamental question of “effective in terms of what?”. Therefore, arriving at a particular definition of effectiveness seems to be the starting point for investigating teacher effectiveness evaluation. This is of an utmost significance as effective teaching can be interpreted in different ways.

It is worth highlighting that there is more to teachers’ effects in an educational system that can meet the eye. Teachers’ role in shaping their students’ future life especially at school level is an undeniable fact that has been more or less agreed in the literature. However, the traditional yet predominant conception of teachers’ effects is often confined to students’ learning outcomes which should be demarcated from teachers’ wider influences on students' social lives. Creemers (1994, cited in Kyriakides, Campbell, & Christofidou, 2002, p. 291) asserts that the activities practiced in the classroom can shape students’ academic outcomes. The review of the literature testifies to a transition from the so-called traditional and unidimensional conception to a rather sophisticated and multidimensional understanding of teacher effectiveness, from a single-criterion evaluation to multiple-criteria assessments, and finally from teachers’ intra-classroom practices to teachers' multifaceted professional responsibilities and activities.

The issue of teacher effectiveness has kept the attention of researchers and practitioners in the last decade especially in primary education. It is argued that such willingness on the part of researchers and practitioners emanates from change of needs and thereby policies. As Ingvarson and Rowe (2008) contend, the global economic, technological and social changes have led to an imperative need for high-quality teaching (p. 6). However, as stated previously, the true meaning of teacher effectiveness seems to be shaped by other components of educational systems such as teacher education programme, teacher development programme,
teacher evaluation system, teacher recruitment, etc., all of which form a wider
domain of educational effectiveness. For instance, as Ellett and Teddlie (2003, p.
103) assert, teacher effectiveness appears to be inevitably interconnected with
teacher evaluation and school effectiveness.

It is worth mentioning that the issue of ‘quality’ rests at the heart of the above-
mentioned notions, albeit their differences. Nevertheless, since each of these
concepts and notions carries different connotations, I am inclined to use with
cautions concepts such as quality which as Sayed and Ahmed (2011, p. 103)
contend, remains yet less well-understood.

3.3. Teacher effectiveness, a multidimensional phenomenon

Teacher effectiveness research has always been a multifaceted phenomenon
throughout its history. Investigating TE from different perspectives through different
lenses and from different angles would help us better understand the underlying
processes of teacher effectiveness in different contexts. ‘Context’ does not
necessarily mean two different countries, given the fact that even within a similar
geographical territory, e.g. same country or even same city, our understanding of
TE could be quite different from one teaching context to another. For instance,
teacher effectiveness could be understood and hence interpreted differently in
public vs. private, primary vs. higher education contexts. As Darling-Hammond
maintains, teachers who teach students at public schools with less access to
educational resources might require more skills (2009, p. 1).

And this is why the

use of the so-called ‘one-size-fits-all’ approach is expostulated in the literature.

Kyriakides et al. (2002, p. 299) criticise the use of students’ progress as a measure
for identifying the qualities of an effective teacher, on the understanding that
effective teachers can serve other purposes such as contributing to national policy
development. Therefore, in pursuance of a deep understanding, it seems wise to
investigate TE from different but nonetheless interrelated perspectives, viz. subject
matter, level of education, measure of evaluation (sources of information), context
of the study, etc., each of which is to be briefly introduced in this section with a
follow-up detailed explanation in forthcoming chapters. It might be worth
highlighting that such categorisation is for ease of a better understanding of the
literature; perhaps some other elements could be added to the proposed classification.

3.3.1. Subject area

Having reviewed a bulk of literature and having taken the focus of this study into account, I am inclined to classify the entirety of literature on teacher effectiveness into two major categories. These include teacher effective research in mainstream education and second/foreign language education. The comparative study of TER in mainstream education and L2 education would bring up a myriad of interesting questions and points for further consideration as to, for instance, the extent to which the implications of each of these categories are applicable to the other one, i.e. TEFL vs. mainstream education. This category is of an utmost significance as much of our understanding of TER has its roots in mainstream education.

3.3.2. Level of education

Another lens though which research on TE could be reviewed is the level of education within which research has thus far been conducted. As it will be argued later, we will notice that the bulk of research in this domain has been much dedicated to primary and secondary education with few studies addressing the issue in higher education contexts. In other words, the literature would suggest a thought-provoking gap between research on TE in primary and secondary education and that of higher education (HE), thereby raising the question as to whether or not the obtained findings and hence their pedagogical implications could be generalised to HE context. Magno (2009) reminds us of the exceeding sensitivities of evaluating teacher performance at college level in which the decisions might bear professional consequences such as hiring and tenure (p. 76). I will argue then how such sensitivities may vary and perhaps intensify with regard to the level of education.

3.3.3. Measures of evaluation (source of information)

Measures of evaluation are an important dimension of TE research. These measures are in fact sources of information based on which teachers’
effectiveness is often evaluated. This aspect of teacher evaluation has been the focus of discrepancies among researchers and practitioners. With regard to the contentious nature of this area, I may refer to the old yet seminal question as to whose voices need to be included in the appraisal. Moreover, much of the debate surrounding measures of evaluation proceeds from the priority that should be given to each measure in the overall assessment. For instance, whereas some researchers place emphasis on students’ ratings as the most important criterion for judging teachers’ effectiveness, other scholars give more weight to other measures such as students’ learning outcomes, teachers’ self-evaluation, peer supervision, etc. I will come back to this important aspect of TE in details shortly.

3.3.4. Western vs. non-Western (Middle-eastern) education

The review of literature has brought to my attention the idea that the majority of leading research on TER has been conducted in Western educational systems with few studies in non-Western societies, let alone in Middle-eastern contexts. It is unfortunate that such an important aspect of educational research has been given scant attention in some Middle-eastern contexts. This would sow seeds of concerns and doubts about the underlying reasons for such negligence in some contexts which appears to have its roots deep in political, socio-cultural and socio-economic conventions, values and norms. This is extremely important as TER is highly imbued with contextual values. I would suggest that the existing Western-oriented understanding of TE and appraisal models for TE cannot be simplistically transferred to non-Western contexts. In other words, it is unwise to adopt a model from one country and put it into practice in another country. Indeed, it is unlikely for such a model to accommodate the contextual underpinnings such as the socio-cultural factors associated with teacher evaluation in the target country.

3.3.5. Other less well-investigated areas

As stated earlier, the above-mentioned perspectives are deployed for ease of reviewing the literature based on the proposed research questions and hence other comparisons could be drawn therefrom. One of these elements pertains to methodological orientations of research on teacher effectiveness. The majority of
studies conducted so far are either qualitative or quantitative with little space for other possible methodological stances which in turn confines the interpretations of the gained results. This dimension will be discussed in details in the upcoming chapter of methodology. Teacher effectiveness in developed vs. developing countries could be another exemplar of other less well-investigated and grey area in the literature. For instance, Mangiante cast some doubts on whether or not students in low-income communities could have access to qualified teachers (2011, p. 41). The challenges of setting standards/criteria for teacher effectiveness and considering the impacts of non-teacher-controlled factors on teacher effectiveness, amongst others, are other critical areas which are yet to be thoroughly scrutinised in different contexts including Iran.

### 3.4. Teacher effectiveness in mainstream education

It has been argued that teacher effectiveness research is a rather sophisticated area in educational research. Given the fact that research on TEFL lecturers’ effectiveness is highly informed by that of mainstream (general) education and hence is mostly underpinned by similar theories, the literature reviewed in this section generally appertains to general education embracing both theoretical and experimental studies.

Teacher effectiveness research has widely been investigated in the literature (e.g. Campbell et al., 2003, 2004a, 2004b; Cheng & Tsui, 1996, 1998, 1999; Coombe, Al-Hamly, Davidson, & Troudi, 2007; Darling-Hammond, 2009; Ellett & Teddlie, 2003; Kyriakides et al., 2002; McBer, 2000; Muijs, Campbell, Kyriakides, & Robinson, 2005; Ramsden, 1991; Robinson & Campbell, 2010; Rockoff & Speroni, 2010; Stronge, Ward, & Grant, 2011). Accordingly, several models and schemes have been proposed thus far for evaluation of teacher effectiveness which will be reviewed in details in Section 3.12. However, TER has obtained much of its reputation within the realm of general education. It is worth noticing that even within mainstream education per se, different conceptions could be drawn from the central question of ‘what counts as an effective teacher’ on the grounds that different higher education institutions might have a different understanding of TE and perhaps different expectations of their academic staff. As Magno (2009)
maintains, different institutions might adopt different methods for measuring their teachers’ performances such as classroom observation, supervision, peer and student feedback (p. 75). Whatever the procedure, Magno believes such measurement should be in line with the institution’s mission and vision (p. 75). Drawing on Scriven (1969) and Stufflebeam’s (2000) conception of metaevaluation (evaluation of an evaluation), Magno (2009, p. 78) conducted a metaevaluation study on teacher performance system in Manila, the Philippines to explore the extent to which the evaluation system has taken into consideration the standards of feasibility, utility, propriety and accuracy (p. 75). The results suggested that whilst the metaevaluation standards of utility, propriety and feasibility were fair, the accuracy standard was poor (p. 90).

The qualities and characteristics of an effective teacher, e.g. academic qualification, have always been one dimension of research on teacher effectiveness. In a large-scale study with 100,000 participants including 10,000 Australian school teachers and 90,000 students, Leigh (2010, p. 480) explored teacher effectiveness with a panel data driven from students’ test scores in a two-year time span. Whilst teacher experience was found to exert the strongest effect, the results showed no evidence to teacher higher qualification as an indicator of students’ achievements (p. 480). Similarly, in their review of the literature, Chingos and Peterson (2011) scrutinise the prior research on teacher effectiveness from a number of perspectives. They address six elements that might have impacts on teacher effectiveness including “Pre-service teacher preparation, University selectivity, on-the-job training (teacher experience), teacher examinations, certification by the National Board of Professional Teaching Standards and finally Master’s degree” (pp. 450-452). Rejecting the significant impact of teachers’ type of qualifications and the universities from which they have been graduated, Chingos and Peterson claim that teacher’s experience, what they call ‘on-the-job training’, could be relatively influential in evaluating teacher effectiveness, though it might have some negative impacts at some point in future (p. 464). They finally conclude that identifying effective teachers would be more straightforward than training less effective ones to become more effective (p. 464).
Teaching 'certification' has also been a point of contention in TER history. While many institutions across the world are aimed at certifying novice teachers for teaching careers, Jacob and Lefgren (2008) cast doubts on the axiom that certified teachers do better and are more effective than their uncertified counterparts (p. 103) and challenge such mentality among educators. In their study, Jacob and Lefgren tried to identify the extent to which principals could identify more effective and less effective teachers. Although their findings indicated that principals have the potential in distinguishing the teachers on the two extremes of the effectiveness continuum, i.e., teachers with largest and smallest achievement gains, they fail to identify teachers who are in the middle, i.e. those who are relatively effective (2008, p. 103). In this study, principals measured teacher effectiveness based upon students' math and reading achievement (p. 103). This raises some reservations about the gained findings, given the fact that the use of students' outcomes as a single measure could not provide a comprehensive picture of multidimensional aspects of teacher effectiveness as well as the biases associated with principals' evaluations.

While 'context' as an influential factor in teacher effectiveness has received undivided attention, its multifaceted realisations seem to be less well-investigated. In a multivariate study on teacher effectiveness, Teddlie and Liu (2008) examined teacher effectiveness from two perspectives, namely the community type (rural/urban schools) and level of effectiveness (more and less effective schools) (p. 387). The results provided evidence for a relationship between effective schools and more effective teachers and also between rural schools and more effective teachers (p. 401). In another study on Singapore primary teachers, Kelly, Ang, Chong and Hu (2008) investigated the attributes of teacher performance appraisal system and tried to identify the extent to which such attributes can influence issues such as job satisfaction and motivation (p. 39). They examined five variables viz. “fairness, clarity, controllability of the appraisal system, teacher participation in appraisal development and the appraiser-appraisee relationship and the appraiser’s credibility” (Kelly et al., 2008, pp. 43-44). The findings suggested that all variables except teacher participation in developing the appraisal system tend to exert positive influence on issues such as enhancing teacher's satisfaction,
motivation and collegiality and also lessening the stress with the appraisal system (pp. 50-52). Similarly, Elizabeth, May and Chee (2008) conducted a study to investigate how what they call ‘outstanding’ teachers perceive teacher effectiveness and endeavoured to devise a model for teacher success (p. 623). Based on the data collected from 15 primary and secondary teachers in Hong Kong, they concluded that the concept of ‘teacher success’ should be approached from a broader rather than a narrowly defined perspective (p. 631). In other words, the coexistence of personal attributes, professional qualities and contextual factors need to remain at the heart of teacher success (pp. 631-633).

As mentioned earlier, the appraiser-appraisee relationship is one of the dimensions of teacher effectiveness research that has been less well-investigated so far. Chow, Wong, Yeung and Mo (2002) in a study conducted in an Elementary school context in Hong Kong, explored teachers’ perceptions of such a relationship and identified the effects which different appraisers can exert on teachers as appraisees (p. 98). While Chow et al.’s (2002) study focused on teachers’ perceptions, Slate, LaPrairie, Schulte, and Onwuegbuzie (2009) conducted a study to investigate college students’ perceptions of their best and poorest professors. Based upon the obtained findings, 15 themes (e.g. communicative, fun, motivating, etc.) were developed for effective teachers the majority of which targeted cognitive aspects of a teacher (p. 75). On the other hand, 12 themes (e.g. uncommunicative, boring, uncaring, etc.) were identified as pertinent to poor or ineffective teachers which tended to be mostly the opposite traits of effective teachers (p. 75).

Teachers’ characteristics such as gender, age, etc., could be perceived as a dimension of teacher effectiveness which is still in dispute. Based upon UNESCO advocacy brief (2006) which advocates the positive impacts female teachers can exert on girls’ achievements, Chudgar and Sankar (2008) conducted a study to examine the nature of the relationship that exist between teacher’s gender and students’ achievements in Indian context. While the obtained findings gave evidence to a difference between teaching practices of male and female teacher, they gave little evidence to the influence that teachers’ gender could exert on students’ learning outcomes, given that such a relationship tends to be influenced
by other variables such as the subject taught (pp. 627-639). The relationship between teachers’ behaviour and students’ learning outcomes has also drawn researchers’ attention. In their investigation into the effect of teacher behaviour embracing eight factors, viz. “orientation, structuring, questioning, teaching modeling, application, management of time, teacher role in making classroom a learning environment, and classroom assessment”, on student achievement, Panayiotou et al. (2014, pp. 1-2) noticed a positive relationship between teacher behaviour and student achievement.

What I presented in this section was merely the gist of the research on teacher effectiveness in different parts of the world. A detailed review of the literature on TER along with a review of some appraisal models and frameworks proposed in the literature (e.g. Campbell et al., 2004a; Cheng & Tsui, 1999; McBer, 2000; Middlewood & Cardno, 2001; Piggot-Irvine, 2003; Poster & Poster, 1993; Wragg, Wikeley, Wragg, & Haynes, 1996) will be addressed in Sections 3.10, 3.11 and 3.12.

3.5. Teacher effectiveness research in higher education

As discussed in Section 3.3.2., much of the literature on teacher effectiveness research pertains to schools, given that teacher effectiveness research in higher education is relatively sparse. Perhaps this is due to the fact that research in higher education is behind that of primary/secondary education. Nevertheless, as the literature suggests, there has been growing public attention to higher education in the last two decades especially on the part of governments (El-Khawas, 2007). It has been argued earlier that the very concept of teacher effectiveness tends to be more or less similar in most educational contexts. Despite all the similarities between TER in schools and universities, however, it is important to identify the differences between what counts as ‘effective teaching’ as well as ‘teaching excellence’ at different levels of education, i.e. primary, secondary and higher education. Such recognition appears to be an important aspect of researching into teacher effectiveness and teacher evaluation, inasmuch as teachers’ responsibilities, challenges and struggles at different levels of education may vary. Examples of such varied responsibilities with which teachers tend to struggle in
higher education, amongst others, are pressure to publish, grantsmanship, large classes with greater student diversity, etc. (see Knight, 2002, p. 215). Notwithstanding lecturers are subject to such various duties and responsibilities, a review of the literature reveals that university teachers tend to have fewer training opportunities than their counterparts at schools. Such lack of professional preparation for EFL lecturers seems to be a true reflection of the context of this study, i.e. Iran, in that rarely are lecturers required to participate in initial training programme. The paucity of professional training requirements for university academics as compared to school teachers (Laurillard, 2002, p. 12) clearly highlights the significance of research into TE in higher education.

Another point of difference between teaching at universities and schools relates to the expectations of the society from each of these educational environments. There has been an increasing demand in many educational contexts for universities to go beyond their routine educational activities and become more involved in other sectors such as culture, industry, etc. Such expectations have led universities and hence lecturers to take on more demanding roles. However, it seems that with the emergence of ‘mass higher education’, the traditional conception of higher education which considers universities as “cultural institutions” carrying “intellectual, academic, and national traditions” has been eroding (Välimaa, 2008, pp. 9-10). Noteworthy among the concerns surrounding teacher effectiveness research in higher education are the issues of ‘managerialism’ and ‘accountability’. Universities’ accountability towards high quality education has become more concrete over the past decade. The emergence of “performance indicators (PIs)” in higher education appears to be a response to governments’ willingness to increase universities’ accountability (Ramsden, 1991, p. 129). Morley (1997, p. 234) maintains that what counts as effectiveness seems to be a matter of “political judgment” and is one of the focuses of new managerialism in higher education.

‘Goals’ of higher education are another important piece of the jigsaw in building an accurate understanding of teacher effectiveness in universities. Indeed, evaluating teacher effectiveness in higher education in any particular context tends to be contingent upon the understanding of the goals of higher education which from
Forest's (2007, p. 351) viewpoint range widely, “from discipline-based factual knowledge to critical thinking, and from moral and ethical behavior to civic management”. Such awareness is important as our understanding of the 'goals' of higher education influences our understanding of the duties of lecturers which, in turn, shapes our understanding of the characteristics of an effective lecturer. For instance, the literature shows that the focus of research on teaching in the last couple of decades has been mostly directed towards the 'learner' rather than teaching (McKeachie, 2007, p. 458). I think this is why, as Atkins and Brown (2002) assert, effective teaching is sometimes equated with student learning indicating successful teaching. However, as Atkins and Brown continue, in addition to 'success', effective teaching should be concerned with some “appropriate values” (p. 5). It appears that such values are mostly, if not determined, at least influenced by the society within which teaching takes place. As Ashwin (2009) argues, there is a strong evidence in the literature that societies shape their higher education systems. Therefore, it seems essential for researchers and practitioners to pay more attention to teachers/lecturers who play an important role in high quality education.

Teaching effectiveness in higher education has been investigated in many educational contexts even though at a smaller scale as compared to teacher effectiveness in schools. However, similar to that of schooling system, research on teacher effectiveness in higher education system is subject to two important concerns, i.e. characteristics of effective teaching and measures of evaluation. As Murray (2007, p. 145) contends, awareness of the characteristics contributing to effective teaching has two advantages, i.e. “a better theoretical understanding of teaching” and “the development of improved programmes for faculty selection, faculty evaluation, and faculty development”. It has been argued that qualities such as knowledge of subject matter, ability to transfer knowledge, etc. tend to be characterised as traditional conceptions of effective teaching. However, research on TER in higher education has revealed that such conceptions can be influenced by some misleading factors such as “Dr Fox Effect” (see Ware & Williams, 1975; Williams & Ware, 1977). In their study on the impact of lecture fluency on students’ metacognitive knowledge, Carpenter, Wilford, Kornell, and Mullaney (2013) found
that students’ perceptions of their lecturer’s effectiveness were based on lecture fluency rather than their own actual learning. This finding shows that lecture fluency in academic settings could be deceptive, thereby misleading students’ perceptions of their own attainment and hence their lectures’ effectiveness (Carpenter et al., 2013, p. 1354).

In addition to the aforementioned question of the ‘qualities of an effective lecturer’, there is another important question whose answer helps better understand how teacher effectiveness research can be approached in higher education. The question then is whose voices need to be heard in teacher evaluation? This important aspect of TER will be discussed in details in Section 3.10. It is worth reiterating that researchers’ understanding of teacher effectiveness in universities is somehow consistent with that of schools in terms of its theoretical underpinnings. However, due to increasing pressure and more demanding roles with which lecturers tend to struggle in higher education, it is important for researchers and practitioners to appreciate the unique challenges of teaching in universities and turn their attention to lecturers who play a leading role in high quality education, and consider their wants and needs.

3.6. Teacher effectiveness in ESL/EFL education

It is argued that language teacher research tends to be a microcosm of general or mainstream teacher research with similar underlying principles and underpinnings, though each of which has its own idiosyncrasies and exceptions. Similar to research in mainstream education, TE has been researched in the language education domain, even though on a far small scale (e.g. Bailey, 2006; Borg, 2006; Coombe et al., 2007; Nerenz & Knop, 1982; Pennington & Young, 1989). Investigating into teacher effectiveness is a daunting task especially in cross-cultural contexts (Bailey, 2006, p. 217). The theory-practice nexus and teachers’ transition from what they have been taught theoretically during their academic education to practical intra-classroom activities have been investigated throughout the history of language pedagogy. Moreover, the use of concepts such as theory and practice has often been a matter of sensitivity. Ellis (2010, p. 185) reminds us about the difference between practice of researchers with that of teachers as well
as the distinctive understanding of *theory* between researchers and teachers. Ellis proposes two ways to fill such lacuna, namely the use of ‘research done by teachers for teachers’ (Stewart, 2006, p. 422, cited in Ellis, 2010, pp. 185-186) and the transfer of researchers’ technical knowledge to teachers’ practical knowledge’ (p. 186). Perhaps, identifying the intricacies and peculiarities inherent in language teaching could help us better understand the similarities and differences between teaching in its generic sense and in that of language education. Not only do language teachers have all general teachers’ concerns such as ‘subject-matter knowledge’, they tend to encounter L2-specific challenges such as ‘teacher language awareness’ which as Andrews (2003, p. 81) claims is an important issue in teacher professionalism.

The relationship between teachers’ expertise and their effectiveness is another piece of TER which has been investigated by several researchers (e.g. Kreber, 2002; Palmer, Stough, Burdenski, & Gonzales, 2005). The concept of ‘expertise’ has received considerable attention in the literature upon which as Murray (2001) states there exist some differences of views (Cited in Katz & Snow, 2009, p. 73). Tsui (2003) in her book ‘Understanding Expertise in Teaching’ has addressed this issue from different perspectives. Tsui introduces experience and practice as the major cornerstones of acquiring expertise (2003, p. 20). Goodwyn (2011, p. 9) maintains that the concepts of expertise and expert teaching cannot be addressed in a vacuum and introduces knowledge, skill and values as elements which are associated with expertise. Based upon her review of the characteristics of novice and expert teachers, Tsui concludes that expert teachers are more efficient and selective in processing information with better improvisational ability and tend to be more autonomous and flexible in planning and teaching (2003, pp. 36-41). Tsui proposes three dimensions entailing critical differences in expert teachers’ understanding as follows: ‘The extent to which teachers integrate or dichotomise different aspects of teacher knowledge, the extent to which teachers relate to specific contexts and ‘situated possibilities’ and finally the extent to which they are capable of theorizing their practical knowledge and practicalise their theoretical knowledge (pp. 247-253).
Another important issue that has been especially of ELT researchers and practitioners’ interest is the highly reputed notion of ‘reflective practice’. Farrell (2008, p. 1) maintains that the ‘constructivist learning theory’ in which reflection is a core element, feeds into the revitalisation of reflective practice. Introducing three approaches towards reflection, namely ‘action research, teaching journals, and teacher development groups, Farrell concludes that reflective practice provides teachers with an opportunity to deepen their understanding of teaching, evaluating their professional development and engendering decision-making skills which can reinforce their confidence (pp. 2-3). Teachers’ self-improvement of their teaching practices through reflecting upon their experiences critically and systematically (Farrell, 2007, 2013), can eventually enhance their teaching effectiveness. Such reflections on teachers’ own practices along with other techniques for promoting teacher effectiveness can help teachers evolve more smoothly and improve their teaching practices with higher level of confidence.

In this section, I will briefly review a number of experimental studies on teacher effectiveness in L2 context which can deepen our insights into the status of TER in the language education realm. In a two-year study aimed at investigating teaching effectiveness in Turkey, Eken (2007) addresses what she calls ‘less easily definable aspects of teaching’ (p. 176). The findings suggest that the challenges with which teacher trainers tend to be confronted emanate from three major areas, viz. teachers’ personal qualities, communication skills and the interactions they tend to have with students and finally the use of affective techniques (Eken, 2007, p. 177). It is worth noticing that the concepts of teacher effectiveness and effective teaching are often perceived differently on the part of different stakeholders, even within the same context, e.g. one single classroom. Different stakeholders could have different expectations of teacher effectiveness. This dimension of teacher effectiveness has been examined by Park and Lee (2006) in Korean context. They investigated the characteristics of effective English teacher from both teachers and students’ points of view (p. 236). The results from a self-report questionnaire focusing on three major characteristics of effective teaching, viz. English proficiency, pedagogical knowledge and socio-affective skills suggested that there is a significant difference between teachers’ perceptions of effective teacher and
those of students (Park & Lee, 2006, p. 246). Interestingly, the findings indicated that while ‘English proficiency’ was ranked the highest by teachers, ‘pedagogical knowledge’ was perceived as the most important element by students (p. 246). Based upon such findings, Park and Lee advocate the centrality of promoting English proficiency to teacher education programmes and call for teachers’ acquaintance with L2 acquisition theories, teaching methods and testing (p. 246).

3.7. Teacher effectiveness research in Iran

I begin this section by recapitulating my earlier argument that the existing literature on teacher effectiveness in most Middle-eastern countries including Iran does not suffice for understanding the minutiae of such a complex phenomenon. Given the lacunae in teacher effectiveness research in Iran, it is my contention that building on the previous limited works, this research can conduce to raising the awareness of teacher effectiveness and teacher appraisal. The problem tends to be further aggravated, on account of a dearth of research on EFL teacher effectiveness in higher education context with very few studies. Prior to reviewing some of the studies on teacher effectiveness in Iran, it appears wise to briefly explain teacher recruitment process in the Iranian higher education system. To qualify as a university lecturer, the applicant is required to hold at least Master’s degree from one of the universities and higher education institutions accredited by MSRT. Nonetheless, due to the current accelerating demand for running postgraduate programmes, universities are currently interested in employing PhD holders or PhD candidates instead. Having met the essential general and scientific (interview) criteria successfully (MSRT, 2011, p. 409), the applicant is thereafter allowed to begin his/her academic career. In fact, there is no compulsory Initial Teacher Education Programme (ITEP) for the novice teachers who might be devoid of experience. While a paucity of such a preparatory programme seems to be less inimical to EFL applicants (they are taught more or less the pertinent theoretical and pedagogical skills during their academic education, i.e. MA, PhD), it can have detrimental consequences for those applicants in other academic fields such as Physics, Biology, etc., who have never been taught teaching and learning theories. This is of high importance as becoming an academic in universities in Iran carries
with it different types of responsibilities including teaching and research. I will return to this point shortly in the upcoming chapters. As to the status of TER in the Iranian context, the following section will address some pertinent studies.

Salsali (2005) conducted a study in the Iranian nursing education context in order to explore nurse educators and students’ perceptions of teaching effectiveness appraisal (p. 1). The participants of the study included 143 nurse educators, 40 undergraduate and 30 graduate students whose perceptions of teaching effectiveness were identified by administering a researcher-made questionnaire (p. 1). Based upon the findings of the study, Salsali (2005) advocates the inclusion of continuous and systematic appraisal coupled with staff development as the objectives of staff evaluation (pp. 1-8).

Teacher emotional intelligence (EI) has been recently investigated by a number of Iranian scholars. Following their review of the literature, Moafian and Ghanizadeh (2009) maintain that teachers’ emotional intelligence and the extent to which it could have an impact on teachers’ attitudes and beliefs and level of self-efficacy have not been well-investigated in EFL context (p. 709). Moafian and Ghanizadeh (2009) conducted a study on 89 Iranian EFL teachers to explore the relationship between emotional intelligence and self-efficacy of Iranian EFL teachers (p. 708). The gained data gave evidence to a positive relationship between EI and teaching efficacy (p. 714). Moreover, whereas two components of emotional quotient (EQ), namely “interpersonal relationship, and problem solving” were identified as having the highest positive relationship with teacher efficacy, “emotional self-awareness” was found to be a negative predictor of teacher self-efficacy (p. 715).

The role of higher education leadership in developing a positive and supportive climate in which faculty members can accomplish their teaching effectively is undeniable (Mohammadkhani, 2010, p. 3086). The extent to which emotionally intelligent leadership (EIL) could contribute to faculty effectiveness was the objective for which Mohammadkhani (2010) conducted a study on 351 faculty participants (p. 3092). The findings of the research led to the emergence of a 10 principle-model of faculty effectiveness e.g., self-leadership, moral, trust, etc. (See Mohammadkhani, 2010, pp. 3093-3094). Similar to other studies conducted in this
domain, he fails to elaborate on the multidimensional nature of the concept of ‘effectiveness’ addressing different characteristics of an effective teacher.

EFL teachers’ effectiveness in terms of different communities within which teaching occurs, i.e. public and private school, is another aspect of teacher effectiveness that has been recently investigated by Rahimi and Nabilou (2011). They conducted a study to explore the EFL teachers’ effectiveness in two contexts, i.e. public vs private schools in Iran (p. 67). Adopting external observations and self-evaluation as two measures of evaluation, they collected data from 83 teachers in 76 schools and analysed the data using a seven-factor scale, viz. “presentation, pedagogical skills, methodology, personal characteristics, teacher/student interaction, interpersonal skills and caring behavior” (2011, p. 74). The results of the study suggested that teachers of private schools are more effective than their counterparts in public schools (p. 74). Moreover, teachers’ experiences and age were found to be significantly influencing their effectiveness (p. 67). Rahimi and Nabilou’s (2011) notion that private schools in Iran tend to employ more effective language teachers (p. 74) seems to be an oversimplified conclusion, given the dearth of research in this area. There are a number of questions appertaining to their study as to whether teachers in private school were effective prior to employment or they become effective during their career as a result of the acquired teaching experiences. Akbari and Dadvand (2011) conducted a study to explore the relative effect of teacher education (academic qualification) on teachers’ use of pedagogical thoughts in their classrooms (p. 45). The results obtained from 8 teachers including 4 teachers with BA degrees and 4 teachers with MA degrees revealed that teachers with MA degrees tended to produce more pedagogical thoughts in comparison with their BA holder counterparts (p. 55). While teachers with MA degree exhibited 5.18 pedagogical thought units per minute, teachers with BA degrees produced 2.58 units per minute (Akbari & Dadvand, 2011, p. 44). I would suggest that studies similar to that of Akbari and Dadvand (2011) need to be approached with enough cautions as it might lead to a false impression that the higher the teachers’ academic qualifications, the more their effectiveness.
3.8. **Teacher education, recruitment, development, performance and appraisal**

As to the literature reviewed on TER, it is my contention that teacher effectiveness is one link in the chain of teachers’ professional journey. As the title of this section propounds, teacher effectiveness appears to be indeed a successful denouement of other interrelated activities such as successful pre-service teacher education programme (TEP) and in-service teacher professional development programme (TPDP) which tend to operate within a cycle. In other words, findings in teacher appraisal can be potentially pregnant with some information germane to the overall educational effectiveness of a context. Thus, in order to understand teacher effectiveness in a context, it seems wise to understand teachers’ education, recruitment, development, and performance as well as appraisal system in the first place. This is exceptionally important given the fact that some current practices germane to teacher evaluation are simply imported from other contexts drawing on commonsense approaches towards teacher appraisal. In other words, not all countries have been able to devise their teacher evaluation system based on nationwide empirical research. This sound wise as designing a national system is a daunting and tedious phenomenon. As an exemplar, I can refer to Chile which spent ten years on designing, piloting and conducting its evaluation system (OECD, 2011a, p. 85). Although TEP and TPDP are not the focus of this study, it seems inevitable to refer to these two important programmes as they can have some bearing on TE. Actually, the related literature gives clear evidence to direct and indirect constructive consequences of these two programmes for teacher effectiveness and the considerable influences that they can potentially exert upon effective teaching. This is why TEP and TPDP have been continuously investigated more or less throughout the history of teacher education research.

As lifelong learners, teachers tend to quest for recognition of their learning needs (Troudi, 2009, p. 65). Pre-service and in-service teacher trainings are the two important opportunities by which administrators can accommodate such indisputable demands. Based upon substantial evidence in the literature, Darling-Hammond (2000a, p. 166) asserts that teachers who have had teacher preparation
programmes are likely to be more successful than those with no or little experience of the programmes. She further comes to the conclusion that teacher with greater knowledge of teaching and learning tend to be regarded as more effective (Darling-Hammond, 2000a, p. 167). Despite the extensive research on TE, as stated earlier, the dynamics of the interrelationships between teacher effectiveness and other areas such as TEP, ITEP and TPDP have seemingly escaped researchers’ attention, thereby remaining less clear. Undeniably, this concern raises some questions as to the extent to which our understanding of a successful TEP and TPDP along with recent developments such as reflective practice could contribute to improvement of teacher effectiveness. As Elizabeth et al. (2008) propound, a good conceptualisation of teacher success is the cornerstone of teachers’ professional development (p. 623), in that educators will be able to design and develop teacher development programmes that can best accommodate teachers’ needs and support their development.

3.9. Teacher Evaluation: the challenges of standards

Teacher effectiveness research has been reviewed in previous sections from three distinct yet related perspectives, viz. TER in mainstream education, ESL/EFL education and the Iranian higher education context. The majority of the aforementioned studies indeed sought to explore the means by which teacher effectiveness can be assessed and hence improved. It is argued that such endeavours, among others, have been made in response to the accelerating concerns of policymakers, administrators and governments over the issue of ‘quality’. In many countries, governments compel their ministries of education or higher education to establish and run programmes for the purpose of promoting teacher effectiveness and education quality. The National Board Certification (NBC) in the United States which is offered by National Board for Professional Teaching Standards (NBPTS) could be one exemplar (NBPTS). According to NBPTS, the standards for teaching are based on five core propositions and the one-year valid NBC is designed for most school teachers (NBPTS). However, the central question of the relationship between professional certification and effective teaching has been a point of contention. Chingos and Peterson (2011, p. 450)
warn about the impacts of such certification, given the contradictory findings of different studies. The assessment or evaluation of teacher effectiveness has always been a high stakes issue which can be discussed from at least two angles. On the one hand, such evaluation often brings about job-related consequences such as tenure, promotion, contract renewal, higher pay, etc., and on the other hand, as the review of the literature suggests, teachers have usually been reported to be unsatisfied with the mechanism through which they have been evaluated. They usually complain of a paucity of transparent and clear criteria or standards for evaluation. This is why some researchers such as Scriven (1981, p. 244, cited in Quirke, 2007, p. 90) consider teacher evaluation as a ‘disaster’ which as Quirke asserts is in dire need of change and improvement (p. 90).

It is worth highlighting that the main challenge here lies with the notion of standards/criteria per se, on the grounds that different stakeholders can have different impression of quality/effectiveness standards. There are several studies assenting to teachers’ discontent with the extent to which their voices are reflected in teacher evaluation system. For instance, Elizabeth et al. (2008, p. 623) maintain that teachers’ voices are not well-heard by administrators in Hong Kong. Korthagen (2004, p. 77) also call for further attention to the essential qualities of a good teacher and the processes through which one can become a good teacher. Interestingly, Korthagen admits in advance that it is unlikely to provide such questions with fixed and definitive answers, given different conceptions of a good teacher in different contexts (p. 78).

Moreover, it can be argued that different constituents of a typical educational system tend to interrelate with one another in such a way that any alteration in one constituent would exercise influence on other components. As Kelly et al. (2008, p. 39) contend, appraisal systems can influence teachers’ behaviours and attitudes which in turn can exert impacts on teachers’ performances and learners’ achievements thereupon. A quick review of different national schemes for measuring teacher effectiveness across different contexts around the world bring to light the fact that the majority of evaluation systems and schemes are more or less based upon some basic standards which are mostly extrapolated from some
widely accepted standards. This is in line with Goodwyn’s notion that introducing standards into teaching is a ‘global phenomenon’ (2011, p. 26).

Another concern pertaining to teacher effectiveness research is the core question of ‘what counts as effectiveness and ineffectiveness?’ This concern raises a number of points and questions as to whether or not researchers can benefit from investigating ineffective teaching instead or besides effective teaching or, for instance, whether or not a comparative analysis of teachers’ outperformance and underperformance may yield more comprehensive results thereby helping us achieve profound insights into teacher effectiveness research. The research on teacher evaluation suggests that teacher ineffectiveness and the challenges of novice teachers in the very beginning stages of their teaching career should be also given some weight in teacher effectiveness research. In their search for investigating the qualities of effective teachers, researchers and practitioners have proposed and assigned a number of terms to effective teachers, viz. effective, qualified, excellent, proficient, outstanding, expert, competent, etc. In the same vein, different terminologies have been utilised by researchers to refer to ineffective teachers, vis-à-vis their effective counterparts, from the euphemistic notion of “poor performers” (Fidler & Atton, 1999) to Jones, Jenkin, and Lord’s (2006, p. 15) classification of different ranges of ineffectiveness, viz. “ineffective, struggling, under-performing, sliding, sinking, stuck, or incapable”.

3.9.1. What counts as teacher effectiveness

As stated earlier, one fundamental challenge in research on TE is the concept of ‘effectiveness’ per se which is sometimes used equivocally by researchers and practitioners. For instance, Campbell et al. (2004a) assert that in spite of their interrelatedness, school effectiveness, teacher effectiveness and educational effectiveness are being used inconsistently in the literature (p. 2). However, such incertitude can be avoided once researchers reach at a consensus on what makes an effective teacher. As Arthur et al. (2003, p. 275) point out, we need to define “effective in terms of what”. As discussed earlier, researchers and practitioners’ conceptions and understanding of teacher effectiveness have mostly evolved during the last two decades. Casey, Gentile, and Bigger (1997) maintain that while
there is more or less agreement on what is ‘bad teaching’, the concept of ‘good teaching’ remains yet unclear which in turn raises questions as to whether or not it is equivalent to ‘quality teaching’ or ‘effective teaching’ (p. 461). As Cheng and Tsui (1998, p. 39) claim, the traditional approaches towards research on teacher effectiveness investigate individual teachers with a focus on within-classroom activities and pay little attention to the elements that could exert influence on teachers’ performances such as the intricacies of organisational environment. Therefore, it can be argued that discrepant views among researchers and practitioners would hardly come to an end unless the very nature of the concept of ‘effectiveness’ becomes transparent to all researchers so that a similar and less contradictory understanding of TE could be drawn, even though this might be practically impossible. Teachers tend to feel more confident and as Kelly et al. (2008, p. 43) posit, tend to welcome appraisal systems that are fair. Therefore, teachers need to possess a conviction that the mechanism of their appraisal is fair and unbiased. Reaching such a transparent and consistent understanding is not as simple as it might seem, given the multidimensionality of teacher effectiveness and the contentious nature of teacher evaluation (e.g. Curtis & Cheng, 2007; Robinson & Campbell, 2010). As a consequence, the need for devising a set of parameters or criteria seems to be inevitable.

It is widely accepted that no single definition can suffice to comprehensively address the notion of teacher effectiveness, for different definitions could be drawn from the concept of ‘effectiveness’ with respect to the context within which the concept is being used. As Olivares (2003, p. 235) put is forward, teacher effectiveness needs to be grounded in a theory which can elucidate the relationships between teacher effectiveness and other constructs as well as teachers’ observable behaviours. However, from among current definitions of teacher effectiveness in the literature, I am inclined to refer to the definition proposed by Campbell et al. as follows:

“The power to realise socially valued objectives agreed for teachers’ work, especially, but not exclusively, the work concerned with enabling students to learn” (Campbell et al., 2003, p. 354; 2004a, p. 2).
3.9.2. Teacher effectiveness and quality in education

The literature on teacher effectiveness resonates with the concept of ‘quality’ to such an extent that it seems unwise to discuss teacher effectiveness and teacher evaluation without pondering over teacher quality and a wider realm of educational quality. Based on their review of the literature appertaining to educational effectiveness, Ingvarson and Rowe (2008, p. 6) believe that the need for establishing standards of instructional effectiveness has been given scant attention. Learners’ access to high-quality teaching irrespective of their backgrounds (OECD, 2001; 2005, cited in Ingvarson & Rowe, 2008, p. 6) is a prerequisite of amelioration of the quality of learners’ achievements (Ingvarson & Rowe, 2008, p. 6). Darling-Hammond (2009, p. 2), also reminds us of the difference between teacher quality and teaching quality. Whereas she conceives teacher quality as an amalgamation of ‘personal characteristics, skills and understanding’ which are mostly brought into the career by teachers (p. 2), teaching quality pertains to what she calls ‘strong instruction’ which accommodates students’ needs and meets educational objectives, and to that end the quality of teaching is partly a function of the quality of teachers (Darling-Hammond, 2009, p. 3). Finally, Darling-Hammond concludes that ‘high-quality’ teachers in one context might not show the same quality in other contexts (p. 3). In addition, building on the difference between “teacher quality” and “teacher practice”, Robinson and Campbell (2010) draw a distinction between the concepts. From their perspective, as a “normative” concept, teacher quality bears upon teachers’ roles in promoting students’ educational attainments (p. 674). With regard to teacher practice, Robinson and Campbell (2010) introduce it as a “descriptive concept” embracing teachers’ not only intra-classroom activities but also those outside the classroom (p. 674). As to the status of teacher practice and quality in teacher evaluation, Robinson and Campbell (2010, p. 674) argue that teacher quality has been given more attention.

3.10. Forms of teacher evaluation (Measures of evaluation)

Obviously, different educational systems adopt different approaches towards teacher quality as a ‘high-stakes’ issue each of which might specify different sets of
standards or criteria for teacher evaluation. This is why it might not be possible to prescribe a set of criteria and generalise or extrapolate them to other contexts around the world. The criteria based upon which effective teaching could be identified have always been arousing controversy. Such discrepant stances have led to concerns and questions as to whose voices should or should not be considered in teacher evaluation or whether the focus should be on effective teachers or less/ineffective teachers. The literature on TE suggests different parties who have a stake in the phenomenon and hence their voices need to be included in teacher appraisal, e.g. students, teachers, administrators, policymakers. However, the weight or priority that should be given to each party in the overall appraisal scoring system is yet a point of contention. For instance, Harris, Ingle, and Rutledge (2014, p. 73) argue that policymakers tend to attach more stakes to students and observation-based measures. Nevertheless, the literature on teacher evaluation assents to a truism that there is an increasing tendency towards the use of multi-measure approaches towards teacher evaluation, even though how to apportion the scoring/rating system among different measures remains yet unanswered. According to Jacob and Lefgren (2008, p. 105), principals collect data pertinent to teachers’ performances from three major sources, namely ‘formal and informal observation as well as parents and students’ achievement scores’. Referring to the issues associated with the sole use of students’ ratings such as raters’ bias, Burden and Troudi (2007, p. 163) call for a comprehensive approach, what they call ‘teacher-centered, teacher-led evaluation’.

Referring to the most important measures of evaluation which can provide administrators (appraisers) with information about teacher effectiveness, the following sections will also touch on the merits and demerits associated with each and every measure, in brief.

3.10.1. Students’ evaluations of teaching (SETs)/Students’ ratings

As mentioned earlier, students’ ratings or SETs have been consistently conceived of as an important source of information for evaluating teacher effectiveness throughout the history of TER. Several researchers have investigated the status of SETs in teacher effectiveness evaluation (e.g. Greenwald, 1997; Greenwald &
Students’ ratings as a symbol of students’ satisfaction with the educational systems have always been central to policymakers and administrators. Nonetheless, the genuine objectives of students’ ratings, i.e. enhancing teachers’ effectiveness and hence students’ learning, sometimes get lost. Instead, SETs tend to be viewed in some contexts rather as an opportunity to satisfy or please students as customers as part of educational bureaucracy. It seems that such fallacious mentality occurs in educational systems within which education is regarded as business and students as customers. I am not inclined to critique such an argument in this section; however, it is my contention that administrators’ stances towards SETs, especially their interpretation of SETs, are as important as SETs per se. Given its established status as a measure of teacher appraisal in most educational systems including that of Iran, SETs have been given particular attention in this study. With respect to two of research questions proposed at the outset, the pros and cons of SETs particularly in Iran will be extensively investigated in this research.

Much of the studies and hence the literature germane to the validity of SETs have been done during the 1970s and 1980s (Nilson, 2012, p. 213). Whilst student rating was hardly debated in the 1970s, it became plausible as a valid measure in the 1980s (Greenwald, 1997, p. 1182). Stark-Wroblewski, Ahlering, and Brill (2007, p. 403), maintain that student evaluation of teaching yet remains controversial (p. 403). In his analytic critiques of student ratings, Olivares asserts that “academic control and authority have been placed in the hands of students” (2003, pp. 242-243). Burden and Troudi (2007) maintain that students’ ratings are not ‘evaluating’ as students just ‘rate’ a set of narrowly defined teachers’ practices of which the interpretation of the results tends to remain unclear to students themselves (p. 153). They raise some important questions about students’ ratings as to whether or not novice students do have enough knowledge to evaluate their teachers (p. 153). There are some other concerns that are associated with such questions. Given the fact that students’ ratings are highly informed by their understanding of an effective teacher, there seems an urgent need to explore what students have in mind for
effective teaching in the first place. As Spencer and Schmelkin (2002, p. 397) contend, despite the burgeoning literature on students’ ratings, there is a dearth of research on students’ perceptions of their ratings. In their study on students’ perceptions of course and teacher, Spencer and Schmelkin (2002) realised that students were in favour of doing the ratings and most interestingly had no fear of repercussions (p. 397). Nevertheless, their findings provided some evidence of students’ reservations about the extent to which their ratings are taken into consideration by administrators (p. 397). To better depict the picture, Spencer and Schmelkin (2002) call for further research on the faculty and administrative perspectives on SETs (p. 406).

Stark-Wroblewski et al. (2007, p. 403) are under the impression that SETs as indicators of teacher effectiveness are employed for occupational purposes such as hiring, promotion and tenure. However, it is worth highlighting that different countries have different schemes for evaluation of teacher effectiveness within which different weights tend to be given to SETs. Despite its overall acceptance, SETs, as stated earlier, has also been a point of severe criticism. Whereas some researchers advocate the validity of SETs as a measure of teacher effectiveness (e.g. Cohen, 1981; Marsh, 2007), others adopt rather cautious approaches (e.g. Eiszler, 2002). For instance, in his large scale study on 50,000 college courses, Centra (2003) noticed that whereas students’ learning outcome positively affected their ratings, there was a low correlation between students’ expected grades and their evaluations of instruction (p. 495). By contrast, Eiszler (2002, p. 483), for instance, found in his study a predictive relationship between student ratings and grade inflation. Stark-Wroblewski et al. (2007, p. 404), also bring to attention the threat of higher grade exchange for better ratings and the likeliness of consumer-oriented teaching (p. 404). To find better solution to the aforementioned problems, they conducted a quantitative study with pre and post measures of learning. The data driven from 165 students suggested that use of related yet independent measures of SETs and learning outcome evaluate distinct aspects of TE, thereby none of them should be divorced (Stark-Wroblewski et al., 2007, p. 410). Similarly, based upon their review of the voluminous literature, Marsh and Roche (1997) introduce a number of potential biases associated with students’ ratings e.g. “Dr.
Fox Effect” which as they define is “the overriding influence of instructors’ expressiveness on SETs ...” (p. 1193). In other words, ‘seductiveness can influence students’ evaluations of instruction and achievements’ (Ware & Williams, 1975, p. 149).

To sum up, students’ evaluations of teachers (SETs) or students’ ratings tend to have both merits and demerits. SETs could be useful in evaluating teacher effectiveness if (a) different dimensions of effective teaching are identified and incorporated into students’ evaluation questionnaire or any other evaluating tools, (b) students are taught the required skills for evaluations and (c) the potential biases are recognised, minimised or eliminated. Due to a multidimensional nature of teacher effectiveness and SETs, single-criterion approach towards effective teaching should be avoided (Marsh & Roche, 1997, p. 1187).

3.10.2. Students’ learning outcome/student achievement

It is worth distinguishing students’ ratings from students’ learning outcomes as a criterion for teacher effectiveness. The literature on teacher effectiveness attests to the strengths of students’ achievements in teacher appraisal upon which there have been little discrepancies among researchers. There is a growing body of research which is suggestive of the importance of students’ learning outcomes as a predictor of TE. For instance, based on the findings of his large scale study, Centra (2003, p. 495) suggests that students’ learning outcomes tend to exert positive impacts on students’ ratings. Despite the general consensus on the inclusion of learners’ outcomes as an important criterion in measuring teacher effectiveness, Curtis and Cheng (2007) maintain that student learning is not always under the control of teachers (p. 57). Sanders and Rivers (1996) conducted a study to explore teacher effect on students’ future achievement. The results of their research provided evidence for teachers’ additive and cumulative effect on student academic achievement (p. 6). However, similar to other measures, students’ learning outcome as a criterion for teacher effectiveness is not free from threats and criticism.
3.10.3. Peer evaluation

Given the multidimensionality of teachers’ behaviours and practices in the classroom, it might not be wise to rely on any single-criterion evaluation model. In other words, each evaluation method tends to have its own merits and demerits. More importantly, it is my contention that the knowledge, skills and trainings of evaluators/raters who are not necessarily appraisers have thus far escaped the attention of researchers. This is an important point, given the fact that assessment, supervision and evaluation in any format do require trainings and experience both theoretically and practically. What I would like to argue is that students, peers and whosoever may do the evaluation need to be cognisant of faculty evaluation. Undoubtedly, such deficiencies on the part of evaluators can impact on the results of evaluation and hence misleadingly affect the consequent policies. Peer evaluation is another measure of teacher evaluation which has gained credit in the literature on TE in mainstream education. However, this topic has been given scant attention in the realm of language education, despite some published studies (Bailey, 2006, p. 189). A review of the literature reflects the views of both proponents and opponents. Weller and Weller (2000, p. 234) maintain that teacher appraisal can provide a holistic framework for teacher performance, even though its cost and objectivity tend to be the points of criticism. There is also a group of researchers and practitioners who cast doubts on the usefulness of peer evaluation (e.g. Marsh & Roche, 1997). They challenge the reliability of colleagues and administrators’ evaluations as there is no systematic correlation between such ratings and those of students and other measures for teacher effectiveness (p. 1189). I would then argue that such explanations seem rather implausible. First, such differences of views among evaluators and even colleagues appear to be natural, given the fact that they might view teaching practices from different perspectives via different lenses. Such discrepant views do not necessarily mean peer evaluation is not reliable. Thus, in lieu of such an understanding, I think, differences of views assent to the multidimensional and dynamic nature of teaching and teacher evaluation. On the other hand, the systematic correlation between peer evaluation and SETs (Marsh & Roche, 1997, p. 1189) is not a tenable indicator for the usefulness of peer evaluation. Peer evaluation, amongst other
measures, could be a reliable measure for evaluation, provided peers are possessed of prerequisite skills and informed of transparent criteria/standards. As a developmental element, peer evaluation not only can serve as measure of evaluating teachers, but also as a precious opportunity for learning to teach through observing and reflecting on one’s teaching practices as well as exchanging ideas and expertise among teachers in a collegial environment.

3.10.4. Teachers’ Self-evaluation

Teacher themselves can be a source of information about their evaluation (Bailey, 2006). Informed by qualitative methodologies, teacher’s self-evaluation tends to be an informal type of evaluation contributing to teachers’ professional development (Robinson & Campbell, 2010, p. 675). Given the notion that teacher self-evaluation can be perceived as a microcosm of a broader teacher reflective practice, it can be argued that teachers have been doing such evaluation strategy since the emergence of reflective practice in the 1980s (e.g. Schön, 1983). As suggested by the literature, self-evaluation can be performed through a number of different ways including questionnaire, journal, teacher’s portfolio, etc. It seems that such a measure is more useful for formative rather than summative purposes. For instance, as Yaode and Murphy (2007, p. 124) maintain, portfolio could be adopted in teacher development programmes in which teachers would have the opportunity to refer to their own practices and identify their strengths and weaknesses, and thereby distinguishing more effective teaching strategies from those with little or no effectiveness. Within such a secure and relaxed atmosphere, teachers can also promote their self-confidence. Marsh and Roche (1997, p. 1189) consider teacher self-evaluation persuasive and beneficial, for it can be implemented in all educational settings. Kyriakides et al. (2002, p. 301) maintain that teachers as ‘natural learners’ can recognise their effective actions. Teacher’s self-evaluation tends to provide teachers’ with the chance to identify their good/effective teaching practices (Kyriakides et al., 2002, p. 302). In other words, teacher’s self-evaluation can be considered as a booster for teacher’s authentic self-development. As Campbell et al. (2004b, p. 453) note, teacher self-evaluation is an opportunity to delve into the values in teacher effectiveness.
Despite the widespread support for self-evaluation in the literature, it is unfortunate that this measure is rather disregarded in universities in Iran. I will report and discuss lecturers' perceptions on this measure in chapters five and six. It is also worth mentioning that the abovementioned measures for evaluating TE are among the most important measures that have been suggested by the literature. Obviously, at the heart of each of these measures, there rests a set of criteria. Whilst it can be argued that the aforementioned measures are more or less similar across different disciplines and contexts, the detailed criteria functioning within each measure could hardly function in the same way in different contexts. This is due to different expectations of an effective teacher in different contexts. In other words, it might not be a good idea to propose a set of predefined, fixed and cliché criteria against which teacher effectiveness could be assessed in different contexts. Interestingly, even within a same context, these criteria could also change over time. This, in turn, emanates from the changes which occur in stakeholders’ needs, i.e. teachers, students, administrators, etc.

3.10.5. Appraiser’s observations

Supervision, appraisal, observation, evaluation and even audit are amongst terminologies that have emerged throughout the literature on TE to refer to a process in which an outsider monitors teachers’ practices in the classroom. There are a wide range of studies that have addressed the observation model for evaluating teacher effectiveness. While some researchers and practitioners advocate the use of such a method, others call for more caution with the interpretations of data driven from it. In other words, it seems less feasible for an observer to evaluate a teacher’s various classroom behaviours in finite occasions. As Leigh (2010) asserts, the observer will face some challenges as to comparing different teachers’ practices (p. 480). Jacob and Lefgren (2008) introduce a number of features for principals at schools including interacting with teachers, parents and students, reviewing lesson plans, observing classes and having access to students’ scores (p. 102). Based on their gained findings, Jacob and Lefgren (2008) contend that policymakers should give more weight to principals' evaluations in teacher compensation and promotion programmes (p. 103).
3.11. Types of evaluation models

As mentioned earlier, different models for evaluating teacher effectiveness have evolved throughout the history of TER some of which are currently adopted in different countries. Notwithstanding these models and schemes might be different in their forms, they are inherently similar by nature and are proposed by policymakers as indices for making decisions upon teacher effectiveness. It is worth highlighting that concerns as to how to accommodate teachers’ shifting needs, wants, ideals, etc., as well as those of other stakeholders have been foremost on researchers’ minds. According to Wright (2010, p. 264), practices, theories and research germane to second language teaching have witnessed dramatic changes since 1985. In response to such a dynamic nature of teacher evaluation, different types of models have emerged especially during last two decades. Drawing on the literature, this section will address different types of teacher evaluation models. It is worth noting that the concept of ‘type’ per se is rather less consistent in the literature as different classifications can be proposed. However, for the purpose of this study, I am inclined to review teacher evaluation from four perspectives, viz. purpose, context, methodological orientations and data interpretation.

3.11.1. Formative and summative approaches

*Formative* and *summative* evaluations are two major purposes of evaluation models introduced in the literature on TE. It is worth mentioning that the formative-summative dichotomy is mostly viewed as a purpose rather than a type of evaluation model, on the grounds that different types and forms of evaluation models can be used for both purposes. However, the review of the literature on teacher evaluation testifies to the interchangeable use of concepts of *forms*, *types* and *purposes* of teacher evaluation. Whereas formative evaluation is considered as feedback to shape teachers’ performances and develop their practices, summative evaluation is mostly utilised for job-related decision making purposes (Peterson, 2000, p. 63). Robinson and Campbell (2010) propose three purposes for teacher evaluation systems. Informed by what they call “performativity ideology”, the first purpose is an endeavour to develop teachers’ accountability and
is approached within a wider domain of education system evaluation (p. 675). The second purpose as they continue is an official evaluation aiming at supporting teachers’ professional development and promotion (p. 675). Finally, emphasising the centrality of teacher evaluation to the evaluation of an educational context, e.g. school, Robinson and Campbell (2010, p. 675) view teacher evaluation as a formative tool to improve school effectiveness. However, it is worth referring to the importance of different value assumptions underpinning formative and summative evaluation for which different criteria might be deployed (Creemers & Kyriakides, 2008, p. 249). Perhaps, this is why Peterson (2000, p. 63) considers mixing summative and formative purposes in teacher evaluation a mistake.

Therefore, identifying the theory or theories which inform an appraisal model seems to be a major step towards understanding the nature of that model. Indubitably, these theories give rise to different goals and purposes which would thence have quite different bearings on teachers’ professional career. While summative evaluation is used for hiring, promotion, contract renewal, tenure or dismissal purposes, formative evaluation is used for teachers’ self-awareness of their teaching skills and is used partially as a blueprint for designing a suitable compensatory professional development programme. Such a standpoint is focused on teachers’ learning from a professional development perspective which as Borko (2004, p. 3) contends helps teachers improve their instructional practices and thence student learning. Kyriakides and Demetriou (2007, p. 46) assign the formative and summative nature of evaluation to two functions of evaluation, “improvement” and “accountability”, respectively. Although summative and formative evaluations are consistently used in TER to refer to the goals of teacher evaluation, there exist some other terminologies proposed by researchers which refer to the same concepts, e.g. Curtis and Cheng’s (2007, p. 69) notion of developmental and evaluative goals of teacher appraisal. More importantly, the review raises some concerns with regard to the objective of evaluation as to the priority/weight attached to each of these models and the way their findings can be linked. Rejecting the so-called single-criterion approaches towards teacher evaluation, it is widely accepted that summative or formative evaluations per se
cannot suffice for the purpose of teacher evaluation and hence need to complement each other.

In reviewing the purposes and types of teacher effectiveness evaluation in the literature, one can easily notice the overuse of terms such as teachers’ accountability, teacher performance, teacher quality, context (school) effectiveness, etc., which could emanate from what Ingvarson & Rowe (2008, p. 6) call “outcomes-driven economic rationalism”. Moreover, the exact relationship between ‘accountability’ and effective teaching needs to be addressed with more scrutiny. As Goodwyn (2011, p. 15) points out, there are some questions as to the extent to which different levels of accountability could exert developmental influence on effective teaching. The concepts of evaluation and accountability became buzzwords in educational reforms era in the 1980s (Ellett & Teddlie, 2003, p. 106). Assigning three purposes to teacher evaluation internationally, viz. accountability, promotion and staff development, Ellett and Teddlie (2003, p. 102) maintain that it is hardly used for development purposes.

Formative evaluation is one key area which is thoroughly addressed in this study, given the fact that such an understanding of teacher appraisal is conspicuously absent in most universities in Iran. It is worth re-emphasising that it seems futile to design an evaluation model for assessing teacher effectiveness without specifying a purpose for it beforehand. In other words, the consequences of teacher appraisal tend to revitalise the significance of the research on evaluating teacher effectiveness. As suggested by the literature, issues such as tenure, career promotion, reward or dismissal are the consequences that are associated with summative evaluation whilst enhancing teachers’ pedagogical self-awareness and improving their teaching skills are among the ones that are associated with formative evaluation.

3.11.2. Generic vs. differentiated models

The need to distinguish between generic and differentiated models of evaluation (e.g. Campbell et al., 2003, 2004a; Muijs et al., 2005; Walker & Dimmock, 2000) seems to be self-evident in that generic models would not be contextually viable in
different contexts. This is important as the literature is replete with studies indicating the impacts of contextual values such as socio-cultural and socio-economic ones on educational policies. As Walker and Dimmock (2000, p. 155) argue, it appears to be unwise to ‘clone’ and ‘transplant’ appraisal systems into quite different contexts and cultures. Robinson and Campbell (2010, p. 676) distinguish differentiated models from the generic ones whereby they mean appraisal models which are accorded with individual teachers’ needs in different contexts. A review of research on appraisal models for teacher effectiveness provides evidence for a shift from the so-called “one-size-fits-all” and generic to a contextualised and differentiated understanding of teacher appraisal (e.g. Walker & Dimmock, 2000).

However, such a differentiated model should not be merely relegated to a set of fixed values. It is my contention that not only should contextualisation encompass social, cultural, economic and political considerations, it needs to consider the intricacies and peculiarities associated with, for instance, L2 education which besets EFL lecturers with challenges of teaching in a language other than students’ mother tongue.

3.11.3. Qualitative vs. quantitative models

Whereas teacher evaluation has received a wealth of attention in terms of purpose and form of evaluation, I am slightly convinced of a dearth of research on methodological orientations of teacher evaluation and the philosophical and theoretical underpinnings of appraisal models. The methodology and methods used for data collection are significant aspects of teacher appraisal. As Popham (1988) argues, teacher evaluation may require different techniques and even different personnel with respect to its purposes (Cited in Peterson, 2000, p. 63).

In their proposed distinction between self-evaluation and external evaluation, Robinson and Campbell (2010, p. 676) maintain that quantitative methodologies tend to be adopted usually in external evaluation using a variety of rating scales including “teacher quality, classroom observation, attainment test and the value added to students’ attainment”. On the contrary, as they continue, qualitative
Methodologies are adopted in self-evaluation that is usually conducted in an informal environment (pp. 675-676). It might be worth noticing that qualitative and quantitative methodologies could not be confined solely to self- and external evaluation. Indeed, they can be utilised in different forms and with different functions for different purposes. Obviously, there are some critical considerations associated with each and every of these forms. Perhaps, the evaluation conducted by government at national level thereby covering a large number of participant stakeholders including teachers, students, administrators, policymakers, parents, etc., need to adopt quantitative methodologies for the ease of conducting the research. Since such evaluation schemes are often used for decision making purposes or policy reforms, they might need to be evidence-based in order to meet the funding requirements put forth by governments. On the other hand, the evaluations conducted at institutional level, e.g. at college or university level, can adopt qualitative methodologies and thereby giving more weight to teachers’ voices. These methodological issues will be elaborated on in details in the next chapter.

3.11.4. Subjective vs. objective models

Drawing on the categorisation presented in 3.10.3., this section will review teacher appraisal models based on their epistemological frameworks. Notwithstanding the concepts of subjective and objective evaluations are clear enough to all researchers in the field, different researchers and practitioners tend to assign different methods to objective and subjective evaluation. However, the majority of research on teacher appraisal appears to take a back seat to the importance of policymakers and administrators’ epistemological approaches towards teacher evaluation. Whereas much of research on appraisal models are in the spirit of objective and subjective approaches, there is a dearth of research informed by other epistemological stances such as ‘constructivism’ whereby teachers’ ‘learning’ can be more investigated. The literature on teacher effectiveness shows some studies on the contribution of subjective and objective teacher evaluation. For instance, in their study on students and teachers in public schools, Rockoff and Speroni (2011, p. 687) tried to explore the extent to which subjective and objective
evaluations of teacher effectiveness could predict the achievements of teachers’ future students, thereby distinguishing effective teachers from the ineffective ones. The results indicated the predictive power of subjective evaluation for teachers’ future success in terms of their students’ achievement (p. 695).

3.12. An overview of some schemes and models developed for teacher appraisal around the world

The central ideas and major trends in teacher effectiveness and teacher evaluation have been reviewed in the previous sections. It has been argued that these notions have been given scant attention in some Middle-eastern contexts, thereby suggesting the need for further investigation into TER in these contexts. Whereas some countries (mostly Western) have developed their own national schemes and have passed Acts and brought in some legislation addressing teacher appraisal and also constantly substantiate the veracity of their proposed evaluation systems, others tend to develop and/or use appraisal models mostly extrapolated from other contexts. Irrespective of the context within which it is implemented, teacher evaluation embraces a number of principles and concerns which tend to exist in most educational contexts. Therefore, it might be a good idea to review some of the evaluation schemes and appraisal models in use in different parts of the world together with their theoretical justifications. This, I think, gives us some deep insights into the nature of EFL teacher appraisal, thereby allowing a fairer understanding of the strengths and weaknesses associated with teacher evaluation in Iran.

Teacher effectiveness and teacher quality are inextricably interwoven in education research to such an extent that they tend to be used interchangeably in the literature. Promoting teacher effectiveness and thence the quality of education, amongst others, have always been a major concern for educational policymakers and a high priority for governments. Such concerns over the quality of education have led to some Acts and national schemes in some countries which have been proposed for TE, e.g. “The No Child Left Behind Act” (NCLB) in the United States throughout which the centrality of ‘highly qualified teacher’ is highly stressed (The
No Child Left Behind Act (NCLB), 2001). However, to the best of my knowledge, such Acts and legislation mostly focus on primary and secondary education.

In addition, the quality of education in some countries is regularly checked by some non-governmental nonpartisan organisations and institutions which are authorised by governments to monitor the overall quality of education in schools and universities. Organisation for Economic Co-operation and Development (OECD) headquartered in Paris, France, National Board for Professional Teaching Standards (NBPTS) in the United States, Higher Education Academy (HEA) and the Office for Standards in Education (Ofsted) in the United Kingdom are amongst several organisations involved in research on teaching quality and teacher effectiveness in some way. As explained in chapter two, The Council of Supervision and Evaluation of Higher Education (CSEHE) in Iran is a ministerial council which monitors the quality of higher education. Although all the above-mentioned organisations and councils have common goals, i.e. improving the quality of education in primary, secondary, tertiary and higher education, they have different duties, responsibilities and activities which are mostly indicated in their mission and vision. It is worth highlighting that the adequacy and utility of such (non)ministerial, (non)governmental organisations, offices, councils, committees, etc., are contingent upon the enforcement power attached to their decisions. Otherwise, they will turn to something of bureaucratic centres. This is of central importance in the policy-implementation nexus which based upon the findings of this study is a point of contention.

In order to have an idea of the activities of quality assurance organisations, NGOs, agencies, etc., I will briefly introduce OECD as an international economic organisation with 34 member countries (OECD). Established in 1961, OECD seeks to “to promote policies that improve the economic and social well-being of people around the world” (OECD). As part of OECD, The Centre for Educational Research and Innovation (CERI) conducts extensive research on learning and the future of schools and universities (CERI). Perhaps, one of the most important activities of OECD focusing on quality of teaching and learning environment is “OECD Teaching and Learning International Survey (TALIS)”. According to OECD, TALIS
is the first international programme aimed at investigating “learning environment” and teachers’ “working conditions” in schools and comparing education systems among different countries, thereby bridging the information gap (OECD, 2008, p. 3; 2011b, p. 4).

The literature on TER offers a number of prevailing teacher evaluation models in some countries. Transparency, fairness and credibility, amongst other characteristics, are found to be the prerequisite elements for a successful teacher evaluation system whereby the voices of all leading stakeholders are heard. Kyriakides and Demetriou (2007, p. 43), maintain that a valid teacher evaluation system would have its underpinnings in Teacher Effectiveness Research (TER). Pointing to the impacts of ‘political dynamics’ exerted on evaluation system, Kyriakides and Demetriou (2007, p. 43) call for stakeholders’ clear perceptions of theoretical assumptions underpinning teacher evaluation system as well as their subsequent concerns. However, the overriding concern with teacher appraisal relates to its subsequent sensitivities, given the fact that appraisal per se, in any formats, tends to be a contentious phenomenon by its very nature. Such sensitivity emanates from both direct and indirect bearings teacher evaluation could have on lecturers’ professional career. Thus, reaching a comprehensive model encompassing all aspects of teacher effectiveness especially ESL/EFL teacher effectiveness seems to be just utopian optimism. In order to have an idea of some current teacher appraisal frameworks and models, the following section will review few of them, in brief:

Referring to the limitations of the traditional conception of teacher effectiveness, Cheng and Tsui (1996, p. 12) propose a new multi-levels and multi-domains conceptual framework of what they call “total teacher effectiveness” whereby the process and development of teacher effectiveness can be better investigated. As shown in Figure 3.1, their framework assumes teacher effectiveness at three levels and within three domains and involves teaching and learning at four layers (Cheng & Tsui, 1996, p. 13; 1998, p. 41).
Figure 3.1. “The structure of total teacher effectiveness model” (Cheng & Tsui, 1996, p. 13; 1998, p. 41)

The “individual, group and school” levels of teacher effectiveness and the “affective, cognitive and behavioural” domains of teacher effectiveness and performance are the aspects assigned to teacher effectiveness (Cheng & Tsui, 1996, 1998). According to their proposed model, teaching and learning processes pertinent to teacher effectiveness operate at four layers, namely the layers of “teacher competence, teacher performance, student experience and student learning outcome” (Cheng & Tsui, 1996, 1998). In order to have a deep understanding of teacher effectiveness, they suggest what they call “the whole structure strategy” whereby all constituents of their proposed framework, viz. layers, arrays, cells, etc., are taken into account (Cheng & Tsui, 1998, p. 44). Their developed model of total teacher effectiveness presents some interesting and informative pieces of evidence to the complex nature of teacher effectiveness in its modern realisation. However, their proposed model does not account for the
subcomponents within each domain of effectiveness. Moreover, the weight/priority attached to different constituents in the framework remains rather unclear. Notwithstanding their proposed model embraces important general issues with respect to teacher effectiveness, it does not refer to specific criteria/standards for evaluating teachers’ cognitive, professional, etc., abilities based on which evaluation can be implemented. It is worth reminding that for the purpose of evaluation, there should be a set of transparent criteria under the rubrics identified, e.g. affective domain. Although they argue that the relationship between teachers’ “static quality” (competence) and their “dynamic quality (performance) could be influenced by external elements such as organisational factors, leadership, etc. (Cheng & Tsui, 1998, p. 40), the dynamics of such interplay between teacher effectiveness and contextual variables remains less well-explained.

Following their proposed model of teacher effectiveness in 1996 and 1998, Cheng and Tsui (1999) propose their new conception of teacher effectiveness which consists of seven models including:

- “The goal and task model
- The resource utilization model
- The working process model
- The school constituencies satisfaction model
- The accountability model
- The absence of problems model
- The continuous learning model” (p. 142)

Criticising the traditional conception of TE which focuses on classroom level, Cheng and Tsui (1999, p. 150) maintain that their proposed models help understand the multifaceted conception of teachers’ work and performance. They introduce multimodels of teacher effectiveness from three perspectives, namely “conception of TE, condition for model usefulness and areas for assessing and monitoring” (p. 142). According to Cheng and Tsui (1999, pp. 142-144) multimodels of teacher effectiveness, teachers are regarded effective if they can:

a. ‘Achieve pre-planned goals and tasks in line with school’s goals
   (The goal and task model)

b. Maximise the use of resources and support in order fulfill their duties
   (The resource utilisation model)

c. Ensure the quality of their smooth teaching and working process
d. *Meet main school constituencies’ needs and expectations*  
(The school constituencies’ satisfaction model)

e. *Show accountability and reputation and competence*  
(The accountability model)

f. *Do their duties with no problems and weaknesses*  
(The absence of problems model)

g. *Adapt themselves to the external and internal challenges*  
(The continuous learning model)

Cheng and Tsui’s (1999) seven-model conception of teacher effectiveness is indeed a step forward in deepening the insights into teacher effectiveness. Nonetheless, similar to their earlier total TE model (Cheng & Tsui, 1996, 1998), it remains yet unclear from some aspects, e.g. how to apportion the weights among different models or whether administrators should use more than one model for teacher evaluation or not. Given the teacher-centered nature of the proposed models, there raises a concern as to how such an approach can control or account for the undesirable effects of the so-called non-teacher-controlled variables.

Although it is very unlikely that teachers’ future effectiveness can be predicted from their achievements in teacher education programme, I think, the seeds of teacher effectiveness partially emerge during their pre-service and in-service education whereby successful teacher education and development programmes can help teachers maximise their true potentials. Korthagen (2004) interestingly addresses the issue of teacher effectiveness from another perspective, i.e. teacher education, thereby probing the central question of “the essential qualities of a good teacher?” (p. 77). In his review of the literature, Korthagen (2004) refers to the movements and breakthroughs in teacher education research, i.e. the emergence of ‘performance-based’ or ‘competency-based’ model in the mid-20th century which called for ‘observable behavioral criteria’ as the basis for teacher education and also the ‘process-product model’ which advocated teaching behaviours (concrete competencies) that highly correlated with students’ achievements both of which more or less led to ‘fragmentation of the teacher’s role’ (p. 79). As Korthagen continues, later on in 1970, the Humanistic Based Teacher Education (HBTE) movement focusing on the ‘person of the teacher’ did appear (2004, p. 79).
Korthagen (2004) appropriately identifies major problems associated with such models in that they fail to take into consideration the fact that good teaching cannot be described or as I suggest prescribed in terms of a set of specific criteria which I think tend to be mostly stereotypes and cliché. As discussed elsewhere, it seems unwise to merely evaluate teachers against a set of predetermined and generic criteria or standards, even though the need for such basic standards for teacher evaluation is inevitable. What I would like to argue, however, is that any sets of standards or criteria need to be contextualised and differentiated in a way that they can accommodate not only teachers and students’ needs but also those of administrators and policymakers, i.e. their expectations and demands. Critiquing the traditional perceptions of a good teacher, viz. competency-based approach vs. approaches towards teacher’s self, and drawing on Bateson’s model, (Dilts, 1999 cited in Korthagen, 2004), Korthagen (2004) proposes his so-called ‘onion model’ as an attempt to address the question of ‘a good teacher’ (pp. 79-80). As shown in Figure 3.2., the model depicts different levels of changes that tend to occur to individuals.

"This image has been removed by the author of this thesis/dissertation for copyright reasons"

Figure 3.2. “The Onion: a model of levels of change” (Korthagen, 2004, p. 80)

Korthagen (2004, p. 80) argues that only the outer levels of ‘environment’ and ‘behavior’ could be observed directly. According to the model, teacher knowledge, e.g. subject knowledge rests in the level of ‘competencies’ which per se is informed by teacher’s ‘beliefs’ (p. 80). The level of ‘identity’ refers to teachers’ beliefs about
themselves, i.e. their professional identity (p. 80). Finally the level of ‘mission’
pertains to the roots of teachers’ “personal inspiration” (p. 80). It is my contention
that onion model can deepen teacher educators' insights into how to design and
develop teacher education and development programmes. In addition, it raises
administrators' awareness of the fact that not all teachers’ good qualities are
observable and concrete. Therefore, the contributions of the model are twofold, i.e.
teacher education and teacher evaluation. As Korthagen (2004, p. 87) posits, the
question of ‘good teacher’ cannot be simplistically answered by a list of teacher’s
competencies as one cannot expect ‘good teachers’ always to exhibit ‘good
teaching’. Korthagen (2004) maintains that the awareness of different levels of
change in the onion model could help better understand the multifaceted aspects
of good teaching (p. 87).

It is worth noting that the review of literature on TER shows an emphasis on
teacher effectiveness with few studies on appraisal effectiveness per se as though
the effectiveness of appraisal models is inferior to that of teachers. It is axiomatic
that valid and reliable information about teacher effectiveness can be garnered
through an effective teacher appraisal; otherwise the obtained findings would be
fundamentally flawed. Therefore, the extent to which an appraisal model can
effectively assess teacher effectiveness is a requisite concern. Drawing on data
from three studies, Piggot-Irvine (2003) introduces basic features for an effective
teacher appraisal as shown in Figure 3.3 as follows:
Piggot-Irvine (2003, p. 177) brings to attention the idea that such an appraisal needs to be developed in a wider context of “performance management”. More interestingly, as she postulates, such features cannot be simply “turned on”, on the grounds that they are mostly informed and influenced by a wider cultural context (p. 177). She also makes an eloquent reference to what she calls “educative process” on which there has been a dearth of research (p. 176). Indubitably, Piggot-Irvine (2003) features of effective appraisal are some generic but nonetheless important rudiments of an effective appraisal model which can be applied even to non-educational contexts. Nevertheless, I think, there are more to be conceived of as features of an effective appraisal model as discussed in the earlier sections.

Perhaps, one of the most established endeavours to investigate teacher effectiveness by which this study is informed is Hay McBer’s research report entitled ‘Research into Teacher Effectiveness, A Model of Teacher Effectiveness’ in the UK (Department for Education and Employment, 2000; McBer, 2005). As indicated in Figure 3.4., McBer (2000) introduces three major measures of teacher effectiveness each of which significantly contributes to learners’ achievements and none of which per se can lead to value-added teaching. According to the report,
these elements include “teaching skills”, “professional characteristics” and “classroom climate” (McBer, 2000, p. 6) as shown in Figure 3.4., as follows:

As the report suggests, an assemblage of teacher’s subject knowledge and teaching methods coupled with pedagogical and professional skills will lead to students’ progress (McBer, 2000, p. 8). Surprisingly, based on the collected data, McBer (2000, p. 8) claims that teachers’ biometric (demographic) data such as experience is not predictive of their effectiveness. Similarly, no specific evidence was adduced in respect of the predictive value of school context for students’ achievements (McBer, 2000, p. 8). Finally, McBer (2000, p. 9) suggests that the aforementioned elements, viz. teaching skills, professional characteristics and classroom climate could predict over 30% of the variance in students’ achievements.

McBer (2000) offers several constituents for each of the above-mentioned three elements involved in promoting students’ progress. With regards to ‘teaching skills’, he classifies 35 micro-behaviours under seven Ofsted inspection rubrics as depicted in Figure 3.5., in the following:
With respect to teachers' ‘professional characteristics’ which are at the root of teachers’ motivation thereby shaping their performance, McBer (2000, p. 19) proposes 16 characteristics under five rubrics as illustrated in Figure 3.6., as follows:
Finally, according to the report, the third measure of classroom climate refers to students’ perceptions of being a student in a classroom whereby their motivation and performance are influenced (McBer, 2000, p. 27). Hay McBer’s proposed dimensions for classroom climates include: “clarity, order, standards, fairness, participation, support, safety, interest and environment” (McBer, 2000, pp. 27-28). The analysis of the gained data revealed that those teachers who are possessed of high levels of professional and teaching qualities are deemed to positively affect their students’ progress (McBer, 2000, p. 34). Although with Hay Maber’s Model, some valuable and deep insights into teacher effectiveness such as professionalism might open up, my earliest concerns over the extent to which such findings could be applied to higher education context on the one hand and L2 education context on the other hand remain open to question.
3.13. Reflections on the existing lacunae (Support for a differentiated model)

The crux of the research on teacher effectiveness and teacher evaluation has been reviewed in this chapter. As expected, the review offered evidence to substantiate the idea that no single method, model, etc., for teacher evaluation can suffice for meeting the multidimensional issue of teacher effectiveness appraisal. In order to address the reviewed literature from a rather critical point of view, I am inclined to critique the reviewed literature from two interrelated yet distinct perspectives: the need for context-specific and subject-specific teacher appraisal model.

Campbell et al. (2004b, p. 451), maintain that much of the research into educational effectiveness appertains to the United States, the Netherlands and the UK with a focus on student achievements (p. 451). In a similar vein, Walker and Dimmock (2000, p. 155) express their reservations about what they call “penetration of Western policies and practices including teacher appraisal into Asian cultural context”. They also remind us of the influences that ‘culture’ can exert on the implementation of teacher appraisals across different contexts (2000, p. 159). Such an idea has been more or less reiterated by other researchers in the literature. Taking such a rather axiomatic criticism into consideration, it seems unwise to adopt an evaluation model for teacher effectiveness which is planned, designed, developed, and implemented in one context, e.g. Western countries and put it into practice in another context without considering its contextual underpinnings. By contextual underpinnings, it is meant all political, cultural, economic, and social norms, conventions and values that tend to be attached to any personnel appraisal model. In other words, although the model can be transferred, its contextual values which are mostly determinant would never be transferred. This does not mean that all countries need to develop their own teacher evaluation systems, given the fact that the development of a teacher evaluation system is a time-consuming phenomenon thereby taking several years to complete. Instead, the argument focuses on the extent to which the appraisal model is tailored to the individual needs of teachers in a particular context whereby it can accommodate the needs of all stakeholders as much as possible. Otherwise,
even a successful teacher evaluation model in one context might be nonfunctioning or malfunctioning in another context. Considering ‘politics’ as the element that forms “the character of personnel evaluation”, Bridges and Groves (1999, p. 321) propound that such an influence is often detrimental. The idiosyncratic nature of politics from one context (country) to another also makes clear the need for developing context-specific evaluation models. As Stoynoff (2007, p. 110) claims, teacher evaluation could not be implemented regardless of the social and political contexts within which it operates, on account of the influence they exert on the evaluation system. In other words, the characteristics or qualities of an effective teacher in one context are not necessarily guaranteed to be interpreted likewise in another context. For instance, the kind of communication and discourse between a teacher and his/her students with the opposite gender could be interpreted differently in different contexts considering the cultural and/or religious issues.

I have argued that the dictum that an effective teacher is effective regardless of what he or she teaches may not always be true. Moreover, EFL lecturers’ duties and responsibilities were found to be far different and to some extent more demanding than those of their counterparts in general and mainstream education, given the double-edged nature of learning English as a foreign Language. This is line with Troudi (2005, p. 125) who calls for ESL/EFL teachers’ “critical knowledge” of learners’ cultures and learning experiences. Reminding us of the uniqueness of foreign language teaching and criticising the adaptation of findings in general education to language education, Park and Lee (2006, p. 236), in a similar vein, advocate the need for in-depth investigation of the characteristics of effective foreign language teaching. Fradd and Lee (1998) argues that ESOL teachers tend to face the challenges of diversity emanating from students’ different cultural and educational backgrounds as well as different levels of achievement, more than any other educators (p. 761). According to Fradd and Lee (1998), such complexity necessitates having sufficient knowledge base for ESOL teachers in order to teach effectively (p. 761).

The current teacher evaluation system in the Iranian context similar to many other countries around the world is struggling with some challenges. As a part of
Educational policy, teacher evaluation systems need to be informed by what Pennycook (1990, p. 310) calls “empowerment of teachers” instead of viewing them as “classrooms technicians” thereby providing teachers with opportunities to reflect on their practices (also see Giroux & McLaren, 1989). To do so, not only should teachers’ voices be heard in their appraisal by policymakers and administrators but also valued. Moreover, power relationships which are mostly easily accepted (Burden & Troudi, 2007, p. 152), need to be critiqued and justified for a critical teacher appraisal system. Burden and Troudi (2007, pp. 163-164) also remind us of the misuse of evaluation on the part of administrators for purposes such as coping with overstaffing and call for ethical consideration inclusion in the evaluation scheme.

Drawing on the criticism attached to the traditional conception of TE and building on the recently differentiated conception of teacher appraisal (e.g. Campbell et al., 2004a; Walker & Dimmock, 2000), this study endeavours to inform the Iranian teacher evaluation system by lecturers’ perceptions of teacher effectiveness evaluation. Not only does this study call for the need for context-specific teacher appraisal, it advocates for discipline-specific appraisal model for ESL/EFL lecturers. It is hoped that such a differentiated model can answer some central concerns raised in the literature e.g. the ‘one-size-fits-all’ approaches towards teacher evaluation. In other words, the main question here is ‘how is it possible to evaluate ESL/EFL lecturers through an appraisal model used for lecturers from other departments and subjects, e.g. Biology or Physics. As sated earlier, such a model is indeed a generic appraisal model which is tailored to meet ESL/EFL lecturers’ individual needs.

3.14. Summary

In this chapter, I endeavoured to present a review of the most relevant research and to critique the literature with regard to my research questions with a particular focus on the existing gaps in the literature. The chapter started with an introduction to teacher effectiveness research. Then, the multidimensionality of teacher effectiveness research was explicated from a number of perspectives. These aspects included subject area, level of education, measures of evaluation,
geographical context of the study. In order to cast light onto the peculiarities and idiosyncrasies underpinning research on TE in different contexts, the literature was presented in three different contextual settings, viz. teacher effectiveness research in mainstream (general) education, second/foreign language education and finally the Iranian higher education context. It was argued that the literature on TER tends to have its roots mostly in Western contexts and more importantly in mainstream (general) education with few studies focusing on TER in the Middle-eastern contexts including Iran and, in particular, in L2 education. Afterwards, the problematic and challenging nature of defining the standards/criteria for teacher effectiveness was explained. The chapter thereafter proceeded to elaborate on different forms of teacher evaluation, viz. SETs, student achievement, peer-evaluation, self-evaluation, and finally internal or external observations. Four types of teacher evaluation were the next section which was examined thoroughly in this chapter. The types of evaluation included formative vs. summative, generic vs. differentiated, qualitative vs. quantitative and finally subjective vs. objective evaluations. As an important body of literature, the chapter then reviewed some famous and relevant national Acts, legislations, schemes and models germane to teacher effectiveness and teacher evaluation proposed by governments or researchers around the world. The chapter ended with a brief reflection upon the reviewed literature with a particular emphasis on the identified lacunae.
Chapter Four: Methodology
4.1. Introduction

The *methodology* section of most scientific writing such as theses, dissertations, journal articles, etc., has always been in the foreground. Methodology tends to serve as a medium through which researchers’ approaches towards research are signposted. It reflects researchers’ paradigmatic stances towards the phenomenon into which they investigate and thereby is associated with philosophical and theoretical underpinnings. Methodology is indeed “the philosophy of methods” (Sapsford, 2006, p. 175). To my understanding, methodology is to research as what engine is to a car. Methodology serves as a tool by which researchers can identify how to approach their research. Such awareness is of utmost significance as there is often more than one way to approach a phenomenon. As Pring (2004, p. 33) contends, ‘variety’ is one of the seminal features of educational research, in that different research questions necessitate different kinds of research. However, since researchers bring their worldviews to the research, they need to make their paradigmatic assumptions and frameworks explicit and also be cognisant of their own impacts on the conduct of research (Creswell, 2007, p. 15).

In this chapter, first, the philosophical and theoretical assumptions underpinning the current study will be introduced. To do so, the core concepts of ontology and epistemology will be presented and justified for the purpose of this study. The relationships between theory and practice will be also discussed. Afterwards, the methodological assumptions appertaining to the study, viz. methodology, methods, instruments and procedure will be explicated in details. Having introduced all the above-mentioned methodological aspects of the research, the adopted design for answering my research questions will be introduced. I will thereafter refer to data collection procedures including both quantitative and qualitative data collection procedures. The chapter will then proceed to data analysis section in which a full account of stages involved in the analysis of both types of datasets is provided. The quantitative-qualitative data nexus will be also examined in this chapter. The ethical dimensions germane to this research will be explained subsequently. In addition to the above-mentioned sections, I will recount any foreseen or
unforeseen challenges and limitations that emerged throughout the conduct of study. A brief summary of the whole chapter will be the final section of this chapter.

4.2. Theoretical and philosophical assumptions

Educational researchers have different philosophical stances which can exert impacts on the conduct of research (Pring, 2004, p. 88). In a similar vein, Crotty (1998, p. 1) reminds us of the vague relationship between methodologies and methods and their theoretical elements. According to Crotty (1998, p. 1), even the same terminology can be adopted differently and even contradictorily. And that is why Grix (2004) contends that most of research books tend to either skip this area or try to present the research traditions into a so-called dualistic understanding of research paradigms, i.e. positivism vs. interpretivism (pp. 76-77). As Creswell (2013) eloquently posits, researchers often bring certain philosophical stances to their research whether consciously or inadvertently (p. 15). It is worth noticing that such philosophical underpinnings lie behind our research methodology (Crotty, 1998, p. 7). Perhaps, one prerequisite step in addressing the philosophical dimensions undergirding research methodology is the concept of ‘paradigm’. Drawing on Guba (1990, p. 17), Denzin and Lincoln define paradigm as “a basic set of beliefs that guide action” (1994, p. 99; 2000, p. 157; 2005b, p. 183). Grix (2004) defines paradigm as “the understanding of what one can know about something and how one can gather knowledge about it” (p. 78). Nonetheless, as Scott and Morrison (2006, p. 169) posit, some other terminologies such as “episteme” (Foucault, 1972, cited in Scott & Morrison, 2006, p. 169) and “tradition” (Maclntyre, 1988, cited in Scott & Morrison, 2006, p. 169) have been used in social sciences and educational research to refer to the concept of paradigm. Although there is a general consensus on the components of a paradigm, the classifications are sometimes less consistent. According to Denzin and Lincoln (1994), a paradigm consists of three elements, viz. epistemology, ontology and methodology (p. 99). Nonetheless, they add “ethics (axiology)” as the fourth concept to the constituent elements of paradigm (Denzin & Lincoln, 2000, p. 157; 2005b, p. 183). Whereas Crotty (1998) considers the four elements of epistemology, theoretical perspective, methodology and methods as key components of any research
process (p. 2), Grix (2004, p. 74) introduces “ontology, epistemology, methodology, methods and sources” as key elements of research. Similarly, Scott and Morrison (2006) propose four elements of ontology, epistemology, strategy and method (pp. 85-86). Comparing these classifications, one can simply notice some slight differences among different researchers in their use of research language and terminologies. There are also some other concerns as to whether or not researchers should be in favour of “method-led” or “question-led” (Grix, 2004, p. 68) research thereby addressing the nature of interrelationships that exist among the aforementioned key elements of research. Grix (2004) adopts the analogy of “footings to house” for depicting the centrality of ontology and epistemology to research (p. 57). Taking the above-mentioned philosophical issues into account, it seems exigent to address the two seminal philosophical standpoints, namely ontology and epistemology. As Grix (2004) maintains, these two philosophical positions can lead to different perspectives about the same social phenomenon (p. 64).

As briefly noted earlier, educational inquiry has witnessed some discrepancies on researchers’ part throughout its history. Guba and Lincoln address their “metaphysics of inquiry paradigm” with respect to three elements including ontology, epistemology and methodology (Guba & Lincoln, 1994, p. 108; 2005, pp. 193-195). While Burrel and Morgan (1979, cited in Cohen, Manion, & Morrison, 2007, p. 9) propose a fourth assumption named “human nature”, Crotty (1998) introduces ‘theoretical perspective’ as a fourth element (p. 2) along with those of Guba and Lincoln (1994, 2005). In the same vein, Creswell (2007) introduces five philosophical assumptions two of which are what he calls “axiology” and “rhetorical” (p. 15). Axiology refers to the “role of value in the research” while rhetoric denotes the style of language used in the research (2003, cited in Creswell, 2007, p. 16). According to Creswell, researcher’s stance towards each philosophical assumption tends to have bearing on research design and conduct (p. 15). Having considered the above lines, it seems that the ontological, epistemological and methodological assumptions are indispensable components of research which spawn observable aspects of the study and hence should be given careful scrutiny. In the following sections, I will present an overall picture of the
methodological orientations of this study from ontological, epistemological, theoretical and methodological perspectives along with the methods and instruments utilised for data collection.

4.2.1. Ontological assumption

Ontology is “the study of being” and is concerned with the nature of reality and existence (Crotty, 1998, p. 10). Blaikie (2000, p. 8) defines ontological assumptions as those “concerned with what we believe constitutes social reality”. Although some researchers such as Crotty (1998) are under the impression that ontology and epistemology tend to merge together (p. 10), some others such as Grix (2004) contend that ontology is the starting point of research which logically precedes researcher’s epistemological and methodological stances (p. 59). In the same vein, Scott and Morrison (2006) distinguish ontology from epistemology in that, whereas the former refers to the nature of the reality, the latter denotes how such reality can be known (p. 85). Indubitably, different researchers tend to have different realities (Creswell, 2007, p. 16). However, a paucity of transparent understanding of ontological and epistemological underpinnings of a study might be perceived as inimical to a research project. Indeed, the boundaries sometimes are not clear-cut, thereby shattering researchers’ confidence especially the novice ones. The use of mixed methodology in research projects such as this study aggravates such challenges, given that fact that it gives rise to other philosophical questions and dilemmas as to whether or not one’s philosophical stances are commensurable. Having considered the above lines of thought, the paradigmatic stances adopted in this study are highly underpinned and informed by ‘pragmatism’. Actually, it has come to my attention that distinguishing amongst different philosophical positions towards a research project might not be clearly straightforward. Having read about different schools of thoughts about the nature of reality, my attention turned to pragmatism according to which reality in research is to be “revealed and experienced” (McCaslin, 2008, p. 672). Rejecting the traditional understanding of incompatible nature of scientific and interpretive paradigms, pragmatism, as a type of methodological pluralism, allows the researcher to utilise whatever “works best”, and to that end, is the rationale underpinning mixed methods research (Hewson,
It is my contention that pragmatism gave me the flexibility I needed to better answer my research questions. To my understanding, the ultimate goal of the present research was not merely a matter of explanation (‘Erklären’), but rather to understand the phenomenon (‘Verstehen’) (Schwandt, 1994, p. 119; 2000, p. 191). In other words, although it is contended that ‘reality’ exists, as Bryman (2008, p. 18) reminds us, social constructions tend to emerge from “perceptions and actions of social actors”. Rejecting the pure “naïve realism” advocated by positivism and “relativism” called for by constructivism (Guba & Lincoln, 1994, 2005; Lincoln & Guba, 2000, p. 165), this study is inspired by a more pragmatic ontological worldview. Nonetheless, such pragmatic standpoint does not imply that this study simultaneously adopts two distinct stances as one might argue they might not commensurate with each other. In fact, teachers' perceptions and understanding of teacher effectiveness and teacher evaluation are likely to be constructed in interaction with others.

4.2.2. Epistemological assumption

Blaikie (2000) defines the concept of epistemology as “the theory or science of the method or grounds of knowledge” (p. 8). Epistemology is the “theory of knowledge”; “it is a way of understanding and explaining how we know what we know” (Crotty, 1998, p. 3). According to the Sage Dictionary of Social Research, epistemology is “concerned with the possibility, nature, sources and limits of human knowledge” (Sumner, 2006, p. 92). It indicates the “acceptable knowledge” (Walliman, 2006, p. 15). Epistemology underpins “the nature of the relationship between the knower or would-be-knower and what can be known?” (Guba & Lincoln, 1994, p. 108). How a reality is known (epistemology) tends to be influenced by the beliefs about the nature of reality (ontology) (Scott & Morrison, 2006, p. 85). In line with ontological stances delineated earlier, a pragmatic approach was adopted towards ‘knowledge’ in this study. Creswell (2003) contends that knowledge claims tend to be based upon “pragmatic grounds” (p. 18). Moreover, as Dörnyei (2007) reminds us, mixed methods research is underpinned by “pragmatist position” (p. 30).
It is worth noticing that lots of terminologies under debate are often in the form of traditional ‘dualistic’ approaches towards research. In order to avoid such a dualistic position, the present study, as will be discussed later, is informed by rather pragmatic stances towards the phenomenon. It tries to bring together the views of both camps so that maximum understanding of the researched issue could be achieved. The epistemological standpoint of this study emphasises on constructing rather than exploring or creating the phenomenon. In order to bring the traditional dichotomous understanding together, viz. etic vs. emic, or nomothetic vs. idiographic dimension of the inquiry (Denzin & Lincoln, 1994, p. 100), this study adopts a mixed-methods approach to conduct research. Several researchers contend that ‘pragmatism’ is the best philosophical underpinning for ‘mixed methods’ research (e.g. Teddlie & Tashakkori, 2003, p. 20). Such an approach allowed me to combine survey with in-depth interviews, thereby providing profound insights into lecturers’ perceptions.

4.2.3. Research, theory and practice

The relationship between theory and research is not clear-cut (Bryman, 2008, p. 6). According to Bryman (2008), it is of high importance to identify whether the data is collected for the purpose of testing or generating a theory (p. 6). As the adopted philosophical and theoretical stances of this research project imply, the present study is aimed at building rather than testing theories. The manifestation of the nexus between theory and practice has been somehow challenging. As Creemers, Kyriakides, and Sammons (2010) posit, the relationship between general practices in educational effectiveness research has not always been well established (p. 64). Education as a field which is imbued with values as well as facts adds to the complexity of the theory-practice nexus (Winch & Gingell, 2008, p. 212). However, there are some central lines of ideas as to how the gap between theory and practice could be bridged. The findings of the present study are likely to have some direct and indirect bearing on general awareness of the existing teacher evaluation, teacher education and teacher development policies. It is hoped that this study can offer some theoretical and pedagogical implications for teachers as practitioners, administrators and policymakers. Therefore, it is of seminal significance to look into
the nature of the gap between theory and practice and get to know how one can inform the other thereupon.

4.3. Methodological assumptions

The quality of a piece of research, amongst others, rests upon the appropriateness of the selected methodology. According to the Sage Dictionary of Social Research (Jupp, 2006), methodology is “the philosophical stance or worldview that underlies and informs a style of research” (Sapsford, 2006, p. 175). Creswell and Plano Clark (2007) consider methodology as “the process of research” (p. 23). Research methodology has witnessed diverse philosophical and political views during last forty years (Dörnyei, 2007, p. 20). However, as Blaikie (2000) asserts, methodology and method are sometimes used interchangeably (p. 8). As discussed earlier, the concept of methodology cannot be addressed in a vacuum. Indeed, the very concept of methodology resonates with philosophical and theoretical underpinnings. Indeed, it is a framework enriched with paradigmatic assumptions from which a researcher benefits (O’ Leary, 2004, p. 85). Sometimes, it seems challenging to adjudicate between variants of methodologies, inasmuch as the selection of methodology is intertwined with philosophical underpinnings.

As discussed earlier, this study is informed by mixed methods research assumptions. As Cohen, Manion, and Morrison (2011, p. 115) remind us, no single blueprint can be proposed for planning a piece of research, given the fact that research design is determined by “fitness for purpose”. Drawing on Cohen et al. (2011) notion, it is my contention that the use of a mixed method design for this research project is a response to the purpose of this study. In other words, having considered the philosophical dimensions discussed earlier, I came to the conclusion that the proposed research questions could be best explored and answered through a mixed method study. Design functions as a road map to a researcher as it reflects methodologies, methods, instruments and materials used in a research project. In order to better depict the processes involved in the conduct of this study, a full account of mixed methods approach used in this research project is presented in the next section. Mixed methods research has gained an increasing reputation throughout the last decade. Hodkinson and
Macleod (2010, p. 186) maintain that mixed method research is experiencing a strong movement in the United Kingdom. Amongst several contributions that a mixed methods design brings to research conduct, is the freedom it provides for researchers. Imbued with the notion of ‘what works’, pragmatism is a philosophical assumption undergirding mixed methods research (Creswell & Plano Clark, 2007, p. 23). It is worth noticing that ‘methodological pluralism or eclecticism’ which is associated with mixed method research (Johnson & Onwuegbuzie, 2004, p. 14) will help the researcher meet the weaknesses and threats associated with unidimensional purist stances towards the phenomenon. The pragmatist underpinning of mixed methods research, as discussed earlier, will provide the researcher with, as Kivinen and Ristela (2003, p. 372) contends, the opportunity to cast aside “epistemological speculations”.

4.4. Research design

Bryman (2008) reminds us of the difference between research design and research method in that it offers a framework for execution of methods for data collection and analysis (pp. 30-31). Creswell, Plano Clark, Gutmann, and Hanson (2003, p. 211) define ‘design’ as “a procedure for collecting, analysing and reporting research”. According to Creswell and Plano Clark (2007), methodology can entail different designs which, as they argue, are “various types of approaches to research” (p. 23). With this end in view, design seems to be a pivotal milestone in any research projects. As an overarching concept embracing the whole process of research conduct, methodology is different from research design, for the latter points to the way towards answering research questions (Marczyk, DeMatteo, & Festinger, 2005, p. 22). As discussed earlier, the philosophical and theoretical underpinnings of the present research have fed into a mixed methods design. Such a decision was tough as it tends to be pregnant with several concerns as to how to select a design from amongst different research designs. Probing into this question, Teddlie and Tashakkori (2003) eloquently maintain that the ultimate goal of research is to address its predetermined questions (p. 14), and to that end, I believed that my research questions could be best answered through an exploratory mixed methods design.
The next step after selecting mixed methods research is to decide on what specific variants of design could best fit the research (Creswell & Plano Clark, 2007, p. 58). I suggest mixed methods design can best accommodate the needs of both the researcher and those being researched, and hence could best answer my research questions.

4.4.1. Mixed methods research

In previous sections, a detailed account of the philosophical underpinnings of the present study has been addressed. Since each single-method approach is often associated with some limitations, as Jick (1979, cited in Creswell, et al. 2003, p. 211) contends, using multiple methods for data collection can alleviate such shortcomings. Mixed methods research has kept researchers’ attention in the past decade (e.g. Creswell, 2014; Creswell & Plano Clark, 2007, 2011; Hesse-Biber, 2010; Tashakkori & Teddlie, 2003, 2010; Teddlie & Tashakkori, 2009). The escalating impact of mixed methods research in the past decade is undeniable (Hesse-Biber, 2010, p. 1). According to Dörnyei (2007), the emergence of the concept of ‘triangulation’ was the starting point of the real breakthrough of integrating qualitative and quantitative research (p. 43). Mixed methods research has been considered as the third methodological stance in research during the last 15 years (Dörnyei, 2007, p. 42). Hesse-Biber (2010, p. 1) distinguishes mixed methods from ‘multimethods’ in that the latter is a combination of two or more qualitative methods or two or more quantitative methods in a single study rather than combining two or more qualitative and quantitative methods. As Johnson and Onwuegbuzie (2004) remind us, mixed methods research permits the researchers to utilise methods with regard to research questions rather than fruitless biases pertinent to research paradigm hegemony (p. 23). Mixed methods approach provides the researcher with a better understanding of the phenomenon (Dörnyei, 2007, p. 47). In order to reach a vivid picture of the adopted design, let’s start with the definition of ‘mixed methods research’. Creswell et al. defines it as:

“A mixed methods study involves the collection or analysis of both quantitative and qualitative data in a single study in which the data are collected concurrently or sequentially, are given a priority, and involve the
Likewise, according to the Sage Dictionary of Research Methods (Jupp, 2006), mixed methods research is defined as:

“The combined use of both quantitative and qualitative methodologies within the same study in order to address a single research question” (Hewson, 2006, p. 179).

Mixed methods approaches towards research allow researchers to blend qualitative and quantitative methods (Plano Clark, Creswell, O’Neil Green, & Shope, 2008, p. 364). Other authors have a similar understanding of mixed method research. Nonetheless, one of the most comprehensive definitions of mixed methods research seems to be the one put forward by Creswell and Plan Clark (2007). They define mixed methods research both at the level of methodology which in turn implies philosophical underpinnings and at the level of methods. Creswell and Plan Clark (2007) propose the following definition for mixed methods research is:

“a research design with philosophical assumptions as well as methods of inquiry. As a methodology, it involves philosophical assumptions that guide the direction of the collection and analysis of data and the mixture of qualitative and quantitative approaches in many phases in the research process. As a method, it focuses on collecting, analyzing, and mixing both quantitative and qualitative data in a single study or series of studies. Its central premise is that the use of quantitative and qualitative approaches in combination provides a better understanding of research problems than either approach alone” (Creswell & Plan Clark, 2007, p. 5).

As the literature suggests, there are a number of purposes for using mixed methods design. For instance, Greene, Garaceli and Graham (1989, cited in Hesse-Biber, 2010, pp. 3-5) introduce five reasons for which mixed methods tend to be adopted by researchers, viz. “triangulation, complementarity, development, initiation and expansion”. In the same vein, building upon Sandelowski’s (2003) two main reasons for mixing methods, Dörnyei (2007) adds a third purpose and introduces three lines of rationale for mixing methods, viz. “expanding the understanding of a complex issue, corroborating findings through ‘triangulation’ and reaching multiple audience” (p. 164-166).
A number of designs can be used within mixed methods paradigm which will be briefly discussed in the following sections. However, I will justify why ‘exploratory’ design can best fit the present study. It is worth pointing out that the selection of a particular design mostly rests upon research questions rather than researchers’ favoured options. In view of the centrality of research questions, the proposed questions in this study did convince me to adopt an exploratory mixed methods design in order to better investigate teacher effectiveness thereby better answering the research questions. As Hesse-Biber (2010) asserts, the emphasis on qualitative dimension of mixed methods provides the researcher with an opportunity to view the phenomenon from a broader perspective and to deal with the issue of “social change, power and authority” with a particular emphasis on “multiple subjective realities” (p. 16). This, undeniably, helped me gain a more profound understanding of the event under investigation especially when supported by complementary quantitative data.

Given the variety of mixed methods research designs, it is of high importance to choose the most appropriate variant from among different designs. Creswell et al. (2003, pp. 218-219), propose four assumptions that give researchers some insights into the selection of mixed methods design, viz. “implementation, priority, integration and theoretical perspective”. Therefore, it might be a good idea at this stage to justify the selected mixed methods design for the present study in accordance to Creswell et al.’s (2003) four guidelines:

### 4.4.1.1. Implementation

According to Creswell et al. (2003), the assumption of implementation denotes the sequences of both quantitative and qualitative data collection which could be done ‘concurrently’ (contemporaneously) or ‘sequentially’ (over time) (p. 215). Referred to as “pacing and implementation”, ‘timing’ refers to the time two sets of data are collected and the order in which they are used by the researcher within a study (Creswell & Plano Clark, 2007; 2011, p. 65). In order to better answer the research questions, this study has adopted a sequential mixed methods design. It is worth noticing that in this design, the quantitative phase (questionnaire) is the first stage of data collection followed by a qualitative phase (interview). This design is, in fact,
a customised version of the one proposed by researchers (e.g. Creswell, 2009; Creswell & Plano Clark, 2011; Creswell et al., 2003). As it will be justified in 4.4.2., the QAUL–quan sequence was changed to quan–QUAL sequence.

4.4.1.2. Priority

Priority refers to the importance or weight given to quantitative or qualitative research during data collection phases of research (Creswell, 2009, p. 65; Creswell et al., 2003, p. 219). Referring to the problematic nature of making a decision as to which type of data should be emphasised, Creswell, et al. (2003) maintain that such a decision mostly relies upon the researcher’s convenience (p. 219). And therein lies a challenge, in that it does not seem to be solely a matter of convenience. Further to what Creswell, et al. (2003) call ‘comfort’, I think, such a decision is a matter of research questions. In other words, deciding on the weight given to either of two types of data seems to intrinsically rest upon research questions along with their philosophical and theoretical underpinnings. Of the very primary objective of this research, amongst others, is to explore the EFL teachers’ perceptions and to include their voices which have been less well-heard. As a consequence, such exploration can hardly yield appropriate findings using quantitative data, e.g. a survey study. Instead, what convinced me as a researcher to prioritise and emphasise qualitative data over quantitative data was the opportunity I sought to become an insider so that I could explore and understand participants’ perceptions of effectiveness.

4.4.1.3. Integration

According to Creswell et al. (2003), integration is “the combination of quantitative and qualitative research within a given stage of enquiry” (p. 220). This can occur at different levels of research, viz. “research questions, data collection, data analysis or interpretation level” (Creswell, et al. 2003, p. 220) and also at the level of design (Creswell, 2009, p. 66). In the case of this research which has adopted an exploratory design, as Creswell et al. (2003) point out, the emphasis is on description and understanding the phenomenon under investigation for which open-ended data collection such as interview is contributive (p. 220). Most
noticeably, Creswell et al. (2003) reminds us that the process of integration mostly tends to be brought into effect at ‘data analysis’ and ‘interpretation’ stages (p. 220). Taking the above-mentioned points into account, I came to the conclusion that this study could most benefit from integrating two datasets at the levels of data analysis and interpretation.

4.4.1.4. Theoretical perspectives

According to Creswell et al. (2003), theoretical perspectives refer to the lens via which researchers view the phenomena which could be explicit or implicit (pp. 222-223). This dimension of research design has been fully discussed in the earlier sections pertinent to philosophical underpinnings of the current study. However, it is worth restating that this study adopted an exploratory approach towards the inquiry and tried to bring together both objectivist and subjectivist understanding of the phenomenon.

4.4.2. Sequential Exploratory Design

In previous sections, different design variants appertaining to mixed methods research have been proposed. There are four major research designs within mixed methods research, viz. “the Triangulation Design, the Embedded Design, the Explanatory Design, and the Exploratory Design” (Creswell & Plano Clark, 2007, p. 59). Given the nature of the research questions set forth in this research project which were to explore EFL teachers’ perceptions of teacher effectiveness and to delve into their understanding of teacher appraisal system, I was convinced that sequential exploratory design could best fit the present research. The use of such design allowed me as the researcher to identify the expanded ideas and to “explore the phenomenon” thereupon (Creswell et al., 2003, p. 227). Once again, it is worth mentioning that the adopted design for this study was a two-phase design whose first stage was a quantitative data collection phase preceding the next stage of qualitative data collection with a special weight given to qualitative data. According to Creswell (2003), the integration of the findings takes place in the “interpretation” phase (p. 215). Here, the goal is “exploration” (Creswell, 2003, p.
The plan of the whole design in a nutshell is depicted in Figures 4.1., and 4.2., which are adapted from Creswell et al. (2003, pp. 224-225):

![Diagram of Sequential Exploratory Design]

**Figure 4.1.** The adopted research design based upon Creswell et al. (2003, pp. 224-225) four criteria

### 4.4.2.1. A few remarks on the design differentiated for this study

In previous sections, an overview of different mixed methods designs and sequential exploratory design in particular has been presented. Nonetheless, there appear some subtleties associated with Creswell, et al.’s sequential exploratory design. As Creswell and Plano Clark (2007) remind us, researchers should not expect rigorous findings unless they adopt a well-designed research procedure (p. 79). Given the researchers’ willingness to alter some aspects of their four major designs, Creswell and Plano Clark (2007) encourage researchers to select a single design that can best fit the their research (p. 79). This is a commonsense approach in that no single design can suffice for researchers’ purposes. This is why I adapted the sequential design to make the most of the advantages afforded by quan-QUAL sequence. In other words, the collection of quantitative data as the first stage of data collection provided me with valuable insights into the areas which required further in-depth investigation in the subsequent stage of qualitative data collection, i.e. interviews. Indeed, fastidious attention was paid to identify those areas which received either full or little support from respondents. To do so, the QUAL-quan sequence was tailored to quan-QUAL sequence. In addition, I postponed the full data analysis until both sets of data were collected. The sequential exploratory design used in this study can be portrayed as follows:
4.4.3. Participants (Sample of the study)

In addition to the term ‘sample’, there exist a number of alternative terms such as *informants* and *interlocutors* which have more or less been adopted by researchers to refer to people who take part in a research project. However, some of these terms connote quantitative or qualitative underpinning, e.g. *participants* in qualitative research and *subjects* in quantitative research. Perry (2005) defines ‘*sample*’ as the source of data used for answering research questions or testing research hypotheses (p. 55). Perry views sample consisted of ‘*cases*’ which as he argues, are formed by ‘*subjects*’ and more recently ‘*participants*’ (pp. 55-56). In quantitative studies, as Dörnyei (2007) asserts, *sample* refers to a group of participants who take part in an empirical investigation, whereas *population* denotes “the group of people whom the study is about” (p. 96). As the two-stage design of this study suggests, two groups of participants took part in the research phases. The concepts of sample and sampling strategies tend to be different in qualitative and quantitative research. According to The Sage Dictionary of Social Research Methods (Jupp, 2006), ‘sampling’ refers to the “techniques used to select groups from a wider population” (Davidson, 2006, p. 271). However, as Dörnyei (2007, p. 125) reminds us, participant sampling tends to be approached quite
differently in qualitative and quantitative research. According to Kemper, Stringfield, and Teddlie (2003), sampling is of two different types, viz. ‘probability’ and ‘purposive’ sampling (p. 277). As Dörnyei (2007) posits, most researchers in the field of Applied Linguistics use purposive or non-probability sample (p. 98). He further adds that ‘convenience’ or ‘opportunity sample’ is the most common type of sample in L2 research (2007, p. 98). In such sampling, participants are selected based upon certain important criteria appertaining to the purpose of the study (Dörnyei, 2007, p. 98). It is worth highlighting that, ‘non-representative’ sampling is not considered as a problem in qualitative research (p. 98). As exploratory mixed method research, this study similar to qualitative studies is not interested in the representativeness of the sample and hence generalisability of the findings. Instead, the goal of this study is to reach rich and deep insights into the phenomenon inherent in research questions for which ‘purposeful’ or ‘purposive’ sampling seems to be the best approach (Dörnyei, 2007, p. 126). According to Silverman (2005), purposive sampling provides the researcher with an opportunity to critically select participants on the basis of a feature or process in which the researcher is interested (p. 129). Reminding us of the concerns pertinent to the selection of interviewees, Esterberg (2002) places emphasis on the selection of participants who can provide the researcher with the most possible insights into the research topic (p. 93).

Given the exploratory nature of this study, purposive, non-probability and in particular ‘criterion based’ sampling strategies were employed at both stages of the research. According to Creswell (2003), ‘purposeful sampling’ tends to be used for qualitative data collection in that it allows the researcher to select the individuals who ‘have experienced the central phenomenon’ (p. 220). *Purposive* sampling is the sampling strategy that has been used in the qualitative phase of this study. In purposive sampling the researcher uses a criterion or purpose in selecting the samples (Kemper et al., 2003, p. 279). Moreover, as Patton (1990, p. 169, cited in Kemper, et al., 2003, p. 279) contends “selecting information-rich cases for study in depth” is the underlying rationale for purposive sampling. Inasmuch as the present study was aimed at exploring the participants’ perceptions in depth and hearing their voices, purposive sampling seemed to fit the research objectives. Needless to
say that such a sampling strategy tends to be associated with some limitations, given Dörnyei’s (2007) notion that any general claims of the results should be approached with caution (p. 99). Another aspect of sampling is the sample size which is significantly varied in quantitative-qualitative continuum. While quantitative social scientists are much interested in studying a large sample to legitimise the generalisability power of their findings, qualitative researchers are more interested in studying small purposive sample (Schrank, 2006). Creswell (2013) introduces a general rule for the sample size of qualitative study in that not only should it be focused on a few individuals, it needs to extensively investigate each individual in depth (p. 157).

Having taken into consideration the above-mentioned issues, I invited 15 lecturers to participate in a pilot study. In the main round of research, a purposive and criterion based sample of 43 lecturers were selected to take part in the quantitative data collection phase. In a similar vein, a purposive, criterion based sample of 14 lecturers were requested to participate in interview stage. I deliberately selected the respondents and participants based on the assumption that they all had a set of particular features and characteristics, thereby enabling me for better exploration and understanding of my central themes and questions (Ritchie, Lewis, & Elam, 2003, p. 77). As the foremost criterion, I selected the participants who were EFL lecturers. However, in order to have a more comprehensive understanding, I intentionally selected the interviewees from both genders with different experiences, levels of academic qualification and rank, major, and employment status, affiliated to different types of universities and higher education institutions. Whereas the majority of participants completed their academic programmes in TEFL, there were also lecturers with academic degrees in other areas, i.e. English Language and Literature, and Translation. However, all lecturers work in ELT departments at their respective universities. As to the level of academic education, three groups of lecturers were identified, i.e. lecturers with MA or PhD degrees as well as those who were classified as PhD candidates. Table 4.1., shows some background information of interviewees as follows (for questionnaire respondents, see appendix 3):
<table>
<thead>
<tr>
<th>Pseudonym</th>
<th>Gender</th>
<th>Qualification</th>
<th>Academic major</th>
<th>Experience (years)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Ali</td>
<td>Male</td>
<td>PhD candidate</td>
<td>TEFL</td>
<td>0-5</td>
</tr>
<tr>
<td>2 Soroush</td>
<td>Male</td>
<td>PhD candidate</td>
<td>TEFL</td>
<td>11-15</td>
</tr>
<tr>
<td>3 Sohrab</td>
<td>Male</td>
<td>PhD candidate</td>
<td>TEFL</td>
<td>0-5</td>
</tr>
<tr>
<td>4 Sarah</td>
<td>Female</td>
<td>MA</td>
<td>TEFL</td>
<td>0-5</td>
</tr>
<tr>
<td>5 Majid</td>
<td>Male</td>
<td>PhD candidate</td>
<td>TEFL</td>
<td>0-5</td>
</tr>
<tr>
<td>6 Parham</td>
<td>Male</td>
<td>PhD candidate</td>
<td>TEFL</td>
<td>6-10</td>
</tr>
<tr>
<td>7 Sepehr</td>
<td>Male</td>
<td>PhD candidate</td>
<td>TEFL</td>
<td>16-20</td>
</tr>
<tr>
<td>8 Mersedeh</td>
<td>Female</td>
<td>MA</td>
<td>TEFL</td>
<td>0-5</td>
</tr>
<tr>
<td>9 Thelma</td>
<td>Female</td>
<td>MA</td>
<td>TEFL</td>
<td>0-5</td>
</tr>
<tr>
<td>10 Rima</td>
<td>Female</td>
<td>MA</td>
<td>TEFL</td>
<td>0-5</td>
</tr>
<tr>
<td>11 Saman</td>
<td>Male</td>
<td>MA</td>
<td>Translation</td>
<td>6-10</td>
</tr>
<tr>
<td>12 Amir</td>
<td>Male</td>
<td>MA</td>
<td>TEFL</td>
<td>0-5</td>
</tr>
<tr>
<td>13 Armin</td>
<td>Male</td>
<td>PhD candidate</td>
<td>TEFL</td>
<td>6-10</td>
</tr>
<tr>
<td>14 Niloofar</td>
<td>Female</td>
<td>PhD candidate</td>
<td>Literature</td>
<td>11-15</td>
</tr>
</tbody>
</table>

Table 4.1. Interviewees’ background information

4.4.4. Methods and instruments

Cohen et al. (2007), define methods as ‘the range of approaches used to gather data for purposes of inference and interpretation, explanation and prediction’ (p. 47). As Cohen et al. (2007) continue, the concept of method is traditionally perceived as ‘those techniques associated with positivistic model’ (p. 47). According to Johnson and Turner (2003), a method of data collection refers to the ‘technique that is utilised for collecting empirical data’ (p. 298). Whatever the definition, the concept of method is imbued with other concepts such as instruments, tools, techniques and strategies all of which should be well-considered prior to conduct of a study. Nonetheless, as Creswell (2003) puts it forward, selection of methods is based upon a prerequisite step, i.e. whether the goal of study is to specify the type of information in advance or to let the data emerge from participants (p. 17). The methods used in the present study for collecting the two sets of data, i.e. quantitative and qualitative data, were informed
by research questions. In other words, the selection of the appropriate methods to collect data lies with the nature of research questions and how the proposed questions could be best investigated. Informed by such philosophical standpoints, this study benefited from two major methods, namely questionnaire and interview. However, while methods refer to general concepts and tools for data collection, instruments refer to research-specific methods adopted for collecting data. Instrument specifically refers to any devices used for data collection (Perry, 2005, p. 52); it is a ‘mechanism for measuring phenomena’ (Colton & Covert, 2007, p. 5). Therefore, deciding on the appropriate instruments and tools for collecting different kinds of data which can best fit the adopted research design is the next step a researcher needs to consider. A full account of the processes involved in designing, developing, piloting and administering the instruments used in this study will be discussed in data collection section shortly in this chapter. However, the two main methods utilised for collecting data are briefly introduced as follows:

4.4.4.1. Questionnaire

Rugg and Petre (2007, p. 141) maintain that questionnaires are rarely used well. Based on what a questionnaire targets to measure, three different types of data can be obtained, viz. “factual, behavioral and attitudinal” (Dörnyei, 2003, p. 8). The questionnaire developed and used in the present study is an attitudinal questionnaire aimed at exploring the respondents’ feelings of and attitudes towards teacher effectiveness and its pertinent evaluation system. There are a number of different question and response formats in questionnaires (Cohen et al., 2007). Two main types of questionnaires, namely close-ended and open-ended questionnaires were developed for the purpose of this study to collect quantitative and qualitative data respectively. Moreover, the questionnaire was supplemented with a section for demographic and factual information which were further used for descriptive statistics. The details of different sections of the questionnaire and processes involved in designing and developing the instrument will be fully elaborated in Section 4.5.1, i.e. quantitative data collection.
4.4.4.2. Interview

As it was stated earlier, this study is aimed at exploring lecturers’ feelings and ideas about teacher effectiveness and teacher evaluation. The philosophical and theoretical underpinnings of such an aim give weight to the idea that such reality is at least partially socially constructed on the part of participants. To explore such idiographic rather nomothetic behaviour (Cohen et al., 2007), this study benefited from semi-structured interviews. Qualitative interviews help the researcher better “understand experiences and reconstruct events” (Rubin & Rubin, 2005, p. 3). Creswell (2013) introduces the “mechanics” of conducting interview as a challenge in qualitative interviewing (p. 172). Nonetheless, the very challenge seems to have its roots in researchers’ epistemological considerations. In other words, there tends to be different epistemological standpoints as to whether interview should be considered as an opportunity to ‘collect’ knowledge or ‘construct’ knowledge (Kvale & Brinkmann, 2009, p. 48). Therefore, it might be possible to extrapolate the way a researcher conducts his/her interviews based on his/her epistemological stances towards the interview.

Conducting a good interview is demanding whereas spoiling it is a simple task (Rugg & Petre, 2007, p. 137). Semi-structured interviews have a number of advantages for both interviewer and interviewees. Not only does semi-structured interview give the interviewer the ability to guide the conversation more flexibly to research questions, it provides the respondents with some latitude for expressing their interests (Hesse-Biber & Leavy, 2006, p. 125). A detailed account of different stages involved in developing the interview guide and conducting the semi-structured interview is given in the upcoming sections.

4.4.5. Procedure

The procedure adopted for the conduct of this study is explained in a visual model adapted from Creswell, et al. (2003) as shown in Figure 4.3. The following schematic figure illustrates different stages involved in the conduct of research, e.g. participant selection, piloting, etc. As Creswell (2003) maintains, checking the validity and accuracy of both quantitative data and qualitative findings is another
dimension of data analysis in a mixed methods study (p. 221). This is why prior to the selection of participants I ensured that the instruments were reliable through a pilot study. Drawing on Creswell et al.'s (2003) proposed model for sequential exploratory design, I conducted this study through the following stages:

![Diagram of Sequential Exploratory Design](image)

**Figure 4.3. The procedure for Sequential Exploratory Design of the study (adapted from Creswell et al., 2003, p. 225; 235)**

### 4.5. Data Collection

The process of data collection, as Creswell (2013) notes, appears to be “a series of interrelated activities” practiced by researchers in order to better answer their research questions (p. 146). According to Creswell’s “data collection circle”, there exist a series of activities researchers need to get through to collect their data.
These include “locating site/individual, gaining access and making rapport, purposefully sampling, collecting data, recording information, resolving field issues and storing data” (p. 146) some of which have already been mentioned. In this section, the processes involved in collecting two sets of data will be presented. Taking the characteristics of sequential exploratory designs outlined earlier into consideration and based upon Creswell et al.’s (2003) proposed design, data collection was conducted in a two-phase sequential fashion. Although Creswell (2003) suggests a qualitative-quantitative sequence, as justified earlier, data collection was implemented the other way around. Indeed, it started with a quantitative phase whereby questionnaires were administered to a group of participants. After gathering the data and doing the preliminary data analysis, the second phase of data collection, i.e. interviews were conducted. This sequence provided me with an invaluable opportunity to gain insights into aspects that were of more significance to participants. Having assured myself of the breadth of the data garnered in quantitative stage, I thereafter explored in depth the ideas accentuated by respondents in the following interviews. Adopting such an approach helped me not to miss out any undeclared and unresolved ideas on the part of the participants. In the following sections, I will present a detailed account of quantitative and qualitative data collection along with different methods and instruments used for collecting the data, viz. closed-ended and open-ended questionnaires as well as semi-structured interview.

4.5.1. Quantitative data collection

As stated earlier, in order to collect quantitative data, a researcher-developed questionnaire was adopted in this study. However, it is worth referring to different activities and stages involved in developing a questionnaire. These stages and activities are briefly introduced in the following sections whereby I will explain the processes involved in designing, developing, piloting, revising and administering the questionnaire.
4.5.1.1. Questionnaire

Questionnaire is perhaps the commonest method of data collection (Johnson & Turner, 2003, p. 303) for which there are a number of different names such as “inventories”, “opinionnaires”, “scales”, etc. (see Dörnyei, 2007, p. 102). However, constructing a reliable and valid questionnaire is rather a challenging task. Johnson and Christensen (2000, cited in Johnson & Turner, 2003, p. 303) introduce thirteen principles for questionnaire development. Given the purpose of the study and the type of data a questionnaire is about to collect, different kinds of questions and statements can be interpolated in a questionnaire. Generally, three different types of data can be gained through a questionnaire, viz. “factual”, “behavioural” and “attitudinal” data (Dörnyei, 2003, p. 8; 2007, p. 102). Based upon Dörnyei’s (2003) categorisation, the questionnaire developed for this study embraced both factual and attitudinal questions. It is worth emphasising that the attitudinal questions interpolated in the questionnaire covered different concerns appertaining to the respondents’ ‘attitudes, opinions, beliefs, interests and values’ (Dörnyei, 2003, p. 9).

The developed questionnaire included three major sections, namely closed-ended questionnaire, open-ended questionnaire and a section for demographic information. In this study a 63-item Likert-scale close-ended questionnaire was developed and administered to 43 participants (appendix 1). ‘Likert scale’ is the commonest type of closed-ended questionnaire in which the five responses range from ‘strongly agree’ to ‘strongly disagree’ (Dörnyei, 2003, pp. 36-37; 2007, p. 101). Given Dörnyei’s (2003) notion that questionnaire is not suitable for exploratory research (p. 47), an open-ended section was added to compensate the paucity of ‘openness’ in the questionnaire. Dörnyei (2003) introduces three sources for creating questionnaire items and what he calls ‘item pool’, namely the researcher’s own thoughts, the informants’ (interviewees) data, and established questionnaires which have frequently been adopted in the literature (p. 52). In line with Oppenheim’s (1992) conception of ‘pilot work’, the questionnaire developed for the purpose of this study went through different stages of construction including composing, revising, trying out, refining an improving (p. 47). Indeed, the
processes involved in the development of questionnaire in this study included *designing* and *developing* the first version of the questionnaire, *piloting* the questionnaire, *refining* and *designing* the final questionnaire with established validity and reliability. In the following sections, each of the aforementioned stages is explained in brief.

### 4.5.1.2. Designing a researcher-developed questionnaire

After developing an item pool for the questionnaire as discussed earlier, I attempted to develop a questionnaire that has clear objectives and a logical structure with sections (Lewin, 2005). The developed questionnaire had three major sections, namely closed-ended, open-ended sections as well as a section for participants’ demographic information. The initially developed questionnaire included 98 statements which was eventually reduced to a 63-statement questionnaire in a Likert scale format with five responses, viz. strongly disagree, agree, neither agree nor disagree, disagree and strongly agree. Since the Likert scale is a battery of questions with identical responses (Bryman, 2008, p. 222), a horizontal answer format with abbreviation was used. The open-ended sections, in contrast, contained 6 questions which were thematically linked to my research questions. In formulating the questions, special attention was paid to avoid ambiguity, complexity, double-barreled and double-negatives statements and questions, thereby providing the participants with clear and polite instruction as to how to complete the questionnaire (Lewin, 2005, p. 220) and also not to cramp the overall presentation of the whole questionnaire (Bryman, 2008). The integration of two types of questionnaire, i.e. close-ended and open-ended sections contributed to both *depth* and *breadth* of the instrument. While the close-ended section contributed to cover a broad range of questions, the open-ended as an ‘information revealing’ section (Perry, 2005, p. 110) allowed the respondents to express their thoughts and beliefs in a deeper sense.

‘*Validity*’ and ‘*reliability*’ are two main concepts indicating whether or not an instrument yields credible and accurate information (Colton & Covert, 2007, p. 65). Being around for a long time, the concept of validity has been the grounds for discussion among researchers who have proposed their definitions of the concept
Validity refers to the extent to which an instrument measures what it is supposed to measure (e.g. Dörnyei, 2003, p. 110). As a matter of degree, validity has different types one of which applies to the questionnaire developed for this study, i.e., ‘content’ validity. Content validity refers to the extent to which “an instrument is representative of the topic” (Colton & Covert, 2007, p. 68). Since it might not be feasible to identify and/or include all factors germane to the focus of the study in the questionnaire, in order to operationalise the construct, the researcher can consult ‘content experts’ about the relevance of the items (Colton & Covert, 2007, p. 68). Having adopted a ‘qualitative’ approach towards the instrument, I made every attempt to design and develop the questionnaire based on the reviewed literature as well as the feedback received from two topic experts in the field which as Colton and Covert (2007, pp. 68-71) assert, demonstrate content validity of the instrument, thereby indicating whether or not it measures the construct. Therefore, the instrument was deemed to have acceptable validity.

Special attention was given from the outset to develop the questionnaire with regard to research questions. The statements in the close-ended section and the questions in the open-ended section were thence thematically linked to my research questions. Such a strategy helped me ensure that all aspects of research questions were covered by the questionnaire as much as possible. Also, this helped me in the subsequent stages of data analysis in which an intricate network of interrelationships among two datasets and research questions was established, better interpret the entire data. As stated, the very first developed questionnaire consisted of 98 statements in close-ended section and 6 questions formulated for open-ended section as well as 7 questions focusing on respondents’ demographic backgrounds. To establish whether respondents were willing to take part in the second phase of the research, i.e. interview, a short note explicating the follow-up phase of research along with its purpose, thereby inviting them to participate, was appended to the questionnaire. In the second phase of developing the instrument, I sent the questionnaire to my supervisors for their comments and feedback. After making the amendments and refining the first questionnaire, a second version of the questionnaire was designed with 72 statements in the closed-ended section
and 7 questions in the open-ended section for the pilot study which will be explained in the following section.

4.5.1.3. Piloting the questionnaire

Conducting a pilot study is a desirable and important part of research (Bryman, 2008). Dörnyei (2007) urges the researchers to pilot their research instruments and procedures prior to the commencement of research conduct (p. 75). Piloting provides the researcher with insights into the validity, reliability, and feasibility of a construct. Being more central to quantitative data, piloting qualitative data is different from that of quantitative data (Dörnyei, 2007, p. 75). There are a number of advantages associated with using a pilot study prior to main research data collection amongst which are to establish the adequacy of the given instructions, to detect the questions that tend to be less realised by participants which need to be asked in the follow-up interview and to see if the instrument as a whole functions properly and meets the instrument objectives (Bryman, 2008, pp. 247-248). Piloting also helps the researcher identify the potential pitfalls associated with the questionnaire (Lewin, 2005). As an integral part of developing a questionnaire, ‘field testing’ can help minimise the effects of actual wording of the items on responses (Dörnyei, 2003, p. 63) and if properly done, piloting, as a stepwise process, might require several weeks to conduct (p. 65). In this study, the pilot study was implemented at two stages, namely (a) piloting of the item pool, (b) final piloting (dress rehearsal) followed by item analysis (Dörnyei, 2003, pp. 66-68).

The whole process of piloting the instrument indeed lasted more than what I initially expected. Interestingly, the pilot study turned out to be a precious opportunity for me whereby I could discuss some important issues pertinent to the questionnaire with my colleagues. Many of them expressed their interest in the topic of the research, i.e. teacher effectiveness. Being inspired by the ideas in the questionnaire, some of the respondents confessed that they had never thought about their appraisal critically before. Piloting helped me identify those items which were proved to be verbose, ambiguous, double-barreled, etc. However, the majority of the participants were happy with the thoroughness of the questionnaire.
believing that most aspects of teacher effectiveness they could imagine were reflected in one way or another in the questionnaire.

As stated in the previous section, the original questionnaire developed for pilot study consisted of 98 statements and six questions in close-ended and open-ended sections, respectively. The questionnaire also included a section for participants’ demographic information which was included for further statistical analysis. In the first stage of the pilot study, 15 participants were selected to participate in the study. All participants were fully informed of the purpose of the study and their rights as the participants of the study. The process of administering the questionnaire was done during four weeks. While the response rate was of an ideal range, some participants left the open-ended section of the questionnaire either unanswered or partially answered. In the item analysis of the questionnaire, three major points were given full attention including ‘missing values, range of responses and internal consistency’ of the instrument (Dörnyei, 2003, pp. 68-69). A number of statistical analyses were conducted for the obtained data such as Tests of Normality, Kurtosis, Skewness, etc. However, the most important statistical analysis at this stage was establishing the reliability of the scale, i.e. questionnaire. It is worth emphasising that the present questionnaire is a researcher-made questionnaire developed for the purpose of this study for which as Dörnyei (2003) reminds us, it is not feasible to establish indices of every aspect of reliability (p. 110). As Pallant (2007) contends, one of the main issues here is the scale’s internal consistency which indicates the extent to which the scale’s items ‘hang together’ (p. 95). According to Pallant (2007), Cronbach’s alpha coefficient is one of the most conventional indicators for identifying the internal consistency of a scale which should ideally be more than 0.7 (DeVellis, 2003, cited in Pallant, 2007, p. 95). According to Dörnyei (2003), the researcher can feel fairly safe provided the questionnaire has the internal consistency (p. 110). The obtained Cronbach’s alpha coefficient was 0.844. Drawing on the findings of the pilot study, a new 72-item questionnaire was prepared. In order to explore the extent to which the questionnaire has been refined and improved in terms of the Internal Consistency Reliability, a second round of item analysis was performed. The obtained Cronbach’s alpha coefficient for the newly prepared scale reached 0.910 which
seemed to be ideal for the scale. Moreover, the Mean Inter-Item Correlations has increased from 0.069 in the pilot study to 0.136 in the refined questionnaire. However, in order to remove the statements with negative values and hence to increase the homogeneity of the scale, seven statements were removed from the questionnaire culminating in a 65-item questionnaire which eventually changed to a 63-item questionnaire.

4.5.1.4. Constructing the final questionnaire

After gathering the relevant information from the pilot study, the main research questionnaire was developed which was actually a more refined and polished version of the first questionnaire. Throughout the process of developing the questionnaire, special attention was given to the reliability and validity of the instrument. Based upon the above-mentioned results of the pilot study, a final version of the questionnaire was developed. The questionnaire then had three major sections including closed-ended, open-ended and demographic information sections. Moreover, an introductory paragraph about the aims of the research and also a clear instruction on how to complete the questionnaire were included at the outset of the questionnaire. As I expected some participants might be interested in the second phase of the research which was an interview, a paragraph appreciating participants’ interests in completing the questionnaire and thereby gently inviting them for the next research stage was appended to the end of the questionnaire. As stated in the previous section, the final instrument used for main research consisted of 63 statements in the close-ended section, 6 questions in the open-ended section and 7 factual questions germane to participants’ demographic information.

4.5.1.5. Administering the questionnaire

The questionnaire developed for this study was ‘self-administered’ or ‘self-completion’ questionnaire (Bryman, 2008, p. 216). There are a number of different ways to administer a questionnaire, viz. ‘face to face, telephone, post or electronically’ (Lewin, 2005, p. 220-221). In this study, the questionnaires were administered in two ways. Mainly the questionnaires were handed out to
participants in a face to face mode and were collected once they had been completed. Such one-to-one administration helped me build rapport with the participants, explicate the purpose of the research and encourage them to participate in the study (Dörnyei, 2003, p. 81). Nevertheless, a number of participants completed and sent back their questionnaires via email. I was a bit flexible with participants’ requests for completing and returning the questionnaires in the following days. The participants were mostly informed about the purpose of the study and were asked in advance whether or not they were happy to take part in the study. Moreover, they were requested to make their choices as to the date, time and venue for meeting at their convenience. Generally, the data were mostly collected in English Language Departments. The response rate of the questionnaires was relatively high as most of them were administered face to face. As participants were fully briefed on the questionnaire, few missing values were found in the closed-ended sections. Moreover, a gentle follow-up reminder was sent to those participants who preferred to complete the questionnaire via email, thereby asking them to return the questionnaire at their earliest convenience.

4.5.2. Qualitative data collection

As Marczyk et al. (2005) maintain, the choice of data collection strategy depends on the nature of research questions and variables in the study. The actual types of data and the adopted procedures for collecting data tend to be the typical reaction to qualitative data collection (Creswell, 2013, p. 145). As Creswell (2013) adds, data collection embraces a series of actual processes including ‘gaining permission, doing a qualitative sampling strategy, preparing means for recording and storing the data, and being prepared for any unforeseen ethical issues that may arise during the research’ (p. 145). Since this research was highly informed by exploratory underpinnings, a second phase of qualitative data collection was devised for answering the research questions. In this phase, two methods were utilised for qualitative data collection, i.e. interview and open-ended questionnaire. A detailed account of the stages involved in developing the adopted methods and instruments will be presented in the following sections.
4.5.2.1. Semi-structured Interview

Interview is the second commonest method for data collection (Johnson & Turner, 2003, p. 305) and a crucial component in most qualitative research (Marczyk et al., 2005). A review of the books and articles written on qualitative data collection introduces a number of strategies and procedures for doing the interviews (e.g. Creswell, 2013; Kvale, 2007; Kvale & Brinkmann, 2009; Rubin & Rubin, 2005). Interview is actually ‘inter-view’ and ‘inter-action’ between the interviewer and interviewee through which knowledge tends to be constructed (Kvale, 2007, p. 1) and follows a structure and purpose (Kvale & Brinkmann, 2009). Radnor (2001) proposes three actions to introduce interviewing, viz. ‘ask, listen and record’ (p. 59) which clearly indicate different processes involved in an interview. Radnor (2001) further introduces ‘active listening’ as a major skill needed for semi-structured interview, on the grounds that this will allow the interviewer to explore an overall picture of the topic under investigation from interviewees’ perspectives (pp. 60-61). While probing participants is not possible in questionnaire, interviews provide the researcher with an opportunity to probe the interviewees for illumination even though it is more expensive to administer (Johnson & Turner, 2003, p. 305). Esterberg (2002) contends that interview rests at the heart of social research (p. 83). Interviewees’ perceptions are the objective for which many interview studies are conducted (Silverman, 2005, p. 48). According to Dörnyei (2007), ‘semi-structured interview’ is amongst the commonest interviews in the field of Applied Linguistics (p. 136). Semi-structured interview benefits from both ‘pre-prepared guiding questions and prompts’ as well as ‘open-ended format’ in which the interviewees can elaborate on raised issues (Dörnyei, 2007, p. 136). Flick (2006) introduces three types of questions for semi-structured interviews; viz. “open questions, theory-driven questions and confrontational questions” (pp. 156-157). Interview is quite different from other data collection techniques in that it requires higher interpersonal skills (Oppenheim, 1992, p. 65). Oppenheim (1992) further introduces two kinds of interviews, namely ‘standardised and exploratory’ interviews (p. 65). While the former tends to be used in large-scale surveys, the latter tends to be conducted in research in which the goal is to explore and understand how people think and feel about the issue under investigation (p. 67).
He further assigns the notion of ‘idea collection’ rather than data collection to such in-depth interviews (p. 67). Semi-structured interviews allow the researcher to explore the topic under investigation and to elicit the interviewees’ ideas in their own words (Esterberg, 2002, p. 87). Dörnyei (2007) also encourages researchers to prepare an interview guide which can provide them with the best possible assistance (p. 137). Proposing two main characteristics for a good qualitative interview, viz. the natural flow of the interview and its richness, Dörnyei (2007) draws our attention to the importance of the researcher’s neutrality during the conduct of the interview (p. 140).

The first round of quantitative data collection helped better understand the participants’ interests and gave me deep insights into the areas which seemed to be of participants’ interest. Indeed, it allowed me to better recognise those issues which are of a typical EFL lecturer’s primary concerns. Similar to the processes involved in developing and constructing the final version of questionnaire, preparing reliable and comprehensive interview questions was not an easy task. In other words, the final interview guide (protocol) was the final product of a series of challenging processes, form the first intention of developing a tentative protocol, to receiving the feedback from experts, and to piloting and constructing the final version. As Creswell puts forth, it might be regarded ‘a series of steps’. A review of the literature clearly reflects different stages of doing an interview proposed by several researchers and scholars (e.g. Creswell, 2013; Kvale, 2007; Kvale & Brinkmann, 2009). Kvale’s seven-stage framework of planning an interview is amongst the most well-known dynamics of an interview journey. An interview inquiry is a linear progression through seven stages viz., “thematising, designing, interviewing, transcribing, analyzing, verifying and finally reporting” (Kvale, 2007, pp. 35-36; Kvale & Brinkmann, 2009, p. 102). Kvale places particular emphasis on the pre-interview stages of thermalizing and analyzing as they help the researcher for high-quality interviews (2007, p. 50). Likewise, Creswell (2013) introduces nine steps that need to be considered by the researchers in conducting interviews including:

- “Deciding on research questions
- Identifying interviewees
Having considered the above-mentioned lines of ideas, several stages involved in developing and conducting the semi-structured interviews used in this research project will be introduced in the following sections.

### 4.5.2.1.1. Interview guide/schedule

Interview guide ‘structures the course of the interview’ (Kvale & Brinkman, 2009, p. 130). An interview guide or interview protocol is an outline of the questions that are to be asked from participants. These questions tend to be the sub-questions in a research study that could be best understood by interviewees (Creswell, 2013).

Dörnyei (2007) asserts that an interview guide helps the researcher make sure about his or her interviews in a number of different ways including:

- “Everything is covered and nothing is missed out by accident
- Appropriate question wordings
- List of probing questions
- A template for opening statements
- Comments to be considered” (p. 137)

Dörnyei further suggests different sections that can be fit into an interview guide viz. the first few questions which tend to be factual or personal, the content questions, probes and the final closing question (2007, p. 137).

Being informed by common templates suggested for an interview guide (e.g. Bryman, 2008; Dörnyei, 2007; Kvale, 2007; Kvale & Brinkmann, 2009), an interview guide or schedule was developed in this study (appendix 2). The adopted interview protocol started with some personal and factual questions deemed to function as a warm-up and introductory stage in interview. Afterwards, a number of questions which corresponded to research questions thematically were formulated. Some probing questions were also developed for further use in interviews in case the need arises. There were also some closing questions which asked the
participants to express whatever they would like and also to mention any further points that from their points of view had been ignored or less focused in the interviews.

4.5.2.1.2. Piloting the interview guide

Similar to the processes involved in developing the questionnaire as discussed in the previous sections, the interview guide and procedures adopted for conducting the interviews were also refined through a small pilot phase. Not only did this pilot phase helped me distinguish the questions which seemed to be ambiguous or difficult to answer by participants, it brought to my attention my own questioning style and interview strategies for which I needed more rehearsal. In fact, I practiced the habit of being a good listener with a lesser speaking role. Moreover, I learned how to make an interview different from an oral questionnaire. This helped me a lot in the next real research interviews in which I provided the interviewees with more opportunity to talk. Additionally, I could explore the ways probing questions can make the flow of conversation on the right and desirable track.

4.5.2.1.3. Constructing the final interview guide

After piloting the initial version of the interview guide, a new ready-to-administer interview guide was developed. In the newly refined version, I made every attempt to make the interview conduct as smooth as possible. Based upon the pilot study of the interview guide, I tried to remove the questions that tended to be ambiguous or complex or double-barreled. The goal was to make the questions brief and simple (Kvale & Brinkmann, 2009). The interview started with some warm-up and introductory questions including the interviewees’ career experiences and interests proceeded by the core research questions. Moreover, some probes were developed to ensure that interviews will be done on the right track. Use of probes helped me not to allow the interviewees to diverge or avoid from the topic of the interview. Although in the main interview guide there were some specific types and a limited number of interview questions, the actual type and number of questions in interview varied across different interviews. In other words, while the core research questions were consistent in all interviews, different sub-questions emerged
throughout each interview which in some cases were specific to a particular interview. Interestingly, most of different types of research concepts proposed by Kvale and Brinkmann, namely ‘introductory, follow-up, probing, specifying, direct, indirect, structuring, interpreting questions as well as silence’ (2009, pp. 135-136) were more or less used in the interviews. These types of questions are more discussed in the forthcoming sections.

4.5.2.1.4. Conducting the interviews

The number of participants and the sampling strategy were fully discussed in the previous sections. Kvale (2007) contends that the number of participants in a study heavily ties in with the ‘purpose of a study’. However, the number of interviews in common interview studies tends to be around 15±10 (Kvale, 2007, p. 44; Kvale & Brinkmann, 2009, p. 113). Having gone through all previous prerequisite stages, I conducted fourteen interviews twelve of which were one-on-one interviews and two of which were conducted online via Oovoo (an online video chat program). As a variation of computer-assisted learning, chat interviews tend to be more synchronous in time and similar to face-to-face interactions (Kvale & Brinkmann, 2009, p. 149). During interviews, I tried to be a good listener rather than a frequent speaker (Creswell, 2013), although this position varied a lot depending on the interviewees. Indeed, the characteristics of interviewees seemed to be an important factor in conducting the interviews. For instance, there were interviews in which I spoke less as the participant was motivated enough to talk. Nonetheless, I experienced interviews in which I had to be more active as the interviewees constantly obliged me to use probes and questions in order to make them talk. Therefore, Creswell’s notion of a ‘good listener’ seems to partially tie in with the interviewees themselves. I really found the very beginning few minutes decisive (Kvale & Brinkmann, 2009, p. 128). I tried to establish a friendly environment and build trust so that interviewees feel at ease and feel free to share their feelings and insights with me as a researcher. Throughout interviews I tried to show my attentive listening and to declare that I was interested in their responses. As a general guideline, the interviews tended to commence with a briefing stage (Kvale & Brinkman, 2009, p. 128) in which the interviewees were fully briefed on the
purposes of interviews and the contribution they could bring to the research. The interviewees were also informed of the issues of confidentiality and privacy and their rights for withdrawal. The interviews were continued with introductory or let’s say warm-up questions whereby the interviewees were asked to talk a bit about themselves, their career, etc. Then, a number of core research questions were asked supplemented with some probes and other types of questions if needed. The purpose of using different types of questions was to keep the interviews on the right track. Questions in semi-structured interviews might not necessarily follow what is outlined in the interview guide (Bryman, 2008, p. 438). Since the interviews conducted in this study were of semi-structured nature, I was slightly flexible in asking the questions that were not included in the interview guide but rather emerged due to the interviewees’ responses (Bryman, 2008). In practice, I noticed how interviews turned out to be richer in terms of the different flexible questions one after another.

The interviews were recorded on a portable digital recorder (Dictaphone). After each interview, I created a backup copy of the interview audio files on my laptop. This was to avoid the loss of the interview recordings by accident. Moreover, using a professional Dictaphone has some other advantages as it could filter out the background noise and record the interviews with a superior quality. The interviewees were informed briefly in advance of the purpose of the study and the possible length of the interviews. Most of the interviews were conducted in the lecturers’ offices at their relative universities. In fact, most of them were happy to be interviewed at their universities. I was quite aware of the demanding and challenging nature of doing interviews especially with lecturers and therefore made every endeavor to predict any unforeseen challenges that could happen during the interviews. However, I faced some unforeseen limitations which will be discussed later on in this chapter.

4.5.2.2. Open-ended questionnaire

Although open-ended questionnaire cannot really be considered as a separate instrument, for ease of presentation, I explain it in this section as an instrument used for collecting qualitative data which was integrated with the close-ended
section of the questionnaire. The use of open-ended questions allows the participants to respond however they wish (Bryman, 2008). Similar to other instruments for collecting data, open-ended self-completion questionnaires are associated with some advantages and disadvantages. As Bryman eloquently points out, open-ended questions allow the respondents to freely express their thoughts and feelings beyond the boundaries that they might be trapped in close-ended questions. Without imposing any particular answering pattern, open-ended questionnaire can help the researcher explore new ideas as well as the areas of which the researcher is less aware (2008, p. 232). The crux of the matter throughout designing and developing the questionnaire laid in how the issues of depth and breadth could be best met. The integration of open-ended questions and closed-ended questions bestows confidence on the researchers to address the issues of depth and breadth in data collection.

Having considered the above important guidelines, seven open-ended questions were developed for the open-ended section of the first version of questionnaire. The seven open-ended questions were revisited after a pilot analysis stage. A pilot study on 15 participants was conducted to illuminate the weaknesses or ambiguity associated with the questions which will be briefly explained in the following section.

### 4.5.2.2.1. Piloting the Open-ended questionnaire

One of the advantages of piloting an open-ended questionnaire was to identify the questions that were problematic to code into categories (Dörnyei, 2003, p. 65). Similar to the results of piloting the close-ended section of questionnaire reported in Section 4.5.1.3., in this section I will present an overview of piloting the open-ended part of questionnaire, in brief. The approaches towards qualitative data analysis seem to be more challenging and time-consuming than those towards quantitative data analysis which are said to be more straightforward. More importantly, how these two sections, i.e. close-ended and open-ended sections, interact and how the extracted data will work in tune with each other seem to be amongst interesting yet challenging points in overall analysis of the whole questionnaire. The participants’ responses were analysed using MAXQDA10
Having done the pilot study, a number of issues and problems were identified. As a consequence, the original seven questions were revised and six questions were reformulated and appended to the final version of the questionnaire.

4.5.3. Trustworthiness

The issues of validity and reliability have already been explained in the section of quantitative data collection. However, in qualitative research in which the researchers “strive for understanding” (Creswell, 2007, p. 201), these concepts are superseded by some alternative terms. “Credibility”, “transferability”, “dependability” and “confirmability” are the qualitative criteria which replace their positivist counterparts, i.e. ‘internal validly, external validity, reliability and objectivity’ (Bryman, 2012, p. 390; Denzin & Lincoln, 2005a, p. 24; Lincoln & Guba, 1985). The central idea here is the notion of ‘trustworthiness’ which, as Given and Saumure (2008, p. 896) contend, allows qualitative researchers to demonstrate the worth of their research. However, each and every of these ideas carries a number of strategies or techniques to operationalise the criteria for quality (Creswell, 2007). Creswell and Miller (2000, p. 126) propose several validation strategies which may vary depending on the researcher’s paradigmatic assumptions, e.g. ‘triangulation’, ‘thick, rich description’, etc. Given Creswell’s (2007) recommendation that qualitative researchers need to engage in at least two of validation strategies, I ensured the quality of qualitative aspect of my research through the following ways:

“Triangulation”, as a technique to improve trustworthiness, involves the use of multiple types and different sources of data through multiple methods and lenses (e.g. Creswell & Miller, 2000; Ritchie, 2003). As a mixed methods inquiry, this study has its roots in the construct of triangulation (Greene, Kreider, & Mayer, 2005, p. 274). The triangulation or more specifically methods triangulation (Hesse-Biber, 2010) adopted in this study, allowed me to use multiple methods to collect multiple types of data, thereby “strengthening” the obtained findings and “enriching” the subsequent interpretation (Rothbauer, 2008, p. 892). The mixed method approach provided me with an opportunity to contrast the methods and the
collected data, thereby increasing my confidence about the findings (Cohen et al., 2007).

“Transferability” implies that research findings can be transferred to other contexts other than that of the study (Jensen, 2008, p. 886). However, as Lincoln and Guba (1985, p. 316) argue, it’s not the researchers’ duty to prove the transferability of their research. Instead, they need to provide the readers with sufficient and rich data based on which the users of research can decide on the extent to which it is transferrable (cited in Cohen et al., 2007, p. 137). Qualitative researchers can use two major strategies to increase this quality, namely ‘thick description’ and ‘purposeful sampling’ (Jensen, 2008). “Thick, rich description” is a validation strategy which increases the credibility of research whereby a qualitative researcher provides a detailed account of the setting, the participants and the themes of the study (Creswell, 2007; Creswell & Miller, 2000, p. 128). In this study, I made every endeavour to present a vivid picture of the context of the study (chapter 2). I also provided a full account of respondents and participants' individual background data including their gender, experience, qualification, etc. (chapter 4). In addition, I tried to purposefully select participants who are most consistent with my research design and aims (Jensen, 2008).

4.6. Data Analysis

Analysing text and deciding on how to present the data is a challenging task (Creswell, 2013). Having collected both sets of data, I faced voluminous data to which I had somehow different feelings. Whereas quantitative data analysis seemed relatively straightforward to me, I was impressed by huge amount of qualitative data. In fact, I needed to decide on the ways that I wanted to approach my so-called messy qualitative data. Rugg and Petre (2007, p. 153) maintain that clear research design and research questions give the researcher an idea as to how to go about doing data analysis (p. 153). As Hesse-Biber and Leavy (2006) posit, there exists no one way for doing the data analysis (p. 344). Data analysis in mixed methods research depends on the research strategies adopted for research procedures (Creswell, 2003, p. 220). As Creswell reminds us, data analysis could
be implemented both *within* and often *between* two data collection approaches, i.e. qualitative and quantitative data (2003, p. 220). Prior to discussing the data analysis procedures being utilised in this research project, it might be a good idea to refer back to some points raised earlier as to how the integration of the two sets of data, i.e. qualitative and quantitative data was done. According to Creswell et al. (2003), data analysis in sequential designs tends to follow quantitative and qualitative stages independent of one another (p. 232). In this study a number of procedures and tools were used for data analysis which are explained as follows:

**4.6.1. Quantitative data analysis**

As an exploratory study with a mixed methods design, collecting and analysing quantitative data can shed some lights on data interpretation and can help the researcher make sense of data and feel assured of its breadth. The collected data were analysed using the SPSS software (Statistical Package for the Social Sciences) Version 18. The use of some statistical procedures such as descriptive statistics, gave me a picture about the interplay among participants' different attributes as well as their attitudes towards the notion of teacher effectiveness and teacher appraisal. For instance, it seems interesting to get to know how participants' various attributes such as gender, academic qualification, field of study, years or experience, employment status and type of their affiliated university tend to interact with their perceptions and understating of an effective teacher. Seemingly, even minor alterations as such can potentially exert influence on the way different lecturers approach the appraisal system through which they are assessed. It also helped me explore the similarities and differences which did exist between the two sets of data, i.e. quantitative and qualitative data which apparently rest at the heart of data interpretation and discussion.

A number of preliminary data analysis procedures were done in order to secure the reliability of the instrument. Therefore, the internal consistency reliability of the questionnaire was calculated using SPSS. As explained in Section 4.5.1.3, the calculated Cronbach alpha value of the questionnaire used for the main study showed a relatively high level of internal consistency. Apart from the preliminary statistics, every endeavor was made to spot on any similarities and differences
between the findings obtained from the two instruments. That's why, as mentioned earlier, the questionnaire was developed thematically. In other words, in the initial stages of designing and developing the questionnaire, I tried to develop the statements on a thematic basis that could best fit the proposed research questions. Furthermore, the same strategy was adopted while designing the interview guide. A detailed account of some descriptive statistics will be presented in the next chapter.

4.6.2. Qualitative data analysis

The emergence of a huge amount of data is among the difficulties researchers tend to face in doing qualitative data analysis (Bryman, 2008, p. 538). However, as an exciting stage, qualitative data analysis requires an amalgamation of “creativity, inspiration and diligent detection” (Spencer, Ritchie, & O’Connor, 2003, p. 199). As Radnor (2001) reminds us, the ultimate goal of qualitative data analysis is to see whether the collected data can answer the research questions (p. 69). She further contends that data analysis tends to be followed by interpretation of the findings which in turn allows the researcher to develop theoretical explanations (p. 69). Qualitative data as compared to quantitative data is often more voluminous and messy and as the literature suggests, there are a number of different approaches to its analysis. Such different approaches tend to be shaped by epistemological underpinnings attached to qualitative inquiry and the researchers’ roles thereupon (Spencer et al., 2003, p. 200). The trade-offs among different approaches towards qualitative data analysis depend on the way they address different issues such as “the status of the data, the primary focus of analysis, the way data are reduced, the kinds of concepts generated, the way concepts are applied to the data, … the place of the researcher in the analytical account”, etc. (see Spencer et al., 2003, pp. 202-206).

The commencement of qualitative data analysis as Coffey and Atkinson (1996) remind us, resonates with identifying ‘key themes and patterns’ (p. 26). As stated earlier, in the early stages of data analysis I felt less confident in approaching qualitative data analysis as compared to quantitative data analysis. However, I found the processes involved in various stages of qualitative data analysis quite
interesting. After collecting and preparing the data, I prepared a report on the qualitative data analysis which as my first attempt contained a long list of categories, thereby embracing a myriad of subcategories and codes. Upon the receipt of my supervisor’s feedback, I tried to polish and cut down the number of categories by merging the categories which tended to be of similar nature or repetitive. A detailed account of different stages of analysing qualitative data including analysing the semi-structured interviews and open-ended questionnaires are discussed in the forthcoming sections in this chapter.

4.6.2.1. Interviews data analysis

A review of the literature on different stages and strategies involved in the process of data analysis suggests a number of general guidelines or strategies for data analysis (e.g. Creswell, 2013; Hesse-Biber & Leavy, 2006; Kvale, 2007; Kvale & Brinkmann, 2009; Marczyk et al., 2005). For instance, Kvale and Brinkmann (2009) introduce some key approaches to interview analysis, viz. ‘analyses focusing on Meaning, analyses focusing on Language and General analyses’ (p. 197). In their proposed “the analytic hierarchy”, Spencer et al. (2003) suggest a “conceptual scaffolding” which drives the researcher from initial ‘raw data’ to ‘findings’ in an iterative manner and involves three stages of ‘data management, descriptive account, and explanatory account’ (pp. 213-217).

Despite some variations in the literature, data analysis tends to involve the following stages (e.g. Creswell, 2013, pp. 180-187; Kvale & Brinkmann, 2009, p. 201):

- “Data management: preparing and organising the data
- Data exploration: Reading and memoing
- Data reduction: describing, classifying and coding, condensing the data
- Data interpretation: Representing and visualizing the data”

As Creswell further adds, the processes of collecting, analysing and reporting data are interrelated and might be done simultaneously (2013, p. 109). Likewise, there were times in this study in which one stage of the data analysis overlapped with another stage. The data analysis procedures used in this study evolved through four major stages including ‘data management, data exploration, data reduction,
and finally data interpretation’ which are discussed in the following sections. Technically, having collected the data, I assigned three different files on my computer to each and every one of my instruments adopted for data collection, viz. close-ended questionnaire, semi-structured interviews and open-ended questionnaires. The next step was to prepare the data for the analysis and to transcribe all the interviews thereupon. Having transcribed the interviews and prepared the respondents’ answers to open-ended questionnaires, I had to decide on selecting and using suitable data analysis software for the analysis of qualitative data. Although I felt computer-literate enough to use technological advancement for data analysis, I had some reservations about the feasibility of doing the analysis by NVivo, given the fact that I had never used it beforehand. Having decided on doing the qualitative data analysis using NVivo instead of doing it manually, I attended some workshops on NVivo and fortunately soon after I was able to use the software quite well. The use of NVivo allowed me to easily assign any descriptions, memos, annotations, attributes, etc., to my participants, which in turn helped me have a more profound understanding of the data. As mentioned earlier, the stages involved in the analysis of qualitative data were more or less similar to those suggested in the literature of which a detailed account is presented as follows:

4.6.2.1.1. Data Management: preparing and organising the data

Data management which is actually preparing and organising the data is the first step in data analysis. There are some actions that should be taken prior to this stage for which some choices need to be made; for instance, deciding on the mode of analysis, use of equipment, etc. Transcribing the interviews and the quality of transcriptions are undeniably important elements at this stage. The quality of transcription has received little attention as compared to interviewing (Kvale & Brinkmann, 2009, p. 177). Transcribing an interview is a process in which an interview is transformed from an oral format to a written one. The interviews in this study were all verbatim transcribed. It was endeavoured throughout the data management process to increase the reliability of the interview transcription by listening to the interviews again in order to double check for any discrepancies in
the transcripts. In the same vein, special attention was given to the punctuations in the transcripts as they could potentially bring about different interpretations. As Kvale and Brinkmann (2009) assert, even the exact same written words could have different meanings based upon the punctuation used by the researcher (p. 185). Although the analysis of interviews was mostly done using NVivo, I managed to create some data files equivalent to those of NVivo’s in a Microsoft Office Word Document format on my laptop.

4.6.2.1.2. Data exploration: reading and memoing

Hesse-Biber and Leavy (2006) contend that data exploration is a stage in which the researcher ‘thinks about’ the data and tends to evolve with a subsequent step of data reduction (p. 347). Amongst various advantages of the use of NVivo to analyse data is the researcher’s ability to add memos, annotations, and descriptions to the generated themes, categories and subcategories. After preparing and organising the data, the next step was to read interview transcripts and respondents’ responses to the open-ended questionnaires. Reading the materials was not simply reading the transcripts; instead, I read between the lines to capture and draw out any neglected or less noticed ideas on the part of participants. Such a strategy helped me form some core concepts from the transcripts which were of two major types. Whereas some ideas were rather new, some others served as supporting statements for the pre-developed concepts or categories. As Maxwell and Miller (2008) state, a number of different labels have been coined for this process. ‘Units of data’ (Maxwell & Miller, 2008, p. 465), ‘unitizing’ (Labov & Fanshel, 1977, pp. 38-40; Lincoln & Guba, 1985, p. 344, cited in Maxwell & Miller, 2008, p. 465), ‘segmenting’ (Coffey & Atkinson, 1996; p. 26; Tesch, 1990, p. 91, cited in Maxwell & Miller, 2008, p. 465) are among the labels reflecting this process. As mentioned earlier, the use of NVivo brought a number of advantages to data analysis. As I read the transcripts, I could easily add memos and annotations and highlight segments, thereby allowing me to make a deeper sense of data in later stages.
4.6.2.1.3. Data reduction: describing, classifying and Coding the data

Generating codes and categories is the cornerstone of qualitative data analysis in which themes are developed (Creswell, 2013, p. 184). Search for ‘meanings’ rests at the heart of data analysis for which ‘coding’ and ‘memoing’ are two important techniques (Hesse-Biber, 2010, p. 190). As the most used categorising strategy (Maxwell & Miller, 2008, p. 465), ‘coding’ requires the researcher to aggregate the text into small categories, documenting and labeling the generated codes (Creswell, 2013, p. 184). Coding embraces a range of approaches to organise, retrieve and interpret data (Coffey & Atkinson, 1996, p. 27). As Scott and Morrison (2005) put it forward, generating codes can be challenging from both conceptual and methodological perspectives (p. 46). Coding process can commence as soon as the researcher begins to collect data (Hesse-Biber & Leavy, 2006, p. 347). However, coding and categorising seem to be adopted interchangeably by different researchers. Coding differs from categorising in that while the former is done by attaching a keyword to a test chunk, the latter tends to be ‘a more systematic conceptualisation of a statement’ (Kvale, 2007, p. 105; Kvale & Brinkmann, 2009, pp. 201-202).

Coding is practically an amalgamation of ‘data reduction and data complication’ (Coffey & Atkinson, 1996, p. 31). There are two approaches to coding, namely “concept-driven” and “data-driven” coding (Kvale & Brinkmann, 2009, p. 202). Concept-driven coding implies researcher’s use of pre-developed codes obtained from literature or the material, whereas data-driven coding refers to the codes that tend to develop and emerge through reading the materials (Kvale & Brinkmann, 2009, p. 202). In a similar vein, Grabtree and Miller (1992, p. 151) introduce ‘prefigured’ and ‘emergent’ categories in their proposed coding continuum (cited in Creswell, 2013, p. 185). The categories and codes generated in this study were of both types. However, the majority of categories and codes did emerge through the analysis of the data. Creswell (2013) further draws attention to another aspect of coding, namely “code labels”. He refers to three sources for labeling the generated codes including ‘in vivo codes’ which are the exact participants' words, code names obtained from social sciences, or the names that researcher develops.
which can best describe the ideas' (Creswell, 2013, p. 185). Given the above-mentioned ideas, this study witnessed similar strategies for generating codes. In other words, there were times in which participants explicitly referred to a concept which per se could best depict the ideas they tend to convey, whereupon I used them as labels. By contrast, I also could not exclude myself from some pre-existing ideas and concepts from the relevant literature and hence had some pre-developed categories and codes which, as stated, were informed by the literature. However, there were also some occasions in which I had to develop a category or code that could best reflect what the participants intended and meant to say. That is why I can say the codes and categories generated in this study were an amalgamation of all these types of coding strategies.

4.6.2.1.4. Data interpretation: representing and visualizing the data

As Coffey and Atkinson (1996) remind us, coding is not solely a series of classification, categorisation, coding or collating data; instead, it tends to be a process whereby a social phenomenon is represented and reconstructed (p. 108). Data representation is the final stage of data analysis which represents the data in different forms, i.e. ‘in text, tabular or figure’ format (Creswell, 2013, p. 187). As mentioned earlier, the processes involved in data analysis might not necessarily be in an interlinear fashion. However, data presentation and interpretation were the final stage of data analysis in this study. At this stage, the researcher tends to go beyond the surface meaning of statements, and tries to ‘re-contextualise the statements that were de-contextualised by categorisation’ (Kvale, 2007, p. 108). As a fluid process, research requires the researcher to constantly get involved in different stages of research in a dynamic way and try to get back and forth throughout the research process and search for new connections thereupon (Hesse-Biber, 2010, p. 194). Nonetheless, it is worth considering Coffey and Atkinson’s notion of the ‘integrity of analysing and theorising’ in that these two aspects of research are not separate stages (1996, p. 139). This stage of data analysis is rather challenging as I had to play around with myriad generated ideas to identify the best possible combinations for which I had to merge and/or collapse the subthemes, categories and subcategories recurrently. Apparently, the use of a
mixed methods design adds some new orientations to the way data were interpreted and represented. This will be more discussed in the upcoming section of quantitative-qualitative nexus in this chapter.

4.6.2.2. Open-ended questionnaire data analysis

The open-ended questionnaires were analysed much similar to the way that semi-structured interviews were analysed. In other words, all documents from open-ended questionnaire were transferred to NVIVO software similar to those of the interviews. It is worth pointing out that I did the analysis of qualitative data three times. After doing the preliminary and prerequisite stages of preparation of data, as discussed earlier, I first started with the analysis of semi-structured interviews, on the grounds that I was somehow confident about the richness of data in interviews as compared to open-ended questionnaires. After preparing the first round of data categorisation which was done based on the data obtained from semi-structured interviews, I began to analyse the open-ended questionnaires for which I allocated a specific folder in NVIVO. In my third attempt to review and analyse the data, I used the data extracted from interviews as a platform in my total qualitative data analysis to which I could add the categories or subcategories from open-ended questionnaires. Two strategies were involved in this process. First, there were some ideas generated from open-ended questionnaires that were absolutely new with no reference in semi-structured interviews, therefore they were added as new categories or subcategories. Second, the ideas were either repetitive or similar by nature with those of the interviews; therefore they were merged with the existing themes, categories and subcategories so that the ideas and codes were supported with both quotes from interviews and open-ended questionnaires.

4.7. The qualitative-quantitative data nexus

A detailed account of different stages involved in the conduct of this study, from methodological assumptions, to data collection and then to data analysis procedures, was presented throughout this chapter. Having collected and analysed the data, the journey arrived at the point of reporting and writing up the findings. There is a grain of truth in the idea that this point could hardly proceed unless the
researcher has gathered and analysed the data (Bryman, 2008). Reporting and drawing inferences and conclusions from data seem to be a challenging and demanding stage of a research. Nonetheless, the use of mixed methods design which in turn imposes two sets of data on a researcher adds to such challenges, for the researcher then needs to decide on how to relate and marry different kinds of data.

Reporting and reflecting on voluminous data obtained from questionnaire and interview is a formidable task and hence should not be taken for granted. As Kvale and Brinkmann (2009) eloquently remind us, interview report is a ‘social construction’ and such reporting is not simply re-presenting the interviewees’ views along with the researcher’s viewpoints (p. 267). Reporting qualitative data was a more delicate and unenviable task as compared to that of quantitative data. While presenting quantitative data tends to be more straightforward by using statistical software, presenting the results of interview studies turns out to be more challenging for which there are no standard ways (Kvale & Brinkmann, 2009, p. 279). Indeed, there appear to be different ways to disseminate the findings each of which might target specific types of audience (Rubin & Rubin, 2005, p. 247). Kvale and Brinkmann (2009) introduce ‘interview quotations’ as a common mode of presenting the result of interview studies which tend to range from ‘precise verbatim’ to ‘narrative restructuring’ (p. 279).

Further to the details of sequential exploratory design discussed earlier, it has been argued that mixing can happen at different stages throughout the conduct of the study. The central line of idea at this stage was how to relate these two sets of data and make sense of the entirety of data thereupon. Having considered all aspects undergirding the present study, I came to the conclusion that the interpretation phase could be perhaps the best stage for relating the findings. In fact, special attention was given to develop the instruments thematically. Not only did I develop the questionnaire thematically based upon the research questions, I did design the interview guide thematically. This helped me a lot to juxtapose qualitative data with quantitative data and ponder over the entire data for spotting on any similarities and/or differences. Moreover, special endeavor was made to organise and
articulate the arguments throughout the presentation of findings in a smooth and logical order, from idea to idea (Kvale & Brinkmann, 2009, p. 254).

4.8. Ethical considerations

As a “moral enterprise”, special attention should be given to both the means and the ends of an interview (Kvale, 2007, p. 23). Ethical issues may arise at any stage throughout an interview inquiry. Given the fact that research involves others, Curran (2006) reminds us of the ethical apprehension that a researcher may inevitably confront in the field (p. 198). Similarly, Cohen et al. (2007) draw attention to the “costs/benefits ratio” which, in their words, refers to the ethical dilemma researchers tend to encounter in achieving a balance between their own demands and their participants’ values (p. 51). The more the research becomes particular and concrete, the more the ethical considerations tend to exacerbate (Cohen et al., 2007, p. 51). It was tried throughout the conduct of this study to give a rigorous scrutiny to ethical issues especially those appertaining to interviews. It is also worth pointing out that this research project was informed by British Educational Research Association’s (BERA) Ethical Guidelines for Educational Research (BERA, 2011). Some of the epitomes of ethical considerations germane to the present research are outlined as follows:

4.8.1. Informed consent and withdrawal right

Informed consent is perhaps the most crucial dimension of ethical considerations that needs to be taken into consideration by researchers. There has been an exceeding tendency among researchers to get the participants to sign the informed consent forms (Bryman, 2008, p. 122). According to article 15 of BERA’s guidelines, participants should be aware of their right of withdrawal at any time for any or even no reason (2011, p. 6). Creswell introduces six elements that are required for a consent form, namely the participants’ voluntary right to withdraw, the main purpose of the research, protection of participants’ confidentiality, the known risks of participation, the expected benefits of participation and the signature of both participants and the researcher (2013, 153). Drawing upon Diener and Crandall’s definition of informed consent, Cohen, et al., (2007)
introduce four elements involved, viz. “competence, voluntarism, full information and comprehension” (p. 52). As Cohen et al. (2007) point out, while in some contexts it is of high stringency, informed consent might be approached less tightly in other contexts (p. 52). Participants’ right to self-determination is the underpinning of the principle of informed consent (Cohen, et al., 2007, p. 52). Further to Cohen et al.’s notion that the stringency of informed consent could be varied across different cultures, it is worth re-emphasising that the present study is highly informed by BERA’s *Ethical Guidelines for Educational Research* (BERA, 2011, p. 5) in which ‘Voluntarily Informed Consent’ is the first responsibility to participants (p. 5, articles 10, 11, 12, 13). Therefore, particular stress has been laid on obtaining the participants’ consent. To do so, an informed consent form (appendix 4) developed by Graduate School of Education (GSE) at the university of Exeter was adopted for this study. Having endeavoured to make the participants’ apprised of their rights during the conduct of the research, the participants in both stages of data collection were fully briefed on the purposes of study and informed of their rights as respondents and participants. Having been informed of the purposes of the study and their roles in this research project, all participants including questionnaire respondents and interviewees were asked to sign the informed consent form thereupon.

4.8.2. **Institutional ethical approval**

The ethical code of ethical approval was also obtained prior to the conduct of research (appendix 7). it is worth highlighting that the present research project was rigorously in line with ‘Ethics Committee Guidelines’ of Graduate School of Education at the University of Exeter (“Ethics Committee Guidelines,”) and was granted the Certificate of Ethical Research Approval prior to the commencement of the fieldwork. Obtaining the university’s ethical approval, indeed, testifies to my particular attention to the sensitive ethical issues pertinent to the research.

4.8.3. **Privacy, confidentiality, anonymity, data storage and disclosure**

The issues of confidentiality and anonymity are considered as the norm in conducting a research (BERA, 2011, p. 5). As Bryman (2008) puts it forth, privacy
is invariably interconnected with confidentiality and anonymity (p. 124). Given a particular emphasis placed by the code of ethics on protecting the identities of both participants and the locations (Christians, 2005, p. 145) and in line with BERA’s Ethical Guidelines for Educational Research (2011, articles 25 to 31), this study sensitively observed the participants’ privacy, confidentiality and anonymity. Anonymity and confidentiality are two important issues highlighted by most ethical codes for which researchers tend to use pseudonyms (Ogden, 2008, p. 692). In order to protect their identity, all participants were given pseudonyms. Pointing out the imperfectability of pseudonyms as they tend to be identified by insider researchers, Christians (2005, p. 145) reminds us of the impossibility of ‘watertight confidentiality’. As the present study is a PhD thesis project, there was no liability on the part of the researcher to disclose the information to any third parties. Hence, I assured the participants of their secured privacy and confidentiality. The participants were insured that their data will be securely stored and will not be disclosed to any third parties unless their written permission is obtained.

4.8.4. Accuracy

Christians (2005, p. 145) reminds us of the centrality of ‘Accuracy’ to research. In order to ensure this important ethical code, any fabrications and use of fraudulent data must be strictly prohibited. Similar to other ethical codes, this dimension was also taken rigorously into consideration in the present study. Similarly, special attention was directed towards presenting the research objectives to participants as what they really are with no covert method for collecting the data.

4.9. Limitations and Challenges of the study

It seems rather unlikely for a researcher not to confront any limitations and challenges during his/her research conduct. These potential challenges mostly are of practical nature. Many of these challenges and limitations could be anticipated by the researcher prior to the conduct of study. By contrast, some unforeseen challenges could also happen throughout data collection process some of which might be beyond researchers’ control. Therefore, it is the researcher’s responsibility to minimise the likelihood of such challenges as much as possible.
One way to deal with this is to predict and recognize the possible scenarios apropos of data collection, from the early steps of entry and access to any potential ethical considerations that may arise during the research (Creswell, 2013, p. 171).

The challenges I experienced in this study were of two types, i.e. methodological and topic-related limitations. Although every endeavor was made to predict the ethical challenges ahead of the research in advance and prior to the fieldwork, some unforeseen challenges happened during the project which will be discussed shortly. It is worth noticing that some of these problematic aspects appertaining to the construction and administration of questionnaire and interview are inevitably beyond the researcher’s control. Fortunately, the response rate of the questionnaires in this study was relatively high in that more than 80 percent of the participants completed and returned the questionnaire, although some of them left the open-ended sections unanswered. This is an acceptable rate, given Gillham’s notion that over 50 percent of response rate is an acceptable range (2000, p. 9).

One of the reasons for such a high rate can be due to one-to-one administration of the questionnaires in which rapport and trust were built with each and every participant whereby they were assured about their confidentiality and privacy. Moreover, the participants’ willingness to respond to the questionnaires might have its roots in what Gillham calls ‘real importance’ of the topic of the questionnaire such as ‘job organization, status and salary’ (p. 10). This is of high significance since ‘the market is questionnaire saturated’ (Gillham, 2000, p. 10).

Timing and preparatory activities preceding the fieldwork were also among the challenges I faced during data collection. Since my data collection coincided with final examinations at universities in Iran, it was difficult for me to find EFL lecturers thereby requesting for appointments for interviews. Many of them were either busy with their students’ final exams and the pertinent paperwork or had planned for taking some days off. The main reason for this was my little awareness of the importance of building rapports with participants some weeks or even months before the fieldwork. In other words, I started from the scratch with no prerequisite contact with the interviewees whereby I could have made at least some tentative arrangements.
The second type of challenges and limitations I confronted, as stated earlier, pertain to those inherent in the subject under investigation in this study. This group of challenges, I surmise, mostly emanate from the focus and the context of this study which per se are challenging to tackle especially when it comes to ‘teacher appraisal’. Actually, it came to my attention that some of the participants were more or less reluctant to criticise the existing appraisal model in details albeit their little satisfaction with the evaluation system. Therefore, I had to utilise different strategies to establish mutual trust through which I could avoid their guarded and cautious attitudes towards freely criticising the existing model. The next major challenge I realised during data collection was the areas that were selected to investigate in this study simultaneously, i.e., teacher effectiveness and teacher evaluation (appraisal). Although these two areas are inherently highly interrelated with one another, each of these aspects per se, I suggest, deserves to be investigated as a separate PhD project.

4.10. Summary

The philosophical, theoretical and methodological assumptions adopted for this study were addressed in this chapter. The chapter illustrated the philosophical underpinnings and the theoretical and conceptual frameworks by which this research project is informed. The nexus between research, theory and practice along with the socio-political context of the research were also discussed. The methodological assumptions which in turn fed into the methods utilised for data collection were justified afterwards. The chapter then introduced the design of the study referring to its constituents, viz. participant selection, instruments and procedure. It has also been argued how an exploratory mixed method design informed by pragmatism within an interpretive framework could best fulfill this research thereby better answering the research questions. Additionally, a detailed account of different processes involved in different stages of data collection and data analysis was presented. The chapter thereafter proceeded to ethical considerations pertinent to this research project. The limitations and challenges occurred throughout the conduct of this study were explicated and discussed as the final section of this chapter.
Chapter 5: Data Analysis & Findings
5.1. Introduction

This chapter reports on the data analysis and the findings of this study. Throughout the chapter, the major themes, categories and subcategories are presented and subsequently supported by some evidence extracted from interviews and/or questionnaires. Since the two instruments, i.e. interview and questionnaire, were formulated and developed thematically in accordance to research questions, it provided me with a precious opportunity to get to know how different lines of thought and ideas generated during data analysis from each of these constructs interact with one another. More interestingly, it seems of high significance to identify and pick up any similarities and/or differences among participants based on their responses to each instrument. The realisation of such interrelationships was found to be of three different orientations. First, the emerged ideas and categories were positively supported by both datasets, i.e. interviews and questionnaires. Second, the ideas and categories were supported by only one source, i.e., either interviews or questionnaires. Third, the ideas and categories were negatively or contradictorily supported by the two instruments. For instance, whereas a category received strong support in the interviews, it was either unnoticed or given scant attention in the questionnaire. Nevertheless, the underpinning reasons for such confirmatory or contradictory relationships will be justified and discussed in the next chapter of discussion.

As its starting platform, this chapter starts with the core concept of teacher effectiveness in higher education whereby the major themes and their pertinent categories and subcategories will be presented. Lecturers’ perceptions of teacher effectiveness and the qualities of an effective teacher will be presented as the first theme of the study. I will explain shortly why these two initially distinct themes were eventually merged. The chapter will then proceed to the ‘measures of evaluation’ and will focus on different sources of information based upon which teacher evaluation can be implemented. In this section, different stakeholders who are perceived as having a stake in the phenomenon will be addressed. ‘Different ways to improve and accelerate teacher effectiveness’ will be the next theme to be reported. I will then have a look at another important aspect of this research
project, i.e. the ‘Iranian appraisal model and evaluation system’. Lecturers’ understanding of the current evaluation system in the Iranian universities and higher education institutions through which lecturers tend to be assessed each semester or academic year along with their perceptions of the strengths and weaknesses associated with such evaluation will be discussed afterwards. I will then argue that deciding on one’s teaching effectiveness can be easily misconstrued unless the influence of some non-teacher-controlled intervening factors are considered in the evaluation process. Indeed, teachers’ effectiveness can be potentially vulnerable to some internal and external variables such as the effectiveness of curriculum which are rather beyond teachers’ control. It will be then explicated how teachers, as one of the main stakeholders, operate within a wider overarching domain of educational effectiveness. Finally, the chapter will end with the last theme of the study, i.e. lecturers’ ideal/critical appraisal model. I will examine some qualities which are suggested by participants to be incorporated into the appraisal system so that teachers consider it as a fair realisation of their potential and actual capabilities. The evidence used for supporting these ideas (themes, categories and subcategories) are of two forms: the verbatim quotes from interviews and/or descriptive statistics from quantitative dataset, i.e. questionnaire. Juxtaposing and integrating qualitative findings with those of quantitative data, helps us better understand the depth and breadth of the ideas.

5.2. Six Major Themes

Having considered the research questions set forth in this study and given the findings garnered from the data, seven major themes were initially developed which were then reduced to six major themes each of which embraces a number of categories and subcategories. As illustrated in Figure 5.1., the themes under the overarching domain of EFL lecturers’ effectiveness include ‘qualities of an effective lecturer, measures of evaluation, ways to improve teaching effectiveness, the Iranian appraisal system, the non-teacher-controlled factors affecting teacher effectiveness and lecturers’ idea/critical appraisal mode’. 
Figure 5.1. Major themes generated for teacher effectiveness in HE

Since themes 1 and 2 were highly interrelated, they were merged into one major theme thereby eliminating their entailed repetitive ideas. For this purpose, I categorised the whole data driven from the participants into six major themes each of which embraced several categories and a number of subcategories which will be presented in the following sections shortly. Nonetheless, in order to have a picture of the data, a brief overview of the number of sources and references used for the analysis of the major themes generated from qualitative data, i.e. interviews and open-ended questionnaires are presented in Table 5.1., as follows:
<table>
<thead>
<tr>
<th>Theme</th>
<th>Sources</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. General perception of TE</td>
<td>19</td>
<td>44</td>
</tr>
<tr>
<td>2. Qualities of an effective teacher</td>
<td>36</td>
<td>260</td>
</tr>
<tr>
<td>3. Measures of Evaluation</td>
<td>14</td>
<td>80</td>
</tr>
<tr>
<td>4. Ways to increase TE</td>
<td>21</td>
<td>65</td>
</tr>
<tr>
<td>5. The Iranian appraisal system</td>
<td>27</td>
<td>63</td>
</tr>
<tr>
<td>6. Non-teacher-controlled factors affecting TE</td>
<td>24</td>
<td>62</td>
</tr>
<tr>
<td>7. Lecturers' ideal/critical appraisal model</td>
<td>12</td>
<td>40</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>153</strong></td>
<td><strong>628</strong></td>
</tr>
</tbody>
</table>

Table 5.1. Frequency of Sources and References (Themes, Categories and Sub-categories)

* Source refers to research materials, i.e. interviews and open-ended questionnaires

5.3. General perceptions of TE & Qualities of an effective teacher

This section attempts to answer the question of TESOL lecturers’ understanding and perceptions of language teacher effectiveness in the Iranian higher education context which serves as a platform for other research questions. Both interview and questionnaire were used to address and investigate this question. Nonetheless, having worked on the analysis of data, I soon came to the conclusion that this question highly overlapped with the next question of the qualities and characteristics of an effective teacher. Therefore, I decided to merge these two questions so that a better and more in-depth understanding of the concept of teacher effectiveness and effective teaching could be explored. The analysis of the two sets of data, i.e. qualitative and quantitative data obtained from interviews and questionnaires led to the emergence of some sub-themes and categories which are depicted in the Figure 5.2., as follows:
This section tries to answer the question of teachers’ perceptions of the qualities and characteristics of an effective teacher. In fact, the notions of teacher effectiveness and effective teaching tend to be slightly illusive and vague as the concepts are not consistent in the literature. Moreover, as it was mentioned in the previous chapter, we need to address the question of “effective in terms of what” in the first place. Indeed, the analysis of these two research questions gave birth to a myriad of ideas some of which were in line with those of the literature and some other which were relatively new and genuine in the field. The responses offered a wide range of qualities and characteristics for effective teaching. The generated ideas were divided into five categories including Personal, Cognitive, Pedagogical, and Professional qualities as indicated in Figure 5.2. In the following, each of these classifications will be elaborated on along with the associated constituents.

5.3.1. Personal qualities

There is a wealth of literature on the importance of teachers’ personal characteristics (e.g. Graham, 1999; Meijer, Korthagen, & Vasalos, 2009; Stronge et
al., 2011). Stronge, Tucker, and Hindman (2004, pp. 30-31), propose six indicators for the quality of teacher as a person including ‘caring, fairness and respect, attitude towards the teaching profession, social interactions with students, promotion of enthusiasm and motivation for learning and reflective practice’. Having been asked to express their ideas and feelings towards the qualities of an effective teacher, the participants referred to a number of characteristics which I classified under the rubric ‘personal’ qualities as illustrated in Figure 5.3.

![Figure 5.3. Theme I. Category I (Personal qualities)](image)

Although most of the categories were generated from the interviews, the analysis of quantitative data also provided strong support for the importance of personal traits in teacher effectiveness. Prior to reporting the subcategories, it is worth highlighting that personal qualities, in a general sense, were perceived by the participants as a key element in their teaching effectiveness. As the results suggest, the majority of the respondents (95.4%) agreed to a statement whereby teachers' personal traits were considered generally influential in their effectiveness.
Nonetheless, it is worth noticing that personal traits seem to be a matter of judgment moving along a continuum. In other words, one can hardly find a teacher lacking the entirety of, for instance, friendly personality. Whereas some teachers may show high level of friendliness, others may express such a human intrinsic behaviour at a lower level.

5.3.1.1. Friendly and Approachable

As Stronge (2007, p. 116) contends, teacher’s personality is one of the first qualities to consider for an effective teacher. Teachers’ personality and the way they behave in the classrooms and treat students were found to be of paramount significance. Personal qualities were remarkably mentioned by almost all participants. Amongst others, qualities such as friendliness, honesty, sense of humour and interpersonal skills were the characteristics the participants assigned to an effective teacher. For instance, Parham highlighted the role of friendly personality as he believed:

“He should be approachable; students can access him very easily, be friendly with the students, constructs a friendly atmosphere, a stress-free atmosphere in the class in order to motivate them, encourage them to study, be honest to students, and treat them as his brothers and sisters, that’s it”.

In the same vein, Rima also accentuated the relationship between teacher and student in a friendly environment as indicated in the following quote:

“and also should have a very good relationship with the students and also to try to provide a friendship or friendly atmosphere in the class and never focus just on some kind of students, don’t pay attention to criteria that are not important such as for example race or ages or gender”.

The qualitative findings were confirmed by the statistical results obtained from the questionnaire in that 93.1% of the lecturers were in favour of the contributions of teachers’ friendly personality to their effectiveness. Such a remarkable tendency towards teacher-student relationship may have its roots in the cultural underpinnings.
5.3.1.2. Informative & Dedicated

Connected with the previous sub-category, the qualities of being informative and having a sense of dedication to students were among participants’ proposed characteristics for effective teachers. For instance, Amir placed emphasis on such characteristics as clarified in the following extract:

“An effective teacher should have, uh Ok, so I think the first point is here, teacher dedication. If the teacher dedicates himself to his or her students, I think yes I think he will be an effective teacher, here is the point”.

Armin also believed that an effective teacher is expected to be as informative as possible as indicated in the following quote:

“So I think for them … they expect a teacher to be as informative as possible”.

Although the concept of ‘dedication’ seems to be problematic from which different interpretations could be drawn, it convinced 81.4% of the questionnaire respondents of the importance of such a trait as an element of effectiveness. Stronge also considers the trait of being dedicated as a quality for ‘caring teachers of at-risk students’ (2007, p. 32). Therefore, these traits seem to be of central importance in teaching students in need of more support.

5.3.1.3. Respectfulness

As Stronge et al. (2004) contend, respect and fairness require teachers to behave in a balanced and open-minded way in all circumstances (p. 30). Highly intertwined with friendly personality, qualities such as respectfulness and fairness were postulated by interviewees as prerequisites for effective teaching. Niloofar, one of the interviewees, drew attention to what she considers as ‘principles of morality’ as follows:

“… in teaching different subjects, so many things may interact. For example, I would say that the society, the people to whom I teach the receivers, in my comparison the receivers are also important but the minimum level of expectation is that teachers must consider principles of morality as a human being in the class special a unique case to apply power is given to them and they must always remind themselves not to misuse or abuse that special position”.

Ali, in a similar vein, highlighted the role of ‘respect’ and said:
“... you have to be somebody who students respect”.

Parham also maintained that

“... the second one is um considering the student's rights and also their needs”.

Respectfulness was also firmly confirmed by statistical results as 97.7% of the respondents declared for it. It is worth noting that, respect as a personal trait tends to be bound up with cultural norms of the society to which special attention needs to be paid thereupon. In other words, it is unwise to expect lecturers’ understanding of respect in Iran similar to that of lecturers teaching in the UK. For instance, whereas calling a teacher in his/her first name is plain commonsense in the UK, it tends to be interpreted as an insolent behaviour in Iran.

5.3.1.4. Inspirational and Self-motivated

This category received strong support from the data as most participants referred to the importance of motivation to effective teaching. Motivation tends to be a core element of success in almost every career including education. Motivation and enthusiasm play an important role in teaching and learning effectively. However, not only do teachers need to give encouragement and try to raise students’ engagement in the classroom activities, they themselves are deemed to be self-motivated. In other words, one can hardly be expected to be inspirational whilst lacking enthusiasm about his or her own teaching career. Effective teachers tend to be possessed of motivation, enthusiasm and confidence in their practice. Teachers’ ability to instill motivation and engendering incentives among learners for learning is among other personal characteristics introduced by the participants. For instance Thelma stated:

“For first of all, a person who is inspirational, a person who can motivate students, never gets tried ... a person who actually wants the learners to learn something, that gives the thing that the learners want, not the thing that he or she thinks is true.”

Nonetheless, as proposed by some other participants, affective teachers not only should inspirationally instill enthusiasm and encouragement into their students but also they need to be self-motivated and enthusiastic enough about their own teaching career. Whereas teachers nowadays are deemed to accept new roles,
some of the participants still had the so-called traditional stance towards their role, i.e. teacher as a model. For instance, Sepehr capitalised on this expecting his students to follow him as a model:

“A good teacher should … a good teacher is the person who can have some kind of influence on students regarding encouragement, positive attitude and something like a model. A good teacher can be a model and should function as a model for the students to follow him”.

Students’ engagement in classroom activities as pedagogical quality is highly imbued with the personal quality of being inspirational. Being self-motivated and inspirational, a teacher should be enthusiastic about raising his/her students' motivation, giving them encouragement and engaging them in classroom activities. Effective teachers make every endeavour to engage their students in classroom activities. However, Referring to the prominence of student encouragement, Niloofar cautioned against an unfair approach towards teachers’ involvement in such activities:

“They must always believe that motivation in the greatest way and if they put barrier on this motivation the results may not be so satisfactory and I want to say that some are too idealists, some want to make it quite a miracle and some are quite bored and exhausted with years of struggle, something in between something in average”.

As to the previous personal trait, motivation seems to be a matter of degree and hence cannot be easily gauged.

5.3.1.5. Confidence

Teachers’ sense of confidence and empowerment was also referred to by interviewees as an empowering quality for their teaching practices. Thelma believed that students’ confidence tend to be engendered by that of their teachers as shown in the following extract:

“Self-confidence is very important. Until you don’t have the self-confidence you cannot make your students be confident. You have to be confident so the students can see you are confident they can learn from that”.

Maria adopted a similar stance and asserted that:

“Taking relevant measures to boost my confidence as a person and of course as a teacher teaching subjects that I am interested in”.
Teachers’ self-confidence seems to be intertwined with other qualities such as their personality and self-esteem. Ali views what he calls teachers’ *strong* personality as the grounds for students’ reliance on their teachers:

“The effective teacher has got a strong personality, so strong to have deep effect on students’ lives. Students should accept him as a reliable person.”

However, it is worth highlighting that confidence can be potentially influenced by intra-classroom factors. It seems that participants’ demand for confidence was due to insufficient training they had received which can be partly redressed through initial and then in-service teacher education programmes.

### 5.3.1.6. Patience

Patience is a quality which can be considered as both personal and professional characteristic. Many language teachers as human beings tend to have different patience threshold. Patience is perceived as a crucial element in effective teaching especially for *language* teachers who teach in a language other than students’ mother tongue imbued with its peculiar intricacies. There is a wealth of literature expatiated upon the quality of ‘Role of caring’ an element of which tends to be teacher’s patience (Stronge et al., 2004, p. 32). As a personal characteristic, patience received support from participants. The following quote from Majid indicates how he perceives patience as follows:

“The teacher should be open to him all her (student) when anything happens and to have some have good ears, to be patient enough”.

Mersedeh adopted a similar stance and stated:

“S/he must be very patient and stimulate and encourage the students to have interaction with each other.”

The statistical results also gave full support as one of the statements of the questionnaire which specifically inquired into respondents’ views on patience was approved by the majority of the respondents (95.4%).

### 5.3.1.7. Openness & Adaptability

Teachers’ openness to *other* stakeholders including students could be a pathway for better student achievement. There are several lines of thought pertinent to the
quality of openness. Depriving oneself from others’ (students, colleagues, etc.) suggestions and experiences found to be highly deprecated by the interviewees. Teachers need to be open to their colleagues, students and all other qualified others and consider their contributions as a developmental pathway into which they need to step in. In response to an open-ended question, Shahab stressed teachers’ learning:

“Engendering within himself/herself a strong desire to learn from the learners.”

Sarah, similarly, added weight to teachers’ openness and willingness to welcome new ideas and suggestions:

“… always should be open to students’ ideas and suggestions, and try even to use other experiences, other colleagues’ experiences”.

Referring to the same issue, Amir also emphasized that a teacher needs to listen to the voice of other stakeholders:

“Teaching as you know is not just a science; it can be an art too. So the teacher needs to have some traits, some characteristics that are innate in him or her..... I think for example the teacher should be open, should have a very good talent in listening to others, to understand what the other the other person ....”.

The statistical findings also corroborated the idea in that 86.1% of respondents were in support of teachers’ openness. Indeed, the concept of openness here connotes teachers’ positive attitude towards a wide range of issues including openness to the feedback and criticism they receive from others. Such feedback may target teachers’ personal behaviour, pedagogical practice or professional characteristics and hence tends to function at a more holistic level which will be discussed in details in the next chapter of discussion.

5.3.2. Cognitive qualities

Further to qualities classified under the rubric ‘personal’ traits, a number of other qualities and characteristics were also highlighted by the participants, which can be construed as ‘cognitive’ qualities. This category includes language proficiency, subject knowledge, curriculum knowledge and ICT literary as indicated in Figure
It is worth mentioning that cognitive qualities discussed here refer to different forms of *teacher knowledge*. In other words, from participants' perspective, teacher knowledge seemingly embraces a broad spectrum of knowledge. In fact, the literature also resonates with such an understanding of teacher knowledge. In his categorisation of knowledge base for teachers, (Shulman, 1987) proposes seven types of knowledge including:

- "content knowledge"
- *general pedagogical knowledge*
- *curriculum knowledge*
- *pedagogical content knowledge*
- *knowledge of learners and their characteristics*
- *knowledge of educational contexts*
- *knowledge of educational ends, purposes, and values, and their philosophical grounds*” (p. 8)

Notwithstanding the participants more or less referred to most of generic attributes of teacher knowledge, they gave full support to teachers’ ‘Language Proficiency’, inasmuch as they considered English language teaching. The four major aspects of teachers’ knowledge are presented in the following:

![Figure 5.4. Theme I. Category II (Cognitive qualities)](image)

**5.3.2.1. Language proficiency**

Teachers’ General English knowledge or language proficiency was among the cognitive attributes the participants were strongly in favour of. An example of
participants’ insistence on teacher’s English language competence can be discerned in the following quote from Niloofar:

“full mastery I can say in four skills, well, if you want to, from the macro level in this major students of English, they are expected to gain I mean they are expected to be fluent in four skills from the macro level of course it is divided into different categories based on different subjects. We have got different subjects students learn but in all those subjects no matter whether it is oral production, short stories, translation or grammar, fluency in four skills the teacher must have as a role model for students, the one who want to observe, the one who can detect the problem, must have full fluency”.

In the same vein, Rima and Ali maintained that EFL teachers need to be possessed of knowledge of language:

“And a good teacher a proficient teacher should have the ability in order to express him or herself to the students and make them understand for example it means to have a good way of teaching or to make the students understand about the subjects that mentioned in the class”. (Rima)

“He/she should have good command of English (especially good command of spoken English)”. (Ali)

Findings from the survey also corroborated the ideas generated from interviews as 83.8% of respondents were opposed to the idea that language proficiency is not beneficial to their effectiveness. A review of the literature also gives clear evidence to teachers’ English language knowledge. As Richards (1998, p. 7) puts it forward, non-native speaker language teachers (NNS) need to have certain threshold of language proficiency so that they can teach effectively. He further draws attention to a two-facet concern germane to teacher’s language proficiency, i.e. the most needed component of language proficiency and the dynamics of the interaction between language proficiency and other dimensions of teaching skills (p. 7). The former has been addressed by some of the participants. As stated in the above quotes from interviews 14 and 1, Niloofar and Ali perceived fluency and verbal abilities as important aspects of language proficiency. However, entangled with teachers’ confidence, language proficiency tends to be an intricate issue which can impose a lot of stress on teachers especially the novice ones in Iran. Teachers’ challenges for fluency in Iran seem to emanate from insufficient training they received in their BA and MA programmes. This is of an utmost importance in Iran as an EFL context. These areas will be more discussed in the next chapter.
5.3.2.2. **Subject Matter knowledge**

Subject matter knowledge has been extensively investigated in the literature. According to Richards (1998), it relates to ‘what second language teachers need to know about their subject’ (p. 8). It refers to teachers’ knowledge about what they teach rather than about teaching per se which is distinct from that of teachers from other disciplines (p. 9). Teachers’ subject matter knowledge has also been picked up by some of the participants in this study. Not only have they mentioned teachers’ knowledge about the course they teach, but also they referred more specifically to other aspects of knowledge such as TEFL-oriented theoretical knowledge as well as knowledge of curriculum development which, as they believe, can help teachers teach more effectively. As suggested by most interviewees, teachers should have the minimum knowledge base for a wide range of areas interacting with their teaching practices. For instance, Ali contended that:

> “and the third one can be he should be a knowledgeable person in his own field ok as you put it correctly subject matter should be a knowledgeable person in that field to for example to be able to answer the students’ questions and deep knowledge of his or her subject matter because you know that if the students want to ask you some other questions which are not related directly to the books that your teaching so if you don't have the deep knowledge of …”.

Knowledge of subject matter and language proficiency are indeed two different but complementary domains of teachers’ knowledge without each of which teachers’ effectiveness is very unlikely. However, the quantitative findings revealed that 67.4% of respondents regarded subject knowledge as a central factor in effective teaching. Although two-thirds of the respondents concurred with the idea, this simply indicates that, from Iranian lecturers’ viewpoints, knowledge of subject matter cannot guarantee their teaching effectiveness. With regard to the Iranian context, the only way for making decisions on lecturers’ subject knowledge rests upon academic departments for which there is no clear mechanism.

5.3.2.3. **Knowledge of curriculum & syllabus**

As mentioned earlier, an effective teacher is deemed to have an acceptable knowledge base of both subject matter and language general proficiency. Nevertheless, effective teachers should have such a knowledge base in other
areas such as curriculum. As professionals, teachers seem to be required to have the minimum knowledge of curriculum development which is needed for improving their teaching practices. Curriculum knowledge entails how and where the subject sits in the whole curriculum spectrum as well as how it is assessed throughout various stages in education (McGregor, 2011, p. 10). The history of curriculum development is imbued with the notion of syllabus design even though the current understanding of curriculum development dates back to the 1990s (Richards, 2001, p. 14). The analysis of quantitative data also testified to the importance of teachers’ knowledge of curriculum and syllabus. Indeed, 88.4% of the respondents considered teachers’ knowledge of curriculum development, syllabus design and lesson plan as a necessity for effective teaching.

The following extract is from Sepehr who referred to teachers’ abilities to prepare and develop teaching materials:

“Try to design and prepare teaching materials according to any given teaching context, considering students’ cognitive, social, and emotional characteristics.”

Parham adopted a similar position and stated:

“... the curriculum and syllabus of the course is very important. Because it guides the lecturers, what is expected them to do so what the outcome of the course is, they will be informed.”

Ali refers to not only teachers’ knowledge of material development but also the power they need:

“... as a teacher they should know at least something about material development. But when they have the power to change the curriculum ...”

Interlinked to the previous aspects of teachers’ knowledge, it seems that teacher have not yet obtained what they really need in their classrooms neither in their pre-service education nor during their in-service trainings. A rich teacher development programme can help teachers overcome such challenges. These areas will be more explicated in the forthcoming chapter.
5.3.2.4. **ICT Literacy**

Use of technology enhanced learning (TEL) in teaching was one of the findings of this study. By the outset of the 21st century, EFL teachers similar to those of mainstream education have stepped into a new pathway in which Information Technology (IT) became a turning point in language learning and teaching. Suffice it to say that this movement has been hastened by policymakers as part of their educational policy towards quality assurance. The pace of the technology movement in education has been to such an extent that Information Communication Technology (ITC) is becoming an index of high quality teaching. In recent years, the notions of ‘e-Learning’ and ‘Virtual Learning Environment’ (VLE) are becoming exceedingly a commonplace in most higher education contexts. Pritchard introduces a number of features for the use of ICT in educational setting including ‘Speed, Capacity, Automations, Communicability’, etc. (See Pritchard, 2007, p. 19). The findings indicated a rather surprising contradiction between the two datasets. Armin was the only interviewee who explicitly emphasised the significance of technology literacy:

“and I think you know just the ways asking and ask questions or something like this using the facilities ok can be very useful because you know she should be informed on the new technologies and try to bring them into the class I think that’s very important and yeah maybe criticize not just use the computers you know the new like Oovoo we are using right now, the new technology literacy I think is very important in the class, to make to go to the what the learners are dealing with.”

However, Armin’s idea was firmly confirmed by the analysis of the survey which indicated that 90.7% of the respondents place emphasis on the contributions of technology in education. In recent years there has been an exceeding tendency towards the use of technology for educational purposes in Iran which has been successful to some extents in some universities. However, it is worth highlighting that such literacy tends to revolve around two axes. Whereas the physical realisation of technological literacy has been highly improved in recent years in Iran, teachers’ mentality towards the use of technology in their classroom is yet under question.
5.3.3. Pedagogical skills

The next category which embraces a number of qualities for an effective teacher is what has been named hereinafter as ‘pedagogical skills’. Given the diverse nature of the pedagogical skills, one can find different elements under the overarching concept of pedagogical skills. For instance, from McGregor’s (2011, p. 10) point of view, pedagogical knowledge is a multifaceted issue including what is seen as ‘practical teaching knowledge’, ‘beliefs about teaching’ and ‘understanding of learners’. Since English language lecturers in Iran include those who major in both TEFL and non-TEFL disciplines such as English Language Literature and Translation or Linguistics, pedagogical skills tend to act more influential. In other words, the non-TEFL English language teachers are at high risk of treating students as 'slow learners' (Carder, 2007, p. 383) which emanates from their relatively little knowledge of pedagogical skills as compared to that of their colleagues who major in TEFL, TESOL, Applied Linguistics and pertinent areas. Novice teachers’ awareness of ‘pedagogical resources’ helps them teach concepts and skills to students in a more appropriate way (Johnson, 2009, p. 18). The aspects of pedagogical skills that have been identified in this research are shown in Figure 5.5.

![Figure 5.5. Theme I. Category III. (Pedagogical skills)](image)
Each of these subcategories will be discussed in the upcoming sections. Nevertheless, a quick glance over the respondents’ responses to the questionnaire testifies to lecturers’ preference for high quality pedagogical skills as an element of teacher effectiveness. As the results suggested, 88.4% of the participants in the survey advocated the centrality of pedagogical skills, in a general sense, to teacher effectiveness.

5.3.3.1. **Instructional planning and alternative instruction**

Instructional planning here refers to teachers’ plans and activities needed to fulfil their educational objectives. However, the concept has turned out to echo other similar labels and terminologies. While ‘didactics’ and ‘teaching methods’ are well-known in European contexts and English-speaking contexts, respectively, ‘instructional design (ID)’ has seemingly gained global popularity during last 50 years (Seel & Dijkstra, 2004, p. ix). As Rima put it forward:

> “A good teacher or as you mentioned an effective teacher, for example, at first should have a some kind of criteria and maxims, the same as Kumara [Kumaravadivelu] mentions, some kind of determined maxims for him or herself, ok, and also should have a plan before starting the class or the course, should have a very predetermined syllabus”.

Soroush also made the point referring to teacher knowledge of language teaching methodology:

> “Erudite: knowledgeable both in General English and methodology”.

Benefiting from proficient students was introduced by Majid as one of his instructional strategies as indicted in the following quote:

> “Sometimes I ask my students who are proficient enough or they are more proficient, ask them to change their seats, help their friends, what I just did it today.”

Although Rima and Soroush somehow touched upon Instructional planning, I was quite surprised that nobody else raised the issue in interviews. I wouldn’t be surprised if most interviewees were not acutely aware of the importance of instructional planning. This may link in with my earlier point about the quality of teachers’ in-service training.
There is a general consensus that language teaching has gone through an evolutionary progress throughout its life, from the so-called pre-method to method and more recently to post-method era. As Kumaravadivelu (2006) puts it forward, by the outset of the third millennium, teachers have been inevitably facing both new challenges and opportunities to go beyond the confines of language teaching ‘methods’ (p. 161). In the following quote from Sarah, she depicted how she believed that use of different sources and teaching strategies can better accommodate students’ needs:

“Sometimes I have to study more, and yes work, for example, in one semester I just introduced George Yule for Linguistics. In another semester, I tried to use two sources, from Rodman and Julia Falk for Linguistics. I tried to use the harder books and more than one book. And sometimes I have to change my way of teaching. For example, I added some lectures to this class that the students should have as lecture in the class”.

She reiterated her ideas in the open-ended questionnaire:

“…[Effective teachers] Have a lesson plan, be familiar with new sources.”

Similarly, Sepehr referred to the use of different teaching strategies as a quality of effective teachers:

“…[Effective teachers] Make use of a wide range of instructional strategies.”

Shahab also asserted:

“… Providing real-life examples from a diversity of contexts in his/her teaching.”

Focused and organised instruction was also referred to in the interviews. Being professionally organised seems to be highly interlinked with being personally organised. As Stronge et al. (2004) maintain, the way a classroom is organised tends to exert influence on the behaviour in it (p. 68). Hence, it seems that teachers’ organised behaviour can dictate students’ behaviour which in turn may instill such an approach into students. As shown in the following quotes, Ali and Kasra touched on the point as follows:

“Keep everything in order. The students should be, what do you call, should be aware of the way in the way of your teaching, I mean, from where you want to start, which books they were supposed to study, what would happen..., a kind of syllabi … for that special course or syllabi for all the
courses they are having …. . They should know the map. They should know the time of, the exact time of the midterm. The final and the books they are going to be taught.”

“… providing clear and focused instruction.”

It is also worth referring to the lecturers’ responses to one of the statements of the questionnaire whereby 76.7% of the respondents agreed with the view that an effective EFL teacher should have a TEFL-driven understanding of teaching. This interestingly corroborates the above-mentioned ideas which are rather the mainstay of teaching English as a foreign language. Given the burgeoning literature on alternative instruction, I expected more participants to refer to this notion in the interviews. The possible reasons will be discussed in the next chapter.

5.3.3.2. Simplicity/Tailoring materials to students’ needs

Effective teaching does not mean to teach in an arrogant sophisticated style. Indeed, simplicity in presenting the material seems to be a key factor in making teaching more effective. Simplicity draws the attention (Mortiboys, 2010) and helps learners feel more at ease. Simplicity may happen at different stages, i.e. in preparing material, hand-outs, PowerPoint slides, presenting the session. Majid in his interview stated:

“One of the indicators that I usually try to employ is ‘simplicity’. I try to simplify myself as much as possible, making a kind of connection”.

Given the fact that different interpretations could be drawn from the notion of simplicity, it is worth illuminating that simplicity does not refer to the content of the subject being taught. Instead, it refers to how a teacher presents teaching materials to his/her students. Therefore, even complicated subjects/topics can be presented in a simple and easy to understand fashion. Simplicity appears to be an important element in pedagogy in that some teachers appear to be unconscious of the gap between a teacher and his/her students which, if not treated appropriately, can hinder students’ learning.

Nonetheless, simplicity needs to be combined with teachers’ ability to tailor the material to their students’ needs. There has been a large discrepancy upon the status of course books in language teaching. Such discrepancies, as Nation and
Macalister (2010, p. 159) posit, give evidence to the importance of a flexible approach and a course book which allows the teacher to adopt such a flexible stance. As it was mentioned earlier, one of the most important qualities through which teachers can enhance their effectiveness is the ‘knowledge of curriculum and syllabi’. Teachers’ awareness of material development skills can help them tailor the curriculum/syllabi to accommodate students’ needs which per se imply a flexible approach towards teaching. These aspects have been highlighted by Mersedeh as follows:

“I think the teacher must change the syllabus of the class sometimes. For example, the book that you’re teaching your students sometimes is very difficult and maybe you must change the book or you must try to change the lessons, they must be very digestive [easy to understand] for the student”.

Moreover, Souroush and Mersedeh referred to the use of L1 as a teaching technique which helps them in teaching more effectively. The extent to which first language (L1) tends to be used by teachers in foreign language classrooms is widely debated (Ellis, 2012, p. 128). Deciding on the weight that should be given to the use of L1 and keeping the balance between the target language (TL) and L1 is a contentious issue with which teachers tend to be confronted (Littlewood & Yu, 2011, p. 64). Regardless of its confirmatory or contradictory nature, the relationship between teacher beliefs about and their actual use of L1 in foreign language classroom have been widely addressed in the literature (e.g. Cook, 2001; Ellis, 2012; Littlewood & Yu, 2011; Moore, 2013). Soroush touched on this issue and said:

“I remember, it was in the beginning of my teaching that I was teaching, of course, an ESP course, I think ESP for ‘Accounting’, that I have to confess that I overestimated my students’ abilities and their, for example, familiarity with this lexical terms. I was trying to teach in English or, in other words, my language of instruction was English but I bumped into their confusion. They were adults they were confused and they couldn’t understand anything. I tried more to explain this kind of, for example, text. For example, if the text was about financial statements, I tried to explain it by similar terms and structures but I found out that it was not effective either so it was at that time that I planned to make some kinds of changes. So, in fact, in advance, I overestimated them but later I came to the conclusion that it was not the case and I planned to use, of course, Farsi in the classroom as a kind of instruction language”.

Mersedeh also posited:
“… sometimes I ask … to write the new words with Persian equivalent for the students and that the students can comprehend and understand it better…….. …. Sometimes, for example, it is necessary that you translate in Persian, you are teaching in English and you must translate in Persian or sometimes for example based on the topic, first of all you must talk about the topic in Persian and then start in English, it must be very digestive for [digested into] the student.”

The status of use of L1 in L2 classrooms in Iran is in harmony with that of other EFL contexts around the world. While the advocates view L1 as a panacea, opponents call for highly limited use of L1 some of whom tend to take their attitudes to extremes dismissing the idea.

**5.3.3.3. Needs analysis & student engagement**

Tailoring the materials to students’ needs requires, at its heart, students’ needs analysis. As part of pedagogical knowledge, an effective teacher should be keen on students’ individual needs, thereby acting as a so-called needs analyst. This is a prerequisite step for the further stage of student engagement as mentioned earlier. Rima asserted that students learn at a different pace for which teachers need to be prepared:

“**An effective teacher should have a right knowledge about what he wants to do. Besides … s/he should try to meet students’ needs and recognize that students learn at different rates.**”

In the same vein, Sarah called for teachers’ familiarity with students’ needs:

“**Teachers should be familiar with the students’ needs, you know, to know what they want and give that. … I think … the teacher can be effective in this way.**”

Interestingly, Ali moved beyond students’ learning needs and considered an effective teacher as the one who can predict students’ future needs in society:

“**The second one is considering the student’s rights and also their needs. What they need, as a, as an instructor, you can predict when they enter the business, when they enter or when they graduate from their studies, what they need or want their right is. So an effective teacher or a good teacher should predict what their needs are in that special courses and try to link or make a link between what there’s studying and what they are expected to do in the future. So a practical type of teaching.**”

Needs analysis is among the areas with which lecturers with non-TEFL/TESOL backgrounds such as Literature, Translation, Linguistics, etc., tend to struggle.
Thus, administrators need to backup those lecturers who are not academically familiar with language teaching and learning theories and practices.

As to teachers’ next step, they need to create opportunities for their students based on their needs to get involved in classroom activities. Student engagement as a pedagogical quality received widespread support from respondents. Indeed, the majority of the lecturers (97.7%) called for student engagement as an attribute of an effective teacher from whom 60.5% strongly agreed with the idea.

5.3.3.4. Attentive to affective filter

The notion of ‘affective filter’ in second language acquisition was introduced by Stephen Krashen in the 1980s. Indeed, the ‘affective filter hypothesis’ was among five major hypotheses building up Krashen’s (1989) theory of second language acquisition. Being caused by lack of motivation and self-esteem, increased students’ anxiety can lead to the affective filter which per se tends to function as a learning barrier (Krashen, 1989, p. 11). As the findings suggest, some of the participants believed that teachers should adopt a ‘humanistic’ approach towards their students and observe affective variables in their teaching setting. For instance, Sohrab contended that:

“It is related to many factors or, let’s say, it is influenced by many factors. For example, students satisfaction comes from both the relationship of the students, I mean, between the students and the teacher, the friendly atmosphere, the atmosphere in which there is I mean the affective filters are low, okay, their understanding of the lesson, okay, there are so many factors that come together to make the students satisfied. Clear? The second one is the notion of affective filter. I preferred to have a lively enjoyable class in which they they look at the English matter, English subject not is a let's say duty but as a let's say fun or something that they enjoy.”

Likewise, in response to the open-ended questionnaire, Shahrdad maintained that an effective teacher brings a sense of joy into classroom:

“Deriving joy and pleasure from learning language”

The idea was also validated by statistical findings from the survey as 88.4% of the lecturers confirmed their need for a stress-free environment in the classroom.
5.3.3.5. **Transferring knowledge**

Perhaps, the concept of ‘transferring knowledge’ as an attribute of an effective teacher is one of the most traditional conceptions of teachers’ responsibilities. The hidden dimension of such conception will be me more illuminated in the next chapter as it emanates from deep contextual underpinnings leading teachers to such a philosophical worldview about their career, i.e. transferring knowledge and hence should not be neglected. As Hiramatsu (2005) puts it forward, even competent teachers are prone to inhibitions due to what he calls as the fear of a traditional understanding of teaching as a process of transfer of knowledge (p. 132). Nevertheless, it seems that some of the participant lecturers believed that transferring knowledge is a prominent priority for an effective language teacher. For instance as Ali asserted:

“An effective teacher is the one who is able to transfer his knowledge to his students. I know some people who are very knowledgeable but don’t have the skill of transferring their knowledge to their students. So, one aspect of teaching effectiveness for me is ability to transfer your knowledge.”

Saman also drew a distinction between teachers’ knowledge and their ability to teach in that knowledgeable teachers are not necessarily effective teachers as they may not be able to transfer their knowledge to their students:

“Use your knowledge, you know the Knowledge to transfer it, you know. We have or there are a lot of teachers or instructors which who we consider knowledgeable but they cannot transfer what they have in their mind you know or they cannot...”

In a similar vein, Parham posited:

“An effective teacher is the one who is able to transfer his knowledge to his students. I know some people who are very knowledgeable but don’t have the skill of transferring their knowledge to their students. so one aspect of teaching effectiveness for me is ability to transfer your knowledge.”

Despite being rather traditional, it seems that such perceptions of teaching yet continue to be somehow a predominant understanding of teachers’ duties in Iran.

5.3.3.6. **Assessment**

Use of appropriate assessment strategies and, in particular, dynamic assessment was evident in lecturers’ understanding of effective teaching. The notions of
'scaffolding' and 'dynamic assessment' which were directly mentioned by interviewees are supported by a wealth of literature (e.g. Lantolf, 2009; Poehner, 2008). As a pedagogical asset, teachers' awareness and utilisation of assessment techniques which can fit course objectives have been picked up by some of the participants. For example, Sohrab stated that:

"Yeah, actually I have the opinion that the objectives should be based at some parts on the level of the students. If I didn't change the materials and my expectations, I know, at the end of the term they did not know anything... Okay... They just memorise something and verbalise something and but they didn't get anything. Okay... So this notion of scaffolding comes in mind. I try to help them come gradually step-by-step I mean forward" 

Rima and Soroush raised the same point whose quotes are as follows, respectively:

"According to Vygotsky theory (ZPD), a teacher should scaffold students’ learning effectively and promote effective interaction in the class".

"Well familiar with dynamic assessment".

It was interesting to see how Kian touched on the importance of assessment ‘for’ learning as shown in the following quote:

"It appears that assessment should not only be “of learning”, but also “for learning” for it to be effective."

The contributions of assessment for learning to students’ learning outcomes seem to be an important aspect of teachers' knowledge and skills (see, e.g. Black, Harrison, Lee, Marshal, & Wiliam, 2003; Black & Wiliam, 1998) .

As depicted in the above quotes from participants, teacher’s awareness of assessment theories alongside those of teaching and pedagogy will help them promote their effectiveness. Teachers’ awareness of assessment strategies was also supported with the statistical analysis, in that 93% of respondents concurred with the idea. It is worth highlighting that very few participants mentioned the strengths of dynamic assessment. Hence, I am afraid, the practical applications of dynamic assessment principles in universities in Iran are far less than I expected.
5.3.4. Professional skills

In the previous sections, a number of qualities of effective teachers, namely personal, cognitive, metacognitive and pedagogical qualities, have been presented and discussed. In this section, the findings on the last category, i.e. teachers’ ‘professional skills’ will be reported and interpreted. Given the fact that hardly can teachers gain what they need at pre-service stages and also due to the dynamic nature of language teachers' knowledge base, an on-going renewal of teacher professional skills tends to serve as an important element in teacher development programmes (Richards & Farrell, 2005, p. 1). This is why the findings reported in this section are perceived to be highly significant.

Possessing a set of highly professional skills provides teachers with an invaluable opportunity to better resolve problems and mediate conflicts in their classroom (Landau, 2009, p. 743) and minimise and alleviate students’ dissatisfaction with the adequacy of teacher’s pedagogical skills. It is worth highlighting that, as to the subcategories reported in this section (professional qualities), it is not uncommon to consider some of them as teachers’ pedagogical skills as well. The analysis of the data fed into the emergence of a number of subcategories as illustrated in Figure 5.6., as follows:
5.3.4.1. **Sense of responsibility & accountability**

Teachers’ sense of responsibility and accountability as professional qualities were the ideas explored in the interviews and the survey respectively. Explaining his sense of responsibility towards his students, Saman maintained that:

“I think as a teacher I am responsible to my students. In this case I try to first, if for example stuck in the class, I try to ask my students why for example what kind of problem do you have? What’s wrong with you? I try to ask about myself, for example, I ask them Am I right? Or am I clear for you? Can I teach well or I cannot teach well”?

Indeed the concept of ‘accountability’ has become the mantra in many countries around the world especially in Western contexts. The Act named ‘No Child Left Behind Act’ in the United States as fully discussed in chapter three is among different signposts showing an increasing global demand for accountability and quality teaching. However, further to the analysis of quantitative data, 81.4% of
respondents maintained that an effective teacher should possess a sense of accountability towards other stakeholders including students and administrators. However, similar to some other qualities discussed so far, to the best my knowledge there is seemingly no specific mechanism to measure one’s sense of responsibility and accountability, and to that end it is strenuously challenging to distinguish responsible teachers. I will come back to this point later on in the next chapter.

5.3.4.2. Creativity & Innovation

The notions of creativity and innovation were highly referred to by the participants in this study. As pointed out by Armin, effective teachers are innovative and creative for which they need to be given some sort of flexibility by the curriculum:

"Yeah, you can label it creativity ok or innovation in teaching that’s very important and … I think, yes. Because you know we should not limit ourselves to the curriculum or … impose on our side that’s very important. Innovation I think is a criterion for a making a teacher to have effective teaching".

Similarly, the following quote from Amir, one of the interviewees, portrays the participant’s aversion to what he calls consistent use of the same teaching methods:

"Yes, the teacher always should focus on his teaching; he shouldn’t be always in the same way. He should change himself, … reutilization is a very big danger for teachers. Reutilization means for example the teacher always goes to classes in the same by the same methods, by the same techniques, saying the same things all days and nothing changed, so it is not actually … it can be the death of the teacher. Teacher should always make himself to progress, should always be progressing actually, I think”.

From Thelma’s point of view, teachers should creatively monitor and meet their learners' needs rather than their own needs:

"A person who is creative, we need creativity, and well as it goes down and it boils down to a person who actually wants the learners to learn something, that gives the thing that the learners want, not the thing that he or she thinks is true".

Interestingly, the analysis of survey data showed that 100% of participants extolled the virtue of innovation in effective teaching. Such unusually high level of
consensus among lecturers about the necessity of innovative teaching practices delineated the centrality of innovations and creativity in teaching.

5.3.4.3. Authority & Management Skills

As a ‘social construction’, authority and what it refers to have been extensively disputed in the literature (Pace & Hemmings, 2006, p. 1). It tends to be a contentious grey area which needs to be identified from other resources that a teacher uses to control the classroom, e.g. ‘coercive power’ (p. 5). Based on Metz (1978), Pace and Hemmings (2006, p. 5) differentiate between authority and other resources such as ‘coercion’, ‘exchange’ and ‘influence’ in which a teacher may make use of ‘threat of punishment’, ‘offering incentive’ and ‘personal relationship’, respectively, to control the classroom. However, the participants’ readings of the notion of authority were different. Whereas some of the participants maintained that a teacher should have the ability to ‘affect others’, i.e. students, others considered teachers as managers with management skills such as time management. Referring to teacher’s ability to control the classroom, Ali commented on management skills stating that:

“For example if for example you don’t have a strong personality in order to have some effects on people’s lives okay then how can you be a good teacher? … Well I can say authority but I have to define authority first of all. I mean, by management and authority I mean that, for example, you’re dealing with 30 students in the class okay, they’ve got 30 different personalities, okay, even at this level, University students, you may face some problems in the class for example some of the students are not always satisfied with the way you teach, satisfied with even sometimes with the environment they are learning something even with the for example, the administrators and so many other things so they show their dissatisfaction in the class and you have to control them you have to control them you have to manage the class”.

Mersedeh similarly asserted:

“… a teacher must control and handle the class. I mean, based on the characteristics and behaviors of the students, there are many different factors.”

As to teacher authority, the participants were requested to express their views on whether or not teacher authority exerts influence on their effectiveness. The result was slightly sobering in that only 55.9% of respondents concurred with the idea
that authority is a keystone of lecturer effectiveness. It sounded as though authority was not considered as an issue in education in from lecturers' viewpoints.

5.3.4.4. Building Rapport & Trust

The qualities of honesty and trust were introduced in the previous category of personal attributes. However, it is argued that an effective teacher should be able to build rapports with her students. Amongst others, ‘lecturer-student rapport’ has been found as a recurring element in flourishing student independence (Ramsden, 2003, p. 74). As a professional characteristic, such a quality needs to be acquired in teacher education and teacher development programmes. In other words, as a professional skill, this is beyond the simple personal trait, in that not all trusting teachers can be successful in building rapport with their students. The following quotes from interviewees and respondents testify to lecturers’ reliance upon this quality as a professional characteristic. For instance, Maria made a clear reference to the idea and stated:

“… and at the same time building the kind of rapport with the students that can make learning and teaching pleasant and successful.”

In a similar vein, Rima pointed to teachers’ pastoral care activities:

“When I ask them about their problems ok I think it’s a good way to motivate them because they can understand oh … teacher understands us.”

5.3.4.5. Fulfilling course objectives/Intended learning outcomes

As expected, one of the issues highlighted by some of the participants was their understanding of course educational objectives. From their perspective, teacher fulfillment of the course educational objectives can translate into effective teaching. This is of paramount significance and implications pertaining to students’ learning outcome, in that such fulfilment may not necessarily lead to higher student achievement. This will be more discussed later on. The following extract from Maria portrays how teacher effectiveness entails achieving the intended learning outcomes as follows:

“I believe teacher effectiveness on the whole means producing the intended result which is reaching the aims established by the curriculum”.

Soroush also had a similar position stating:
"I want to say that teacher effectiveness refers to the extent to which a teacher is successful in obtaining the course let’s say final goals and changing the behaviour".

Sepehr adopted a similar stance and maintained:

"Teacher effectiveness is, therefore, limited to the extent to which they could have been able to cover the materials assigned to them, students’ final term performance at exams, and organizational level of satisfaction".

Fulfilling educational objectives is indubitably a widely accepted ideal to which lecturers have to adhere. With regard to the Iranian context, lecturers are usually informed of the educational objectives asserted in the curriculum by course/module manual.

5.3.4.6. Improving Students’ Performance

In line with the accountability feature discussed earlier, and highly connected with the previous subcategory, teachers are deemed to improve students’ performance. As posited by one participant, an effective teacher needs to bring about changes in students. Amir had an oblique reference to this point stating:

"Teaching actually has the notion of learning with itself and when we talk about teaching, learning comes with this so I think an effective teacher would be a teacher who makes his or her students [to] learn English. If they can learn English so the teacher would be effective, yes."

In line with Amir, Rima considered student achievement as an indicator of teacher effectiveness:

"Teachers can have large effect on student achievement. I also think one can differentiate between those students who have the most qualified teacher and those who have the least qualified one according to their achievement." (Open-ended questionnaire 9)

However, Niloofar viewed this issue from a rather different perspective and contends:

"Indeed, conceptualization of teaching effectiveness has been under the influence of the existing psychological and educational models and theories. For example, in the era of behaviorism, teaching effectiveness was evaluated in the light of teachers’ pre-defined behaviors and students’ achievements. Under the influence of cognitivism, effective teachers were evaluated based on the process of teaching and learning rather than the prescribed and observable product. As a result, various affective, cognitive
and social characteristics were incorporated to the definitions and measures of effective teaching.”

There is a lot to say about such an understanding of teacher effectiveness. Universities’ enthusiasm for high student achievement is a hidden agenda to which lecturers are obliged to conform. Whereas proponents link one’s teaching effectiveness with students’ achievement, opponents argue that student learning outcomes tend to be influenced by a multitude of implicit and explicit factors. These areas will be more discussed in the next chapter. The findings of the survey also indicated that only 51.2% of the respondents believed that student achievement can be a good indicator of their effectiveness.

**5.3.4.7. Teachers’ Networking**

Perhaps a sense of collegiality and collegial networking is a precious endowment for a teacher which can boost his or her teaching effectiveness. As Milem, Sherlin, and Irwin (2001, p. 147) put it forward, collegial networks, as a social support network, can facilitate career success. In their analysis of the literature on collegial networking, Milem et al. (2001) introduce three aspects for such networking, namely ‘type of individual outcomes’ which could be either personal or job-related’, ‘type of network which can be locally or nationally based’ and finally what they call ‘access’ to such networks (p. 147). Cultivating collegial relationships, as Stronge (2007, p. 103) asserts, helps novice teachers raise their awareness and foster behaviours and qualities of effective teachers. Nevertheless, it is worth noticing that such collegial networking can never supersede teachers’ responsibility to promote themselves. Indeed, although teacher development programme is more or less practiced in universities in Iran, effective teachers are responsible for keeping themselves informed and updated of the latest theoretical and pedagogical advancement in language teaching. The following quote from Kian indicates his views on the advantages of sharing experience with colleagues:

“A life-long learner, willing to learn from experience and share it with colleagues, continual self-assessment, monitoring what truly works and what does not, evaluating the quality of lessons being delivered, reflecting on and improving the teaching practice on a regular basis, with a wide range of pedagogical skills”.

Saman also commented on the importance of updating as follows:
“He/she should interact with his own group such as colleagues and the department. He/she must always try to improve and update the criteria needed for the current education system”.

Despite the overall acceptance of teachers’ networking as a precious opportunity to improve teacher effectiveness, it seems that such a professional behaviour has not gained ground in the Iranian context yet. Referring to the beauty of networking among teachers, Sohrab contended that teachers’ networking is currently missing in Iran as indicated in the following quote:

“The second one is the teachers’ network something that especially in Iran which we are missing is teachers’ network, getting I mean one another’s experience of teaching. This may be a good I mean habit if you want to for example teach Pragmatics okay you ask other colleagues to know about their way of teaching, the I mean materials they are introducing to the students and many other factors that can be solved by teachers networking or let’s say teaches network”.

5.3.4.8. **Overcoming Problems**

An effective teacher should have the ability to confront unforeseen challenges and overcome any problems that may arise during his/her teaching sessions in the classroom. I think this is an important quality for an effective teacher, for teachers are very prone to encountering challenging moments in their classrooms about which they might never thought about in advance, e.g. bullying. This can be elicited to some extent in the following quotes from Ali:

“My answer according to what I have learned from my professors is that you should be an artist at the first stance. I mean you have to know how to overcome the problems, how treat the students, how for example, what do you call … use your even very limited facilities around you to make that class as active as possible. This is an art, I believe in that”.

Ali’s view was strongly corroborated with statistical findings. As confirmed by the quantitative data obtained for the survey, 97.7% of the respondents agreed upon considering teachers’ ability to confront unexpected challenges in the classroom as a quality of an effective teacher.

5.3.4.9. **Fair evaluation of Students**

The quality of being fair with students as an element of interpersonal communication was discussed earlier in the category of personal attributes of an
effective teacher. However, it is worth highlighting that the notion of fairness needs to continue to permeate at all layers, from teaching to assessment and to the ways teachers socialise with their students. A number of dimensions were assigned to the quality of fairness by some of the participants. Shahab voiced his concerns for different forms of unfairness as follows:

“Avoiding all forms of discrimination, sarcasm, belittlement, etc.”.

Referring to teachers’ power, Niloofar cautioned against what she considered as teachers’ abuse of power:

“The minimum level of expectation is that teachers must consider principles of morality as a human being in the class special a unique case to apply power is given to them and they must always remind themselves not to misuse or abuse that special position”.

Fairness is one of the professional attributes which tends to have worldwide appeal and Iran is no exception.

5.3.4.10. Reflective practice

The notion of reflective practice has its roots in John Dewey’s works and became more popular by Donald Schöen in 1983 (Craig, 2010, p. 189). Whereas the concept of metacognition and metacognitive strategies are relatively well-researched as students’ learning strategies which boost learners’ success, teachers’ awareness of such strategies as mediums though which they can ‘plan, monitor and evaluate’ (Murray & Christison, 2011) their own practices has been little investigated. Reflective practice and critical thinking were among the ideas proposed by some of the participants. An indispensable dimension of reflective practice is the notion of critical thinking. Whereas critical thinking has been widely investigated in the literature, it has been less frequently considered as a professional asset for language teachers themselves. In other words, language teachers, similar to mainstream education teachers, are required to instill critical thinking as a skill to their learners. Hence, not only should teachers teach critical thinking as skills to students, they do need to flourish their own teaching criticalities through similar skills. Critical thinking and reflective practice among different
qualities were touched upon by some of the participants. For instance, Soroush stated:

“I think, one of the most important part of, let’s say, characteristics of a good or let’s say an effective teacher is his critical thinking and reflective thinking. Whatever I am talking about experience, I say that experience and even effectiveness is not necessarily related to the number of years a teacher is teaching. But it actually refers to the number of hours that a teacher has thought about his teaching and made some meaningful and logical changes in his teaching. So to me an effective teacher is a critical thinker and a reflective one. Whatever we are going to start the classroom we have a lesson plan .... in the classroom we are teaching but at the same time, we are receiving different kinds of feedback from our students. As soon as I leave the classroom, I try to reflect what happens in the classroom. So to me, if you’re going to, let’s say, make some significant changes in our effectiveness we should let’s say be a, let’s say, a reflective critical teacher and thinker and this is actually the kind of source that can inspire us to go further and make some meaningful changes in the process of our teaching”.

Closely connected with critical self-examination, reflective practice tends to lay the ground for decision making, planning and action (Richards, 1996, p. ix). Kian, one of the respondents, considered this a tool through which he can improve his practices regularly:

“A life-long learner, willing to learn from experience and share it with colleagues, continual self-assessment, monitoring what truly works and what does not, evaluating the quality of lessons being delivered, reflecting on and improving the teaching practice on a regular basis, with a wide range of pedagogical skills”.

Bahar adopted a similar position and stated:

“An effective EFL teacher should be a good critic to her/his performance”.

Despite the popularity of reflection and critical thinking, their implementation in universities in Iran is highly sceptical. This seems to partly emanate from teachers’ lack of training about teaching and learning about teaching which as Loughran (1996) puts it forward is a formidable task as they are imbued with various convoluted thoughts and actions (p. 3). It is worth emphasising that such professional attributes seem to be less well-appreciated in Iran to such an extent that many of the participants were not cognisant of the privileges of reflection and self-evaluation. Indeed, teachers need to engender critical thinking and effective practices in their teaching career for which they do need training and development programmes. Hence, one can argue that universities should also be accountable
for instilling such analytical approaches towards teaching into their lecturers’ minds.

5.4. Measures of Evaluation

Lecturers’ perceptions and their understanding of effective teaching have been presented so far. A number of ideas emerged from the data which were classified into different major categories and subcategories. Having explored participant lecturers’ understanding of teacher effectiveness, I moved beyond their perceptions and asked them to share their ideas about measures of evaluation or sources of information based upon which one’s teaching effectiveness can be established. Indubitably, this section is among the milestones of this research study, in that it could disentangle teachers from the current practices in teacher appraisal in the Iranian context by providing them an opportunity to have their say and express their ideal appraisal practice and evaluation system. As Danielson and McGreal (2000) contend, the evaluative criteria based upon which teachers are evaluated serve as the cornerstone of any teacher appraisal systems (p. 32).

Similar to many other contexts, as Kyriakides (2005, p. 44) put it forwards, a valid personnel evaluation system through which teacher performance can be established and their professional development can be boosted are among the concerns in most educational contexts. In educational contexts, research on teacher appraisal can entail widely diverse yet related domains including the roles of teachers, management, classroom behaviour, etc. (Wragg et al., 1996, p. 18). Some key concerns and questions as to how one’s teaching effectiveness can be established, whose voices should be heard and included in the appraisal model or the weight that should be given to each measure of evaluation through which appraisers can judge teacher effectiveness were fully discussed in the previous chapters. I also examined different models and frameworks that are more or less practiced across the world. The ideas emerged from the analysis of the data were of two major types, i.e. the ideas which resonated with those of the literature echoing the previous works and the ideas which were relatively new as to teacher evaluation in foreign language teaching context. These subcategories are illustrated in Figure 5.7., as follows:
5.4.1. **Student Evaluation of Teachers (SETs)/Student Ratings**

There is wealth of literature on student ratings and more recently student evaluation of teachers (SETs) and its importance in teacher evaluation and appraisal system. Indeed, SETs or students ratings have tended to be used interchangeably in the literature. As discussed earlier in the chapter of literature, research on teacher evaluation has witnessed contentious discrepancies over the inclusion and the weight of students' ratings. Whereas the advocates of SETs, consider it as an indispensable part of evaluation measures, critics shed some
doubts on the reliability of students' ratings as a source of information based on which one's effectiveness can be established. As opponents argue, such a measure is easily prone to some hidden threats and intervening variables. Such potential threats range from student-related factors such as level of attainment to teacher-related characteristics such as teachers' personality, gender, reputation, etc. Indeed, given the globally acclaimed movement for quality assurance, student ratings tend to be of paramount significance in most teacher appraisal schemes all over the world. Student evaluations of their teachers have been highlighted by some of the participants in this study. Parham, an interviewee, had a deep insight into this issue reminding the prerequisites needed for useful SETs:

“At university level, especially EFL, teaching English to foreign language students, I think the students can be a good source of feedback but not the only one in just provided that they are instructed how to evaluate their teachers”.

Soroush adopted a similar stance stating that:

“You know, for example, in reality in my country, the appraisal is based on students’ evaluations. I do believe that we have to look at it from different perspectives. You know students that have the right and should have their voice in this process and they are one of the most important reliable, of course, source of information in this regard”.

The findings from the survey confirmed these results in that 79% of the lecturers considered student voices, among other measures, as an element of teacher appraisal model. In spite of the above-mentioned support, some of the participants expressed serious reservations about the reliability of such measures. A number of drawbacks associated with student evaluation were also emerged from the analysis, e.g. the relationship between students’ marks and their ratings.

Referring to some drawbacks inherent in students' ratings such as grading bias, lack of training, etc., for instance, Niloofar posited:

“... the one I remember was that the students are asked without any preparation to answer some questions and evaluate the instructor while the instructor was present in the class; the problem was that if the evaluation was before the mid exam, the result was different, if it was after the midterm exam it was different, if it was during the last sessions of the semester the result was different .... confused. It is always me who is always taking pains and is struggling to teach but the receivers evaluate consider some other things for example he is too strict and sometimes too kind if the midterms
Another example is the following quote from Mersedeh, one of the participants, who maintained that students are not well-trained to do the ratings and referred to the interrelationship between teachers’ scoring/marking, appearance, etc., and their students’ ratings:

“Unfortunately, nowadays it is very common the universities provide some questionnaires for the students that the students evaluate their teachers, but I think this is completely wrong. I think the level of the students is not enough to evaluate, yeah, and assess their teachers… You know that for example sometimes some students get good grades from that teacher or for example emotionally they have a good feeling about that teacher, I don't know, the appearance, something like this, it can have effect on their for example learning but sometimes for example, some students don't have a good feeling about a teacher, and I don't know for example sometimes they mention I hate this teacher based on their appearance and ... “

Additionally, 95% of the lecturers concurred with the view that students need to be informed of the criteria for effective teaching from among whom 55.8% strongly agreed. Such strong convictions on the part of the respondents carry meanings which are grounded in the context of the study which will be discussed in Section 5.6. It seems that lecturers’ attitude towards SETs is partly shaped by the existing appraisal system in Iran in which SETs, from participants’ viewpoints, tend to outweigh other measures. These areas will be addressed in Section 5.6., which specifically concentrates on teacher appraisal in universities in Iran.

### 5.4.2. Students’ learning outcomes/achievement

In the previous section, I examined how student ratings or student evaluation of their teacher has been found to be one of the criteria which need to be considered by policy makers and administrators in appraising teachers. However, students seem to be of more weight in teacher appraisal, in that ‘student learning outcome’ or ‘student achievement’ tends to serve as a measure of evaluation.

As the bulk of literature on measures of teacher appraisal suggests, student achievement or learning outcome has given rise to contentious discrepancies upon the nature and the extent to which such measures should be considered by
administrators’ teacher evaluation agenda. Tucker and Stronge (2005) highlight teachers’ accountability for learners’ learning outcomes as well their own teaching (p. 16). Whereas some researchers go to such an extreme that equals teacher effectiveness to student achievement, others consider it as part of the appraisal system. There are also some more critical aspects to learners’ achievement such as the connection between students’ learning outcome and teacher performances which will be discussed in the next chapter.

As indicated in the following quote from Sam, he explicitly related students’ achievement with teacher performance and considered students’ achievement as an aspect of effective teaching:

“To me an effective teacher is someone who is or the extent someone is successful in his career, in his job. I mean, the results, the extent students are successful, can pass, the extent the students are satisfied with the performance of the teacher, the extent the teacher himself or herself is satisfied with his or her performance. The extent the students can achieve the course and this is my understanding of teacher's effectiveness.”

Similar to little support this measure received from the interviewees, the statistical analysis of the survey surprisingly divulged the relatively low status of student learning outcome as an indicator of teaching effectiveness with nearly half of respondents (51.2%) expressing their support with only 4.7% strongly agreeing with the idea. One more plausible scenario that lecturers did not care about student learning outcomes can be due to the fact that this measure is not currently included in the teacher evaluation system in Iran. It is worth reporting that 60.5% of the respondents believed that an effective teacher with one group of students can be less effective with another group of students. More interestingly, the same percentage of lecturers (60.5%) assented to the idea that this measure can be easily affected by the factors which are beyond teachers’ control (e.g. student-related factors).

5.4.3. Peer evaluation/observation

Peer assessment, peer evaluation, peer appraisal, peer review and peer feedback are among various terms which have been used in the literature as one of the sources for information in teacher appraisal. Teacher appraisal refers to an
appraisal in which two colleagues of equal rank appraise each other (Wragg et al., 1996, p. 16). Similar to the any other measure, there emerges a question as to for what purpose such information will be used. In other words, type of interpretations and consequences brought into the results, i.e. summative or formative purposes, indubitably, tend to make a difference on the whole structure of the appraisal. Moreover, given the nature of appraising a teacher colleague, gives rise to another concern as to whether or not the appraiser colleagues are critical enough or possess the minimum training to do faculty evaluation. These aspects will be fully discussed in the next chapter. Soroush viewed peer evaluation as a learning opportunity and posited:

“I do believe that peer assessments can help us to learn from each other, to know about our, let’s say, positive and negative parts, and to accelerate the process of effectiveness better. It can provide us with a very good picture of effectiveness but it is not of course done. But sometimes when some colleagues are somehow intimate to each other, yeah, they do the same thing. It is not something formally”.

In the same vein, Sepehr maintained that peer evaluation is a collegial and non-threatening opportunity in which teachers can share their experiences with each other:

“Yes, your colleagues, coworkers, colleagues yes, of course it is in the condition that there is a context of friendship with your colleagues. You can share your ideas and you know what is happening in your colleagues’ classes and they know what’s happening in your class. Yeah, in this situation colleagues can be a good source of making decisions of if ... someone … is or is not an effective teacher.”

The analysis of statistical data also testified to the importance of peer evaluation in teacher effectiveness with 79.1% of participants supporting the idea 25.6% of whom strongly agreed with peer evaluation. It is worth mentioning that peer observation is not currently included as a measure in teacher evaluation system in most universities in Iran. As a consequence, lecturers’ little awareness of peer evaluation as a measure could be likely grounds for such a relatively low rate of strong support for peer evaluation. It is worth adding that only 39.5% of the respondents believed that this measure can be used for summative purposes.
5.4.4. **Self-evaluation**

Self-evaluation, self-assessment, self-appraisal and self-review are also among different jargons used in teacher appraisal context to indicate the process in which a teacher evaluates him/herself and reflect on his/her teaching practices. As Poster and Poster (1991, p. 124) put it forth, effective teachers tend to be in the habit of ‘self-evaluation’ in which they endeavour to continuously address their teaching consequences and its associated merits and demerits. Self-evaluation was also picked up by some of the participants as indicated in the following quote from Thelma:

“I can assess myself and for example get feedback from the students, form my self-assessment and try to edit my problems and try to improve my problems in the class”.

Sepeher also maintained:

“… it is and … it should be a fair kind of evaluation because no one is more fair than you to yourself.”

Self-evaluation is consonant with the notion of reflective practice as explicated earlier in 5.3.4. As to the findings obtained from the survey, three statements were formulated to elicit lecturers' viewpoints germane to the need and purposes of self-evaluation. The results echoed what I expected. Not surprisingly, 95.3% of the respondents expressed their support for self-evaluation as a tool to reflect on their practices. Nevertheless, whereas 74.4% of lecturers regarded self-evaluation suitable for *formative* purposes, 48.9% of respondents were under the impression that this measure can be used as a *summative* measure with only 4.7% who strongly agreed with it. Similar to peer-evaluation, it seems that self-evaluation has not been yet given the attention it deserves in teacher evaluation system among EFL lecturers in Iran. These areas will be more discussed in the next chapter of discussion.

5.4.5. **Observation (external/internal)**

As a measure of evaluation, ‘observation’ has proved itself as an indispensable part of most teacher evaluation schemes around the globe. Observation, as a way of looking at classroom teaching practices, has become the cornerstone of many
appraisal systems (Campbell et al., 2004a, p. 99). Nevertheless, whereas observation is highly reputed in most countries, it is somehow belittled if not missing in some contexts, e.g. the context of current study. In a similar vein, observation as a measure of evaluation has turned out to be of a broad realisation in practice, from internal observation to external one, from observers within the same academic discipline to those with other educational backgrounds. These dimensions are of an utmost importance and will be fully addressed in the discussion chapter. The significance of observation is shown in the following quote from one of the participants:

“No, it really needs observation, really needs observation, they should be in the class in order to understand that how, for example, whether he or she is an effective teacher or not”. (Ali)

Moreover, the analysis of the data gave rise to a number of features that participants assigned to observation such as:

- Board of experts (accredited observers)
- Throughout semester
- Observers’ knowledge of pedagogy and TEFL
- Quality (Standard) Centre

These features are of seminal importance as they can cast light on the evaluation process making it more credible. From the participants’ perspectives, observers should be ‘qualified’ appraisers, the ones who are possessed of the required knowledge of pedagogy and TEFL and the ones who are well-informed of how to do a faculty evaluation. In addition, the time of evaluation should be also considered in order to minimise the effect of some other intrusive variables. Critiquing different aspects of teacher observation, Parham extensively commented on the necessities of a successful teacher observation as follows:

“and also I agree with some observations, there should be a board at the University, board of expertise who have some criteria, who have some accorded criteria based on the teacher, I don't know, getting together, the instructors getting together, form a criteria based on, for each of the courses, for instance, reading courses or reading modules may be different from writing modules or conversation modules.” “Then select someone or a few people, two or three, depending on the size of the university to evaluate their works, at the beginning, middle, at the end”.

Ali had a similar view and stated:
“… it really needs observation, really needs observation, they should be in the class in order to understand that how for example whether he or she is an effective teacher or not.”

Rima also made the point that appraisers should be acquainted with the EFL domain:

“Those who are skilled at pedagogical courses … pedagogical majors or for example, in … also they are skilled in EFL teaching …”.

The statistical findings also revealed some interesting points which are worth considering. Around half of the participants (46.5%) agreed with the inclusion of observation in their appraisal only 2.3% of whom strongly agreed with the notion. In response to another statement which asked lecturers whether they are not willing to be evaluated by an external examiner, only 55.8% of participants concurred. Perhaps, the paucity of such a measure in the universities in Iran has made it an elusive idea for some of the participant lecturers. Some other possible reasons will be discussed in the next chapter.

5.4.6. Some further points to consider

Five major sources for information or measures of evaluation by the use of which administrators can decide on their lecturers’ teaching effectiveness have been addressed. However, a number of ideas emerged from participants’ words which, I think, might be worth considering.

Highly interlinked with the next quality of being unbiased and fair, power issues and the so-called hidden relationships between teachers and administrators such as head of department, dean of faculty, etc., should not be underestimated as they can easily affect the results of teacher appraisal. In other words, power relations and the so-called hidden agendas can potentially dominate the knowledge in teacher effectiveness appraisal. This is clear in the following quote from Niloofar in which she referred to an imbalance between power and knowledge:

“Ideally the one who has got the knowledge but based on what the reality is the one who has got the power, the one who has got the power can judge but ideally the one who has got the knowledge who can evaluate who the teacher is? Who has got the knowledge about teachers’ effectiveness? But I want to say that the gap is between power and knowledge; the one has
got the knowledge does not have the power and the one who has got the power I cannot see any knowledge in them”.

As to the statistical findings, 58.1% of the lecturers maintained that power relations can influence teacher appraisal. As it has been stated earlier, the majority of the participants complained about being given a vague briefing on the dynamics of teacher appraisal in their respective universities. As a general rule of thumb, when processes involved in a phenomenon are not well explained to its stakeholders, rumors begin circulating. This simply necessitates the need for a transparent appraisal system. By transparency, not only do I mean clarity of the processes involved in appraisal, I want to place emphasis on administrators’ responsibilities to regularly brief their academic staff on the pertinent procedures and keep them informed of any revisions to the codes of practice. These issues will be more addressed in the forthcoming sections.

The notions of fairness and impartiality are not simplistically confined to administrators’ appraisal of teachers. Suffice it to say that all measures discussed so far even that of self-evaluation should be based on fair and unbiased evaluation of one’s teaching effectiveness. In addition to the aforementioned features, ‘flexibility’ was also mentioned by some of the participants. From their perspective, the flexible nature of teacher appraisal embraces different elements of flexibility, i.e. from the time of evaluation throughout the semester to the weight given to each measure. I will elaborate on this issue in more details in the next chapter. It is also worth referring to the statistical findings which revealed that 93% of the respondents maintained that they need to be convinced of the fairness of appraisal system through which they tend to be assessed. Interestingly, more than half of lecturers (53.5%) strongly demanded fair appraisal:

Similarly, some of the participants perceived subjectivity as a threat entrenched in administrators’ evaluation of teachers which arises from affinity and enmity among teachers and their colleagues as appraisers. In the following quote from one of the participants, she raised the point:

“And also the head of the department ... Well, I want to be polite and nice but we have friends and we have enemies and that is again would be entrapped in the trap of subjectiveness, that is again subjective. I have
been years I have been friends with Mr. ….. for years so I would not consider any students’ complaints in that regards, I say no, Mr. ….. is doing his best, that was just as an example in terms of … some people are not fair. If everybody… can everybody be a good judge? I am asking you? Can everybody be a good judge? … of course not. It’s some personal characteristics, a fair-minded a fair-minded individual is needed to be a judge and that is a very a very shaky position. It’s thought something that you can just be relaxed and do that job; it is a very shaky position, it cannot be that fair-minded”.

The last point that deserves thorough attention is the role of universities (policymakers and administrators) in the appraisal. A number of concerns might arise at this stage as to how evaluation is implemented or what consequences teachers are likely to face after their appraisal or who interprets the findings of evaluation. The analysis of the data shed some light on the complex duties and responsibilities of universities in promoting their staff effectiveness. These aspects will be discussed in the next chapter especially the possibility of any conflict between universities’ goals stated in their vision and mission and those of their academic staff.

Having considered the above-mentioned measures of evaluation, it is worth referring to participants’ preference and demand for a multi-measure rather than a single-measure approach towards teacher evaluation which is self-evident in their responses to statements 1 and 7 which advocated for a multi-measure teacher evaluation entailing different stakeholders’ voices. This is of high significance and will be more discussed in the next chapter. As the results suggested, the idea received strong support (100%) from all the participants.

5.5. Ways to improve Teacher Effectiveness in HE

Having explored the participants’ understanding of teacher effectiveness and their perceptions of the measures of evaluation, I tried to bring to light their ideas and insights into the ways and opportunities by which lecturers can enhance their teaching effectiveness. A number of methods and mediums were suggested by the participants to increase the effectiveness of their teaching practices. As expected, some of these ideas overlapped with the qualities of an effective teacher discussed earlier. However, the analysis of the data gave rise to the emergence of some new ideas which will be discussed in the flowing section.
Prior to addressing the emerging categories and subcategories, it is worth emphasizing the need for an on-going continual teacher development programme. It might be argued that organising teacher development programmes and workshops tends to be among administrators’ responsibilities. However, this does not exempt teachers themselves from their responsibility for self-improvement. In other words, further to the training teachers usually receive in their universities, they are accountable for their self-promotion. Since this section is not aimed at addressing teacher development programmes designed for in-service teachers (will be discussed in the next sections), it focuses on the ways, strategies, tools and whatsoever opportunities through which teachers can increase their effectiveness and boost their effective teaching practices. These generated categories and subcategories are also presented in Figure 5.8., as follows:

![Figure 5.8. Theme III (Ways on increase TE)](image-url)
5.5.1. Experience

In line with many attributes and characteristics which were found to contribute to teacher effectiveness, there is strong support in the literature on the importance of experience. While the relationship between teacher experience and student achievement has been suggested in some studies (e.g. Nye, Konstantopoulos, & Hedges, 2004), other research studies have cast doubts on the predictably power of attributes such as teacher experience (e.g. McBer, 2000). Following the analysis of both sets of data, among other elements, experience was found to be an important factor in teaching effectively. In the following quote from one of the participants, he clearly recounted how he evolved throughout years of teaching and gained experience:

“I always liked teaching I always liked teaching I remember the first time I started teaching I really didn’t know what to do. I experienced a lot of stress really and little by little I learned. I learned what to say, how to say and when to say. I learned how to treat students how and I learned for example I know a lot of things about grammar but in this class I have to present this part of my knowledge and not for example half of the knowledge I have got about grammar. I should be as simple as possible these things happened in a for example eight or nine years period really, so experience to my opinion had really big effect great effect on me. Right now I understand or I feel that ok I have become a much more better teacher in comparison to eight years ago seven years ago six years ago even last year and I am still learning by experience”.

Niloofar also commented on experience and declared:

“Effectiveness in teaching skills can be obtained through cultivating those skills regularly throughout one’s teaching career. Experience when combined with high quality training and effort can foster effectiveness”.

Participants’ standpoint was also confirmed by the survey as 81.4% of the respondents pointed to the values of experience. The value of experience is appreciated in the Iranian universities as it is, among other factors, a contributing factor in teachers’ salary, promotion, etc.

5.5.2. Technology Enhanced Teaching and Learning Strategies/ ICT Literacy

ICT literacy has been discussed in the previous section as one of the qualities of an effective teacher. Education and language teaching in particular is now such imbued with the use of computer technology and internet that teachers, seemingly,
are no longer able to abdicate their responsibilities in incorporating the Technology Enhanced Teaching and Learning Strategies (TETLS) in their teaching career. The emergence of such new technologies along with their coined jargons in the literature depicts a new trade-off between traditional and cutting-edge teaching practices. However, there has been some argument upon the balance between the so called 'brick and mortar' and 'click and mortar'/brick and click' educational contexts (e.g. Luke, 2006; Thomas & Thomas, 2012). As it is evident in the following extract from Parham, it seems that he went to extremes maintaining that the use of ICT in teaching in Iran is in its primitive stages:

“They are rudimentary and they are the basics of providing a good atmosphere and context for teaching English but unfortunately, for instance, nowadays web-based technology or web-based teaching is being well-received in developed countries but in our country for instance using internet and using computer is just in the primitive stages we don’t make use of it at all or for instance making use of PowerPoint and I have never seen them in our classes but they are very common here”.

Sepehr also placed value on technological development as an inevitable aspect of teaching in academic context and asserted:

“I will also try to make use of different presentation strategies according to the academic context in which I am teaching. Using technological developments in academic context ... today seems to be a must and therefore I think any effective teacher working in such a kind of context must enhance their ability in using different technological devices in their classrooms”.

The statistical findings also enunciated lecturers' viewpoints on the contributions of Technology Enhanced Teaching and Learning to teaching effectiveness. As the analysis suggested, 90.7% of respondents assented to the notion with 27.9% who strongly agreed with it.

5.5.3. Knowledge of Assessment

Although this subcategory was introduced previously as a pedagogical quality, it continues to be an important aspect of teachers’ quest for improving the quality of teaching. As Coombe, Troudi, and Al-Hamly (2012, p. 20) contend, teachers need to be assessment literate, otherwise they will not be able to help their students with their academic attainment. Effective teachers should consider 'alternative
assessment’ strategies and utilise them in gauging their students’ learning outcomes. This can help teachers have a more accurate understanding of their learners’ achievement. Imbued with higher order skills, the use of alternative assessment strategies can help teachers better explore real life conditions and hence it has become of high repute albeit the disadvantages such as being time-consuming (Farrell & Jacobs, 2010, p. 99). Another important dimension of teachers’ knowledge of assessment is the purpose of assessment. Referring to the influences of assessment on teaching, Kian placed emphasis on the importance of assessment FOR learning as explained in the following extract:

“Washback effect of tests coming from the curriculum and syllabi is an undeniable part of teaching. For example, if the exercises in course books are designed and/or treated in a way that makes students feel they have to prove themselves and their intelligence, it can encourage the either-or type of thinking in learners; this win or lose mentality can hinder effective learning and teaching, because it can be discouraging to some students. It appears that assessment should not only be “of learning”, but also “for learning” for it to be effective”.

As to the quantitative data from the survey, the analysis of one of the statement which asked respondents whether or not they agree with the need for teachers’ knowledge of different assessment strategies, 93% of the lecturers assented to the idea with 46.5% who strongly agreed with it. Although the majority of the respondents appreciated the significance of assessment knowledge, few of the participants referred to the centrality of assessment in teaching. Hence, I wouldn’t be surprised if many of them were not well-cognisant of the benefits of using alternative assessment strategies in their classrooms.

5.5.4. Self-Regulation & Reflective skills

The importance of reflective practice has been partly discussed in the section of qualities of an effective teacher. Since the outset of the third millennium, teachers have been expected to accept new roles as compared to their traditional role for student achievement. Nowadays, in line with the use of information technology as discussed earlier in almost most educational contexts around the world, teachers are now compelled to be more conscious with their use of educational resources. The concepts of reflection and reflective practice seemingly imply reflection on
every aspect of teaching. As Richards and Lockhart (1996) remind us, teachers tend to come across with events in their classrooms and teaching environment by the use of which they can enrich their understanding of teaching through critical reflection (p. 6). Some of the participants in this study referred to different terminologies such as self-regulation, reflective teaching, self-monitoring, etc., to enunciate the need for a continual self-appraisal. As to one question in the open-ended questionnaire which inquired into how to improve teacher effectiveness, Sohrab and Soroush referred to the notion of ‘reflection’ and stated:

“I think reflective teaching can be the best way to enhance teaching effectiveness. This can be done by peer-evaluation, students’ feedback and so on.” (Sohrab)

“Cultivating reflective and critical thinking and teaching” (Soroush)

Kian made a similar remark contending:

“Whether an EFL teacher is effective in what they do, I think, should be sought in (a) their skills in performing their job and (b) their skills in examining and monitoring the quality of their performance to make amendments for their future performances. Such skills are a potential which arguably every EFL teacher possesses. Any deficiency in fulfilling it can be attributable to internal and external factors which guide their EFL career”.

In line with teachers’ reflective practices, teachers may avail themselves of self-regulatory learning skills as proposed by Niloofar:

“I assume, teacher’s self-regulation can be an indicator for teacher’s success. I also agree with the assumption that in the domain of teaching effectiveness, teacher emotional intelligence plays an influential role in fostering not only teachers’ performance but also learners’ cognitive and affective achievements. So, developing EFL teachers’ self-regulatory skills is linked to the enhancement of their success. This in turn should encourage teacher educators, administrators, and policy makers to introduce self-regulated learning strategies to teacher training programs. In line with the current trends in teaching effectiveness, these programs are expected to undergo a shift from curricula pivoting around solid basis of content area knowledge to those equipping teachers with regulating their actions, thoughts and emotions.”

There is a lot to say about the beauty of reflective practice. It is fortunate that the idea has gained a relative acceptance on the part of Iranian lecturers who have never been instructed in reflective practice by administrators.
5.5.5. Keeping updated

State-of-the-art knowledge, updating, reading new articles and books on TEFL/Applied Linguistics, etc., among others, were the notions the participants referred to as the qualities of an effective teacher. At the core of these ideas is teachers’ need for keeping professionally updated adopting a lifelong learning approach towards teaching. The need for teachers’ continual updating can be identified in the following extracts. For instance, Sepehr maintained:

“In an EFL context I will try to get myself as aware of various teaching methods and techniques as possible. Familiarity with various teaching techniques will help me in choosing the most appropriate techniques, activities and tasks according to any given educational context”.

Saman, in a similar vein, drew attention to updating and stated:

“In order for a teacher to be effective and successful, he/she can do the best to be skilful, technical and knowledgeable. He/she must be up-to-the-minute, be aware of all developments in the specialized field. He/she should interact with his own group such as colleagues and the development. He/she must always try to improve and update the criteria needed for the current education system”.

Katayoon and Parham also made the same remarks whose excerpts are presented in the following, respectively:

“A good teacher is always engaged in updating him/herself. S/he is always busy studying the new relevant materials and keeping in touch with the latest developments in the field”.

“Reviewing and reading the new achievements and research findings in the field”.

The notion of ‘updating’ partly rests upon other prerequisites such as facilities, successful teacher development programme and other opportunities of which teacher can take advantage to promote their knowledge. Sepehr introduced ‘seminar’ as an opportunity to keep teachers up-to-date as follows:

“At this point, I believe teachers can develop their knowledge of their career through studying more and more and get themselves up-to-date via participating different relevant seminars”.

While lecturers in Iran are currently provided with some opportunities such as TPDP, some of the participants complained about inadequate resources and
demanded for better facilities such as a rich electronic library, seminars, etc. These areas will be discussed in the next chapter.

5.5.6. Use of Students' feedback

One of the criteria based on which teachers can evaluate their own teaching effectiveness is students’ feedback. Hence, not only do teachers need to provide their students with feedback, they themselves seem to be much in the need of their students’ feedback. In other words, feedback can be perceived as having reciprocal contribution to both students’ learning outcome and teachers’ teaching effectiveness. As Irons (2008) maintains, not only should formative feedback focus on student learning, but also it can be used as an input to teacher development, thereby providing a rationale for modifying teachers’ practices (p. 106). The following quotes show how the participants consider students’ feedback as a source by the use of which they can enhance their effectiveness. Hossein, one of the survey respondents, had an explicit reference to the idea and stated:

“A good teacher is always engaged in updating him/herself. Such a teacher tries to make the most of the situation, especially students’ feedback”.

Majid had a similar stance towards student feedback and asserted:

“I usually try to reconsider my teaching strategies through students’ feedback”.

As to the statistical findings, 86.1% of respondents maintained that effective lecturers should be open to students’ voices.

5.5.7. Universities’ accountability for promoting TE

The majority of methods and mediums addressed in this section through which teachers can promote their effectiveness, pertained to teachers’ responsibilities. However, as highlighted by some participants in this study, some of these opportunities are beyond the hand of teachers. In other words, universities and administrators are also deemed to be accountable for their staff (lecturer) effectiveness, in that they need to provide opportunities and facilities and enough incentives for their staff to actively engage in staff professional development programme, training workshop, etc. Amongst others, such incentives can include
higher payment, academic rank promotion, etc. Although such opportunities potentially can help teachers promote their teaching practices, there are some contentious concerns about the real contributions they can bring about in the Iranian context and herein lies the challenge. As Richards and Farrell (2005, p. 23) posit, such workshops are rather approached as 'hit-or-miss' affairs lacking a solid prerequisite planning. These dimensions will be fully discussed in the forthcoming sections. However, it is worth referring to some of the participants’ demand for appropriate professional development opportunities such as seminars, teacher training courses (TTC), teacher professional development programme (TPDP), etc. Surprisingly, the analysis of both sets of data revealed a widespread discontent at the quality of the existing training programmes in universities. Warning about the gap between lecturers' needs and what they are given by universities, Parham levelled some serious criticism at the existing teacher development opportunities in Iran as follows:

“What are the inputs of instructors? Are they well-trained or not? Are they ready to teach this special course or not? If they are not ready, if they are not well-trained, so how we expect them to deliver a good class or I don't know how to teach effectively? So they should have an input or we should have a input, we should have a good educational system out of which effective teachers are produced or are presented and then expect to talk about teacher effectiveness. When there is a gap and we don’t know what’s happening in teacher training courses. Of course, there is no teacher training course there as far as I know. So we cannot expect too much form the teachers themselves.”

The following quotes are from Majid and Sepehr who raised similar points and added different types of opportunities from which teachers can benefit as presented in the following, respectively:

“As you just said, that universities should provide situations for teachers to be up-to-date, for example, holding workshops, inviting great figures in language teaching .... Yes, holding conferences, seminars and workshops... “(Interview 5)

“I believe that attending various professional workshops, seminars, and conferences will help me get up to date and familiar with the latest innovations in the field.”

In addition to the qualitative findings, the analysis of statements 61 and 63 of the questionnaire which addressed the significance of teacher education programme (TEP) and teacher professional development programme (TPDP), gave clear
evidence to the findings obtained from interviews, i.e. universities can improve the teaching quality through designing and administering such programmes for pre-service and in-service teachers. As the results suggested, 93% and 83% of the respondents were in favour of suitable TEP and TPDP, respectively.

Interestingly, 93% of the respondents would prefer to have an office/unit/department at their universities from which teachers can ask for both general and technical advice about improving their effectiveness. Moreover, 69.8% of the respondents disagreed with the view that an accredited professional preparation programme will not help teachers gain the skills they require. These figures simply convey different messages to policymakers and administrators as to lecturers’ rising expectations of high quality in-service training opportunities as well as continuous professional support.

5.6. Lectures' perceptions of teacher evaluation in Iran

In this section, another important research question pertaining to lecturer appraisal system in Iran will be discussed. So far, teachers’ general perceptions of the notions of teacher effectiveness and the qualities of an effective teacher as well as the measures of teacher evaluation and finally the ways by which teachers can promote themselves and enhance their effectiveness have been discussed. As mentioned in the above lines, the theme to be discussed in this section is germane to teachers’ understanding of Iranian appraisal system and endeavours to focus on the strengths and weaknesses associated with the current practice and the existing scheme for teacher appraisal adopted by the Iranian universities and higher education institutions. This section embraces two parts, namely the perceived challenges associated with the Iranian appraisal system and lecturers’ suggestions and recommendations to improve the existing teacher appraisal policies adopted by universities in Iran as indicated in Figure 5.9.:
As shown in the above figure, this theme gives the answer to one of the research questions appertaining to the Iranian teacher evaluation system. This theme entails two parts including exploring lecturers' understanding of the existing evaluation system in Iran through which teacher are assessed each semester or academic year and lecturers' recommendations and suggestions for improving the appraisal approach.

However, prior to reporting the particulars, it is worth referring to the lecturers' general attitude towards the existing teacher appraisal adopted by their universities. Only 11.6% of the respondents were happy with the appraisal 2.3% of whom were strongly happy with the appraisal.

5.6.1. Some perceived challenges

In general, the analysis of the data testified to participants' little satisfaction with the results of their appraisal which, from their points of view, is due to the absence of a comprehensive model. Some of the main problems associated with teacher appraisal in Iran are as follows:
5.6.1.1. One-size-fits-all approach

The notion of ‘one-size-fits-all’ approach is well-known in the literature which implies at its core, assigning a set of generic characteristics for an effective teacher regardless of the peculiarities of teaching in different contexts. Indeed, research into the status of teacher effectiveness has progressed in a rather generic sense, in that it advocates the so-called ‘one-size-fits-all’ approach which per se underestimates the influential role of students, contexts and subjects (Campbell et al., 2004a, p. 50). As Campbell et al., continues, there raises a need for researchers to move beyond such a simplistic understanding of teacher behaviour with a particular attention to multidimensionality of teaching (p. 73). This has also been found to be a part of participants’ utterances pertaining to their understanding of teacher evaluation in universities and higher education centres in Iran. With regard to an open-ended question about the status of teacher evaluation in Iran, Kasra made the following comment:

“Within Iranian EFL context, teacher effectiveness has not been paid due attention except by some private institutes seeking excellence within competitive markets. Notions such as reflective teaching and collaborative/team teaching are often downplayed within this one-size-fits-all context of English language teaching in Iranian public schools”.

Moreover, 65.2% of the respondents believed that EFL lecturers should be evaluated based on TEFL-specific criteria. Similarly, 65.1% of the lecturers maintained that the appraisal model should be informed by the Iranian contextual specificities, e.g. socio-cultural factors.

5.6.1.2. Power issue: Conflicts among stakeholders' interests/Subjectivity

The interference of power issues in staff development and appraisal in educational contexts has been well-investigated and discussed in the literature. Indeed, the exerted influence of power issues sometimes moves towards extremes to such an extent that the hidden conflicts among different stakeholders become more noticeable. Such power relations of quality assessment depict some kind of subjectivity (Morley, 2003, p. 113). Power relation which is usually caused by lack of consistent transparent standards for effective teaching could be easily
contaminated with political agendas. I will return to this important point in the next chapter. However, what seems to be of high importance is the extent to which different stakeholders’ share common interests. This point has been signified by one of the interviewees as follows:

“As a teacher, I must, you know, uh be honest in the message I am sending but their policy might be attracting students so I cannot think of attraction in the class. I can just think of my commitments and the requirements and the necessities, who am I and what am I doing there? That’s all. So in the macro level, the students policy is not related to my policy as an instructor and if they are expecting me to consider their policy they are wrong. Should I always adapt my methodology based on the university’s policy that they want to attract more and more students?” (Niloofar)

Power relations as an intrusive variable were highlighted by Majid who couldn’t conceive of teacher appraisal without the inherent problem of power issues:

“The main problem, I think, is with the subjective aspect of teacher appraisal in which teacher effectiveness is determined by the degree of relationship between the teacher and authorities”.

Niloofar raised a similar point and asserted:

“The presence of power along with its network can never be overlooked in any (educational) community”.

Despite the above quotes from participants, the analysis of the quantitative data did not provided a strong support, even though more slightly more than half of the respondent (58.1%) contended that power relations tend to exert influence on teacher appraisal.

Many participants believed that the appraisal model and the way it is implemented in their universities are subjective and biased. Referring to the effects of ‘grading bias’ on SETs, Sepehr stated:

“I think the existing appraisal model in Iran is too subjective. The students’ scores expectations Influence their ideas and comments made about the teacher in question”.

The statistical findings also divulged some intriguing information in that only around one-third of respondents (30.2%) agreed that their administrators’ appraisals are subjective and biased. Surprisingly, no one strongly agreed with the idea.
5.6.1.3. Non-teacher-controlled intrusive factors

As it was mentioned earlier, teachers are only one constituent of an educational context whose effectiveness tends to evolve within a convoluted network of interactions. In other words, lecturers’ teaching effectiveness appears to be influenced by several explicit and implicit factors. Whereas some of these influential elements are directly or indirectly in the hands of teachers, some others are found to be beyond their control. Some of these intrusive factors, amongst others, which are beyond teachers’ control, are the nature and difficulty level of the course, students’ background knowledge, number of students in the classroom, etc. These dimensions have also been investigated more or less in the literature. For instance, based upon their study on class size and students ratings, Bedard and Kuhn (2008, p. 253) concluded that class size tends to exert negative impacts on students’ evaluations of their teachers’ effectiveness. This aspect has also been declared by some of the participants. Soroush touched upon some factors which from his point of view appeared to be ignored in teacher evaluation as follows:

“So, but I would like to mention a thing, you know, whenever the university is going to provide us with our, for example, evaluation, they give us the final score based on different courses, in front of the courses it is mentioned the number of the students and, of course, let’s say, the kind of the score that you are scored in that course. But generally speaking they come to average. But your evaluation is going to be compared with others. But your colleague, for example, does not have your courses, does not have the same number of subjects, do not hold the class at the same time. So, see, there are solid different various types of factors that are important here but they are ignored in our appraisal model”.

As to one statement of the questionnaire which asked participants for their views about the notion that students’ achievement is highly influenced by student-specific factors, 60.5% of the participants expressed their agreement, even though 11.6% of the respondents were opposed to the notion. Although a number of interviewees were under the impression that the difficulty of the course negatively affects students’ ratings, fewer respondents (48.8%) shared their interviewee counterparts’ views. This category will be more discussed in Section 5.7.
5.6.1.4. Reservations about SETs

Students’ ratings or students’ evaluation of teachers (SETs) along with its associated merits and demerits were fully discussed. Indeed, the related literature hosts a wealth of research and studies on SETs as one of the key measures of teacher evaluation during last four decades (e.g. Balam & Shannon, 2010; Campbell & Bozeman, 2007; Costin, Greenough, & Menges, 1971; Marsh, 2007; McKeachie, 1997; Spooren & Mortelmans, 2006; Stehle, Spinath, & Kadmon, 2012). A major drawback which was mentioned by some of the participants was the overreliance of the existing appraisal on students’ ratings or students’ evaluation of teachers, given the limitations associated with SETs. Indeed, some of the participants had serious reservations about the effects of such a measure based upon which administrators can decide on one’s teaching effectiveness. SETs susceptibility to factors such as teacher’s age, gender, students’ lack of awareness of the goal of evaluation and the kind of statements in the questionnaire were among lecturers’ concerns towards emphasis on students’ ratings.

As to the statistical findings, 55.8% and 58.2% of the participants maintained that students’ ratings are likely to be affected by teachers’ gender and age, respectively. Moreover, the mechanism through which students’ views are collected was thrown into question by some participants. They maintained that the results of students’ ratings are unreliable, given the fact that the majority of students are obliged to do the ratings many of whom tend to do it in a tick-boxing manner. Some of these points are referred to in the following quote from one of the interviewees:

“I would like to talk about the process, you know, the process of teachers’ evaluation by the students is in the final before the final exam through, … a riddle or let's say an electronic one that should be done as soon as possible may be less than one minute ok so this time most of the time they don't care about of course what they're going to say. So their evaluation, of course, is of course not reliable mostly. Most of the time the students think that their evaluation and the kind of score that they give to the teacher has a kind of effect on their final score, that's why they exaggerate most of the time, number two sometimes they're affected a lot by a kind of factor, for example, the kind of relationship with the teacher so their score is somehow contaminated by different types of factors. The fear to the final exam score, the fear from, let's say, let's say, a kind of relationship …” (Interview 2)
'Grading bias' was also among the drawbacks emanated from an overemphasis on SETs which was postulated by half of the questionnaire respondents. With regard to one of the statements which asserted the correlation between teachers’ marking/scoring system and their students’ ratings, 55.9% of respondents expressed their approval. The time and frequency of student ratings have also been speculated as important as the ratings. As suggested by the statistical findings, 74.4% of the respondents maintained that administering the SETs in mid-semester would minimise, if not eliminate, the ‘grading bias effect’.

5.6.1.5. **Lecturers’ lack of awareness**

Similar to students’ lack of awareness of teacher appraisal and its goals, teachers themselves are not very informed of the dynamics of teacher evaluation. From teachers’ point of view, being less well-informed of the appraisal system which in turn leads to their unawareness of the university expectations of a good teaching practice turns out to be problematic. Unfortunately, failure to clearly publicise the maxims and criteria for an effective teacher can end in lecturers’ confusion. The underlying reasons behind why the majority of participants did mention their dissatisfaction with their awareness of teacher appraisal will be fully discussed in the next chapter of discussion. However, it might be a good idea to see some quotes from participant lecturers who openly admitted to their little awareness of the procedure for their appraisal. For instance, the only measure Sepehr recollected was SETs which from his point of view tends to be implemented hastily.

“I am totally unclear about the model used and think it is performed by distributing some questions (questionnaire) among the students at the end of the term and asking them to respond to them. They do it carelessly and hastily. I think there are no criteria at all”.

Ali had a similar complaint and noted:

“Honestly no, I’m not well-aware. I have heard something from my students that when they want to choose the courses they have to evaluate teachers over there on the internet ….. Really when I come to the class I am not aware of their expectation”.

Referring to a lack of a reciprocal understanding, Niloofar remarked upon the importance of a shared understanding of the phenomenon and stated:
“And there is no mutual understanding between what are the characteristics of such this good teacher, there is no mutual understanding. Everybody, everybody’s knowledge, if there is, is in isolation and might be miles apart from my expectations as I graduate in this field.” (Interview 14)

However, the evidence obtained from the survey was to the contrary. The analysis of statement 51 of the questionnaire addressing lecturers’ awareness of the existing appraisal mechanism in Iran, gave equal weight to those who asserted their familiarity with the evaluation system and those who announced their little awareness of teacher evaluation system in their respective universities.

Whereas 39.5% of respondents expressed their awareness of the appraisal system, 39% of lecturers admitted to their little awareness. Although the results may not suggest a critical situation, I think, the fact that around half of lecturers were not well-aware of their evaluation is rather worrying. This area is of substantial implications which will be discussed in the following chapter. Lecturers’ little awareness of teacher appraisal seems to be an outcome of administrators’ failure to publicise their approaches to teacher evaluation as well as a lack of consistent approaches in different universities. These areas will be more discussed in the next chapter.

5.6.1.6. **Reliability and validity**

Taking the above mentioned points into consideration, some participants in this study shed doubts on reliability and validity of the obtained data. Due to the problems associated with the existing teacher evaluation scheme adopted by the universities in Iran, some of the interviewees and respondents had serious reservations upon the appropriateness and effectiveness of the appraisal system per se as seen in the following excerpt:

“It suffers from shortage of reliability and validity. It is just based on mostly students’ comments on a closed-ended questionnaire. Besides, these contaminated data are not subject to interpretation and decision making”. (Soroush)

Participants’ dubious stance towards the reliability of teacher appraisal was recognised through the analysis of their responses to a statement of the questionnaire which asked respondents' views on the appraisal model in Iran as a reliable and valid indicator of teacher effectiveness. As the results revealed, only
16.3% of lecturers agreed with the idea that the appraisal system at their respective universities was a reliable and valid indicator of their teaching capacities.

However, it is worth highlighting that this category merely reflects lecturers’ understanding and perceptions of the extent to which their appraisal were an accurate reflection of their effectiveness. This is of seminal importance as the concepts of reliability and validity tend to be highly imbued with statistical underpinnings.

5.6.2. Suggestions and recommendations

In the previous section, participants’ ideas, complements and complaints about the strengths and weaknesses of the existing appraisal model in Iran were presented. Nevertheless, aside from articulating their conceptions, some participants were keen to discern how teacher appraisal can be optimised and hence proposed their suggestions and recommendations for improvement of the current evaluation system adopted in universities in Iran the most important of which are discussed in the following:

5.6.2.1. Transparent and publicised standards

One of the very first ideas generated from participants’ words was the need for a set of transparent criteria and clear standards for evaluation of teachers. Becoming aware of universities’ expectations, to a great extent, can help teachers better identify the pathways by which they can improve themselves and fulfil the criteria set for effective teaching practices. Some of the lecturers were keen on this issue calling for publicised evaluation framework. They maintain that the details of the appraisal should be publicised so that all parties have a shared understanding of what constitutes effective teaching. A brief glance over the literature also suggests similar demands for explicit maxims or criteria for effective teaching. For instance, as Middlewood (2001) suggests, teacher performance appraisal and its relation with job-related issues such as pay and promotion need to be transparent, and hence its criteria should be explicit to everyone (p. 191). As Ali posited, lecturers
need to be aware of the values against which administrators tend to evaluate their staff:

“When the expectations are clear then your job will become easier. Because you become aware of the values by which you will be evaluated. Yeah, that’s quite important”.

In a similar vein, Hossein put forth:

“The existing Iranian appraisal model, if any, needs to be publicized to all practitioners of the field. For starters, policies must be defined and goals and objectives of teaching English need to be set. Then everything has to be directed towards those policies, goals and objectives.”

These results were firmly corroborated with the findings obtained from the survey. Indeed, 93% of the respondents contended that teacher evaluation needs to be implemented based on transparent criteria/standards. More importantly, 48.8% of this group of lecturers strongly concurred with the idea.

5.6.2.2. Self-evaluation

The importance of self-evaluation and how it can engender reflective practice behaviour among teachers have been fully discussed earlier. As Ross and Bruce (2007, p. 146) put it forward, self-assessment along with other tools such as peer coaching and external observation can contribute to teacher professional growth and hence improve their effectiveness. Whereas such a measure currently seems to be missing in the Iranian context, the majority of participants called for inclusion of self-evaluation. The following quote from Sohrab shows how he maintained that self-evaluation provides administrators with an opportunity to become aware of his values and aims:

“They should know about your aims, your objectives, okay, no matter what results what the results are, okay, perhaps the problem is with the students themselves, that they don't want to study, they don't want to learn but you had a schedule, you were so principled, you were so punctual okay uh I mean it should be included”.

The enviable reputation of self-evaluation was also discerned in the data obtained from the survey. As to one of the statements which asked respondents about their views on the importance of self-evaluation as an opportunity for reflection, the majority of respondents (95.3%) agreed on the idea with no disagreement.
Nevertheless, similar to peer-evaluation, self-evaluation seems to be given little attention in most Iranian universities. It is worth adding that many participants wished they had had such opportunities as part of their appraisal.

5.6.2.3. Formative evaluation

The summative-formative dichotomy in teacher effectiveness appraisal research has been fairly investigated. Such studies have been conducted seemingly as a response to the concern as to the extent to which administrators’ appraisals of language teachers should benefit from summative and/or formative assessment or rather an amalgamation of these two approaches towards teacher evaluation. The growth-oriented and developmental nature of formative assessment can help teachers reflect on their practices and hence promote themselves in a non-threatening professional environment. There is a wealth of literature on the contributions made by formative evaluation towards promoting teachers' effectiveness. The Iranian context is not an exception, in that most of the participants were in favour of formative dimension of evaluation rather than summative evaluation. As Campbell et al. (2004a, p. 57) assert, formative purposes designed for teacher appraisal system can enable administrators to distinguish teachers' development needs even in a particular subject. Indeed, the use of formative evaluation can be used on top of the agenda set for teacher evaluation system. Formative evaluation has received support from participants in both interview and questionnaire.

For instance, complaining about the absence of this measure in Iran, Sohrab added:

“Yeah ... predetermined standards and the second one there is, I mean, a vacancy of formative assessment in Iranian context. Yeah we don't have it. It is in the form of summative at the end of the term at the semester and there should be some gradual feedback, I mean, step-by-step feedback for teachers to know about their difficulties, positive points at any particular point of time in the semester. It may be more helpful to have more effective teachers.” (Interview 3)

Lecturers' relative preference for formative rather than summative evaluation was partly evident in their responses to a statement of the questionnaire which addressed the purpose of teacher evaluation. As the analysis of the responses
revealed, more than two thirds of respondents (69.7%) maintained that teacher evaluation should mostly focus on formative purposes.

5.6.2.4. Need for revision

Having considered the above-mentioned drawbacks associated with the existing appraisal practice in Iran, there emerged a general tendency among lecturer participants of this study for a revisited appraisal model which can accommodate teachers, students, and administrators' needs and fulfil their expectations. The following quotes show how participants called for improvement of teacher appraisal scheme. Shahab who is an experienced assistant professor of TEFL referred to the status of the basics of teacher education which from his point of view is somehow missing in Iran and stated:

“There are many essential concepts which are missing in Iran’s teacher education packages: general management principles, marketing, fundamentals of the service industries, and of course rights-based management. These vital concepts are so alien to ELT in Iran that I can only assume that the overwhelming majority of ELT specialists might ask in confusion: what on earth is the relationship between marketing and ELT?! The answer of course can be vividly detected by attending almost any general workshop on marketing”.

Armin raised a similar point contending that the current evolution is not an accurate reflection of an effective teacher:

“You know about the details I mean ok it should be reproduced redesigned because you know the criteria that are different … cannot be reflective of an effective teacher in my opinion ok”.

Referring to different dimensions of teacher evaluation, Hossein asserted that teacher effectiveness seems to be the product of educational system and added:

“First of all, the educational system has to be modified. The teacher is just part of the system; textbooks need revision; school officials must be briefed on the possible modifications; pre-service and then in-service training should be conducted. The criteria for teacher recruitment must be based on the qualifications expected from an EFL teacher to display in the educational context”.

Moreover, the analysis of the survey proposed a similar idea in that 74.4% of the respondents called for revision of the appraisal model in Iran which testifies to the results obtained from interviews. Although 9.3% of the respondents were opposed
to the need for revision, the obtained findings from interview and questionnaire show a relative discontent at the quality of teacher appraisal in Iran. The underlying reasons for lecturers’ demand for revision will be more discussed in the next chapter.

5.6.2.5. Observation

One of the most interesting and yet thought-provoking findings from both interview and questionnaire pertains to lecturers’ positions towards external observation. Not only none of the interviewees showed any interests in being observed by an external observer/appraiser, less than half of the questionnaire respondents were in favour of the inclusion of this measure in appraisal model. Indeed, about half of the lecturers (51.2%) had either neutral or negative attitude towards external observation.

Similar to some other measures of evaluation, observation is currently absent in most universities in Iran. This is why I had to explain observation to the interviewees most of whom had a vague impression of the notion. A number of factors can be conceived of as the underlying reasons especially the ones emanated from the alienated policies and practices in the context which will be discussed in the next chapter.

5.7. Non-teacher-controlled factors affecting TE

Four major themes have been presented and discussed thus far. This section addresses the fifth major theme of the present study, i.e. the factors that tend to exert influence on teachers’ effectiveness which are beyond teachers’ control. In other words, many of the ideas and issues discussed so far are directly or indirectly pertinent to teachers, students and administrators as the main stakeholders in the phenomenon. Indeed, teacher effectiveness evaluation has turned out to be of a broader scope embracing more intervening variables which seemingly need to be approached and discussed from a more holistic perspective. Suffice it to say that how it is possible to expect teacher effectiveness while some prerequisites needed for effective teaching are not yet provided for teachers. This is why some participants maintained that discussing teacher effectiveness seems to be unfruitful
unless such preliminary stages are met by policymakers. A number of sub-themes and categories have emerged from the data which are illustrated in Figure 5.10:

![Non-teacher-controlled factors influencing TE](image)

**Figure 5.10. Theme V. Non-teacher factors affecting TE**

### 5.7.1. Curriculum & Syllabi

The influential role of curriculum and syllabi was briefly referred to in the previous sections. As mentioned earlier, effective teachers are supposed to possess, along with others the minimum knowledge of curriculum development and syllabus design needed for teaching effectively. However, curriculum can be also a challenge for teachers especially the novice ones who have just stepped into the path of language teaching. Curriculum in Iran similar to many other countries is designed and developed by the Ministry of higher education (Ministry of science, Research and Technology (MSRT)) some of which was developed more than a decade ago. Given the pace of new advancement in language teaching, there appear some serious concerns about the effectiveness of the curriculum and material per se. This potentially could be a barrier to effective teaching for which
prompt measures should be taken by policymakers to make the curriculum more tuned in to both teachers and learners’ needs. *Curriculum* as a medium through which teachers can identify learners’ learners, suitable syllabus, teaching methods, developing materials, etc., as well as *syllabi* through which teachers can specify the content of instruction (Richards, 2001, p. 14) tend to be key variables which can potentially exert influence on lecturers’ teaching effectiveness. Such important aspects of curriculum were more or less mentioned by participants in this study. Increasing confidence, functioning as a blueprint and being flexible were among the qualities that participants assigned to reliable curriculum as indicated in some of the following quotes. For instance, Sohrab considered curriculum and syllabi as blueprints which help teacher find their route:

“curriculum and syllabus can be I mean something like a window, ... I mean, something that shows the way to the teacher and this make them ... not to diverge from the main points from the route that they are supposed to go, the main aim of the teaching. It gives them more self-confidence, okay, and makes them more organised in teaching”.

Highlighting the influences of curriculum and syllabi exerted on teaching, Kian highlighted the importance of washback effect on both teaching and learning. As he furthered added, curriculum needs to be informed by assessment strategies:

“Washback effect of tests coming from the curriculum and syllabi is an undeniable part of teaching. For example, if the exercises in course books are designed and/or treated in a way that makes students feel they have to prove themselves and their intelligence, it can encourage the either-or type of thinking in learners; this win or lose mentality can hinder effective learning and teaching, because it can be discouraging to some students. It appears that assessment should not only be “of learning”, but also “for learning” for it to be effective”.

Criticising the meddling outsiders, Sahand tried to articulate the importance of teachers’ freedom in designing and customising their syllabus based on their students’ needs and stated:

“The teacher is the best syllabus designer. No outside influence should force him into reconsidering his term instruction. The only thing he should be informed about is the overall expectation of the syllabus results”.

The findings obtained from interviews were in line with those extracted from the survey even through the percentage of proponents was relatively average. As to one of the statements of the questionnaire which asked the respondents’ about
their views on the importance of national curriculum and syllabi in promoting teacher effectiveness, 62.8% of the lecturers appreciated the idea. Sahand’s quote, I think, can be a frame of reference for most Iranian lecturers as many of whom tend to face the dilemma of choosing between the so-called top-down and bottom-up approaches towards syllabus design. This is presumably of high significance as teachers in Iran, similar to many other countries, need to use resources and introduce reading lists that are mostly pre-determined by the Ministry of Science, Research and Technology. Although lecturers are deemed to adhere to the suggested curriculum and need to observe its principles, they are usually permitted to use other resources. Nevertheless, the trade-off between these two is less well-understood.

5.7.2. Facilities and equipment

Another variable which can be considered as a seminal infrastructure in effective teaching is the presence of teaching facilities and equipment in teaching environment. Whereas the issue of facilities and equipment is more or less straightforward in most Western countries, it turns out to be a challenge in most undeveloped and developing countries and Iran is not an exception thereupon. In line with the exceedingly accepted role of ICT in educational contexts and students’ success, universities are now facing more pressure from the public to equip their educational settings with technological advancement such as interactive whiteboards, VLE, etc. This has led to a kind of hidden and recently more visible competition between public and private universities in Iran. Indeed, universities now use their facilities and equipment such as well-equipped laboratories, libraries, hospitals (for Medical Schools) along with recreational facilities such as gym and swimming pool to mount propaganda to attract more students and hence making more money. This issue will be more discussed in the next chapter of discussion. The importance of these elements is stated in the following quotes from Ali and Thelma:

“When you don’t have a laboratory [English Language Lab] okay how can you teach your students, for example, Listening courses, good point, how?”

“Facilities are very important. The facilities of each university are very important because when you are learning English, facilities is part of it”.
It is worth noting that the need for facilities and equipment should not be circumscribed by students’ needs such as Language Laboratory, etc. Indeed, such demand tends to move beyond the classroom settings and include teachers’ needs as well. For instance, their free full access to journal databases, internet, printing credit, along with the recreational ones as mentioned in the above lines. Making an investment in such less noticeable yet important issues can potentially yield a reasonable return. As a consequence, facilities and equipment were found to be important factors contributing to teacher effectiveness.

### 5.7.3. Salary & financial status

Financial issues have always been a challenging and yet determining issue in almost all careers including that of education. Salary and financial issues as incentives tend to be of seminal importance whose influences should not be underestimated. Whereas financial issues can give a peace of mind and instil confidence into teachers, they can impose a sheer strain on teachers’ productivity and effectiveness. I will come back to this point in the next chapter. However, the following quote from Majid shows how lecturers’ effectiveness can be hindered by financial issues:

> “And maybe, maybe financial status of students and teachers, both, for example, if I’m going to be an effective teacher, if I am financially independent, I have enough money; I can earn my living easily, so I can spend more time on my teaching. I don’t cut my classes, I pay more attention to quality of my teaching, maybe 10 hours or 16 hours whatever I must, I accept no extra classes. But sometimes due to financial crisis or financial shortage teachers have to teach 50 hours, 60 hours a week, this is a catastrophe so they have to neglect quality because only quantity here works. There is no quality”.

### 5.7.4. Student-related factors

As briefly pointed out earlier, student-related factors such as students’ age, knowledge, gender, socio-economic status as well as number of students in a classroom, among other variables, are the ones that tend to be beyond teachers’ control. The literature also resonates with clear evidence of the relationship between student-related factors and educational output (e.g. Bedard & Kuhn, 2008). While these characteristics are mostly investigated in mainstream
education, their potential influence in EFL context may not be undermined and needs further investigation thereupon. Issues such as ‘adult’ learning, students’ prior knowledge, number of students in the classroom were considered by the participants, influential in learners’ achievements. Mersedeh touches on this issue and listed a number of student-related factors towards which universities and lecturers needs to be oriented:

*And the level of the students is very important too, yeah, for example, whenever you have some syllabus, you design some syllabuses for a group of students, you must look at their background knowledge, it is very important. For example the age of students is very important. I think most of our students want to pass their courses instead of learning, yeah? And I think the number of students that are encouraged to learn English is very low. And a teacher must adapt himself or herself with this situation. And sometimes the sex of the students is very important. Some of the students are comfortable in front of their opposite sex. Some of them are not comfortable”.

Rima also articulated the centrality of students’ self-motivation and stated:

*“Some students are not motivated in learning, ok, I think this criterion this element may impact on your way of teaching because you have to have a kind of challenge with these students, ok, so having motivation in a student is very important”.*

The analysis of statistical data led to interesting results. As to statement 22 which was specifically developed to explore lecturers’ opinions about the effects of student-related factors on their effectiveness, 60.5% of respondents agreed with the notion. Surprisingly, the same percentage (60.5%) of the respondents averred that an effective teacher with a particular group of students may be less effective with another group of students. It seems that such variables, if not predicted and controlled, may account for the consequent unreliable evaluation and its pertinent interpretations. To the best of my knowledge, such student-related factors have not been considered in the evaluation system in Iran. This could be one of the ostensible reasons why the majority of participants had fairly negative attitudes towards their appraisal.

**5.7.5. University's (leadership) policy towards education**

The extent to which universities’ performance is really aligned with their mission and vision, has often been a point of contention in most education systems. In
other words, there is a cliché question as to whether or not quantity should be invested at the expense of missing the quality of education. Educational systems are now facing public demand to possess the ability to amalgamate ‘technical and economic innovation’ with ‘social and cultural conversation’ (Bates, 2008, p. 277). Bates further argues that as a result of the globalisation movement, education is now compelled to enter a ruthless competition to sustain its economic survival (p. 278). This is why many universities try to have quality teaching on top of their agenda. These are some hard questions which will be more discussed in the chapter of discussion. However, some of these ideas were posited by some participants one of whom directly accused the universities of adopting an ethos of business rather than educational approaches towards education. As Hussey and Smith (2010) assert, the outburst of ‘consumerism’ within education has left a number of impacts the most paramount of which has been imposed on students and their relationship to their teachers and universities (p. 45). In the following excerpt, the interviewee refers to an emerging conflict between her understanding of the purpose of education and that of her university:

“... their policy might not necessarily be related to my policy and commitment as a teacher. As a teacher, I must, you know, uh be honest in the message I am sending but their policy might be attracting students so I cannot think of attraction in the class. I can just think of my commitments and the requirements and the necessities, who am I and what am I doing there? That’s all. So in the macro level, the students’ policy is not related to my policy as an instructor and if they are expecting me to consider their policy they are wrong. Should I always adapt my methodology based on the university’s policy that they want to attract more and more students?” (Niloofar)

Moreover, as posited by some participants, they expect their universities to be more active and engaged in running TEFL-specific teacher development workshops, seminars and conferences through which they can get in touch with the latest advancement in the field and build a collegial network with other lecturers from around the country as presented in the following quote from Sam:

“We really need it but nobody has tried to provide me these workshops, yeah, yes. And inviting figures, international known figures or nationally known figures to come and present their, I don't know, papers or...”.
Soroush also suggested that, universities’ policies towards education and teacher appraisal should be informed by cultural, political and social aspects and contended:

“A very good question; because you know teacher effectiveness, as you said, cannot be done in a vacuum. It is affected by the social, cultural, political factors. And of course it is highly affected by the curriculum, by the language policy, by language planning. I think even these kinds of factors that I mentioned have exerted the greatest influence on teachers’ effectiveness. For example, we can measure teacher effectiveness in one country and even we can compare the same teacher effectiveness in another country by different curriculum, by different syllabus, different social cultural and economic situations. So I think that, let's say, any kinds of organizations as organizations throughout the world are affected highly by the social economic political policies, I think teacher effectiveness is highly affected by the same factors along with curriculum the syllabus and the language policy and planning, why? Because these are the, let's say, factors that shape the process of teaching”.

The importance of the leadership was also supported by the results of the survey as 72.1% of the respondents maintained that educational leadership affects teaching effectiveness. It is worth highlighting that policymakers’ policies towards education may not necessarily be in line with those of administrators and particularly those of lecturers, and herein lies the challenge as any contradictions between policymakers/administrators’ ideas and lecturers/students’ actual needs may place teacher effectiveness and hence students achievements in jeopardy. These areas will be more addressed and discussed in the next chapter.

**5.8. Lecturers’ ideal/critical appraisal model**

The categories which are to be reported for this theme are indeed a reiteration of the categories and subcategories which have thus far been reported for the previous five major themes especially those presented in 5.6.2. As the last major theme of this study, lecturers’ perceptions and expectations of an ideal appraisal model which can afford a comprehensible evaluation of teaching effectiveness, is to be presented in this part. However, thinking and talking about such an ideal model tends to be tricky and challenging in nature especially for the concept of criticality at which arriving a consensus seems to be more demanding. This is why I removed some of the statements pertinent to this theme in the close-ended questionnaire and moved them to the open-ended section of the questionnaire and
interviews. This helped me better communicate with participants and explore different aspects of their ideal appraisal model which can critically assess their potentialities and capabilities. Most participants gave an embellished account of their ideal appraisal model which at its core embraced fairness. Indubitably, such a model can minimise the seeds of doubt and disappointment among teachers which in turn can be superseded by a ray of hope for fair evaluation. These aspects will be discussed in more details in the next chapter of discussion. Nevertheless, having analysed the data, I came across with a number of ideas emerging from the participants’ words as illustrated in Figure 5.11., as follows:

**Lecturers’ views on Ideal/Critical Appraisal Model**

- Contextualised (Social, Cultural & Political values)
- TEFL-specific
- Transparent & fair standards
- Reliable
- Appraisers (external/internal observer, peer, students): Knowledgeable, fair-minded, accredited
- Formative evaluation
- Moral Issues
- Open to criticism

*Figure 5.11. Theme VI. Ideal/critical appraisal model*

### 5.8.1. Contextualised appraisal model

The need for ‘contextualisation’ of the model adopted for teacher appraisal was one of the earliest ideas emerging from the data. As Courtney (2008, p. 547) asserts, defining and contextualising what counts as quality education rests at the heart of evaluation. As she further argues, in order to have a more impeccable understanding of the findings, the data obtained from measuring the quality in education should be contextualised within the context from which the data is collected (p. 548). In their proposed ‘contextualisation agenda’, Thrupp, Lupton, and Brown (2007, p. 111) also advocate the centrality of context in school effectiveness and improvement research. Some of the values highlighted by
participants from whose perspectives need to be considered in designing and developing evaluation systems and appraisal models are socio-economic, cultural and political values as indicated in the following quote:

“A very good question; because you know teacher effectiveness, as you said, cannot be done in a vacuum. It is affected by the social, cultural, political factors. And of course it is highly affected by the curriculum, by the language policy, by language planning. I think even these kinds of factors that I mentioned have exerted the greatest influence on teachers’ effectiveness. For example, we can measure teacher effectiveness in one country and even we can compare the same teacher effectiveness in another country by different curriculum, by different syllabus, different social cultural and economic situations. So I think that, let's say, any kinds of organizations as organizations throughout the world are affected highly by the social economic political policies”. (Soroush)

Similar to the data obtained from the interviews, the statistical analysis corroborated the idea as 65.1% of the respondents called for the need for an appraisal model that can accommodate the cultural, political and social specificities and intricacies in Iran.

5.8.2. TEFL-specific

In line with the participants’ demand for a contextualised appraisal model, they did call for a TEFL-specific appraisal scheme. Amongst their reasons for the need for such a discipline-specific evaluation was the idiosyncrasies and delicacies associated with TESOL/TEFL classrooms. From their perspective, EFL teachers have to teach courses in a language other than learners' mother tongue. Although one might argue that in TEFL, medium of instruction does not tend to be a concern and challenge for teachers and learners, the fact that many of the learners are not well-prepared for being taught in English language should not be disdained. This issue will be more discussed in the next chapter. Referring to the differences between lecturers from different academic disciplines, Majid called for what he named field-specific model and stated:

“I think it should be classified according to the major of the lectures. There should be a kind of difference between someone who teaches English and someone who teaches Arts or someone who teaches Engineering. It should be field-specific, I think”.

The analysis of statement 3 in the questionnaire which asked the respondents whether they concurred with the view that evaluation needed to be informed by TEFL-specific subject criteria, 65.2% of lecturers expressed their approval. Although the findings contained much sound common sense, to the best of my knowledge, all academic staff regardless of their academic discipline (specialty) are evaluated through a similar appraisal system. However, this issue is more complicated as it may look in that there exist some other concerns to consider, e.g. the feasibility of devising and utilising different evaluation schemes for different departments. These areas will be more discussed in the next chapter.

5.8.3. Transparent & fair standards

As discussed earlier, many of the participants admitted they were either not or less aware of the mechanism and dynamics of teacher appraisal which in turn tends to leave a vacuum in their efforts for promoting their effectiveness. As suggested by the obtained data, teachers are quite willing to be informed and briefed of their universities’ expectations of effective teachers. Awareness of such expectations allows teachers to disentangle themselves from any sceptical beliefs about what counts as an effective teacher and helps them shape and reach their ideal horizon with less pressure. Indeed, as a rudimentary prerequisite for teacher evaluation, designing, developing and publicizing a set of fair and transparent standards or criteria can help teachers feel more at ease. These issues have been more or less addressed by some participants. For instance, Niloofar stated:

“Exactly, there are two strategies: one is to catch someone red-handed and two is to make improvements. If we are after improvement, everything must be crystal-clear, I must be aware of what are the shortages to improve them but if someone is catching me then this is the way that is done, and this is the strategy”.

In her description of the possible approaches towards teacher evaluation, Niloofar eloquently highlighted the importance of the purpose and clarity of appraisal process. As she added, teachers need to be treated with respect and need to be informed of the privileges of their appraisal so as to provoke them into active engagement. Despite the merits of a transparent evaluation system on which universities need to capitalise, universities in Iran, as stated in Section 5.6.1.5.,
seem to be rather ignorant of the merit of a transparent and fair teacher appraisal. The statistical analysis of the survey also provided compelling evidence as 93% of the respondents demanded clear and fair appraisal. These findings are in accordance with the earlier results in Section 5.6., which addressed teacher evaluation system in Iran.

5.8.4. Raters’ (as appraisers) knowledge

Teacher appraisal and its pertinent models and evaluation system are complex issues. Not only an appraisal model itself should be flawless and as comprehensive as possible, but also stakeholders who have a stake in the appraisal such as students and appraisers themselves should be trained and knowledgeable enough to perform the evaluation. As suggested by most participants, they expect their evaluation to be implemented by appraisers who are knowledgeable, fair-minded and accredited. As they argued, appraisers need to have knowledge of ELT and be fully qualified and accredited by universities. In the same vein, students as evaluators also need to be instructed how to do the ratings. Such training could raise students’ awareness of the objective of their ratings so that they do their ratings more seriously and realistically. These are partly introduced in the following excerpt:

“The one who has got the knowledge, knows the society, knows the receivers, the knowledge of … . Yeah, because honesty would not be sufficient, I am talking about fluency. So how can they judge my fluency, must be from this society [EFL lecturers] to whom I can speak in one language. We would have one I mean common codes to speak about form this society who knows both things; instructors’ potentialities, the instructors’ background as well as the receivers; the receivers the students are also important, the one who is aware of the students’ potentiality that would not consider it too idealistic. I cannot make a miracle with this wall. This wall with the greatest teacher will not be changed, so if the vice-chancellor would expect me to make a miracle in the class, he is totally wrong……..It’s some personal characteristics, a fair-minded a fair-minded individual is needed to be a judge and that is a very a very shaky position. It’s thought something that you can just be relaxed and do that job; it is a very shaky position, it cannot be that fair-minded.” (Niloofar)

A large number of lecturers (95.4%) agreed with the view that colleagues who evaluate a faculty need to be skilled in evaluation 51.2% of whom strongly agreed with it. To the best of my knowledge, this category is rather missing in Iran as
neither the participants nor myself have ever heard about any workshops or training opportunities pertinent to teacher appraisal in which students, lecturers and colleagues are briefed on how to do a faculty ratings. Since raters’ awareness of methods and strategies of teacher appraisal tends to play a seminal role in the overall appraisal scheme, I will return to this point in the next chapter of discussion.

5.8.5. Formative evaluation

Research on the purpose of teacher appraisal has a relatively rich history in the literature. Many scholars and researchers have debated upon the goal and consequences of teacher appraisal. A review of the literature introduces two dominant and yet controversial purposes for teacher evaluation, namely summative and formative objectives. From most of the participants’ points of view, teacher evaluation should include formative purposes which could be used for improving one’s teaching practices as indicated in the following excerpt from Parham:

“It should be formative one in order to help teachers to improve the quality of their work the quality of what they’re doing, their performance, so it should be informative for the teachers themselves not for promotion not for I do know anything else, getting more salaries or improving their status at their college or university, formative purpose”.

The analysis of two statements of the questionnaire addressing lecturers’ views about the importance of formative and summative evaluation also corroborated this finding. Whereas 69.7% of respondents expressed their support for formative evaluation, only 34.9% of the lecturers expressed their concurrence with summative evaluation. Despite the popularity of formative evaluation, there appears that teacher evaluation in Iran lacks such an approach towards evaluation given the fact that none of the participants could recollect such developmental opportunities.

5.9. Summary

A detailed account of the analysis of the two sets of data (interviews, close-ended and open-ended questionnaire) was presented in this chapter. Given the fact that both interviews and questionnaires were developed thematically based on the research questions, I had an invaluable opportunity to investigate any similarities
and/or differences in participants and respondents’ views about an idea based on the data garnered from the thematic and statistical analysis of the interviews and the questionnaire, respectively. However, some categories were found to have their roots in only one set of data, i.e. either interviews or questionnaires. The analysis of the findings suggested six major themes each of which embraced a number of categories and subcategories. The chapter provided an analysis of lecturers’ understanding of teaching effectiveness and their perceptions of teacher appraisal. Based on the analysis of the qualitative findings and the quantitative results, a gap between lecturers’ perceptions of an ideal appraisal model and the existing evaluation system in Iran, among other noteworthy issues, was identified. The underlying reasons for such a gap will be discussed in the next chapter for which a number of solutions will be proposed in the final chapter of conclusions.
Chapter 6: Discussion
6.1. Introduction

Drawing on the findings reported in chapter 5, this chapter discusses the most important dimensions of teacher effectiveness and appraisal in higher education. Aiming at discussing the findings from a rather critical perspective, this chapter endeavours to provide in-depth and profound insights into the very nature of teacher effectiveness and its pertinent appraisal mechanism within a wider education system in Iran. Amongst several themes and categories raised and reported earlier, this chapter will merely focus on the key findings which are, to the best of my knowledge, either given scant attention or have gone completely unnoticed by policymakers and administrators in Iran. The key dimensions of the findings which are addressed in this chapter include the following themes:

- What makes an effective teacher?
- Measures of evaluation,
- Ways to improve TE,
- The Iranian appraisal system; challenges & opportunities
- Some further issues to consider

The proposed appraisal model suggested by the researcher which will be introduced in the next chapter is highly informed by the above-mentioned dimensions of TE in HE. Such a model/framework is indeed a reflection of EFL lecturers’ understanding and perceptions of teacher effectiveness and teacher appraisal in HE. Since the analysis of the data yielded myriad categories and subcategories, it was rather challenging to decide on the categories which are well worth discussing. As a consequence, every endeavour has been made to select those ideas which seem to be of utmost significance in reference to the Iranian context.

6.2. What makes an effective teacher?

The characteristics of an effective teacher are the mainstay of this research. Teacher evaluation as a ‘system’ is an intricate phenomenon the understanding of which requires educators to ponder over their approaches towards teacher appraisal (Peterson, 2000, p. 35). A brief review of the literature reveals that
research on teacher effectiveness and its germane appraisal models have roots in the very questions of ‘what counts as an effective teacher?’ or ‘what makes an effective teacher?’ Since teacher effectiveness varies widely (Rockoff & Speroni, 2010), it seems unwise to answer such questions simplistically based on a list of characteristics. In other words, the central question, as Arthur et al. (2003, p. 275) put is forward, would be ‘effective in terms of what?’. It is a truism to say that an effective teacher should be possessed of a number of qualities and characteristics. However, only after addressing effective teaching in a wider higher education context can one understand the dynamics of teacher effectiveness. This section focuses on the qualities and traits that are perceived as the characteristics of an effective teacher.

6.2.1. Personal attributes and interpersonal skills

The personal traits explored in this study embrace qualities such as friendliness, verbal abilities, fairness, motivation, dedication, support, patience, confidence, adaptability, etc., most of which affirm the literature. Since a detailed account of this category was reported in the previous chapter, I am not discussing any particular traits individually. Yet, much of the discussion in this section focuses on the importance of personal and interpersonal qualities for a highly accomplished teacher. Whereas such characteristics are mostly intrinsic to one’s personality which might not be easily changed (e.g. friendliness), others could be taught and instilled into teachers’ minds (e.g. patience, confidence). Participants’ sensitivity to personal qualities sheds some light on the contentious concern about the traditional yet prevailing prioritisation of teachers’ technical performances over their personal attributes which can equivalently exert impacts on teaching quality. Effective teachers are exceedingly expected to adopt new roles in their career. As Korthagen (2004, p. 82) reminds us, teachers in the newly emerged era are considered as someone who is expected to guide students rather than merely transferring knowledge. It is my contention that hardly can teachers guide their students (e.g. academic advice or pastoral care) without proper interpersonal skills.

In fact, effective teachers take advantage of their own enthusiasm as a tool to motivate their students (Stronge et al., 2004, p. 36). These qualities seem to be
one of the cornerstones of teacher professional success. However, sustaining such a high level of enthusiasm and motivation can be easily prone to vicissitudes during one’s teaching career life and hence is a formidable quality to acquire. As Townsend & Bates (2007, p. 15) argue, positive qualities of pre-service teachers such as optimism and confidence turn into disappointment as they become newly qualified teachers.

One controversial aspect of the personal traits identified in this study which appears to hold true for the literature is a lack of a benchmark for each characteristic. This is of a substantial importance, inasmuch as it may lead to conflicts of interests among different stakeholders. For instance, how much patient, friendly, etc., a teacher needs to be in order to be considered as an effective teacher? Since such norms are rather influenced by the society, e.g. Iran, the role of socio-cultural variables in shaping a common-sense approach towards effective teaching should not be underestimated. While the majority of the qualities of an effective teacher such as the pedagogical and professional skills are of concrete nature and hence can be easily evaluated through appraisal schemes, the personal ones which are less concrete seem to be difficult to measure. Although qualities such as being inspirational and open are less measurable, they were endorsed by most participant lecturers in this study as inseparable constituents of effective teaching. These findings are in line with the literature. For instance, Minor, Onwuegbuzie, Witcher, and James (2002) in their study on 134 pre-service teachers identify “enthusiastic about teaching” as one of the seven characteristics of effective teaching.

6.2.2. Teachers’ knowledge

Perhaps, Shulman’s (1987, p. 8) categorisation of the so-called knowledge base could serve as a springboard for identifying different dimensions of teachers’ knowledge required for effective teaching. Teachers’ knowledge base as a professional asset is an indispensable aspect of effective and high quality teaching. Teachers’ critical L2 knowledge serves as a precious tool to promote students’ awareness of contextual dimensions of L2 learning, i.e., the socio-cultural, political and economic implications (Troudi, 2005, p. 125). As reported in the previous
chapter, several aspects of teachers’ knowledge were identified which were mostly in line with Shulman’s categorisation. Nevertheless, given the focus of this study, there needs to add on top of the aforementioned categories, the notion of ‘Teacher Language Awareness’ (TLA) which as Thornbury (1997, p. x) defines, refers to ‘the knowledge that teacher have of the underlying systems of the language that enables them to teach effectively’ (Cited in Andrews, 2007, p. ix). It is worth emphasising that TLA is different from teacher language proficiency in that it refers to teacher’s ability ‘to demonstrate knowledge of language from the learners’ perspective’ (Murray & Christison, 2011, p. 69). And to that end, TLA for language teachers is equivalent to subject matter knowledge for teachers in other academic disciplines (p. 69)

In addition to the above-mentioned aspects of teachers’ knowledge, the analysis of data gave rise to some other areas of teachers’ knowledge such as Knowledge of and about language as well as ICT literacy. As it was stated earlier, as an amalgamation of different dimensions of teacher knowledge, cognitive qualities were perceived by the participants as precious assets by which teachers can promote their effectiveness.

While an effective teacher is expected to have an acceptable level of knowledge in each of the above-mentioned strands, the data revealed that this is not the case in Iran. For instance, language proficiency or spoken English has been found to be a challenging issue to most participants. This simply corroborates Troudi’s (1998, p. 783) argument that ‘pronunciation’ is rather neglected in teacher education programs. Good command of English highly affects teachers’ confidence, on the grounds that their lack of proficiency especially fluency (spoken English) can easily mask their confidence which per se is a potential threat to their effectiveness. The obtained findings also reiterate Park and Lee’s (2006, p. 246) findings in which teachers’ English proficiency among different characteristics of an effective English teacher was ranked the highest. This is of high significance especially in Iran in which teachers hardly attain such an ideal level of language proficiency, given the fact that, by convention, language teachers are expected to show acceptable command of English. There are a number of underlying reasons the most
important of which emanates from their education at BA and MA levels. It seems that the four-year BA and two-year MA programmes do not tend to provide the would-be-teachers with sufficient opportunities to enhance their language proficiency.

Another important yet contentious dimension of teacher knowledge is ‘content knowledge’ (Shulman, 1987) which is sometimes referred to as ‘subject-matter knowledge’ (Capel, 2010). Indeed, the question of the extent to which a teacher should have the knowledge of what they teach has been long discussed throughout teacher education history yielding up some discrepancies. The pedagogical orientations of subject matter knowledge which in Grossman, Schoenfeld and Lee’s (2005) words, refer to teacher’s ability to anticipate, respond and exemplify topics (p. 201), seem to be of high significance in one’s teaching effectiveness.

Although this study provided evidence for teachers’ subject-matter knowledge, it did not probe into the amount/level of such knowledge. In spite of the overall agreement on the centrality of teachers’ subject matter knowledge in promoting one’s teaching effectiveness, it strained credulity to believe that the higher levels of subject-matter knowledge necessarily lead to higher levels of effectiveness. Indeed a so-called ‘threshold effect’ should be assigned to the role of teachers content knowledge and student achievements. As Monk’s (1994, p. 125) findings indicate, teachers’ knowledge of what they teach do positively exert influence on students’ learning outcome. Yet, the power of such influence tends to decline over time and across different types of learners. Such threshold level might not function in the same fashion for each aspect of teacher’s knowledge. In other words, whereas high level of ‘subject matter knowledge’ might not necessarily lead to better teaching outcomes, high levels of ‘general pedagogical knowledge’ and ‘pedagogical content knowledge’ seem to be the cornerstones of high quality teaching. For instance, whereas gaining pedagogical knowledge through self-study, TEP, TPDP, TTC, etc., was perceived as influential in promoting TE, participants’ level of academic qualification did not associate necessarily with promoting teacher effectiveness. One plausible reason that the participants
endorsed subject knowledge as an important quality, I think, is seemingly due to their lack of competence in content knowledge. My argument is based on the lecturers’ demands for redressing the current professional development opportunities in Iran which from their viewpoints are not enough to equip them with what they need.

The knowledge of curriculum development and syllabus design was another important aspect of teachers’ knowledge explored in this study. Teaching effectively is partly enshrined in teachers’ knowledge of the rudiments of curriculum development and syllabus design. As ‘how-to-do-it’ activity which embraces various areas germane to Applied Linguistics such as teaching methodology, assessment, material development, etc. (Nation & Macalister, 2010, p. xv), teachers’ practices need to reflect the minimum knowledge of the crux of such broad yet interrelated areas. Such expectation is echoed with Connelly and Clandinin’s (1988) conception of ‘teachers as curriculum planners’. Therefore, teachers need to further their knowledge of curriculum. On the other hand, administrators need to identify the aspects of knowledge in which teachers are in most need and provide them with a suitable teacher development programme or training workshops. These aspects will be more discussed in the next chapter of implications.

My findings indicated that teachers’ ICT literacy and Technology Enhanced Teaching and Learning (TETL) are in a considerable state of flux as there was a strong contradiction between the data driven from the interviews and the survey. There is a wealth of literature on the contribution of ICT knowledge to effective teaching. A number of terminologies have been coined to denote the idea of utilising ICT for educational purposes by teachers. Some of these technological advancement include Computer-assisted Language Learning (CALL) (e.g. Fotos & Browne, 2004; Levey & Stockwell, 2006; Levy, 1997), Virtual Learning Environment (VLE), WIKI, ECHO 360, Modular Object-Oriented Dynamic Learning Environment (MOODLE), WEB 2.0, WEBFOLIO, MOO and PODCAST. Nevertheless, what sounds to be given scant attention is the influence that context (environment) can bear upon the use of ICT in education. And this is, I think, one
possible reason for the interviewee’s neglect for ICT literacy despite the strong support from the survey. Although there has been a technological movement in recent years in the Iranian universities towards adaptation of technology for teaching purposes, some lecturers’ overall awareness has not progressed and promoted in parallel to technological advancement. In other words, while the infrastructure required for technology enhanced teaching and learning is much improved in Iran in recent years, hardly has the mentality of lecturers much changed. There appears a kind of hidden resistance among some experienced lecturers towards the use of ICT such as audio-visual aids in their teaching. One possible reason for this could be due to the fact that ICT is not currently integral to the existing curriculum in the Iranian higher education system whereby lecturers do not feel obliged to commit themselves to utilising ICT in their practices.

6.2.3. Pedagogical skills
As the title implies, this category embraces a number of skills that assist teachers with their pedagogy. This category seems to be of an utmost importance especially with reference to the Iranian context, given the fact that currently English Language Departments in Iran recruit lecturers from a wide range of academic backgrounds, i.e., TEFL/TESOL, Applied Linguistics, English Language Literature, English Language Translation and General Linguistics, and herein lies a challenge. Lecturers majoring in English Language Literature, Translation and General Linguistics tend to have fewer opportunities to be acquainted with teaching and learning theories such as SLA as compared to their EFL counterparts. As a consequence, it seems that the responsibility to provide opportunities for this group of in-service lecturers is down to universities. Nonetheless, the intricacies associated with pedagogical skills as a complex network of teaching capabilities should not be underestimated. As McGregor (2011) asserts, pedagogical knowledge is a rather multifaceted area embracing ‘practical teaching knowledge’, ‘beliefs about teaching’ and ‘understanding of learner’ (p. 10) which can potentially shake lecturers’ confidence especially the novice ones who have just stepped into their professional career. As fully explicated in the previous chapter, the analysis of the data gave rise to a number of pedagogical qualities. Nevertheless, from among these qualities, three major areas drew my attention which I think are of
consequential importance with regard to the Iranian context. These include instructional planning & delivery, alternative instructions and assessment.

Instructional planning serves as a road map to teachers and guides both teachers’ teaching and learners’ learning practices. Successful instructional planning and delivery could help lecturers promote their productivity through different activities such as needs assessment and material development (See Seel & Dijkstra, 2004). As stated earlier, some participants endorsed focused instruction and organised teaching. This is of high significance especially for the novice teachers who tend to struggle to get along with how to start, conduct and finish a session, how to engage students, etc. The responses to these questions have their roots in teachers’ rather epistemological beliefs underpinning their teaching methodologies. Teachers’ actual and practical challenges in classroom are less likely to be extrapolated from what they had read and heard about in their Teacher Education Programme. And herein lies some concerns such as the old yet important dichotomy of ‘Direct’ versus ‘Constructivist’ instruction models. For instance, as Richardson (2003, p. 1629) asserts researching constructivist teaching per se tends to be challenging in that, constructivism is a theory of learning not that of teaching, therefore little is known about the elements of effective constructivist teaching. This is of critical importance to the Iranian context in which lecturers’ recruitment is rarely supplemented with on-the-job trainings as a consequence of which lecturers have to rely on their recollection of what they learned during their education. It seems vital for an effective teacher to be confident of his/her adopted instructional planning and teaching methods. Effective teachers can follow the ‘cycle of plan-revise-teach-assess-reflect-adjust’ (Wiggins & McTighe, 2011, p. 8).

Highly connected with the previous category of instructional planning are the notions of alternative instructions and student engagement which were considered by some participants as useful tools for increasing teaching effectiveness. These findings echo those of the literature, i.e. the contributions of alternative instructions and assessment strategies to students’ learning outcomes. For instance, based on their meta-analysis of research projects published in the U.S. from 1980 to 2004, Schroeder, Scott, Tolson, Huang, and Lee (2007) concluded that the use of
alternative teaching strategies as compared to the traditional ones tend to positively influence students’ achievement (p. 1452). An overview of the history of language teaching whereby a plethora of methods and approaches have emerged as a response to the quest for a better or even the best teaching method/approach testifies to such a need. However, contrary to its more or less advocacy, use of alternative instructions was given unusually scant attention in the interviews as though it was alien to most participants. This, I think, is an alarming signal for policymakers whose underlying reasons need to be investigated. To my understanding, there could be two main reasons for such a condition. First, teachers are not well-aware and well-informed of the strengths of using multiple and alternative teaching strategies in their classrooms. Second, even if teachers are aware of the strategies, they do not show the slightest interest in adopting such an approach in their practices which per se can have its roots in university’s flawed policies, e.g., absence of a consistent supervision/monitoring scheme. Therefore, it seems crucial for universities to fill this gap in their in-service teacher development programme which will be discussed shortly.

Similar to alternative instructions, knowledge of assessment strategies was found to be both pedagogically and professionally expedient in this study. Lying at the heart of teaching process, assessment provides teachers with an opportunity to evaluate the extent to which a lesson is effective in terms of students’ learning and engagement, as well as measuring learners’ progress which per se can be used as a basis for continuing instruction (Stronge, 2007, p. 91). Indeed, one very important yet challenging role teachers might need to take in higher education is to raise the quality and timeliness of their feedback from which learners could benefit (Irons, 2008, p. 1). Nonetheless, assessment can play all the above-mentioned roles and functions ideally, only if, as Kozma and Roth (2012, p. vi) contend, it is measuring the right things. In other words, whereas technological advancement has exerted influence on the traditional teaching, assessment strategies need to be, in parallel with teaching strategies, tailored to students’ needs in the newly emerged era in higher education, e.g. E-assessment strategies. Participants’ emphasis on concepts such as dynamic assessment and washback effect, as reported in 5.3.4.6., elucidates the importance of teachers’ awareness of assessment theories.
and practices as a teaching tool. This in turn helps lecturers use of alternative assessment which is nowadays a popular discourse in education whose values are premised upon alterations to assessment practices (Tan, 2013, p. 21). Although these concepts and notions are of high repute in the literature with works of many scholars and researchers (Lantolf, 2009; Poehner, 2008), yet their applications to the Iranian context has been given scant attention by lecturers who are deeply engrossed in their teaching practices. Although the idea received strong support in the survey, few participant lecturers mentioned the contributions of assessment techniques in the interviews. Again, this is partly down to policymakers and administrators in that not only are teachers less well-trained during their pre-service stage, but also, as maintained by a considerable number of participants, they do not receive training opportunities which can accommodate their in-service teaching needs. Whereas most lecturers were familiar with the traditional types of assessment, i.e. standardized and teacher-made tests, such opportunities can help them familiarise with alternative forms of assessment such as “performance” and “portfolio” assessment (Muijs & Reynolds, 2010, p. 266).

6.2.4. Professional skills
Professional skills as compared to the aforementioned qualities, i.e. teachers’ personal and pedagogical skills are less likely to be acquired prior to teacher practical involvement in HE and hence need to be given particular attention by administrators. Given the fact that hardly can teachers gain what they really need for their future career in pre-service education and also considering the dynamic and evolving nature of knowledge base for language teachers, a constant renewal and revision of teachers’ professional skills serve as the cornerstone of teacher development programme (Richards & Farrell, 2005, p. 1). Possessing professional skills is a precious asset for teachers especially the novice ones whereby, as Landau contends, teachers can resolve problems and mediate conflicts in their classrooms (2009, p. 743). From among different qualities raised by participants, the categories to be discussed here include teachers’ innovation and creativity, collegial networking and technological enhanced teaching & learning.
One dimension of professional skills concerns teachers’ creativity and innovation. The participants’ complete support (100%) for teachers’ innovations in teaching illustrate the importance of teachers’ intrapreneurial skills. Whereas the national curriculum and syllabi prescribe teachers ‘what to teach’, deciding upon how to teach and how to manage classroom is often left to teachers (Pritchard, 2007, p. 21) which in turn necessitates questing for a creative and innovative teacher. Enhancing their competitiveness in offering high quality education and hence attracting more students, universities tend to make a sound investment in employing teachers who are creative and innovative in their career. This is of seminal importance to universities from two perspectives. On the one hand, teachers can help universities to increase their annual student intake. On the other hand, they can serve as researchers who can absorb funds from the outside of their universities.

This research did not probe why these qualities are so accepted by lecturers while there is no apparent pressure from their universities. However, some reasons spring to mind. As competition for employment and job promotion intensifies in Iran, lecturers may have endorsed it as a desirable quality which allows them to stand out. It is worth emphasising that engendering such qualities in teachers’ practices is not as simple as it may sound. In fact, a number of prerequisites need to go hand in hand one of which, I think, could be ‘teacher empowerment’. As a key to ‘enabling risk-taking’ (Carl, 2009, p. 4), teacher empowerment contributes to positives changes and grows professionalism among teachers. Indeed, when empowered to ‘share governance and involve in decision making’ (Blase & Blase, 2001, p. 4), teachers can use their creativity to enhance students’ learning outcomes and this is an important dimension of professional skills to which both administrators and teachers need to pay attention.

One of the reasons for which collegial networking has been selected for discussion in this section is the absence of such a professional hallmark among the Iranian lecturers which seems to be mostly due to a false impression, e.g. disclosure of one’ teaching weaknesses. It has been argued in the previous chapter that as a social support network, collegial network can boost one’s career success (Milem et
Availing themselves of collaboration, teachers can benefit from it both personally and professionally (Smith, 2005, p. 201). Collegial networking has gained acceptance in the literature and is associated with a number of equivalent buzzwords in education including peer coaching, peer assessment, peer feedback, peer review, peer evaluation, etc., all of which have at their core, the idea of exchanging and sharing experience among colleagues. Some plausible scenarios could be perceived for why networking is still in a state of flux in Iran. Whereas some underlying reasons are due to teachers themselves, some others are deeply rooted in policymaking and administrative issues. One reason could be due to the fact that peer evaluation (peer observation) is not recognised as a measure of evaluation, thereby not being included in the current teacher appraisal in Iran. Given such devaluation of peer feedback on the part of universities, it is a common-sense expectation that teachers show little respect for considering such opportunity to get to know their strengths and weaknesses. Indeed, by downplaying peer evaluation, administrators unconsciously depreciate the values of collegial networking, thereby sowing seeds of distrust among their academic staff.

Another reason which seems to be playing an important role has its roots in the conventional, cultural and social issues inherent in the Iranian culture. To my own overall understanding, it seems that teachers, in compliance with a wider social context, are entangled with reserved and defensive approach towards networking which may require the disclosure of one’s weaknesses. Indeed, they tend to perceive networking as a threat to their career rather than an opportunity to become acquainted with their merits and demerits. Yet, none of the participants vividly expressed such a position. Considering peers as rivals, it appears that, teachers tend to be onto the defensive when asked to share their problems and challenges with their colleagues.

6.3. Measures of evaluation

Teacher appraisal in education is rather a microcosm of staff appraisal in other non-educational sectors such as management and business and hence is informed
by similar theories and practices. Since a well-designed teacher evaluation system can improve the quality of teaching and hence student achievement (Looney, 2011, p. 440), it is extremely important to identify each and every measure of evaluation and understand how well they can provide useful data on teachers’ effectiveness. Teacher appraisal has received a wealth of attention in the literature; however, what has been less well-accentuated is its practice in EFL rather than mainstream education context. As Buller (2012, p. xi) asserts, it is very hard to apply the same evaluation skills adopted for student evaluation to the evaluation of faculty. As Buller (2012, p. xi) further adds, evaluating a faculty is indeed evaluating a colleague, thence influenced by “collegiality” and “shared governance”. And herein lies a challenge as colleagues may find their appraisal of each other biased and subjective. Although the findings of this study corroborated those already in the literature, they raised some important issues some of which will be discussed in the following sections:

6.3.1. SETs: reliable or not?

Having a relatively old history (e.g. Adams & Umbach, 2012; Brown, 1976; Greenwald & Gillmore, 1997; Marsh, 2007; Marsh & Roche, 1997; McKeachie, 1997), SETs or student ratings have been constantly considered as a key element in teacher appraisal throughout the literature. In spite of the overall acceptance of SETs as an essential criterion for TE, there have been some validity considerations about SETs; from the early stages of development to implementation and then to interpretation of results (See Greenwald, 1997; McKeachie, 1997). Much of the criticism directed at SETs in this study pertained to students’ little acquaintance with the principles of ‘effective teaching’ and ‘faculty evaluation’ as well as ‘disruptive factors’. These findings are somehow echoed with those already in the literature. In spite of such doubts, SETs have been able to retain its centrality in many education systems around the world including that of Iran. Indeed, SETs have proved to be of a substantial help to administrators provided that appraisers take into their account the multidimensional nature of students’ ratings especially in the phase of interpretation. As Marsh (2007, p. 319) proposes, all stakeholders including faculty, students and administrators can benefit from SETs as feedback
in teaching, course selection and personnel decision, respectively. Despite all its privileges, SETs were found to be associated with some potential threats.

One intrusive variable which needs to be controlled by appraisers is the influence of factors other than teacher-related factors such as course-related variables on students’ ratings. Teachers’ marking system, as fully explicated in the previous chapters, has been found a key intervening variable. The clichéd truism that ‘the higher a teacher’s marking, the higher his/her students’ ratings’ has been carried forward by the majority of respondent lecturers. This is why the majority of lecturers agreed with the idea that SETs should be done in mid-semester. Such reservations simply echo grading leniency or marking bias as a major drawback of SETs which has been already identified in the literature. For instance, based on their data from 1244 evaluations, Brockx, Spooren, and Mortelmans (2011) proposed course grades as a strong predictor of SETs (p. 302). Similarly, Griffin, Hilton Iii, Plummer, and Barret (2013, p. 9) found a moderate correlation between the grade point averages (GPA) and SETs across 2073 courses even though they treated their findings with suspicion. The findings revealed that some participants were under the impression that students’ ratings can be potentially influenced by the extent to which a teacher builds a friendly rapport with students. Another plausible explanation for such standpoint on the part of the lecturers could emanate from the “student-as-customer/consumer” mentality which may exist in some universities. However, such deficiencies can be predicted and hence avoided which will be discussed shortly. This is important in the Iranian context, in that, student ratings is a predominant measure, as perceived by the participants, based on which teacher evaluation is implemented. Therefore, it is of high significance for administrators to minimise (if not eliminate) such intervening intrusive factors.

Moreover, from most participant lecturers’ points of view, students are not serious about their ratings and merely perceive it as a rudimentary stage for final term examination and the continuation of their registration. It is unfortunate that this cast doubts on the efficiency and validity of student ratings in the Iranian universities. As it was mentioned in one of the interviews, some students turn such huge responsibly over to others to do the evaluation (which is nowadays mostly online)
on their behalf and show little interest to have their say. This, I think, partly emanates from administrators’ failure in well informing students about the objectives of the evaluation and assuring them, that their opinions will be considered by the university. Being less well-informed of the nature and the goals of teacher evaluation, it is apparent that, students would tend to do their ratings superficially. Despite such a disappointing tendency towards teacher evaluation on the part of students, they should not be much reprimanded as they have never been briefed on the teacher evaluation and trained as to how to do a faculty evaluation. Having considered the above lines of argument, it seems that administrators need to inevitably inform students of the purpose of SETs and provide them with some training opportunities they need so that their evaluation can provide more reliable findings. SETs can serve as a reliable measure of evaluation, provided that such rudimentary prerequisites are contemplated by policymakers and administrators. However, further research is needed to make any assumption about SETs. Drawing on their review of the research reports published since 2000, Spooren, Brockx, and Mortelmans (2013, p. 598) argue that research on SETs has thus far failed to address the validity issues thereby remaining a contentious area in higher education.

6.3.2. Peer observation

Peer evaluation helps administrators gain a more holistic view and model for the appraisal of their teachers’ performances (Weller & Weller, 2000, p. 234). Surprisingly, this measure is not currently practiced in Iran, albeit all the strengths associated with the use of peer evaluation. In spite of its absence in Iran, peer evaluation won the support of participant lecturers. This is in line with the previous findings in the literature. In their study on 163 observees and 343 observers, Kohut, Burnap, and Yon (2007, p. 19) concluded that both observers and observees considered peer observation as a valuable and effective measure of teaching. It is worth emphasizing that the majority of participants in this study called for a formative rather than summative function for peer evaluation. This echoes another study which found that constructive feedback obtained from peer observation of teaching should be non-judgmental and detailed and needs to be equipped with myriad concrete evidence (Carroll & O’Loughlin, 2013, p. 8).
Nevertheless, peer evaluation similar to other measures has got its strengths and weaknesses. In line with the arguments raised for students’ ratings pertaining to the needs for training and briefing students about evaluations, colleagues themselves as evaluators, need to be familiar, well-trained and adroit enough knowing how to do a faculty evaluation; otherwise the results might turn out misleading. Moreover, the relationships between a teacher and his/her colleagues, their visible and often invisible attitudes towards each other, i.e. friendship and enmity, can potentially shed slight doubts on the objectivity of such a measure. Nevertheless, the merits and advantages of inclusion of peer evaluation in the appraisal model outperform its shortcomings.

Being enthusiastic about peer evaluation, as it was reported earlier, some participants expressed their reservations on reliance on peer evaluation as a summative measure due to its hidden dimensions, e.g. issues such as evaluators’ honesty, knowledge of evaluation, etc. Such cautious attitude seemingly seems to emanate from the lack of peer observation in universities in Iran. The findings of this study revealed that never could any participant recollect peer observation either as an observer or an observee. Evolving in a collegial, non-threatening yet constructive environment in which lecturers can safely share and reflect on their experiences and practices, it seems, peer observation as a measure worth being added to the Iranian appraisal model. This can alleviate the paucity of a multi-approach towards teacher appraisal. However, as it was raised earlier, peer evaluation similar to other formats of evaluation should not be approached simplistically as a merely bureaucratic tick-box exercise. Administrators need and should concede the required trainings and instruction to their academic staff thorough in-service teacher development programme. This will be more discussed later.

6.3.3. Self-evaluation

The importance and contributions of self-evaluation has already been investigated particularly in mainstream education. Similar to peer-observation, the current appraisal model in Iran suffers from lack of teachers’ self-appraisal through which teacher can constantly reflect on their own practices. As Poster and Poster (1991,
p. 124) contend, effective teachers tend to be in the habit of ‘self-evaluation’. Requiring lecturers to self-evaluate themselves would endow critical power and engender reflective approach in their teaching practices. In fact, as a process by which teachers can identify their values (Campbell et al., 2004a, p. 92), self-evaluation can raise teachers’ awareness of their practices and deepen their understanding of their weaknesses as it tends to be a sincere and honest mirror of their effectiveness. Ross and Bruce (2007, p. 155) introduce ‘self-assessment’ as a strategy facilitating teachers’ professional growth, provided it is supported with other professional development tools. In addition, it is worth mentioning that self-evaluation is a ‘cost-effective’ measure without the need for extensive prerequisite arrangements as are normally needed for other measures of evaluation such as external observation, SETs, etc. The findings of this study also demonstrated participants’ preference for formative rather than summative self-evaluation. Hence, as called for by some participants, it might be a good idea for the Iranian policymakers to include this measure in their teacher evaluation scheme. Self-evaluation is an opportunity through which lecturers are encouraged to give voice to their hopes, ideals, suggestions, etc. In other words, empowering lecturers through providing them with self-evaluation opportunities will sow seeds of confidence and respect among teachers, even though the measure tends to be of formative consequences.

6.3.4. Observation (External or Internal)

Classroom observation is the only evidence of teacher performance in most top-down evaluation systems (Danielson & McGreal, 2000, p. 5), even though the use of external evaluators is rather rare (Dilts, Haber, & Bialik, 1994). As a way of looking at teaching practices in the classrooms, observation tends to function at the heart of many appraisal systems (Campbell et al., 2004a, p. 99). The findings drawn from the survey demonstrated that some lecturers were rather skeptical about the inclusion of external observation in teacher appraisal. Possible reasons could be the fear of summative evaluation and its pertinent job-related consequence as well as the bias which is often associated with observation. Since observation, similar to self-evaluation and peer observation, is not currently implemented in teacher appraisal system in the Iranian universities, the inclusion of
such a measure can add credibility to the results of the appraisal. It is worth adding that by observation, it is meant administrators’ observation which tends to be different from those of peers discussed earlier.

The four major measures of evaluation which have thus far been discussed unveil some challenges ahead of teacher appraisal in Iran. However, there appear a number of important issues which exist alongside each of these measures whose potential influence on the validity of findings should not be underestimated. As outlined in the following, teacher appraisal is not simply a matter of adding up the results of a number of measures. There is indeed more than meets the eye, given the fact that teacher evaluation can lead to engendering and endangering relationships between lecturers and administrators. In this section, some other influential factors which I think have not been well-investigated in the literature, e.g. the administration and interpretation of teacher appraisal will be discussed:

Having analysed the data I noticed that some lecturers had serious reservations about their appraisal. Apprehension was evident about appraisers’ knowledge and the validity and fairness of their decisions. Participants’ references to notions such as ‘standard centre’ and ‘board of expertise’, etc., reflect some distrust of appraises’ evaluation on the part of teachers which can potentially lead to their frustrations and decreased effectiveness. One reason could be due to the fact that rarely are EFL lecturers observed by an observer with EFL-related background. In his study on the status of teacher appraisal, Odhiambo (2005, p. 413) also found that teachers had concerns about appraisers’ ability to exclude subjectivity and biases in their appraisal. Lecturers expect accredited appraisers, the ones who are well-trained enough to do a faculty evaluation, thereby reviving the lost trust. Such accreditation should be based on both generic and TEFL-specific criteria. Indeed, appraisers’ knowledge of TEFL along with generic pedagogical knowledge help them better appreciate the idiosyncrasies associated with teaching in a language other than learners’ mother tongue. Meritocracy in educational system was another issue raised by some participants, even though through different languages. It has been earlier argued that power relations and the so-called hidden power agenda can influence the overall appraisal system in any educational system. This is line
with Odhiambo’s findings according to which teachers’ most serious concern was about the their poor relationships with their appraisers after the appraisal (2005, p. 413). Therefore, identifying and mediating the unfavourable interference of power relations in teachers’ effectiveness appraisal should be at the heart of teacher evaluation models.

The next issue lies with the clarity of criteria and interpretation of the data. This question raises some concerns as to how the data is collected, who interprets the data and how such interpretation tends to be done. Teachers do care about the outcomes of their appraisal as this may lead to some job-related consequences such as tenure, promotion, higher payment, etc. To some up, the findings strongly suggested that fairness and transparency are the key elements in a successful teacher appraisal system. These findings corroborate other studies such as Kelly et al.’s (2008, p. 39) study in which they conclude that fairness and clarity of appraisal relate to higher level of job satisfaction and motivation.

6.4. Ways to improve teacher effectiveness in HE

Since the ideas to be discussed in this section apply to both lecturers and universities, I will discuss them in two sections. Although the ideas explored under this theme more or less echo those of the literature, their application with reference to the Iranian context was found to be rather contentious. As it will be argued shortly, not only should the responsibility of promoting teaching effectiveness remain with lecturers but also it should involve universities’ active engagement in providing the prerequisites. In other words, not only are teachers deemed to make every endeavour to promote themselves, universities should also feel obliged to provide enough opportunities required for the professional development of their academic staff. Indeed each party's success or failure to meet its commitment for improving teacher effectiveness should not exempt the other party, e.g. teachers from promoting themselves. Teachers, on the technical front, need to be aware of all opportunities through which they can identify their strengths and weaknesses.
6.4.1. Teachers as front-line pathfinders
As the title of this section implies, teachers, to my understanding, are the ones who should be most responsible for their own teaching practices. I deliberately selected this title as the majority of strategies the participants put forward, were in direct relation to lecturers, e.g. experience, reflective practice, etc. They need to rely on their potentials and capabilities to take advantage of whatever opportunities they may come across throughout their professional lives. In other words, a paucity of an acceptable educational infrastructure needed for quality teaching, does not excuse teachers from their liability to promote their teaching effectiveness. Questing for high quality teaching, teachers need to make every attempt to enhance different aspects of their teaching practices, e.g. personal, pedagogical and professional skills through various ways available to them. To do so, teachers need to realise their highly valued positions in higher education institutions within a wider context of the ‘audit society’ (Ramsden, 2003, p. xii). Being an effective teacher, as Ramsden (2003) contends, is an intangible asset for many academics for whom it is a source of deep satisfaction.

6.4.2. Administrators as logisticians
There has been a growing realisation that enhancing teacher effectiveness should not be simplistically conjectured as a prescriptive matter which can be imposed on teachers. Indeed, as stated in the above lines, the infrastructure of high quality and effective teaching needs to be met by universities (policymakers and administrators). Such prerequisites tend to range from simple administrative issues, e.g. a piece of equipment for teachers (e.g. digital library) to those which have their roots deep in policymaking issues (e.g. promotion scheme). As the title of this section implies, universities should play a more logistic role in enhancing their staff effectiveness. The findings indicated that lecturers expect more support from their universities. One immediate reason could be that some lecturers are rather not satisfied with the services, the support, the training, etc., they receive from their universities. With regard to the context of this study, it seems that some administrators are rather engrossed in their management duties, thereby remaining less vigilant about their leadership responsibilities. Shea (1993) assigns different roles to leaders, i.e. the ability to “inspire, think, motivate, initiate change, dictate,
take decisions, set objectives, set the pace, inspire loyalty and be self-sufficient” (Cited in Smith & Langston, 1999, pp. 10-11). Among various ways and tools available to them, policymakers and administrators, need to turn their attention to ‘initiating’ professional changes such as Teacher Development Programme for in-service lecturers and ‘inspiring’ academic staff to actively engage in such events. However, TPDP cannot guarantee ideal results as it is per se associated with a number of challenges and limitations such as teachers’ lack of willingness to attend TPDP and other similar opportunities.

As suggested by the obtained results, some participants maintained that the content of programmes and workshops do not commensurate with their actual needs. As a consequence, there appears a need for policymakers to revisit and increase the support that is currently offered to lecturers in Iran. Teachers, as the main stakeholders, are presumably expected to break through the barriers despite all foreseen and unforeseen challenges. Yet, there exist some obstacles ahead of teachers which are beyond their control for which policymakers and administrators are expected to intervene. Amongst different mediatory roles the universities can play, initial teacher education programme and teacher development programme are of indisputable significance and need to be high on the agenda. As reported in the findings chapter, ITEP and especially TPDP received strong support from both interviewees and questionnaire respondents. Surprisingly, the majority of participants were not satisfied with their teacher development programme considering it rather as hit-and-miss affair. They maintained that the current TPDP has been degraded to some rather low quality workshops in which some cliché and generic rather than TEFL-specific training is offered from which teacher can hardly benefit.

6.5. The Iranian appraisal system

In addition to the above-mentioned issues which are mostly generic to most educational systems, this study offers some noteworthy findings more specific to the Iranian teacher appraisal system. This section begins with critiquing the current
appraisal and then proceeds to discussing some of the participants’ concerns and suggestions for teacher evaluation.

6.5.1. Teacher appraisal: an enigma
The first and foremost problem that needs to be solved is the ambiguous nature of teacher appraisal in Iran, thereby being enigmatic to most participants. As the findings suggested, hardly were the participants well-aware of the processes involved in their appraisal. One plausible explanation may lie in the fact that rarely have the participant lecturers been informed of administrators’ expectations and briefed on the criteria for effectiveness. The majority of the participants’ concerns, I think, emanate from such as a lack of awareness. Indeed, this raises a question as to how a teacher may fulfil the universities’ expectations while she is not well-informed of and well-briefed on such expectations. The main question here is why teachers have been thus far left unaware of such an important issue. Lack of communication between administrators and academic staff, lack of respect for appraisers on the part of teachers and lack of respect for teachers on the part of administrators could be some possible reasons. Hence, policymakers and administrators seem to be partly liable for lecturers’ little knowledge of teacher effectiveness and teacher appraisal. As Middlewood (2001) asserts, embracing both quantitative and qualitative data, data collection process needs to be clear, transparent and open to teacher (p. 191). Nevertheless, obtaining a valid, reliable and transparent evaluation system which is fair and transparent is a gradual and time-consuming process (OECD, 2011a, p. 85) which makes it a formidable task for policymakers. For instance, Chile has spent some ten years to develop its evaluation system (OECD, 2011a). It is my contention that lecturers’ demand for a well-publicised evaluation framework is a minimal expectation. Developing a clear and fair appraisal model will also help all stakeholders involved in teacher evaluation to arrive at a shared understanding of ‘what counts as effective teaching?’ Perhaps, such a clear teacher evaluation scheme can answer most of the drawbacks and concerns such as power relations, hidden agenda, over-emphasis on SETs, etc., as reported earlier and meet lecturers’ expectations of a reliable and valid appraisal. Indeed, such clarity should include all constituents of a teacher appraisal system, i.e. form and type of evaluation, data collection and data
interpretation processes, etc., so that both teachers and administrators can share a common understanding of the dynamics of teacher evaluation. As Matthews (2002, p. 18) argues, teachers’ self-appraisal should be based upon a set of transparent, established and agreed ‘indicators of effective practice’ which need to be itemised in the appraisal. A number of different attributes and characteristics for an effective teacher appraisal have been proposed such as objective, fair, transparent, reliable, etc. Yet, as Piggot-Irvine (2003) reminds us such characteristics should not be treated as a guarantee of success as though the appraisal model will be impeccable provided these criteria are ‘turned on’ (p. 177). As she further continues, such values need to be embedded in a larger cultural context in which they shape social life (p. 177).

6.5.2. Some major challenges
As it has been then argued, the so-called one-size-fits-all approach towards teacher appraisal can potentially place the validity of any findings yielded up by the Iranian teacher evaluation scheme at stake. Describing teaching as ‘juggling act’, Whitaker and Breaux (2013), repudiate one-size-fit approaches towards teaching. Contextual issues such as cultural, social and political underpinnings exert impacts on the efficiency of evaluation systems. This in line with Walker and Dimmock’s (2000) notion in that the ‘penetration of Western policies and practices including teacher appraisal into Asian cultural contexts’ may end in failure (p. 155). Given, the nature of teaching as a multidimensional and multilayers act, it seems to be prudent for administrators to adopt a flexible and comprehensive approach towards teacher evaluation, albeit the challenges. As Campbell et al. (2004a, p. 73) remind us, there is apparently an urgent need for policymakers and administrators to move beyond a simplistic surface understanding of teacher behaviours in order to better address the multidimensional nature of teaching. Indeed, assigning a set of rather generic criteria for an effective teacher seems to be less likely capable of evaluating EFL lecturers. This partly echoes the participants’ points of view about the top-down nature of the Iranian appraisal model. Being under the impression that their voices are not well-heard by administrators, some participants, as mentioned in the previous chapter, felt marginalised in the appraisal process. The participants’ reaction was quite predictable as they had never been consulted
about their appraisal in their universities, given the fact that universities are solely the implementers of teacher evaluation scheme developed and endorsed in the Ministry of higher education.

It is commonly accepted that effective teachers in different academic disciplines, e.g. TEFL, do share a number of similar characteristics as they are all perceived to be microcosms of mainstream generic education. Likewise, one can simply argue that an effective teacher would be effective or ineffective regardless of what he or she teaches. However, this does not exempt policymakers and administrators to consider the peculiarities associated with TEFL in which the medium of instruction is a language other than learners’ mother tongue. The major concern put forward by most participants was how a TEFL lecturer and a lecturer from another department can be evaluated through a same appraisal model. There are a number of reasons why universities use the same appraisal for their academic staff the most important of which, I think, is one of the feasibility. It might not be possible for universities to develop differentiated appraisal model for each and every department. It is worth emphasising that such stances do not mean the need for a totally different and unique appraisal model, given the fact that elements of effective teaching may be more or less similar across different academic disciplines. Yet, making the current appraisal model informed of some intricacies of foreign language education seems to be a step in the right direction.

Further to the issues discussed in Section 6.3.1., on SETs, the findings of this study revealed several some major drawbacks with students’ ratings in Iran. As carriers of ‘power relations’, students tend to give their voices to educational purposes, even though their voices are likely to be appropriated, rejected and selected (Morley, 2003, p. 145). Whereas some scholars have reservations about the reliability and validity of SETs, others strongly advocate the use of students’ ratings in appraisal, on the grounds that it is implemented by students, the ones who are in close contact with teachers’ practices in the classroom. And to that end, it can serve as a reflection of teachers’ skills in which teachers’ weaknesses and strengths tend to manifest. On the other hand, the critics refer to notions such as Grading leniency, grading bias and Dr Fox effect as deep deficiencies generally
ascribed to SETs. Despite all discrepancies, SETs are generally considered as a ‘multidimensional’ and ‘reliable’ (Marsh, 2007, p. 319) source which can provide administrators with better insights into the extent to which their staff are effective and act in line with their policies.

However, due to a lack of lecturers’ awareness of teacher appraisal mechanism in Iran, surprisingly there was an overall agreement among the majority of the participants on what they perceive as ‘over-emphasis’ on students’ ratings. Accusing the appraisal system adopted by their respective universities of dramatising SETs, the participant lecturers presume that their appraisal is somehow faulty. Teachers’ reservation about the adequacy of the weight given to SETs has been found coupled with another problem, namely the appropriateness of the statements in the SETs questionnaires themselves. In other words, some of participants treated the criteria for effective teachers embedded in the questionnaires items with suspicion. This study did not probe into the verification of such criticism. However, one important explanation could be due to the fact that SETs is the only one measure the participants might have ever seen as part of their appraisal. Another reason could be the one of clarity of evaluation system which was fully discussed earlier. Indeed, in none of the interviews did I get a sense that the interviewee is aware of the real weight of SETs.

‘Grading bias’ and students’ little knowledge of the process and goals of evaluation were other reasons which were found to be precluding students’ ratings from gaining credence among lecturers. As some of the lecturers asserted, students’ seem to do the ratings in a tick-boxing fashion paying little attention to the importance and consequences of their evaluation. Despite some validity concerns, the literature on student rating is mostly in favour of SETs as a measure. This study did not investigate whether or not student rating is correlated with teachers’ marking and/or students’ learning outcomes. However, I think, the current status of higher education in Iran and more specifically the so-called “diploma disease” (Dore, 1976, cited in McCulloch, 2008, p. 171) lie at the root of lecturers’ worries about student rating. It is my contention that the current educational inflation has created a mentality among some of the lecturers that students are mostly here to
receive their academic degrees for many of whom it is very important to get high marks.

One of the most thought-provoking findings of this study was lecturers’ lack of respect for appraisers, the authorities who collect and interpret the data. These results echo other findings obtained in other educational contexts. For example, in her review of a recent teacher appraisal policy in Portugal, Flores (2012) arrived at a similar conclusion, i.e. teachers’ scepticism and lack of respect for the appraisers (p. 351). As it was stated earlier, such undeniable misgivings about teacher appraisal appear to spring from the aforementioned drawbacks, i.e. top-down approach, lack of transparency, etc. Other plausible scenarios for this could stem from the lack of alignment between teacher appraisal and the expected professional development programmes as well as the failure of ratings to represent the reality (Frase & Streshly, 1994, p. 47).

As it was reported in the previous chapter, some participants were under the impression that there exists a hidden agenda for teacher appraisal on the part of administrators. Subjectivity, bias and power relations were among the concepts and notions they used in order to demonstrate their sceptical stances towards appraisal models adopted in their respective universities. Having such serious reservations about appraisal system is tragic for any given educational system as it tends to raise lingering doubts at the back of teachers’ minds about the neglect of their sincere diligence. It is fortunate that such hesitancy can be simply met by administrators through a more transparent dynamics of teacher evaluation system. This will be more discussed in the next chapter.

The findings also suggested a sort of conflict of interests between teachers and administrators. This is of pivotal importance as imposing the externally shaped values on teachers might imperil teachers’ commitment to the values with which they pretend to comply (Campbell et al., 2004b, p. 461). Universities are now more than ever concentrating on mechanisms for attracting more students, thereby increasing their annual student intake. However, some teachers maintained that such business-oriented approaches towards education may eventually succeed, yet at the expense of decreasing the quality of education. One tenable explanation
may lie in the fact that universities in Iran have entered an era of competition for sustainable development, fund, etc., to which attracting students is central. This is not particular to Iran as such issues may apply to other educational systems as well. However, what remains most important is to strike a balance among these aspects.

6.5.3. The missing pieces of the appraisal jigsaw
As discussed partially in Section 6.3., the findings of this study revealed that forms and types of teacher appraisal were among the issues receiving criticism from the participant lecturers. Referring to the advantages and strengths of peer evaluation and self-evaluation, some participants maintained that they can avail themselves of the two measures in a non-threatening environment. Indeed, self-evaluation along with other formative techniques such as peer feedback can serve as a confidence booster to teacher especially the novice ones which in turn can promote one’s teaching effectiveness. Effective self-evaluation is the key to improvement in a successful organisation (Creemers & Kyriakides, 2008, p. 257). As a formative evaluation, self-evaluation resonates with reflective practice whose one central idea is one’s self-evaluation of his or her practices. I argued earlier that such a measure is currently missing in the Iranian context, albeit being highly reputed internationally and herein lies the problem. Indeed, as proposed by some participants, lecturers approach self-evaluation as a medium and opportunity through which they can share their values and aims (which to many of them, seem to be not well-heard) with administrators and thereby policymakers supposedly. Self-esteem is perhaps one important conceivable reason why lecturers extolled the virtues of self-evaluation. Indeed, in a context in which rarely are lecturers asked to have their say about their appraisal, it seems of an utmost importance to include self-evaluation. This measure, if conducted appropriately, contributes to lecturers’ self-esteem and encourages lecturers to give voice to the evaluation system by which they are assessed each semester. Moreover, a rigorous self-evaluation model can supersede the external inspection approach (Middlewood, 2001, p. 192), if done and monitored properly.
Similar to self-evaluation, this study showed that lecturers were mostly keen on peer-evaluation. Indeed, it is argued that the realisation of a successful peer evaluation is not necessarily confined to an official observation process through which evaluators provide administrators with summative evaluation. It has recurrently been sated that hardly could any participant recollect peer-observation during their career. Lack of communication, respect or mutual trust between administrators and academic staff and fear of biased evaluation could be some possible reasons why this measure has rather gone unheeded. Peer evaluation, in a form of informal peer review, allows colleagues to share their feedback which is mostly based on formative purposes. I think lecturers feel they need more support for teaching. Providing formative feedback for teacher professional development purposes is the crux of Teacher Review of Teaching (Biggs & Tang, 2011, p. 298).

Highly connected with peer and self-evaluation, the need for formative evaluation was one important finding of this study. Such findings strongly corroborate those of other studies in the literature. Formative evaluation underpinned by growth and development purposes has usually been compared with its rival form of evaluation, i.e. summative evaluation which tends to follow accountability objectives (McGreal, 1990, p. 41). As feedback, formative evaluation can help teachers promote their practices by shaping performance, building new practices and modifying their adopted teaching approaches (Peterson, 2000, p. 63). It allows administrators to evaluate their teachers’ effectiveness and their relations to the context in which they teach (Shinkfield & Stufflebeam, 1995, p. 22). Whereas such a practice is currently absent in most universities in Iran, the review of the literature testifies to the privileges of formative evaluation combined with a summative one. However, the fundamental questions lie with the weight that should be given to each of these evaluations. The review of the literature over the past few decades reveals that teacher evaluation systems with an emphasis on formative purposes have been successful in increasing levels of satisfaction and reinforcing reflective practice among teachers as well as sustaining accountability demands (Danielson & McGreal, 2000, p. 15). Antithetical to formative evaluation, summative approaches towards teacher appraisal mostly have direct bearing on teacher occupational issues, in that they may lead to job-related consequences.
As the findings indicated, the participant lecturers maintained that there is little room for formative evaluation in the Iranian appraisal model whose constructive outcomes have not yet been well-accentuated. Further to my previous argument, I think, the need for (further) support, concerns about the evaluative consequences of summative evaluation and perhaps lack of confidence could be some possible explanations for lecturers’ call for formative evaluation. It is interesting to see that lecturers consider formative evaluation more suitable for particular measures of evaluation. As reported in the previous chapter, they view formative evaluation more suitable for self-evaluation. This shows that they are more or less aware of the nature of different types of evaluation. Nevertheless, it seems unwise to simply consider formative evaluation as the only or more reliable type of evaluation, given the fact that the efficiency of such evaluation rests upon a number of factors one of which could be what Looney (2011, p. 444) calls ‘timely and specific feedback’. Hence, not only should administrators strike a balance between different objectives of evaluation, i.e. summative and formative appraisal, they need to take into consideration the underlying technical and administrative dimensions of teacher appraisal, i.e. development, implementation, interpretation, etc.

6.6. Some further issues to consider

In addition to the issues which have thus far been discussed, some further areas arise from this study which, I think, worth discussing.

6.6.1. Non-teacher-controlled factors affecting TE in HE

This study brings into light a number of factors and limitations exercising influence on lecturer effectiveness which are supposedly beyond lecturers’ control. Some of these factors are discussed in the following:

- **Curriculum and syllabi**

Curriculum as a medium through which teachers can identify appropriate syllabi, teaching methods and material together with syllabi in which content of instructions tend to be specified (Richards, 2001, p. 14), are two important variables whose influential effects should not be underestimated by policymakers and administrators. Curriculum tends to function as a roadmap for teachers. Yet, such a
A roadmap should be constantly revisited and refined reflecting not only learners’ needs but also those of teachers themselves. As a political and subjective process, such analysis, however, is influenced by the analyst’s ideology (Benesch, 1996, p. 736). A glance over the curriculum and syllabi of English language teaching and English literature programme at all levels, i.e. Associate’s, BA, MA and even PhD at the Iranian universities and higher education centres (available on National Education Assessment Organisation website) (Sarfasi), divulges that some of these programmes have been last updated over a decade ago. This is of seminal importance as ineffective curriculum/syllabi can preclude teachers even the effective ones from teaching effectively. Other tenable explanation, however, might lie in the fact that lecturers feel they need more flexibility. Yet, the on-going dynamic nature of TEFL courses and new advancement in English language teaching and learning require state-of-the-art curriculum/syllabi which can contribute to effective teaching.

- **Facilities & equipment**

The idea of mass access to higher education or ‘unprecedented expansion of higher education everywhere’ (Altbach, 2013, p. 8) is no longer alien to policymakers in most educational contexts similar to that of Iran. However, there remains a concern appertaining to the Iranian context as to whether such expansion embraces educational facilities and equipment. The findings showed that some lecturers expressed their discontent with their respective universities’ services including the quality of equipment and facilities as well as financial incentives. The importance of educational facilities and equipment has been elaborated on in the previous chapter. As Pritchard (2007) asserts, new equipment and facilities can even lead to new or different pedagogical approaches teachers adopt in their classrooms (p. 21). As one of the interviewees pointed out, he happened to find himself increasing his working hours merely for financial purposes which from his point of view will sooner or later lead to a burnout. Hence, it is seemingly a simple equation of investment and return. Such potential diminishing return has now compelled the policymakers and administrators’ attention in order to enter a competitive arena towards increasing their facilities and equipment.
These findings are in line with those in the literature even though little is yet known about the exact nature of the contribution of educational equipment to teacher effectiveness and students achievement. There are few studies focusing on this issue most of which pertain to mainstream primary/secondary education. For instance, based on a large scale study conducted by UNESCO researching 180,000 students at 3,000 primary education schools from 15 countries in Latin America, Murillo and Román (2011) conclude that there is a relationship between basic infrastructure, didactic facilities as well as the number of books in the library and computers and student achievement even though their weight varies from one country to another (p. 29).

In recent years there has been a hidden and more recently vivid competition among universities and higher education centres in Iran to attract more students. This relatively hidden agenda has led to an on-going rivalry among universities to promote their facilities and equipment which per se is good news. As a consequence of the globalisation movement, some universities are now seeking more international reputation. Yet, as stated by some participants, they are still waiting for more concrete and conspicuous changes in the level of services they are offered by their respective universities. As reported in the previous chapter, they expect their universities to provide them with better teaching and research facilities such as free access to journal data bases, TPDP, etc.

- **Student-specific factors**

The centrality of effective teaching in students’ achievement is extensively investigated in the literature leading to some questions, e.g. “Do teachers matter?”, “How do teachers matter?”, etc. (see Stronge, 2013). Nevertheless, it is worth emphasising that teachers are just one component of education polygon whose effectiveness requires an amalgamation of several other interrelated elements. The relationship of students’ background variables and students’ outcomes has been researched in the literature (e.g. den Brok, van Tartwijk, Wubbels, & Veldman, 2010; van Ewijk & Sleegers, 2010). Yet, much of research conducted is more focused on teachers’ characteristics rather than those of students themselves. Among different influential factors in enhancing teacher effectiveness, the ones
pertinent to students themselves are undeniably as important as those of teachers. Students-specific factors including their cognitive and personal traits, socio-economic status, etc., tend to exert impacts on one’s teaching effectiveness and hence the effectiveness of the whole educational output. As Campbell et al. (2004a) posit, a teacher might be differently effective with a group of students based on their ‘background variables’ such as gender and what they call ‘personal characteristics’ such as self-esteem (p. 7). Whereas what Campbell et al. (2004a) argue is mainly focused on mainstream education, I am highly convinced that these issues should be considered as the cornerstones of EFL teachers’ appraisal. Indeed, nowadays, it is a conventional wisdom for families living in high-income districts of the cities in Iran to enroll their children in private English Institutes prior or parallel to primary education. Yet, there is little possibility for families living in the low-income districts to benefit from such opportunities. Hence, in case of the EFL classroom in a university in Iran, it is not uncommon for a teacher to be differently effective with the aforementioned two groups of students even doing the same course. Therefore, it seems that the economic inequality among students could be one possible reason behind their different performance thereby affecting teacher effectiveness.

6.6.2. Lecturers’ ideal appraisal model

Indeed, much of the discussion in this section is an attempt to reiterate some of the participants’ expectations of an ideal teacher appraisal model. Among different characteristics ascribed to an ideal appraisal model, four major categories which were perceived as a clear manifestation of a fair, reliable and valid evaluation model are selected for discussion. Accordingly, a newly customised appraisal model will be proposed on the understanding that it may accommodate lecturers’ needs and meet their concerns. As it will be presented in the next chapter of implications, the proposed model seeks to meet lecturers’ contentions from two main perceptive. First it tries to add those measures of evaluation whose inclusion in the appraisal scheme have received lecturers’ strong support. Second, targeting at the mechanics of teacher appraisal in Iran, the model endeavours to present vividly the processes involved in teacher evaluation, thereby responding to issues such as lack of transparency, bias, fairness, etc., which were found to be at the
The heart of lecturers’ concerns. Adopting a more realistic stance, the proposed model will be most focused on practical rather than idealistic expectations of an effective appraisal model. Having considered how the proposed appraisal model could alleviate the teacher evaluation challenges in Iran, this section will address the most seminal dimensions on which there is still room for improvement.

- **Contextualised appraisal model**

The need for a ‘contextualised’ model which can best accommodate lecturers’ needs have been found to be the cornerstone of lecturers’ perceptions of a reliable and valid appraisal model. Lecturers’ appeal for a contextualised evaluation scheme is in line with the literature with strong support from various researchers. Since teaching is best measured based upon teachers’ own objective of teaching, what counts as effective teaching in one context could be considered as ineffective or less effective in another context (Brown & Atkins, 2002, p. 4). Indeed, defining and contextualising what counts as quality education is yet the mainstay of monitoring and evaluation tools (Courtney, 2008, p. 548). Similar to Courtney who calls for contextualising the evaluation within the culture from which the data are collected, Thrupp et al. (2007) draw attention to what they entitle as ‘contextualisation agenda’ giving prominence to ‘context’ (p. 111). The evaluation of the performance of the education system needs to consider the role of social, economic and political context; otherwise such an approach can be thrown into question (Campbell et al., 2004b, p. 452). Imbued with social, cultural, economic and political orientations, the need for a contextualised appraisal model seems to be inexorably inevitable. The preponderance of some teacher appraisal models which have been devised and developed in Western contexts, casts doubt on the extent to which the exercised influences of cultural, social, political and socio-economic considerations could be monitored and controlled by administrators/appraisers. Indeed, decision-making authorities in higher education cannot simplistically adopt a model from a Western context and put into practice in a Middle-eastern context. Repudiating the penetration of Western teacher appraisal policies and practices into Asian context, Walker and Dimmock (2000, p. 175) warn us against adopting appraisal policies and practices from one context and cloning them in different contexts and cultures. Great delicacy and
understanding need to be taken into account in teacher appraisal in contexts in which teacher-student interpersonal relationships are affected by social and cultural conventions. It is worth emphasising that it seems unwise to simply downgrade the importance of contextualisation to Western-Eastern dichotomy. Apparently, teacher appraisal needs to be researched as part of a wider context of teacher education and teacher performance in higher education which per se is a microcosm of a rather broader domain of social context. Hence, the socio-economic context of teaching and the type of educational context tend to be indubitably a significant yet less well-understood issue which needs to be considered by policymakers and administrators. In other words, not only should an appraisal model be informed by cultural and social conventions of the country in which they are implemented, but also they need to be customised based upon the socio-economic context even within the same country, i.e. private vs. state universities.

- **TEFL-specific model**

As reported earlier, further to their demand for contextualisation, most of the participants accentuated the need for a discipline-specific, i.e. TEFL-specific appraisal model which was thwarted by the so-called 'one-size-fits-all' approach. From lecturers’ points of view, administrators should be deemed to adopt an appraisal model that can account for the challenges a language teacher might come across while teaching. Whereas one might argue that an effective teacher would be effective/ineffective regardless of whatever she teaches, it is worth emphasising that teaching in a language other than learners’ mother tongue for sure tends to add new challenges on top of all challenges non-TEFL lecturers may come across during their teaching practices. This is a ‘true’ picture of the Iranian context, in that the majority of students are less well-prepared for being taught in English especially in the undergraduate programmes. This aspect of teacher appraisal seemingly has escaped the attention of authorities. These findings are in line with the need for discipline-specific teacher appraisal in the literature (e.g. Hill & Grossman, 2013). Nonetheless, it is worth highlighting that the dearth of research in the domain of TESOL/Applied Linguistics is another factor extolling the importance of these findings. Indeed, the majority of studies conducted on teacher
appraisal so far have addressed teacher evaluation in mainstream education primary/secondary education with few studies focusing on TESOL in higher education.

- **Ratings/evaluation training**

So far different issues which can help administrators develop a more comprehensive appraisal model have been introduced and discussed. Yet, developing a comprehensive appraisal model considering all the issues raised so far will not suffice to yield impeccable and comprehensive results. Developing a solid appraisal model is one side of the coin of which the other important yet overlooked side tends to be *appraisers* themselves. Whosoever is deemed to do the evaluation in any formats, i.e. external observation, peer evaluation, students’ ratings, and teachers’ self-evaluation, etc., needs to have the minimum knowledge required for doing a faculty evaluation. There has been a growing consensus that faculty evaluation is, in one way or another, a kind of peer evaluation in which teachers’ defensive attitude towards being criticised by a colleague is inherent (Buller, 2012, p. xi). Different interpretations could be drawn from such an approach towards teacher appraisal. As Buller further argues, evaluation can then be interpreted quite differently, e.g. applying a set of standards, offering advice or just inappropriate meddling (2012, p. xi), and to that end, this may lead to an atmosphere of mistrust and misgivings among teachers. Such concerns give rise to some questions as to whether or not students are well-informed of the processes and objectives of teacher evaluation or if they are well-briefed on the consequences of their ratings. Indeed, evaluating a faculty whether as a professionally accredited appraiser or as a student or peer should not be approached simplistically as a matter of tick-boxing a set of pre-defined qualities.

Having considered the above lines, it is worth recapitulating that, such trainings not only need to be offered to those who are directly involved in observation, evaluation, ratings, etc., such as observers, peers, students, but also those who are involved in post-observation/evaluation phases, e.g. academic or administrative body who are involved in interpreting the data, need to be also informed and briefed about ‘how to do a faculty evaluation?’. There are a number
of ways through which policymakers can ameliorate this situation which will be discussed in the next chapter.

6.7. Summary

This chapter was aimed at focusing on the crux of the findings reported in the previous chapter. The obtained findings were discussed in reference to the context of this study and the exiting literature on teacher effectiveness research. The chapter started with discussing the prerequisite question of what makes an effective teacher and then proceeded to the measures of evaluation. The strategies through which teachers can promote themselves along with the so-called non-teacher-controlled factors exerting impacts on teachers’ effectiveness were discussed afterwards. Participant lecturers’ understanding of the strengths and weaknesses associated with teacher evaluation system in Iran followed by their ideal appraisal model were the later themes elaborated on and discussed in this chapter. Notwithstanding the majority of findings more or less were echoed with those of other studies in the literature, there emerged a number of rather new ideas appertaining to the status of teacher appraisal in the Iranian higher education context within a wider socio-cultural and socio-economic context of teaching. It has been argued that teacher appraisal is a rather less well-researched and well-understood phenomenon in many Asian and in particular Middle-east contexts including Iran. Findings of this study suggested that there are still some problems with teacher evaluation. Important amongst others are lack of alignment between teacher appraisal and teachers’ professional development, little transparency, lack of some types and forms of teacher evaluation.
Chapter Seven: Implications & Conclusion
7.1. Introduction

This chapter provides some implications for the main stakeholders in teacher effectiveness evaluation in universities and higher education institutions in Iran. The drawn implications are perceived to be important in the Iranian context in which the history of teacher evaluation research is rather new. Notwithstanding teacher evaluation, as an educational activity, has its roots back in Socrates time, it had not been systematically researched even until the 1960s after which teacher evaluation drew researchers and practitioners’ attentions as a tool for promoting teaching and learning (Danielson & McGreal, 2000, p. 15). It is argued that the TER in Iran, similar to many other countries around the globe, is not yet well-understood, and to this, there is still room for improvement. The implications are of both theoretical and practical (pedagogical) nature and include those ideas whose absence has been found to be conspicuous in the context of the study. Having discussed and reflected on the drawn implications, the chapter will then proceed to my proposed appraisal model or framework which is informed by the findings of the study. The future research agenda pertinent to teacher effectiveness and teacher appraisal as well as some suggestions and recommendations for further research will be discussed subsequently. Finally, the chapter will proceed to a conclusion section followed by a personal reflection on my PhD journey.

7.2. Theoretical and pedagogical implications for main stakeholders

The findings of this study raised some important theoretical and practical/pedagogical implications. Whereas most ideas generated through data analysis were more or less in agreement with the existing literature, some were found to be noteworthy in reference to the context of the study, Iran. In the following sections, I will discuss the implications according to the main stakeholders to which they apply, i.e. policymakers, administrators and teachers. Indeed, the proposed implications are mostly directed towards the future of macro and micro-policy on teacher evaluation in the Iranian higher education system. Indeed, the majority of the implications are perceived to revolve around the
ministerial (policymakers) and managerial (administrators) levels as illustrated in Figure 7.1., as follows:

7.2.1. Implications for policymakers

Teacher appraisals at universities are, in fact, part of a wider teacher evaluation system which is often considered as a national scheme. Such teacher evaluation schemes are devoted, tested and applied to all universities and higher education centres across the country. Adopting a common-sense approach towards teacher appraisal, such a national scheme can minimise, if not eliminate, the discrepancies over ‘what counts as an effective’ and promote a shared understanding and perception of the qualities of an effective teacher. As it has been constantly argued throughout the previous chapter, there do exist some influential variables which are beyond teachers’ control and in some cases even beyond administrators’ control. Moreover, as the findings suggested, not all types and forms of teacher appraisal are currently implemented in Iran. Therefore, it seems inevitable that the Ministry of Science, Research and Technology needs to revisit its current policies, thereby realigning the policies and practices with the needs, interests and wants of all the stakeholders.

![Figure 7.1. Implications for main stakeholders](image-url)
As pointed out in the above section, designing, developing, piloting and administering an evaluation system and proposing a subsequent appraisal model are the prerequisites of a successful teacher evaluation. Yet, such an appraisal needs to be informed by and tailored to the needs of lecturers as the main stakeholders who are most affected by it. A second step towards shaping and making such a system effective is to adopt it as a national scheme with enforcement power. Departments, councils, committees, etc., involved in teacher evaluation should not be downgraded from a powerful organisation to something of a bureaucratic centre, committee, council, etc. As reported and discussed in the previous chapters, some measures of evaluation are completely missing in practice even though they are suggested by MSRT. Drawing on Ahmed and Sayed’s (2009, p. 213) notion of “the policy-implementation nexus”, I’d like to bring to the fore the idea that the adequacy and utility of policies are contingent upon their appropriate implementation. Nevertheless, such a movement does requisite decisive intentions and momentous actions on the part of ministerial policymakers to guarantee its success, albeit the obstacles and challenges. It seems unwise to surmise that reaching such an appraisal scheme is simply a matter of generating some standards/criteria for effective teaching or refining the existing adopted approach. As the literature on teacher effectiveness suggests, designing, developing, piloting and conducting a successful mechanism for teacher evaluation is a time-consuming and formidable issue which requires a huge budget. For instance, it took Chile some ten years to implement its evaluation system (Establishing a Framework for Evaluation and Teacher Incentives: Considerations for Mexico, 2011, p. 85). As a consequence, I would suggest that the organisational bodies in the Ministry of Higher Education such as the Centre of Supervision and Assessment in HE, play more active and influential roles in promoting teacher effectiveness in universities in Iran.

- National Curriculum and syllabi

Although this study did not probe into the effectiveness of curriculum and syllabi, the findings indicated that most participants assented to the centrality of these
elements in their effectiveness. Updating the national curriculum as a macro-policy-oriented issue may leave teachers more room for maneuver. Therefore, it is suggested that administrative bodies within MSRT such as ‘Office of Planning in HE’ improve the curricula and syllabi based on the latest advancement and trends in the field of Applied Linguistics and its pertinent domains, e.g. TEFL/TESOL. Updated, flexible and enriched curriculum and syllabi help teachers rest assured that their needs as well as those of their students would be met by the newly revisited curriculum.

- **Theory-practice praxis**

Further to my interviews with participants, it was brought to my attention that some participants were not satisfied with their postgraduate studies and did maintain that they were not well-prepared for high quality teaching practices. Given the various types of ELT programmes in different countries, e.g. TEFL, TESL, TESOL, Applied Linguistics, etc., which might embrace different courses and degrees, it is worth emphasising that by postgraduate programme, it is meant MA in TEFL whose purpose is to prepare students for their future language teaching career in higher education. The very mission of second language teacher education (SLTE), as it is argued, is to promote the professional development of language teachers (Johnson & Golombek, 2011). Nevertheless, teachers at their pre-service stage might not necessarily be provided with everything they may need to know (Richards & Farrell, 2005, p. 1). This is in line with the current global understanding of teacher education, in that as Johnson & Arshavskaya (2011, p. 168) argue, teacher education programme has been accused of segregating theory from practice. In view of participants’ dissatisfaction with their trainings especially their in-service professional development, it is presumed that current teacher education programme might not be capable of providing teachers with everything they may require in authentic teaching environment.

A rather concrete evidence for the faulty relationship between theory and practice can be traced back to the courses student teachers do in their postgraduate programmes among which *practicum* seems to be of high significance. As a practice of teaching experience arena in which knowledge and skills evolve,
practicum tends to lie at the heart of most teacher education programme (Richard & Nunan, 1990, p. 101). It has been argued that not only should language teacher training programme benefit from education training but also they need to be replete with wealth of practical trainings in what Pennington (1990, p. 134) calls “tools of the teaching profession: in methods, materials, curriculum and evaluation”. As Pennington further reminds us, the separation of these two aspects of teacher education programmes can hinder the effectiveness of theoretical and practical dimensions (p. 134). This is seemingly an accurate reflection of the programme in Iran in that students need to do a single course entitled 'Practicum' at the end of their studies which indubitably would not suffice for their practical needs and herein lies the challenge of bridging the gap between theory and practice. Therefore, more practical teaching experiences should be added to such academic programmes.

Activities such as microteaching can help the would-be-teachers gain both experience and boost their confidence for their future real teaching career. Providing teachers with an opportunity to plan and teach 'mini-lessons', 'microteaching simulation' has been recognised as a standard step towards filling the theory-practice gap (Johnson & Arshavskaya, 2011, p. 168). Nevertheless, it has also witnessed criticism over its authenticity, in that microteaching is not capable of simulating 'real' teaching which per se has its roots in the paucity of social, institutional and historical elements inherent in 'real' teaching (Johnson & Arshavskaya, 2011, p. 169). And this is why Johnson and Arshavskaya call for conceptualising microteaching activities from a sociocultural perspective (p. 185).

7.2.2. Implications for administrators

The importance of academic and professional teacher education programmes in relation to shaping teachers’ professional identity was discussed in the previous section. However, as a single component of teacher effectiveness within a wider complex network of education effectiveness, such programmes need to be supplemented with subsequently continual support from administrators for training opportunities such as Initial Teacher Education and Teacher Professional Development Programme which are often designed by universities with this end in
view. Yet, as it will be discussed shortly, lecturers in Iran are deprived of having such precious opportunities through which they can gain what they need for conquering their fears and challenges in their classroom. In the following section, a brief overview of the mechanism of lecturer recruitment along with the trainings they may receive in universities and higher education centres in Iran are discussed:

- **Initial Teacher Education Programme (ITEP)**

A successful teacher professional development programme, amongst others, provides any educational enterprise with a golden opportunity to improve and maximise the quality of their ultimate goal, i.e. student learning (Johnson & Golombek, 2011). Nevertheless, even such successful programmes can hardly serve all the needs and interests of teachers. As stated earlier, teaching opportunities such as microteaching cannot impeccably simulate the real teaching challenges. Therefore, there has been a growing realisation among universities and higher education centres that teachers especially the novice ones do require some trainings prior to their real teaching practices, i.e. Initial Teacher Education Programme (ITEP). However, the notion of ITEP is highly imbued with school contexts rather than that of higher education. Yet, similar to their counterparts at schools, lecturers also do need such fundamental trainings. Whereas teacher training programmes such as PGCE, PCAP, etc., are straightforward and common with a rich history in most Western countries (e.g. over 150 years at the University of Exeter's St. Luke’s campus ("PGCE Programme,")), this fundamental support is currently missing in universities and higher education centres in Iran. Aside from the above-mentioned reasons for an established ITE, lecturers’ heterogeneous academic backgrounds and qualifications is another piece of the jigsaw necessitating such preparatory courses especially for those lecturers with academic qualifications and background other than TEFL/TESOL. In other words, since lecturers who teach TEFL courses are not necessarily from TEFL background, i.e. majoring in English Language Translation, English Literature and Linguistics, they are more prone to encounter challenges in their practices. As a consequence, designing and developing different courses for teacher applicants seems to be among the urgent steps administrators need to take in Iran.
The centrality of a TPDP has been extensively discussed in the previous chapters. Given the dynamic nature of language teaching profession which rather evolves in response to new educational paradigms in the field as well as the challenges universities encounter which per se emanate from alterations in the curriculum and students' needs, teacher do need to continually promote their professional knowledge and skills (Richards, 2005, p. vii). Having analysed the data especially from the interviews, I noticed a wide gap between teachers’ actual needs and demands and the existing professional development programme. Having had serious reservations about the efficiency of TPDP, some participants maintained that the existing opportunities for their professional development are hardly consonant with their actual needs. Indeed, avoiding a hit-and-miss approach towards TPDP as solely a matter of administrative duty, administrators should do much to advance and deepen their understanding of TPDP, thereby deploying it as a powerful medium to meet their academic staff professional needs. Nonetheless, it seems unwise to expect administrators to openly admit responsibility for such criticism which sheds doubts on the appropriateness of the existing TPDP. In other words, as a source of conflict between teachers and administrators, the very question of what counts as a 'need' tends to remain yet contentious. Consequently, as a gap between ‘what is’ and ‘what should be’ (Witkin & Altschuld, 1995, p. 9), teachers’ professional needs should be highly appreciated by academics/authorities who are in leadership positions in designing, developing and administering any kind of in-service teacher development programme for which ‘needs analysis’ is a prerequisite. According to Richards and Farrell (2005), following a long-term objective, teacher development tends to be a ‘bottom-up’ phenomenon (p. 4). Moreover, as they further remind us, teacher development should be considered as an opportunity throughout which teachers’ understanding of teaching evolves (Richards & Farrell, 2005, p. 4). Having considered the above-mentioned lines of argument, it seems that administrators need to take some rather urgent steps towards tailoring the existing TPDP to their staff needs.

- **Formative evaluation; feedback & feed-forward**
The analysis of data provided some new insights into the nature of teacher evaluation. Having revealed a lack of recognition of formative evaluation and imperfect implementation of peer evaluation in universities in Iran, this study calls for further attention to formative evaluation. As a constructive learning tool, formative evaluation practically can be rendered as informative feedback (learning from previous practices) and/or feedforward (amending future practices) (Irons, 2008, p. 7). It is fortunate that different measures of evaluation can serve formative purposes through various ways. In other words, any evaluative measure discussed in this study, e.g. SETs, peer, observation, etc., can be deployed for formative objectives which is a key to teachers’ professional growth. However, as the results of this study suggested, it appears that the ‘formative’ dimension of teacher evaluation has been rather consigned into oblivion. Thus, one of the most critical implications of this research is germane to better recognition of the contributions of formative evaluation including feedback and feed-forward to effective teacher appraisal. Since lecturers are currently deprived of having such an opportunity, there is seemingly a growing need on the part of administrators to take decisive actions on giving more room to formative evaluation in teacher appraisal system.

It is worth highlighting that introducing formative evaluation requires a mechanism to be put in place. First, it might be a good idea for universities to add teachers’ self-evaluation as a measure to their appraisal system. Second, in addition to their routine summative staff evaluation, universities begin to evaluate teachers using the measures which are inherently congruent with formative purposes, i.e. peer evaluation and self-evaluation, whereby teachers can reflect on each other without fear of the results. This allows the universities to step forward to support their academic staff with proper in-service teacher development programmes which as shown in this study have provoked the participants’ criticism.

- Transparency & Trust

Participants’ scepticism about the accuracy of teacher appraisal was one of the immediate findings of this study. Regardless of the extent to which mistrust exists among stakeholders, it seems quite implausible to expect a successful appraisal system on which teachers show little reliance. Based on my experience throughout
data collection, I think, such misgivings are deeply rooted in the paucity of a clear and transparent mechanism for teacher appraisal. In other words, a less well-explained appraisal system grounds for creating false impressions and scepticism among lecturers. I was quite surprised when the concept of ‘teacher appraisal’ was a rather nebulous concept to many of them. As reported in the previous chapters, SETs-centred, biased, subjective, unfair were among various attributes some lecturers assigned to teacher appraisal. This is why, I think, enhancing lecturers’ knowledge and awareness of the precise mechanism of teacher appraisal prior to the commencement of each academic year through seminars, workshops, etc., and publicising the proportion each stakeholder (as appraiser) contributes to the final results, can meet lecturers’ uncertainties and doubts. As a consequence, superseding the existing perceptions towards appraisal with an atmosphere of trust as well as developing rapport and engendering respects among academic staff seem to be the very steps administrators need to take towards making teacher evaluation as successful as possible. To do so, they may need to keep their academic staff informed of and briefed on the minutiae of the appraisal system through different ways, e.g. publicising on university websites, Initial Teacher Education Programmes, group or individual tutorials, etc.

7.2.3. Implications for teachers

By the outset of the third millennium teachers started to serve as multi-role professionals, i.e. teachers as facilitators, counsellors, pathfinders, researchers, practitioners, leaders, critical thinkers, etc. It has been argued that not only are teachers deemed to engender learning strategies in their students, they need to improve their own professional qualities, thereby getting more involved in activities such as reflective practice. More importantly, other stakeholders’ failure in fulfilling their responsibilities should never exempt teachers from their self-improvement activities. A number of pathways to enhance teachers' effectiveness have been explicated in details throughout this study some of which lie at the heart of teachers’ self-improvement techniques and strategies. In order to enhance their level of effectiveness, teachers need to enter a cycle of self-development strategies.
It is worth referring to the dynamic nature of teacher effectiveness which warns us against adopting the so-called ‘prescriptive approach’ towards teacher education and evaluation. An effective teacher is, literally, the one who can exhibit all the qualities, attributes, characteristics, skills, etc., presented in this thesis as well as a myriad of other qualities which are perceived as indicators of effective teaching. Yet, such an approach seems to be very unlikely to yield any promising results, for such qualities need to be understood and interpreted in relation to the context in which teaching tends to transpire. Picking up some key strategies raised in the previous chapter, this section will take a brief look at reflective practice and teachers’ engagement in professional development opportunities such as formal TPDP as well as their informal collegial networking.

As an invaluable professional asset to language teachers, ‘Reflective Practice’ and its contributions to improving teacher effectiveness have been extensively addressed. It has been argued how ‘reflection’ has been left out in educational authorities’ definition of effective teaching. The recognition of teachers’ strengths and pitfalls and the peculiarities of foreign language teaching through reflective practice would endow teachers with an opportunity for agonising upon their own practices by which they can bolster their professional career. The concept of ‘reflection’, as a way of thinking about practice, dates back to Dewey’s time (Loughran, 1996, p. 3). Despite some criticism, e.g., ‘cautions against putting overmuch faith in reflection’ (Knight, 2002, p. 29), reflective practice has been able to attain acceptance among scholars and practitioners. Reflection is of seminal importance in teaching in that it enables teachers to view problems and approach challenges from different perspectives. (p. 4). It, indeed, broadens and deepens teachers’ understanding of what constitutes effective teaching and helps them make their ways through obstacles. Referring to the nature of reflection as a prerequisite for realistic reflection, Loughran (2002, p. 33) turns the attention to the relationship among ‘time, experience and expectations of learning’ as an important milestone in reflection. Therefore, reflection, in line with other teaching skills, needs to be learned and internalised. Reflective practice can be well deployed by teacher in order to alleviate the pressure and demands with which they are likely to struggle (Norton & Campbell, 2007, p. 140).
Another important aspect of teachers’ cycle of self-development techniques is teachers’ active engagement in professional development opportunities and their collegial peer networking. The contributions of such opportunities have been explained in details earlier. It has been argued that teachers themselves need to update different aspects of their knowledge and practices. It is unfortunate that most Iranian lecturers tend to be rather ignorant of the privileges of peer feedback and other professional development programmes. Lecturers’ dissatisfaction with the content and quality of the teacher development workshops was found to be another important challenge for which teachers need to avail themselves of each other’s experiences. As a consequence, improving the quality of TPDP through revisiting and improving the existing one and publicising the merits and advantages of professional qualities such as peer networking can be one constructive measure administrators need to take.

7.3. Theoretical contributions: A suggested model

In chapter 3, a number of models and frameworks adopted for teacher effectiveness appraisal around the world were presented and examined, i.e. (Campbell et al., 2004a; Cheng & Tsui, 1999; Korthagen, 2004; McBer, 2000; Piggot-Irvine, 2003). Further to different dimensions of teacher effectiveness appraisal explained throughout this study, a model of teacher appraisal which is informed by lecturers’ perceptions of teacher evaluation in Iran is proposed as illustrated in Figure 7.2. Being informed by TEFL lecturers’ understanding of what makes an effective teacher as well as their perceptions of an appraisal model which is fair, unbiased and reliable, a number of components have been embedded in the proposed appraisal model. These include purpose of evaluation, measures of evaluation, priority (weight) of each measure along with different criteria/standards for evaluation. However, there are a number prerequisite questions and concerns which need to be appreciated prior to the implementation of the appraisal, e.g. ITEP and TPDP. As stated earlier, prior to implementation of appraisal, policymakers and administrators need to provide sufficient training and development opportunities for teachers. Moreover, as emphasised earlier, teachers need to be informed and briefed about universities’ expectations and their
proposed standards or criteria for effective teaching. In other words, the questions of ‘effective in terms of what?’ and ‘what counts as an effective teacher’ need to be agreed in advance. Since many teachers, may not have prior teaching experiences, initial teacher education programme (ITEP) needs to be provided to teachers especially the novice ones. Informed by Dewey’s ‘theory of experience’ (Dewey, 1938), ITEP helps universities prepare their academic staff for their actual teaching in the classrooms. It is unfortunate that this important aspect has been rather ignored in most universities and higher education institutions in Iran. Upon the completion of the first step, teachers would then start their practical teaching in the classroom. However, administrators need to engender qualities such a reflective practice in ITEP and later opportunities such as TEPD. Informed by the works of Donald Schōn (1983), reflective practice has been suggested as an opportunity from which teachers can benefit in order to enhance their effectiveness. Teacher in-service professional development is a further step responding to teachers’ professional needs whose successful outcomes highly rely on both involved parties, i.e. lecturers and administrators’ active engagement.

Having considered the above development ladder and prior to the conduct of teacher appraisal, it might be a good idea to consider the following influential factors:

<table>
<thead>
<tr>
<th>Some prerequisite questions to be considered:</th>
</tr>
</thead>
<tbody>
<tr>
<td>- Is the appraisal transparent to all stakeholders involved in teacher evaluation, e.g. lecturers, administrators, students?</td>
</tr>
<tr>
<td>- Are lecturers well-briefed on ‘what counts as an effective teacher?’</td>
</tr>
<tr>
<td>- Do different stakeholder’s doing the evaluations in one way or another e.g. students, appraisers, etc. receive the required training needed for ‘doing faculty ratings’?</td>
</tr>
<tr>
<td>- Is the overall score/points/percentage, etc., apportioned fairly among different measures of evaluation?</td>
</tr>
<tr>
<td>- Is the weight given to the purpose of evaluation, i.e. formative and summative, clear?</td>
</tr>
<tr>
<td>- What are the consequences for teachers?</td>
</tr>
<tr>
<td>- Are the appraisers certified/accredited?</td>
</tr>
<tr>
<td>- Is the appraisal tailored for EFL lecturers so that their TEFL-related challenges could be accommodated?</td>
</tr>
<tr>
<td>- How mediating/intervening variables such as curriculum, the difficulty of the course, etc., are managed and accounted for in the appraisal final results?</td>
</tr>
<tr>
<td>- Are data collection procedures reliable?</td>
</tr>
<tr>
<td>- ...</td>
</tr>
</tbody>
</table>

Table 7.1. Some rudiments of teacher evaluation
**Proposed Appraisal Model**

For detailed account of qualities for each criterion/standard refer to chapter 5
7.3.1. Contributions to the existing knowledge

It is worth emphasising that this study does not claim that the proposed model is totally new or could be generalised to all contexts. The findings reported in this study are indeed a reflection of the participant lecturers’ understanding and perceptions of teacher effectiveness and teacher appraisal in Iran. Deeply informed by TEFL lecturers’ points of view in universities and higher education centres in Iran, the proposed model reflects lecturers’ perceptions of teacher effectiveness appraisal. Indeed, building on the works of previous researchers and scholars, this model has tried to add some elements that, to the best of my knowledge, are either missing or less well-implemented in universities and higher education centres in Iran. Much of the contribution made in the proposed model is perceived as having contextual values. Nevertheless, as it will be discussed shortly, the model will offer some suggestions which deepen our insights into the nature of teacher appraisal from a rather global point of view. This section endeavours to cast some light on both dimensions, i.e. implications specific to the Iranian context and some recommendations for other EFL contexts around the world. The contributions of this study have been discussed earlier in Section 7.2., as theoretical and pedagogical implications. As it has been argued, the two-fold implications of the present study target policymakers, administrators and teacher themselves as the main stakeholders in this phenomenon at two distinct yet interrelated levels of theory and pedagogy.

The findings of this study say a lot for the quality of teacher appraisal and lecturers’ perceptions and satisfaction about their appraisal in Iran. Lecturers’ discrepant and in some cases contradictory views on the very notions of teacher effectiveness and teacher evaluation say it all. Hence, one very preliminary contribution of the study is to increase the awareness of the problem on the part of policymakers and administrators. Further to participants’ views, it is self-evident that administrators have not yet taken even requisite steps to inform academic staff of their expectations of effective teaching and brief them on the existing mechanism for evaluating teachers. In addition, constraints on adopting a multi-measure approach towards teacher evaluation and downgrading it to some rather basic measures
such as SETs, apparently has begun to sow seeds of doubts and misgivings among lecturers. This is why, the majority of participants expressed intuitive reservations about the reliability and validity of their appraisal outcomes.

Goal or purpose of teacher appraisal is another important issue on which policymakers and administrators need to decide. ‘Accountability’ and ‘professional development’ as the main objectives of a teacher appraisal scheme are perceived to have contradictory relationship (Bartlett, 1996, p. 7). Yet, it is argued that the accountability function of the appraisal is likely to dominate its development purposes (Campbell et al., 2004a, p. 93). Concentrating on the accountability functions of teach evaluation may thwart administrators’ plans for designing and developing professional development opportunities for their in-service teachers. Indeed, such development opportunities are a vehicle for improving teacher effectiveness by which teachers can boost their confidence, gain new insights into their career and develop their teaching skills. As a consequence, administrators need to periodically improve and inform their professional development programmes by the results of their evaluations. Whereas summative (accountability) scheme can be implemented once per academic year, administrators may conduct formative (professional development) evaluation throughout each semester or academic year on different occasions.

The analysis of the results from both types of evaluation can give university authorities some deep insights into the areas which need to be reflected on, refined, revised, or even superseded in a forthcoming professional development programme. Such a strategy can presumably better accommodate teachers’ potential and actual needs. As illustrated in Figure 7.3., improving teacher effectiveness and its pertaining appraisal need to be approached in a cyclical process. Once identified through appraisal, teachers’ needs, weak points and challenges can be addressed in a subsequent professional programme. In addition to such developmental approach, administrators might take decisive actions germane to teachers’ career ladder, e.g. promotion, extra incentives, etc.
In addition to the above mentioned issues which are mostly applicable to the Iranian context, this study seems to have some contributions to promoting a wider understanding of teacher effectiveness and appraisal. *Weight or priority* of the measures were among the important yet less well-investigated issues that came to my attention throughout data collection phase especially in interviews. None of the models reviewed in this study explain the weight/priority each and every measure needs to be given in the overall multi-measure appraisal. This is of a substantial significance in increasing lecturers’ awareness of and trust in any teacher evaluation system regardless of the subjects and levels at which teachers teach and the context in which teaching takes place. This study showed that lecturers do care about the status (weight) of each measure, e.g. SETs in the overall appraisal.

Despite the burgeoning literature on teacher effectiveness in mainstream education, the mechanism of interpreting and making sense of collected data in teacher appraisal particularly understanding the role of contextual factors such as TEFL-oriented factors seem to be less well-understood in the literature and hence is likely to be rather limited. As an attempt to highlight the need and importance of a teacher appraisal model in the context of this study which is informed by
discipline-specific factors, the findings of this research call for better appreciation of the centrality of discipline-specific factors to teachers' effectiveness. Addressing the contentious approach towards effective teaching, i.e. 'an effective teacher is effective regardless of what he or she teaches', this study poses a question as to what makes an EFL lecturer different from lecturers in other academic disciplines such as Biology or Physics, and argues that EFL peculiarities and idiosyncrasies need to be reflected in some way in teacher evaluation scheme. This is why the so-called 'one-size-fits-all' approach, which is perceived as characterising the Iranian appraisal system from participants' points of views, received criticism from the lecturers in this study. Given the fact that EFL lecturers in Iran could be more or less typical representatives of the generality of EFL lecturers around the world, their concerns, wants, ideals and expectations can more or less reflect those of their counterparts teaching in other EFL contexts, even though the role of context-specific factors should not be underestimated.

Another important aspect of teacher appraisal which has been seemingly given scant attention in the literature is the dynamics of the interrelationship between each measure of evaluation and the purpose evaluation. As argued earlier, different forms (measures) and types (purposes) of evaluation can interact in a dynamic fashion. However, as suggested by the literature, some measures have been found to be more suitable for a particular purpose. In other words, since some measures, by nature, are oriented towards either formative or summative evaluation, it is unwise to use all measures for summative or formative purposes. As a consequence, a wise selection of different measures from the so-called 'evaluation continuum' which is “fit-for-purpose” (Brown, 1999, p. 6), as shown in Figure 7.4., rests upon administrators’ (appraisers) decision.
7.4. **Policy-implementation nexus**

One of the findings of this study was the imperfect nexus between policy and implementation. It is unwise to decide on the merits and demerits of an education system just based on its approved policies. Rather, the extent to which the policies are properly implemented should be in the foreground. Policymakers need to take decisive steps towards practical implementation of their proposed policies. This study revealed that the current teacher evaluation system in Iran lacks some measures of evaluation such as teachers' *self-evaluation* which has already gained strong support in the literature. Moreover, whereas a commonsense approach to peer evaluation is to require lecturers to observe and evaluate their peers (colleagues), none of the participants of this study could recollect such an experience either as an evaluator or evaluatee. Indeed, peer evaluation has been seemingly translated into college-evaluation and department-evaluation done by (deputy) dean of the college and head of the department respectively. And to that end, peer evaluation, bearing little resemblance to this measure, is almost neglected as a measure which can serve not only summative but also formative purposes. There is more to this than meets the eye, given the fact the evaluation of
policies cannot be conceived of without understanding how such policies are implemented practically (Ahmed & Sayed, 2009, p. 214). Therefore, I suggest that offices for supervision and evaluation at universities in Iran should get more involved in monitoring the implementation of the approved policies for teacher evaluation. For instance, if peer evaluation is approved as a measure of evaluation, it needs to be done by colleagues as well.

7.5. Suggestions for further research

This study was one of the very few attempts to elucidate different dimensions of teacher effectiveness and its appraisal from lecturers’ perspectives in Iran. While the analysis of the collected data attested to a number of thought-provoking ideas and implications for three main target stakeholders, namely policymakers, administrators and teachers, it raised some important yet little-understood questions and concerns for which further research is needed. Due to the limitations of a PhD project such as funding and time, this study focused on lecturers as the main stakeholder. Therefore, in order to have a more comprehensive understanding of the very nature of teacher effectiveness and teacher evaluation system, it might be a good idea to obtain other parties’ views and perceptions. This could be done as a single study for each stakeholder, i.e. policymakers, administrators and students.

Since there is seemingly a growing conflict of interests among different stakeholders’ towards effective teaching and teacher evaluation, further research needs to be conducted to look for any agreements or discrepancies among different stakeholders, i.e. policymakers, administrators, teachers and students, in terms of their perceptions and expectations of teacher effectiveness and teacher appraisal. This will indubitably broaden and deepen our understanding of the extent to which different parties share similar, if not contradictory, values. Therefore, separate studies can be conducted to explore what effective teaching means to students, to identify policymakers’ positions towards quality teaching and teacher evaluation in higher education and to delve into administrators’ expectations of high quality and effective teaching. This is of high importance, given the fact it is sometimes difficult to reconcile university and academic staff
approaches towards education. Whereas some universities may view students as end-user customers, lecturers tend to consider them as learners.

This study tried to investigate the influence of some external variables such as curriculum which can potentially exert impacts on teachers’ effectiveness. However, further investigation is needed to understand how and in what way they exert impact on teacher effectiveness. Moreover, further research needs to be done to elucidate the effect of different internal and external variables in the process of high quality teaching at universities, e.g. the socio-economic status of the universities and higher education institution, the staff-student ratio, universities’ missions and visions, etc. As a clear signal to policymakers and administrators, the findings of this research call for improvement of teacher appraisal system not only at the level of policy but also at the level of implementation. It is hoped that policymakers at ministerial level provide lecturers with sufficient opportunities to make their voice heard in teacher appraisal. Nevertheless, a number of prerequisite steps and actions need to be taken to inaugurate such a nationwide movement which may take several years to yield its promising results.

As my last suggestion, I would like to refer to the position of teachers in an education system in that teachers are only one component whose effectiveness are influenced by other constituents of the system. As it was argued earlier, policymakers need to show proper regard for not only teachers, but also other little understood yet important components operating in most educational systems. Overemphasis on teachers can easily preclude policymakers and administrators from identifying internal and external factors which can potentially exercise influence on teacher effectiveness, either positively or negatively. Nevertheless, given the idiosyncratic nature of each of these variables, the identification of such an intricate network of interrelationships might not be as simple as it may sound. In previous chapters, a number of these influential elements have been introduced including curriculum, syllabi, socio-economic status of students and teachers, etc. Whereas these elements are said to be known, the mechanism through which each component interacts with other constituents has been yet little-known and hence needs further investigations.
7.6. Conclusion

EFL lecturers' understanding of teacher effectiveness and their perceptions of the existing teacher appraisal system in Iran were the focus of this study. This section will present a brief overview of the issues investigated, reported and discussed in this study and thereafter will proceed to some final concluding remarks. Following the analysis of the collected data, six major themes were formulated each of which embraced a number of sub-themes or categories. With regard to the first theme of the study, i.e. teachers' perceptions of teacher effectiveness and the qualities of an effective teacher, five major categories were identified each of which included several subcategories. These included *Personal qualities, Cognitive qualities, Pedagogical skills* and *Professional skills*. Then, the second theme, i.e. measures of evaluation or sources of information based upon which administrators tend to decide on one's teaching effectiveness were examined. Several measures were explored including *Student Evaluation of Teachers (SETs)/Student Ratings, Students' learning outcomes/student achievement, Peer evaluation, Self-evaluation* and *Observation*. Strategies and techniques though which lecturers can improve the quality of their teaching and promote their effectiveness was the third theme identified and explored in this study. It was argued that in most educational systems, there exist some elements and variables whose exerted influences (mostly negative) on teachers' effectiveness are beyond teachers' control, and Iran is no exception. Entitled as 'non-teacher-controlled-factors', these variables included factors such as *curriculum, facilities, financial incentives*, etc. Afterwards, lecturers' understanding of 'Iranian teacher appraisal model' adopted by their universities was explored the analysis of which raised a number of thought-provoking questions and concerns such as a lack of some forms and types of teacher evaluation. The final theme generated was lecturers' perceptions of an 'ideal appraisal model' which, from participants' perspectives, can appropriately reflect their potential and actual teaching effectiveness. A number of notions and qualities were attributed to such an ideal model by lecturers. Nevertheless, the study progressed beyond the aforementioned six themes, inasmuch as the discussion of data illuminated some less well-investigated dimensions of teacher appraisal. In general, this study has tried to deepen and broaden policymakers,
administrators, researchers and practitioners’ insight into the very nature of teacher effectiveness. The most important conclusions drawn from this study, amongst several others, are as follows:

- Given the broad and multi-dimensional nature of teacher effectiveness, there is an urgent need for policymakers to address the question of ‘effective in terms of what and under what circumstances?’

- Effective teaching is a resultant product of an effective educational system. In other words, the process of teacher effectiveness is not likely to evolve in a vacuum. Educators need to train and prepare effective teachers rather than expecting them to emerge on their own. Such preparation may be partially met by improving the existing TEFL-related academic programmes, i.e. BA, MA in universities in which student-teachers’ professional personality tend to take shape.

- Since EFL lecturers in Iran are not necessarily graduates of TEFL/TESOL or Applied Linguistics, an Initial Teacher Education Programme (ITEP) prior to actual teaching practices seems to be an undeniable prerequisite especially for those with little educational background, e.g., Linguistics, Literature and Translation studies.

- Teachers, even the effective ones, do need continual professional development support. Administrators need to provide such in-service opportunities in universities through which teachers can further their knowledge and experience and promote their effectiveness thereupon.

- It has been argued that teachers may not tie their effectiveness to other stakeholders. Indeed, administrators’ failure to fulfill their duties and responsibilities does not exempt lecturers from their responsibilities to improve their own teaching practices. Instead, lecturers are suggested to constantly promote their practices through a number of ways including ‘reflective practice’, ‘active engagement in collegial networking and peer feedback’, etc.

- Transparency and clear standards/criteria need to be a high priority, inasmuch as lecturers’ deep scepticism towards their appraisal seems to emanate from their little awareness of universities’ expectations of effective teaching.

- All minor and major parties who are involved in evaluation, e.g. students, peers, observers, etc., are expected to have the minimum knowledge required for ‘evaluating a faculty’. Ideally, appraisers are expected to be accredited by
academic or professional bodies within the Ministry of Science, Research and Technology (MSRT).

- The so-called ‘one-size-fits-all’ approach does not suffice for EFL lecturers’ needs. The adopted model needs to be informed by the idiosyncrasies and peculiarities of teaching in a language other than students’ mother tongue.

- It has been argued that the burden of quality teaching should not lie only on teachers’ shoulders, given the fact that teachers are only one yet predominant component in an (effective) educational system. Hence, the impacts of other interrelated components need to be appreciated by policymakers, e.g. educational equipment and facilities, curriculum, salary, etc.

- Notwithstanding teacher effectiveness research has come a long way, it has long way to go especially in terms of ‘what counts as an effective lecturer in higher education’, ‘what counts as an efficient, reliable and valid teacher appraisal model in higher education’, etc.

- Teachers need to reinforce their learning to teach as a professional behavior and to that end administrators are expected to appreciate teachers’ needs in designing, developing and organising initial teacher education and teacher professional development programmes.

It is worth highlighting that there are often some ‘unknown unknowns’ in most social phenomena which tend to emerge over time; and teacher effectiveness research is no exception.

7.7. Reflection on my PhD journey

In this very final section of my thesis, I am inclined to refer to my PhD journey, thereby pointing to the lessons I have learned throughout the programme, the challenges and opportunities. I had the great privilege to start my PhD with the MSc programme. It was the MSc programme which opened new doors for my doctoral journey. Given my previous background which was influenced by positivism and scientific approaches towards educational research, the programme was indeed an invaluable opportunity for me to get to know other philosophical and theoretical standpoints in educational research whereby I could gain the knowledge and skills required for tackling a phenomenon at a doctoral level. Although the
research modules deepened and broadened my insights into teacher effectiveness research, much of my understanding and perception evolved and took shape during the PhD programme especially during the conduct of research. Indeed, the most challenging yet interesting parts of the journey pertained to issues as to how to approach a social phenomenon from different philosophical and theoretical perspectives and to relate them thereafter. I feel really fortunate in having the opportunity to learn and use different technological advancement such as SPSS, MAXQDA, NVIVO, EndNote, etc., to analyse data and write up my thesis. From amongst different qualities I have availed myself during my study I would like to extol the virtues of becoming more critical of the givens. It is my contention that my PhD is indeed a turning point in my academic life after which a new era is to emerge whereby I can learn, research and further my academic and professional knowledge, thereby contributing to education and hence the society.
Appendices
Appendix 1. *Researcher-developed questionnaire*

Dear participant,

I would like to take this opportunity and appreciate your kind cooperation in advance. This questionnaire is developed and administered as a part of my research project. Your comments are invaluable in that they shape and add to our understanding of the existing knowledge of EFL lectures’ effectiveness. For your information, the issues of confidentiality, anonymity and privacy are highly taken into regard by the researcher. Moreover, the collected data will be solely used for the purpose of this study. The questionnaire’s responses range from strongly disagree to strongly agree. It is worth emphasizing that there is no right or wrong answer. Please circle your chosen answer.

Once again, thank you for your participation.
**EFL Teacher effectiveness questionnaire**

*About the questionnaire:*

- *When the statement refers to teacher, academic staff or faculty, it is meant EFL lecturer.*
- *When the statement refers to appraisal model, it is meant appraisal model for evaluating EFL lecturers’ effectiveness.*

**Section 1. Close-ended questionnaire**

Please read each statement and put a tick under your chosen response.

<table>
<thead>
<tr>
<th>Item No.</th>
<th>Statement</th>
<th>Response</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Strongly agree</td>
</tr>
<tr>
<td>1.</td>
<td>Different stakeholders’ voices e.g. teachers, students, administrators, etc. who have a stake in teaching, should be heard and incorporated into the appraisal model.</td>
<td></td>
</tr>
<tr>
<td>2.</td>
<td>A friendly personality is important to teacher effectiveness.</td>
<td></td>
</tr>
<tr>
<td>3.</td>
<td>An EFL teacher effectiveness model needs to be evaluated upon TEFL-specific subject criteria rather than generic education criteria.</td>
<td></td>
</tr>
<tr>
<td>4.</td>
<td>Teacher’s appraisal should mostly focus on formative purposes, i.e. professional development.</td>
<td></td>
</tr>
<tr>
<td>5.</td>
<td>Teachers’ appraisal should mostly focus on summative purposes, e.g. promotion, tenure, etc.</td>
<td></td>
</tr>
<tr>
<td>6.</td>
<td>Teacher effectiveness appraisal should mainly focus on teachers’ performance.</td>
<td></td>
</tr>
<tr>
<td>7.</td>
<td>Administrators (e.g. Dean, Head of department) should adopt a multi-measure rather than a single-measure approach towards teacher effectiveness appraisal.</td>
<td></td>
</tr>
<tr>
<td>8.</td>
<td>Appraisal models mainly depend on students ratings with less attention given to other stakeholders such as teachers.</td>
<td></td>
</tr>
<tr>
<td>9.</td>
<td>Teachers’ self-evaluation will help them reflect on their own teaching practices.</td>
<td></td>
</tr>
<tr>
<td>10.</td>
<td>Teachers’ beliefs tend to exert influence on teacher effectiveness.</td>
<td></td>
</tr>
<tr>
<td>11.</td>
<td>Self-evaluation should be used for formative purposes.</td>
<td></td>
</tr>
<tr>
<td>12.</td>
<td>Self-evaluation should be used for summative purposes.</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>---</td>
<td>---</td>
<td></td>
</tr>
<tr>
<td>13.</td>
<td>Peer evaluation contributes to the improvement of teacher effectiveness.</td>
<td></td>
</tr>
<tr>
<td>14.</td>
<td>Colleagues who evaluate a faculty need to be skilled in evaluation.</td>
<td></td>
</tr>
<tr>
<td>15.</td>
<td>Peer evaluation could be used for summative evaluation.</td>
<td></td>
</tr>
<tr>
<td>16.</td>
<td>Teacher’s gender tends to exert influence on students’ ratings.</td>
<td></td>
</tr>
<tr>
<td>17.</td>
<td>Teacher’s age tends to have impacts on students’ ratings.</td>
<td></td>
</tr>
<tr>
<td>18.</td>
<td>Teachers who give high marks tend to be rated as more effective by students.</td>
<td></td>
</tr>
<tr>
<td>19.</td>
<td>Students should be informed of the criteria for identifying an effective teacher.</td>
<td></td>
</tr>
<tr>
<td>20.</td>
<td>The easier the course, the higher the students’ ratings of their teachers.</td>
<td></td>
</tr>
<tr>
<td>21.</td>
<td>It is a good idea to collect students’ ratings in mid-semester in order to eliminate the ‘grading bias’ effect.</td>
<td></td>
</tr>
<tr>
<td>22.</td>
<td>Students’ learning outcome is highly vulnerable to student-specific factors which are beyond teachers’ control.</td>
<td></td>
</tr>
<tr>
<td>23.</td>
<td>An effective teacher might be less effective with a particular group of students or a particular course.</td>
<td></td>
</tr>
<tr>
<td>24.</td>
<td>Students’ learning outcomes (e.g. test results, achievement) can be a good indicator of teacher effectiveness.</td>
<td></td>
</tr>
<tr>
<td>25.</td>
<td>Administrators’ (e.g. Dean, Head of Department, etc.) appraisal is subjective and biased.</td>
<td></td>
</tr>
<tr>
<td>26.</td>
<td>Teachers are not willing to be evaluated by an external observer.</td>
<td></td>
</tr>
<tr>
<td>27.</td>
<td>Teacher effectiveness should be evaluated based upon a set of transparent standards/criteria.</td>
<td></td>
</tr>
<tr>
<td>28.</td>
<td>Teachers need to be convinced of the fairness of the evaluation system through which they are assessed.</td>
<td></td>
</tr>
<tr>
<td>29.</td>
<td>There is a direct correlation between teacher’s level of academic qualifications and their effectiveness.</td>
<td></td>
</tr>
<tr>
<td>30.</td>
<td>Universities from which teachers have graduated are influential factors in their effectiveness.</td>
<td></td>
</tr>
<tr>
<td>31.</td>
<td>An effective teacher has excellent pedagogical skills.</td>
<td></td>
</tr>
<tr>
<td>32.</td>
<td>Teachers’ subject knowledge lies at the heart of teacher effectiveness.</td>
<td></td>
</tr>
<tr>
<td>33.</td>
<td>Effective EFL teachers should have TEFL-driven understanding of teaching.</td>
<td></td>
</tr>
<tr>
<td>34.</td>
<td>Teacher leadership contributes to teacher’s effectiveness.</td>
<td></td>
</tr>
<tr>
<td>35.</td>
<td>Teachers’ personal traits (e.g. patience) play an important role in their effectiveness.</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>---</td>
<td>---</td>
<td></td>
</tr>
<tr>
<td>36.</td>
<td>Teachers’ language proficiency does not contribute to teacher effectiveness.</td>
<td></td>
</tr>
<tr>
<td>37.</td>
<td>Effective language teachers should consider Technology Enhanced Learning (TEL) in their teaching practices.</td>
<td></td>
</tr>
<tr>
<td>38.</td>
<td>Effective language teachers dedicate themselves to their students to the extent that their needs are met.</td>
<td></td>
</tr>
<tr>
<td>39.</td>
<td>Effective teachers are open to their students’ voices.</td>
<td></td>
</tr>
<tr>
<td>40.</td>
<td>Teacher authority is the keystone of the notion of teacher effectiveness.</td>
<td></td>
</tr>
<tr>
<td>41.</td>
<td>Effective teachers are accountable to other stakeholders, e.g. students, administrators, etc.</td>
<td></td>
</tr>
<tr>
<td>42.</td>
<td>An effective teacher respects the students.</td>
<td></td>
</tr>
<tr>
<td>43.</td>
<td>An effective language teacher engages all students in classroom activities.</td>
<td></td>
</tr>
<tr>
<td>44.</td>
<td>EFL teachers should have the required knowledge of curriculum development, lesson plan, syllabus design, etc.</td>
<td></td>
</tr>
<tr>
<td>45.</td>
<td>An effective TEFL teacher is familiar with assessment strategies for assessing learners’ different language skills.</td>
<td></td>
</tr>
<tr>
<td>46.</td>
<td>Effective language teachers tend to be sensitive to important issues such as students’ race, social class, etc.</td>
<td></td>
</tr>
<tr>
<td>47.</td>
<td>Teachers’ experience is a cornerstone of their teaching effectiveness.</td>
<td></td>
</tr>
<tr>
<td>48.</td>
<td>An effective teacher establishes a friendly environment in the classroom.</td>
<td></td>
</tr>
<tr>
<td>49.</td>
<td>An effective teacher knows how to deal with unexpected situations in the classroom.</td>
<td></td>
</tr>
<tr>
<td>50.</td>
<td>An effective teacher should be innovative.</td>
<td></td>
</tr>
<tr>
<td>51.</td>
<td>I am not well-aware of the evaluation system and the appraisal model adopted by administrators for evaluating teacher effectiveness in the Iranian higher education.</td>
<td></td>
</tr>
<tr>
<td>52.</td>
<td>The existing appraisal model used in the Iranian higher education is a reliable and valid indicator of my teaching effectiveness.</td>
<td></td>
</tr>
<tr>
<td>53.</td>
<td>External observation should be considered as a measure of evaluation in the Iranian appraisal model.</td>
<td></td>
</tr>
<tr>
<td>54.</td>
<td>Power relations might dominate teacher appraisal.</td>
<td></td>
</tr>
<tr>
<td>55.</td>
<td>Universities should have units that provide technical and general advice to less effective teachers.</td>
<td></td>
</tr>
<tr>
<td>56.</td>
<td>I am happy with the existing appraisal model adopted in my university.</td>
<td></td>
</tr>
<tr>
<td>57.</td>
<td>There is a need to revisit the existing Iranian appraisal model.</td>
<td></td>
</tr>
</tbody>
</table>
58. Developing an accredited professional preparation programme will not help teachers gain the required skills.

59. Educational leadership tends to exert influence on teacher effectiveness.

60. National curriculum and syllabi are important factors in promoting teachers effectiveness.

61. Designing a good teacher education programme (TEP) for pre-service teachers can contribute to their teaching effectiveness.

62. Staff development programme such as Teacher Development Programme (TDP) can promote teacher effectiveness.

63. The Iranian appraisal model needs to be informed by the political, cultural and social specificities in Iranian context.

Section II. Open-ended questionnaire

1. What is your overall perception of teacher effectiveness in the Iranian EFL context? How do you define it?
2. Please write the characteristics of an effective EFL teacher. You may write as many as you can.

3. Please let me know the ways that you can enhance your teaching effectiveness?
4. How do you feel about the existing Iranian appraisal model? What are the problems associated with the current model of teacher effectiveness?

5. What is your opinion about the influence of elements such as *curriculum*, *syllabi*, etc., on teacher effectiveness? You may write as many elements as you can.
6. Please let me know any additional comments pertaining to EFL teacher effectiveness.

**Section III. Background information**

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>Name (optional)</td>
</tr>
<tr>
<td>2.</td>
<td>Gender</td>
</tr>
<tr>
<td>3.</td>
<td>Qualification</td>
</tr>
<tr>
<td>4.</td>
<td>Academic Major (Area of expertise)</td>
</tr>
<tr>
<td>5.</td>
<td>Employment status</td>
</tr>
<tr>
<td>6.</td>
<td>University</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>7.</td>
<td>Years of experience</td>
</tr>
</tbody>
</table>
## Appendix 2: Interview Protocol

<table>
<thead>
<tr>
<th>Major themes to explore</th>
<th>Questions</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Introducing questions</strong></td>
<td>- Please tell me a bit about your career? Why did you choose TEFL?</td>
<td></td>
</tr>
<tr>
<td><strong>Some major content questions</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Question 1</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Theme I. General perceptions</td>
<td>- Could you please tell me about your overall understanding and perception of teacher effectiveness?</td>
<td></td>
</tr>
<tr>
<td>Theme II. Measures of Evaluation</td>
<td>- What do you think of the source for information (measure of evaluation) that should be used for evaluating teacher effectiveness? Whose voices need to be heard? What about the weight that should be given to each of these parties?</td>
<td></td>
</tr>
<tr>
<td>Theme III. Criteria (standards)</td>
<td>- Please let me know about the qualities of an effective teacher? Do you think an effective teacher should have any specific criteria or standards? If so, what are they?</td>
<td></td>
</tr>
<tr>
<td>Theme IV. Ways to promote TE</td>
<td>- How can you enhance your teaching effectiveness? Are the recourses available to you?</td>
<td></td>
</tr>
<tr>
<td><strong>Question 2</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Theme V. The Iranian Appraisal Model</td>
<td>- Could you tell me how an ‘appraisal model’ can contribute to a fair evaluation of teacher effectiveness?</td>
<td></td>
</tr>
<tr>
<td></td>
<td>- What is your overall opinion about the existing appraisal model used for evaluation of Iranian TEFL lecturers’ effectiveness? Is it convincing?</td>
<td></td>
</tr>
<tr>
<td></td>
<td>- Do you think that the current Iranian appraisal model is able to address such a multidimensional issue, i.e. teacher effectiveness?</td>
<td></td>
</tr>
<tr>
<td><strong>Question 3</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Theme VI. Multidimensionality of TE</td>
<td>- What is your opinion about factors other than teaching factors that can potentially and actually exert influence on teacher effectiveness, e.g. curriculum?</td>
<td></td>
</tr>
<tr>
<td><strong>Question 4</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Theme VII. Critical Appraisal Model</td>
<td>- Could you tell me whether a critical appraisal model can promote teacher effectiveness? N.B. The concept of ‘criticality’ needs to be well-explained.</td>
<td></td>
</tr>
<tr>
<td><strong>Sample probes &amp; final closing questions</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>- ‘You have mentioned the word ‘………’ twice; what do you exactly mean?’ (Dörneyi, 2007, p. 138)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>- ‘Is there anything else you would like to add?’ (Dörneyi, 2007, p. 138)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>- ‘What should I have asked you that I didn’t think to ask?’ (Dörneyi, 2007, p. 138).</td>
<td></td>
</tr>
<tr>
<td>Thank you for your participation in the interview</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Appendix 3: Questionnaire respondents’ background data

### Gender

<table>
<thead>
<tr>
<th>Gender</th>
<th>Frequency</th>
<th>Percent</th>
<th>Valid Percent</th>
<th>Cumulative Percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male</td>
<td>26</td>
<td>60.5</td>
<td>61.9</td>
<td>61.9</td>
</tr>
<tr>
<td>Female</td>
<td>16</td>
<td>37.2</td>
<td>38.1</td>
<td>100.0</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>42</strong></td>
<td><strong>97.7</strong></td>
<td><strong>100.0</strong></td>
<td></td>
</tr>
<tr>
<td><strong>Missing</strong></td>
<td><strong>99</strong></td>
<td>2.3</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>43</strong></td>
<td>100.0</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Qualification

<table>
<thead>
<tr>
<th>Qualification</th>
<th>Frequency</th>
<th>Percent</th>
<th>Valid Percent</th>
<th>Cumulative Percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>PhD</td>
<td>3</td>
<td>7.0</td>
<td>7.1</td>
<td>7.1</td>
</tr>
<tr>
<td>PhD Candidate</td>
<td>18</td>
<td>41.9</td>
<td>42.9</td>
<td>50.0</td>
</tr>
<tr>
<td>Master’s</td>
<td>21</td>
<td>48.8</td>
<td>50.0</td>
<td>100.0</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>42</strong></td>
<td><strong>97.7</strong></td>
<td><strong>100.0</strong></td>
<td></td>
</tr>
<tr>
<td><strong>Missing</strong></td>
<td><strong>99</strong></td>
<td>2.3</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>43</strong></td>
<td>100.0</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
### Academic Major

<table>
<thead>
<tr>
<th></th>
<th>Frequency</th>
<th>Percent</th>
<th>Valid Percent</th>
<th>Cumulative Percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>Valid</td>
<td>32</td>
<td>74.4</td>
<td>76.2</td>
<td>76.2</td>
</tr>
<tr>
<td>TESOL/TEFL</td>
<td>29</td>
<td>52.3</td>
<td>59.5</td>
<td>59.5</td>
</tr>
<tr>
<td>Literature</td>
<td>6</td>
<td>14.0</td>
<td>14.3</td>
<td>90.5</td>
</tr>
<tr>
<td>Translation</td>
<td>1</td>
<td>2.3</td>
<td>2.4</td>
<td>92.9</td>
</tr>
<tr>
<td>Linguistics</td>
<td>3</td>
<td>7.0</td>
<td>7.1</td>
<td>100.0</td>
</tr>
<tr>
<td>Total</td>
<td>42</td>
<td>97.7</td>
<td>100.0</td>
<td></td>
</tr>
<tr>
<td>Missing</td>
<td>99</td>
<td>15.3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>43</td>
<td>100.0</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Employment Status

<table>
<thead>
<tr>
<th></th>
<th>Frequency</th>
<th>Percent</th>
<th>Valid Percent</th>
<th>Cumulative Percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>Valid</td>
<td>38</td>
<td>88.4</td>
<td>90.5</td>
<td>90.5</td>
</tr>
<tr>
<td>Tenured</td>
<td>26</td>
<td>60.5</td>
<td>68.4</td>
<td>68.4</td>
</tr>
<tr>
<td>Non-Tenured</td>
<td>12</td>
<td>27.9</td>
<td>31.6</td>
<td>100.0</td>
</tr>
<tr>
<td>Total</td>
<td>38</td>
<td>88.4</td>
<td>100.0</td>
<td></td>
</tr>
<tr>
<td>Missing</td>
<td>99</td>
<td>16.0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>43</td>
<td>100.0</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Appendix 4: Informed Consent Form

GRADUATE SCHOOL OF EDUCATION

CONSENT FORM

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 at any stage withdraw my participation

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

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)

One copy of this form will be kept by the participant; a second copy will be kept by the researcher(s)

Contact phone number of researcher(s): 07570093356

If you have any concerns about the project that you would like to discuss, please contact:

om218@exeter.ac.uk

OR

Data Protection Act: The University of Exeter is a data collector and is registered with the Office of the Data Protection Commissioner as required to do under the Data Protection Act 1998. The information you provide will be used for research purposes and will be processed in accordance with the University’s registration and current data protection legislation. Data will be confidential to the researcher(s) and will not be disclosed to any unauthorised third parties without further agreement by the participant. Reports based on the data will be in anonymised form.
Appendix 5: An excerpt from a sample interview manuscript

R: Sepehr, thank you very much for participating in this interview I really appreciate your kind participation. Before starting the interview, I want to inform you of your rights as the participant in this interview. You have the right to withdraw at this stage of this interview for any reason or for no reason. Any data extracted from this interview will be treated highly confidential and your privacy will be highly considered by the researcher.

Before starting the interview, could you please tell me a bit about your career? Why did you choose TEFL as your professional career?

P: Ok yes thank you very much Omid. I am really happy to be one of your participants in your research. First of all, I should introduce myself. I am …… and I have been teaching English for about 15 years in Iran as a university lecturer. Well, first of all, I finished my BA teaching English as a foreign language in …… University …… It was I think 19… or … I do not remember. I started teaching English in some high schools in … and then the next year it was about 19… or … I was accepted to do my MA in TEFL teaching English as a foreign language in the same university. I finished my MA in 19… and at the same year I started teaching English at university level first of all with pre-university courses at …… University and from that time up to now I have been a faculty member of … University.

R: Thank you, actually you answered my second question which was supposed to go to the university which you graduated from. Sepehr, you have mentioned that you have been a faculty member or a lecturer at the …… University?

P: Yes.

R: What training program designed for academic staff you have been to so far while you were doing your career in your University? Did you take part in any training programmes which are usually designed for promotion of the academic staff?

P: Yes, I think you know that in many … University … there is a department, promotional department of the faculty members

R: Yes, I think its name is human resources something like this

R: Yes, human resources, yeah, I have been lucky to participate in a number of workshops, seminars and both focus on language teaching and research methodologies in general

R: All right, so they were not designed for TEFL lecturers?

P: Research methodologies, … no, for teachers in general. But foreign language teachers there were two or three yes, three workshops one for academic writing, one for no no for
language teaching I don't remember, for language teaching I don't remember but for academic writing yes, for translation yes there was a workshop for translation

**R:** Ok, how did you find them? Were they helpful to your promotion, promoting your effectiveness in teaching? How did you find these workshops? Were you happy with them? Did the workshop satisfy your needs?

**P:** I think these kinds of workshops will be helpful maybe not to teach us anything new very new theoretically or practically, but I think they are really necessary to some extent in order to retrain ourselves to remember maybe what we have read or studied before. Maybe it is necessary to repeat what we have heard of that before or what we have studied before. Yes, I've been happy with them may be to remind myself what I have had, studied before.

**R:** Okay, thank you, Sepehr, let's talk about your interests, which courses are you interested in both in terms of the studying or in terms of teaching?

**P:** I'm interested in translation courses and writing courses.

**R:** And maybe you have got the expertise in these areas?

**P:** I don't claim to be an expert but I am interested in translation classes

**R:** All right then, Sepehr could you please tell me your overall perception of the notion of teacher effectiveness? In other words, what does the idea of teacher effectiveness mean to you? When I expose you to this concept, what comes to your mind and how you define teacher effectiveness?

**P:** I think this question was also one or two questions open ended questions of your questionnaire and I tried my best to answer those questions. I think, yes, teacher effectiveness to is the point that the teacher can influence on the students and it is not just a language matter. I think it the humanity aspect of phenomena. A good teacher should and a good teacher is the person who can have some kind of influence on students regarding encouragement, positive attitude and something like a model. A good teacher can be a model and should function as a model for the students to follow him.

**R:** All right so what qualities an effective teacher should have from your point of view? Do you imagine any specific characteristics for an EFL lecturer? A specific set of characteristics or criteria that an effective lecturer should have in order to be teaching effectively at university?

**P:** First of all, friendly characteristic is a key point to me

**R:** Good …

**P:** And deep knowledge of the content
R: That’s nice, so you mean subject knowledge?

P: Yes, if you are competent enough at the subject knowledge, yes you can be relied on by the students and I told you motivation, the teacher if wants to be effective, should be self-motivated

R: Ok, yeah

P: Yes, yes that’s it.

R: What about the pedagogical skills? so you mentioned that a good teacher should have sort of good command of the subject knowledge but what if the teacher or the lecturer has got the subject knowledge but doesn't know how to convey his or her knowledge to the students?

P: …

R: …

P: Well, as an English teacher or an effective language teacher should be armed with wide range of theoretical bases as well as practical and different kinds of methodologies. I think principles of language teaching are key points. Yes, when you are competent on the subject knowledge, you I think you have to be armed with a wide variety of principles and different kinds of methods and techniques and strategies to convey your knowledge to your students.

R: Okay yes that's fine. So, Sepehr, so far we talked about the notion of teacher effectiveness and what effective teaching means to you but who should say Sepehr is effective? I mean who are the stakeholders that should have a voice in Sepehr’s evaluation at the end of each semester or academic year? In other words who should say Sepehr is effective or not?

P: Yes, different participants in the whole area of language teaching, language learning context. One group is students themselves.

R: That’s nice, so students …

P: Yes, your colleagues, coworkers, colleagues yes, of course it is in the condition that there is a context of friendship with your colleagues. you can share your ideas and you know what is happening in your colleagues’ class and they know what's happening in your class. Yeah, in this situation colleagues can be a good source of making decisions of if Sepehr, Omid, or someone whoever is or is not an effective teacher.

.  
.  
.  

End of excerpt
Appendix 6: A sample of coding process using NVivo

<table>
<thead>
<tr>
<th>Name</th>
<th>Sources</th>
<th>References</th>
<th>Created On</th>
<th>Created By</th>
<th>Modified On</th>
<th>Modified By</th>
</tr>
</thead>
<tbody>
<tr>
<td>General Perception of TE</td>
<td>19</td>
<td>44</td>
<td>29/11/2012 17:54</td>
<td>U</td>
<td>14/11/2012 13:00</td>
<td>U</td>
</tr>
<tr>
<td>Building rapport</td>
<td>1</td>
<td>1</td>
<td>28/11/2012 13:01</td>
<td>U</td>
<td>12/12/2012 17:46</td>
<td>U</td>
</tr>
<tr>
<td>Fair evaluation of students</td>
<td>1</td>
<td>1</td>
<td>29/11/2012 17:55</td>
<td>U</td>
<td>12/12/2012 17:46</td>
<td>U</td>
</tr>
<tr>
<td>Fulfilling course goals</td>
<td>3</td>
<td>5</td>
<td>29/11/2012 17:54</td>
<td>U</td>
<td>12/12/2012 17:46</td>
<td>U</td>
</tr>
<tr>
<td>Improving students' performance</td>
<td>3</td>
<td>11</td>
<td>29/11/2012 17:55</td>
<td>U</td>
<td>12/12/2012 17:46</td>
<td>U</td>
</tr>
<tr>
<td>Needs Analysis</td>
<td>2</td>
<td>2</td>
<td>28/11/2012 17:54</td>
<td>U</td>
<td>28/12/2012 17:46</td>
<td>U</td>
</tr>
<tr>
<td>Process-based understanding</td>
<td>1</td>
<td>1</td>
<td>28/11/2012 17:54</td>
<td>U</td>
<td>12/12/2012 17:46</td>
<td>U</td>
</tr>
<tr>
<td>Respecting students' rights</td>
<td>1</td>
<td>1</td>
<td>28/11/2012 17:54</td>
<td>U</td>
<td>12/12/2012 17:46</td>
<td>U</td>
</tr>
<tr>
<td>Satisfaction</td>
<td>4</td>
<td>13</td>
<td>29/11/2012 17:54</td>
<td>U</td>
<td>12/12/2012 17:46</td>
<td>U</td>
</tr>
<tr>
<td>Administrators</td>
<td>3</td>
<td>4</td>
<td>29/11/2012 17:54</td>
<td>U</td>
<td>20/12/2012 16:34</td>
<td>U</td>
</tr>
<tr>
<td>Students</td>
<td>5</td>
<td>4</td>
<td>29/11/2012 17:54</td>
<td>U</td>
<td>29/11/2012 16:54</td>
<td>U</td>
</tr>
<tr>
<td>Teachers</td>
<td>3</td>
<td>4</td>
<td>29/11/2012 17:54</td>
<td>U</td>
<td>12/12/2012 17:46</td>
<td>U</td>
</tr>
<tr>
<td>Self-monitoring</td>
<td>1</td>
<td>1</td>
<td>29/11/2012 17:54</td>
<td>U</td>
<td>12/12/2012 17:46</td>
<td>U</td>
</tr>
<tr>
<td>Teachers' authority</td>
<td>1</td>
<td>1</td>
<td>29/11/2012 17:54</td>
<td>U</td>
<td>12/12/2012 17:46</td>
<td>U</td>
</tr>
<tr>
<td>Teachers' knowledge</td>
<td>1</td>
<td>1</td>
<td>29/11/2012 17:54</td>
<td>U</td>
<td>12/12/2012 17:46</td>
<td>U</td>
</tr>
<tr>
<td>Transferring knowledge</td>
<td>6</td>
<td>6</td>
<td>29/11/2012 17:54</td>
<td>U</td>
<td>28/12/2012 17:06</td>
<td>U</td>
</tr>
<tr>
<td>Ideal (critical) Appraisal Model</td>
<td>12</td>
<td>40</td>
<td>29/11/2012 17:54</td>
<td>U</td>
<td>20/12/2012 18:44</td>
<td>U</td>
</tr>
<tr>
<td>Non-teacher factors influencing TE</td>
<td>24</td>
<td>62</td>
<td>29/11/2012 17:54</td>
<td>U</td>
<td>14/11/2012 13:03</td>
<td>U</td>
</tr>
<tr>
<td>Qualities and Standards</td>
<td>36</td>
<td>260</td>
<td>29/11/2012 17:54</td>
<td>U</td>
<td>12/12/2012 17:45</td>
<td>U</td>
</tr>
<tr>
<td>The Irrelevant Appraisal Model</td>
<td>27</td>
<td>63</td>
<td>29/11/2012 17:54</td>
<td>U</td>
<td>12/12/2012 17:45</td>
<td>U</td>
</tr>
<tr>
<td>Ways to Enhance TE</td>
<td>21</td>
<td>60</td>
<td>29/11/2012 17:54</td>
<td>U</td>
<td>14/11/2012 13:03</td>
<td>U</td>
</tr>
</tbody>
</table>
Appendix 7: Ethical Approval Form

Certificate of ethical research approval
DISSERTATION/THESIS

Your student no: 59065576

Title of your project:
TEFL Lecturers' Effectiveness in Iranian Higher Education Context: Towards a Critical Appraisal Model

Brief description of your research project:
The current study endeavours to critically explore TEFL lecturers' perceptions and understanding of the concept of Teacher Effectiveness. Furthermore, this study aims at exploring the participants' perceptions of the existing appraisal model for evaluating teacher effectiveness utilised in Iranian higher education context. The present research project also tries to explore lecturers' opinions about the measures of evaluation of teacher effectiveness, and the extent to which each of these measures should be given weight. Moreover, lecturers' points of view about the standards/criteria for an effective teacher are amongst the areas to be investigated in the present research.

Give details of the participants in this research (giving ages of any children and/or young people involved):
The participants of the study will be all TEFL lecturers teaching at Iranian universities and higher education centres. Both genders, i.e., male and female lecturers will take part in both quantitative and qualitative phases of the research. The participants will be lecturers with Master's or Doctoral degrees.

Give details (with special reference to any children or those with special needs) regarding the ethical issues of:
a) informed consent: Where children in schools are involved this includes both headteachers and parents. Copies of your consent form(s) you will be using must accompany this document. A blank consent form can be downloaded from the GSE student access online documents.

As stated earlier, the participants of this study will be all TEFL lecturers. Prior to the conduct of the study and data collection phase, the participants of both quantitative and qualitative phases will be informed of a detailed account of the research goals. They will be also informed of their rights to withdraw at any stage during the conduct of the study. All the participants will be requested to sign a voluntary informed consent form.

b) anonymity and confidentiality

The issues of anonymity and confidentiality are highly recognised by the researcher. Moreover, the participants will be assured of their privacy. Taking the issues of anonymity and confidentiality into consideration, the researcher will let the participants decide freely to include their names both in the questionnaires and in the interviews.

Give details of the methods to be used for data collection and analysis and how you would ensure they do not cause any harm, detriment or unreasonable stress:

This study is mainly interpretative in nature as it is aimed at exploring lecturers' perceptions. In order to reach comprehensive and rich data, a mixed methods design and, in particular, an exploratory mixed methods design is to be adopted for the present study. Utilising this design, the researcher will have

Chair of the School's Ethics Committee
updated: April 2011
the privilege to adopt a number of methods for data collection. Both quantitative and qualitative methods for data collection will be used in this study. The conduct of the research will commence by the quantitative phase in which 'questionnaire' will be administered. Then, a number of participants will be invited for the second phase of the research in which 'interview' will be used as the qualitative instrument for collecting the data. It is worth mentioning that the participants of this study, as TEFL lecturers, will be all familiar with the academic discourse and the nature of doctoral research projects. Moreover, this study endeavours to explore participants' perceptions and understanding of the issue of teacher effectiveness and its pertinent evaluation system. In other words, this research is not aimed at evaluating teachers themselves which, in turn, will provide the participants with a non-threatening and friendly environment in which they can freely express their thoughts and feelings.

Give details of any other ethical issues which may arise from this project (e.g. secure storage of videos/recoded interviews/photos/completed questionnaires or special arrangements made for participants with special needs etc.):

Although the present study is to be conducted in Iran, it is informed by BERA Ethical Guidelines for Educational Research. Both the completed questionnaires and the recorded interviews will be stored securely by the researcher.

Give details of any exceptional factors, which may raise ethical issues (e.g. potential political or ideological conflicts which may pose danger or harm to participants):

It seems unlikely that this study will have any exceptional factors such as political issues which may raise ethical considerations. However, the researcher will be well-aware of and prepared for any unforeseen circumstances and will make every endeavour to recognise the participants' rights.

This form should now be printed out, signed by you on the first page and sent to your supervisor to sign. Your supervisor will forward this document to the School's Research Support Office for the Chair of the School's Ethics Committee to countersign. A unique approval reference will be added and this certificate will be returned to you to be included at the back of your dissertation/thesis.

N.B. You should not start the fieldwork part of the project until you have the signature of your supervisor.

This project has been approved for the period: oct 2011 until: oct 2012

By (above mentioned supervisor's signature): [Signature]

Date: 13/11/2011

N.B. To Supervisor: Please ensure that ethical issues are addressed annually in your report and if any changes in the research occur a further form is completed.

GSE unique approval reference: 1213

Signed: [Signature]

Date: 13/11/2011

Chair of the School's Ethics Committee

This form is available from: http://education.exeter.ac.uk/ugreform/ethics/ctt

Chair of the School's Ethics Committee

Updated: April 2011
References


Educational Testing Service (ETS).


Ellis, R. (2012). *Language teaching research and language pedagogy*. Chichester, West Sussex: John Wiley & Sons Ltd.


Ware, J. E., & Williams, R. G. (1975). The Dr. Fox effect: A study of lecturer effectiveness and ratings of instruction. *Journal of Medical Education, 50*(2), 149-156.


