The Assessment Practices of In-service Kurdish Tertiary TESOL Teachers and their Cognitions of Alternative Assessment

Submitted by

Dler Abdullah Ismael Ismael

to the University of Exeter as a thesis for the degree of Doctor of Education in TESOL

September 2016

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.
Abstract

Assessment, measurement, testing, and evaluation are similar procedures for collecting information (Bachman & Palmer, 2010), having the same objective of making inferences about students’ proficiencies (Schmitt, 2002). However, assessment differs from testing significantly because of the wide range of aspects that the former includes. This is closely related to alternative assessments (AAs), which are ongoing processes that involve teachers and students in making judgments on the latter’s progress by using non-conventional strategies (Hancock, 1994).

This study is about the current assessment practices and AAs of in-service Kurdish tertiary TESOL teachers and their cognitions of AAs. It was a piece of action research informed by critical theory, and used questionnaires and interviews with 90 survey participants and 12 interviewees. It was conducted in the 18 English departments of the public universities of the Kurdistan Region. It had five research questions divided into two phases. The first-phase questions concerned the current assessment practices and AAs of the participants, their knowledge of AA, and their beliefs about whether AAs are beneficial. The first phase was used as a rationale for a second-phase action research intervention, in which I presented two seminars in two English departments. The third-phase questions examined the influence of the intervention on increasing the participants’ knowledge of AA, changing their beliefs on AA, and their beliefs about the feasibility and challenges of AAs in their professional context.

The study revealed several important findings. Firstly, the assessment practices of the participants included various AAs, but they needed development in terms of using criteria in marking and implementing them more frequently in a longer time. Secondly, the participants’ knowledge of AA needed to be increased due to four factors: (1) AA was not included in their MA or PhD courses; (2) the limited number of resources on AA; (3) the limited number of training courses in assessment or AA; and (4) the participants’ limited knowledge about the recent teaching, learning, socio-cultural and critical theories that underpinned AA. Thirdly, the participants believed that AAs are beneficial but they could not use the full potential of them, mainly due to: (1) their need of more knowledge about AA; (2) their need for training courses to acquire the necessary skills to do AAs; (3) the limited time spent conducting AAs; and (4) the large number of students in their classes. Fourthly, the participants believed that if more AAs were implemented in the classroom, this would
decrease the time spent marking test papers at home. Fifthly, the participants believed that the main challenges to AAs were related to: (1) the English departments’ assessment system; (2) teachers and students’ subjectivity; and (3) managerial and technical challenges such as limited time, high number of students, and provision and maintenance of classroom technical facilities. The findings and conclusions of this study could have pedagogical and teacher development implications for a wider implementation of AAs.
Acknowledgements

I would like to express my profound thanks to my primary supervisor, Dr Salah Troudi, who gave me detailed and constructive feedback on soft copy files, valuable comments during face-to-face meetings, and consistent guidance and support. Significant thanks should also be given to my second supervisor, Dr Susan Riley, who gave me vital feedback on all chapters of my research, which was extremely constructive and helpful. I am also profoundly grateful to my mentor Dr Andrew Richards, who encouraged me a lot in our meetings and who was interested to know how my research was progressing. I am really grateful to all my module tutors, who taught me about TESOL subjects and issues during my first-year EdD studies (2012-2013). I appreciate greatly the help of the staff of St Luke’s Library of the University of Exeter in locating numerous resources, and I also thank the David Wilson library at the University of Leicester, at which I used numerous resources. I am immensely grateful to my sponsor, the Ministry of Higher Education and Scientific Research – Kurdistan Regional Government, for fully sponsoring me to complete my EdD studies at the University of Exeter. Special thanks go to my friends Sarbaz, Qadir and Ali for helping me in obtaining many resources. I am really thankful to my friends Omer Fatehulla, Esmat Khabur, and Karmand, who helped me a lot with the technological requirements of my study and repairing some PC programs. I am tremendously grateful to my colleague and fellow student Khadija Balouchi for helping me perform the Cronbach’s Alpha test for the questionnaires. I greatly thank my friend Hogr for checking the statistical data of the questionnaires. I also warmly thank my close friends Soran and Barham for their continuous encouragement. Very special thanks to all the teacher participants and the heads and deputy heads of the 18 English departments in which this study was conducted for their great cooperation and encouragement. Last but not least, I would like to thank my whole family for continuously supporting me from my home country throughout the duration of my EdD studies at the University of Exeter in the United Kingdom.
Table of Contents

Abstract ........................................................................................................................................... 3
Acknowledgements .......................................................................................................................... 5
Table of Contents ........................................................................................................................... 6
List of Tables ................................................................................................................................... 12
List of Figures ................................................................................................................................... 13
List of Appendices .......................................................................................................................... 14
List of Abbreviations ....................................................................................................................... 15

Chapter One: Introduction .............................................................................................................. 17
  1.1 Nature of the problem ........................................................................................................... 17
  1.2. Rationale for the study ...................................................................................................... 19
  1.3. Aims of the study .............................................................................................................. 21
  1.4 Research questions ............................................................................................................. 22
  1.5 Organisation of the thesis ................................................................................................. 24

Chapter Two: Context of the Study .............................................................................................. 25
  2.1 The Kurdistan Region ........................................................................................................ 25
  2.2 Education in the Kurdistan Region and recent reforms ................................................. 25
  2.3 Higher education in the Kurdistan Region and ongoing reforms .................................... 27
  2.4 ELT in the Kurdistan Region and recent developments .................................................. 29
  2.5 English departments of the universities of the Kurdistan Region and their assessment system ......................................................................................................................... 29
  2.6 Efforts to reform ELT and assessment in the English departments and the influence of Quality Assurance .................................................................................................................. 31

Chapter Three: Literature Review ............................................................................................. 34
  3.1 Introduction ......................................................................................................................... 34
  3.2 Assessment and testing ...................................................................................................... 35
  3.3 Emergence and evolution of alternative assessment.......................................................... 36
3.4 Conceptualisation of alternative assessment ................................................................. 36
  3.4.1 General definitions of alternative assessment .......................................................... 37
  3.4.2 Intricacy and overlapping nature of conceptualisation of alternative assessment .. 37
  3.4.3 Intermixed nature of the types, methods, practices, activities, and strategies of alternative assessment ................................................................. 38
  3.4.4 Some notable characteristics of alternative assessment ............................................. 38
    3.4.4.1 Performance and authenticity .......................................................................... 39
    3.4.4.2 Multiple practices of alternative assessment and a holistic view of language 39
    3.4.4.3 Alternative assessment integrity with instruction ............................................ 40
    3.4.4.4 Duration of performing alternative assessment ............................................... 40
3.5 Theoretical background of alternative assessment ......................................................... 41
  3.5.1 Theories that emphasise the social aspect of learning and learners’ pivotal roles .. 43
    3.5.1.1 Socio-cultural approach .............................................................................. 43
    3.5.1.2 Constructivism ......................................................................................... 44
    3.5.1.3 Multiple intelligences .................................................................................. 45
  3.5.2 Critical theories that emphasise students’ assessment rights and ethicality of assessment ................................................................................................. 46
  3.5.3 Theoretical framework of the study ......................................................................... 50
3.6 Benefits and feasibility of applying alternative assessments: Theoretical and empirical outcomes ........................................................................................................ 51
  3.6.1 Interrelation of the benefits and feasibility of alternative assessments ............... 51
  3.6.2 General benefits of alternative assessments .......................................................... 51
  3.6.3 The wide spectrum of the types, methods, practices, activities, and strategies of alternative assessment ................................................................. 52
  3.6.4 Multi-faceted and comprehensive assessment of students’ knowledge and abilities ........................................................................................................... 52
  3.6.5 Integration of teaching, learning, and assessment .................................................... 53
  3.6.6 Self-assessment and critical thinking ...................................................................... 54
3.6.7 Focus on students’ strengths ................................................................. 55
3.6.8 Democratisation of assessment through alternative assessments ............... 55

3.7 Empirical research on the benefits and feasibility of alternative assessments at the tertiary level ................................................................. 56
3.7.1 Effectiveness of teacher cognition: Empirical studies of teachers’ beliefs on alternative assessments ......................................................... 56
3.7.2 Empirical research on the tertiary TESOL teachers’ beliefs about the benefits and feasibility of alternative assessments ........................................... 58
3.7.3 Additional research findings of the benefits and feasibility of alternative assessments ......................................................................................... 61

3.8 Challenges of the implementation of alternative assessments: Theoretical and empirical outcomes ........................................................................... 63
3.8.1 Challenges of the integration of learning and assessment ......................... 63
3.8.2 Ethical challenges of alternative assessments ........................................ 63
3.8.3 Challenges of class size and limited time .............................................. 64

3.9 Empirical research on the challenges of implementing alternative assessments at the tertiary level ......................................................................... 64
3.9.1 Studies on teachers’ cognition of the challenges of alternative assessments .... 64
3.9.2 Challenges of self-assessment and peer-assessment .................................. 66

Chapter Four: Design of the Study ................................................................. 68
4.1 The informing research paradigm: Critical theory ........................................ 68
4.2 Methodology: Action research ..................................................................... 68
4.2.1 Definitions and aims of action research .................................................. 69
4.2.2 Rationale for choosing action research methodology for this study .......... 70
4.3 Data collection methods ............................................................................. 72
4.3.1 Stages of data collection ......................................................................... 72
4.3.2 First-phase instruments: Survey questionnaires and interviews ................ 75
4.3.3 Second-phase intervention: Presentation of two seminars ...................... 77
4.3.4 Third-phase instruments: Interviews and focus group interviews .................................... 78
4.4 The participants .................................................................................................................. 79
4.5 Data collection procedures ............................................................................................... 81
  4.5.1 First-phase data collection .............................................................................................. 81
    4.5.1.1 Piloting and distributing the questionnaires .............................................................. 81
    4.5.1.2 Piloting and conducting the interviews ................................................................. 82
  4.5.2 Third-phase data collection .............................................................................................. 83
    4.5.2.1 Piloting and conducting the interviews ................................................................. 84
    4.5.2.2 Holding the two focus group interviews ............................................................... 84
4.6 Data analysis procedures ................................................................................................... 84
  4.6.1 Counting and processing the responses of the first-phase questionnaire items ..... 85
  4.6.2 Coding, categorisation, and thematisation of interview transcriptions and responses
      to the open-ended questionnaire question ......................................................................... 85
4.7 Role of the researcher ......................................................................................................... 88
4.8 Credibility and dependability of the findings ..................................................................... 89
4.9 Ethical considerations ......................................................................................................... 90
4.10 Limitations of this study ................................................................................................... 91

Chapter Five: Findings and Discussions ................................................................................. 94

5.1 The assessment practices and alternative assessments of Kurdish tertiary TESOL
      teachers .................................................................................................................................. 95
  5.1.1 Alternative assessments and test questions................................................................. 95
  5.1.2 The marking of alternative assessments and tests ....................................................... 98
5.2 Discussion of the assessment practices and alternative assessments of Kurdish tertiary
      TESOL teachers .................................................................................................................... 101
5.3 Assessment literacy of Kurdish tertiary TESOL teachers in alternative assessment .. 104
  5.3.1 Participants’ knowledge of alternative assessment ....................................................... 104
  5.3.2 Factors behind the participants’ limited assessment literacy in alternative
      assessment ............................................................................................................................... 107
5.4 Discussion of the assessment literacy of Kurdish tertiary TESOL teachers in alternative assessment .................................................................................................................. 109
5.5 Beliefs of Kurdish tertiary TESOL teachers on whether alternative assessments are beneficial .................................................................................................................. 114
  5.5.1 Benefits of alternative assessments to teaching ........................................ 114
  5.5.2 Benefits of alternative assessments to learning ....................................... 115
    5.5.2.1 The improvement of students’ learning by alternative assessments .... 115
    5.5.2.2 Using various alternative assessments to improve learning and assessment accuracy ...................................................................................................................... 117
    5.5.2.3 Students as assessors, critical thinkers, and autonomous learners .... 120
  5.6 Discussion of the beliefs of Kurdish tertiary TESOL teachers on whether alternative assessments are beneficial .......................................................................................................... 122
  5.7 The intervention outcome: Introducing participants to alternative assessment .... 129
  5.8 Discussion of the intervention outcome .......................................................... 132
  5.9 Beliefs of Kurdish tertiary TESOL teachers on the feasibility and challenges of alternative assessments ........................................................................................................ 135
    5.9.1 Feasibility of implementing alternative assessments ............................... 135
    5.9.2 Challenge (1): The assessment system of the English departments and their context ................................................................................................................................. 136
    5.9.3 Challenge (2): Teachers and students’ subjectivity in marking alternative assessments ................................................................................................................................. 139
    5.9.4 Challenge (3): Management and technology ........................................... 141
  5.10 Discussion of the beliefs of Kurdish tertiary TESOL teachers on the feasibility and challenges of alternative assessments .......................................................................................... 143

Chapter Six: Conclusion .......................................................................................... 150
  6.1 Conclusions ........................................................................................................ 151
  6.2 Implications: Pedagogy and teacher development ........................................ 155
  6.3 Recommendations ............................................................................................ 156
6.4 Contribution of the study to the field of alternative assessment: Research and pedagogy ...................................................................................................................... 157

6.5 Suggestions for further research .................................................................................................................. 159

6.6 Personal reflections on the study .............................................................................................................. 159

Appendices .................................................................................................................................................... 161

References ..................................................................................................................................................... 229
List of Tables

Table 1: Phases of the design of this action research ......................................................... 74
Table 2: SPSS Cronbach’s Alpha test for the internal consistency reliability of the questionnaire items ........................................................................................................ 82
Table 3: Alternative assessments used by the survey participants ................................. 96
Table 4: Types of test questions and items used by the survey participants ................. 98
Table 5: The number of marks allocated for alternative assessments and tests ............ 99
Table 6: Survey participants’ readings and knowledge of alternative assessment .......... 104
Table 7: Training courses that the survey participants attended .................................... 108
Table 8: Survey participants’ ratings of the benefits of alternative assessments to teaching 114
Table 9: Survey participants’ ratings of the benefits of alternative assessments to learning 115
Table 10: Survey participants’ beliefs on the accuracy of alternative assessments .......... 118
Table 11: Survey participants’ beliefs on the feasibility of alternative assessments .......... 135
Table 12: Survey participants’ beliefs on the challenges of alternative assessments ....... 137
List of Figures

Figure 1: Classification of the types of theories underpinning, influencing, or interrelated with alternative assessment ........................................................................................................................................ 42

Figure 2: The theoretical framework: Interrelation of the features and benefits of alternative assessment with the two main aspects of its underpinning theories ................................................. 49

Figure 3: Bachman’s action research spiral (Mertler, 2013:18) ................................................................. 70

Figure 4: Riel’s action research model (Mertler, 2013:19) .................................................................... 71
List of Appendices

Appendix 1: List of themes, categories, and sub-categories from the whole collected data. 161
Appendix 2: The findings, intervention, implications, and recommendations according to the phases of this action research ................................................................. 172
Appendix 3: Statistical results of the first-phase questionnaire items ......................... 174
Appendix 4: The themes, categories, and sub-categories of the answers to the open-ended questionnaire question ................................................................. 178
Appendix 5: Types, methods, activities, practices, and strategies of alternative assessment and its different terms in literature................................................................. 179
Appendix 6: First-phase interview questions ................................................................. 186
Appendix 7: Slides used in the seminars .................................................................... 187
Appendix 8: Third-phase interview questions ............................................................... 194
Appendix 9: Issues discussed in the focus group interviews ....................................... 195
Appendix 10: Guide information sheet ...................................................................... 196
Appendix 11: Consent form ......................................................................................... 199
Appendix 12: Application for a Certificate of Ethical Approval................................. 201
Appendix 13: Certificate of Ethical Approval ............................................................... 206
Appendix 14: Transcript of one interview showing how it was coded ....................... 207
Appendix 15: A sample showing the summarising and coding responses to the open-ended questionnaire question ................................................................. 227
Appendix 16: A piloting SPSS Cronbach’s Alpha test for the internal consistency reliability of the questionnaire items ........................................................................... 228
List of Abbreviations

AA  Alternative Assessment
AAs  Alternative Assessment types, methods, activities, practices, and strategies
ALP  Advanced Language Proficiency
CALPER  Centre for Advanced Language Proficiency Education and Research at Pennsylvania State University, USA
CALx  Critical Applied Linguistics
EFL  English as a Foreign Language
ELL  English Language Learner
ELP  English Language Proficiency
ELTPN  English Language Teaching Professional Network
ESL  English as a Second Language
HCDP  Human Capacity Development Programme
KRG  Kurdistan Regional Government
L2  Second Language
MHESR/KRG  Ministry of Higher Education and Scientific Research – Kurdistan Regional Government
(Name, Int1)  (Name of the interviewee, interview 1, the number denotes the first phase)
(Name, Int3)  (Name of the interviewee, interview 3, the number denotes the third phase)
(Name, Foc3) (Name of the interviewee, focus group interview 3, the number denotes the third phase)

RAND Research and Development. The RAND Corporation is a non-profit institution that helps improve policy and decision-making through research and analysis (Vernez et al., 2014).

SP Survey Participant

TEFL Teaching English as a Foreign Language

TESL Teaching English as a Second Language

TESOL Teaching English to Speakers of Other Languages

TFK Translated from Kurdish

TKT Cambridge Teaching Knowledge Test

ZPD Zone of Proximal Development
Chapter One: Introduction

1.1 Nature of the problem

In current times, teaching in many parts of the world is in the midst of a remarkable transformation. Concurrent with this is the continual escalation of teachers’ expectations to get their students to high standards of performance and to ensure their learning abilities (Hargreaves, 2000). Educational systems are under pressure to be accountable for students’ performances and to produce measurable results (Coombe et al., 2012). Alongside changes in education, assessment is in need of reconsideration because it is “an integral part of the teaching-learning process” (Coombe et al., 2012:20), and has a profound influence on improving students’ learning (Earl & Katz, 2006). However, until recently, in many educational contexts, assessment processes have been largely based on standardised ‘pen and paper’ tests, and these continue to cover a significant proportion of assessments (Ataç, 2012). ‘Pen and paper’ tests are part of a traditional assessment paradigm underpinned potentially by positivism that regards language ability as parallel to physical world objects (Lynch & Shaw, 2005), which means that it can be tested and measured as an object rather than a social component.

Along with some researchers and practitioners in the language-teaching field, and specifically those involved in TESOL, I question the educational value of the traditional approach to assessment (see Section 3.6.4). TESOL stands for teaching English to speakers of other languages, which also incorporates the situations of teaching English as a second or foreign language (Nunan & Carter, 2001).

Nevertheless, Brindley (2001) has noted that rapid and sustainable development is evident in the area of assessment because its issues have gained prominence in the agendas of educational authorities worldwide. This highlights an increasing interest in assessment methods (Ataç, 2012), which may be attributed to the impact assessment has on education, because assessment results are not only used to make academic decisions, but also they have powerful implications for improving the quality of teaching and learning (Wolf et al., 2008). Assessment is considered the engine of reform (Winke, 2011) that drives learning (Coombe et al., 2012). There is abundant evidence that shows the influence of assessment on enhancing the learning of English language (Birjandi & Mahnaz, 2014; Ghahremani, 2013; Shabani,
2012; Ajideh & Nourdad, 2012; Ataç, 2012; Grami, 2010; Wolf et al., 2008). Moreover, Black and Wiliam (1998, cited in Earl & Katz, 2006) synthesised more than 250 studies that linked assessment to learning. The researchers covered in the overview concluded that the intentional use of classroom modes of assessment to promote learning improved students’ achievements.

One consequence of the rapid development of assessment is the addressing of some debatable issues within the field of language assessment, such as how and why language assessment forms are used, their underlying societal values, the consequences of assessment, and the ethical responsibilities of test developers and users (Bachman, 2007). These can be possible solutions to the shortcomings of traditional assessment. Such multi-dimensional considerations of language assessment could also be linked to the emergence of alternative assessment (AA) and its foundation theories, in particular constructivism, socio-cultural theory, and critical approaches such as critical language testing (see Sections 3.3 and 3.5.3). These considerations can lead to more democratic testing and assessment methods, whereby test-takers and local bodies are more involved (Dai Quang, 2007), because these theories and approaches emphasise the pivotal role of students and foster the discourse of ethicality in language assessment (see Section 3.5.3). This ethicality could be further enhanced by Shohamy’s principles of critical language testing, which question the values, agendas, needs, purposes, intended goals, and the actual uses of tests in language assessment (Shohamy, 2001b, cited in Bachman, 2005). Assessment ethicality and democratic assessment can be increased by Freire’s anti-authoritarian, dialogical, and interactive approach, which attempts to examine the relational power issues for students, along with his praxis that requires the implementation of educational practices to create a better learning environment (Chandella & Troudi, 2013).

Taking into account these mainstream and critical considerations of language assessment necessitates the establishment of an assessment system with sound measures and continuous validation. The area of assessment that helps establish such a sound assessment system and addresses the shortcomings of traditional testing is AA, which is informed by many theories, for example interpretivism that regards language proficiency as part of the social world (Lynch & Shaw, 2005). Defining language as a construct (whether it is part of the social or physical world) is significant because it accords with Shohamy’s urge to examine critically
the definition of language ability due to its fundamentality in language assessment (Shohamy, 1998, cited in Piggin, 2012).

The fundamental need for sound assessment in education, alongside the high demands for public accountability and the increasing concern with standards, make assessment literacy more essential than ever. Therefore, educational institutions have introduced various assessment guidelines to make decisions on selection, certification, and achievement (Brindley, 2001). This has possibly been fulfilled by committing to many standards and ethical codes of assessment practice (Taylor, 2009). However, despite the significance of teachers’ literacy in assessment, in AA and in related standards, and the need for language teachers to learn the principles and practices of assessment (Brindley, 2001), Coombe et al. (2012) state that almost all teachers perform assessment without having learned the principles of sound assessment. This is in agreement with Taylor (2009), who discovered that many people are involved in assessment processes, but often without acquiring a background in assessment and without the necessary training for this role. Such a need is stressed by Bailey and Wolf (2012), Troudi et al. (2009), and Brindley (2001), who all confirm that local TESOL teachers should be supported in the field of assessment literacy in order to broaden and improve their assessment practices. In order to do so, TESOL teachers should follow a variety of AAs to ensure fair assessment for students (Troudi et al., 2009), which would enable students to demonstrate their knowledge (and abilities) in many ways via a range of assessment approaches and multiple measures, thus establishing a composite picture of their proficiency (Earl & Katz, 2006). In this study, the plural form of AA, ‘AAs’, is intended to mean the various types, methods, practices, activities, and strategies of alternative assessment.

1.2 Rationale for the study

The rationale for this study has developed from the doubt surrounding the need for increasing the assessment literacy and developing the assessment practices of the majority of in-service Kurdish tertiary TESOL teachers. Such a doubt was based on my reflection after completing the EdD TESOL modules, particularly critical issues in ELT, when I became more aware of the nature of traditional assessment methods and the advances achieved through AA. After reflecting on my previous experiences as a test-taker, a teacher, a deputy head and head of an English department, and a coordinator of quality assurance, I began to question: (1) the use of
traditional testing in the English departments of the public universities of the Kurdistan Region; (2) the validity and ethicality of the assessment forms conducted by their teachers; (3) teachers’ assessment literacy; and (4) students’ dissatisfaction about the test and assessment results. Previously, students did not have the opportunity to critique their teachers’ assessment methods, but recently the Quality Assurance Programme, launched by the Ministry of Higher Education and Scientific Research - Kurdistan Regional Government (MHESR-KRG), has achieved some improvements in the universities of the Kurdistan Region (including the English departments). As such, students are now given the power to evaluate the teaching, learning, and assessment methods of their teachers and the right to see samples of test questions prior to exams (MHESR-KRG, 2010b).

However, despite these developments, AAs require further development to be performed more adequately in the English departments. This is primarily because AA is not incorporated adequately in their current assessment system (see Sections 5.1.1 and 5.1.2) and teachers’ need to increase their assessment literacy in AA (see Sections 5.3.1 and 5.3.2). According to their assessment system, teachers are allowed to use around 10% of the marks for AAs but it is not compulsory and not explained in the guidelines of the English departments (see Section 5.1.2). In these English departments, where traditional tests are mostly known and implemented, teachers may need to be introduced to the advances achieved in the field of language assessment, in particular through AA and assessment standards, in which assessment use, its societal values, its consequences, and ethicality are considered (Bachman, 2007). Exploring this issue of teacher literacy constitutes a core element of the rationale for this study, since more assessment literacy in the developments of assessment practices and standards can certainly influence ELT positively. At present, however, this requires development, which could possibly lead to the teachers’ improved integration of multiple methods of AA into teaching and learning. Hence, there is a need to introduce the participants to AA and its standards and requirements in order to increase their assessment literacy in AA (see Section 5.3.1). An additional complication is associated with the feasibility and challenges of the performance of AAs. Some AAs are perhaps feasible for the tertiary EFL students in the Kurdistan Region, yet several challenges may emerge in conducting many of them (see Section 5.9). Therefore, another influential aspect of the rationale for this study is to investigate the extent to which AAs are practical alongside the associated challenges.
The field of AA in the English departments of the universities of the Kurdistan Region constitutes a significant gap of the existing academic research on language assessment, and as such it needs to be investigated in many dimensions. Broadly speaking, Brindley (2001) highlights numerous research gaps in the area of language assessment, clarifying that despite having a considerable amount of research into teachers’ assessment practices and professional development needs, relatively few investigations have been conducted into those issues in the context of language teaching programmes. This is obvious in the context of this study, because according to my thorough search of the extant literature, there has not yet been any research conducted solely on assessment practices. There is therefore a need for further research on the assessment of English language proficiency in the English departments of the universities of the Kurdistan Region.

The current study is of significance because it explores assessment practices and AAs, and the need for increasing the assessment literacy of tertiary TESOL teachers in AA, along with their beliefs about the feasibility and challenges of conducting AAs. That is why one of the goals of this study is to equip Kurdish tertiary TESOL teachers with some introductory information on AA and related standards, thus making this study pedagogically significant for developing TESOL teachers’ implementation of AAs that ensures fair assessment for their students (Troudi et al., 2009). This is essential because teachers cannot afford to ignore the latest developments in AA and how they are related to TESOL International Association standards and critical language testing principles (see Section 5.4). The TESOL standards have reached nearly universal status (Cumming, 2009), have thrust assessment into a more prominent position in the learning process, and have influenced assessment in language teaching considerably (Grabin, 2009).

1.3 Aims of the study

The topic of this study was explored through an action research methodology in three phases. The aim of this study in the first phase was to explore the use of AA in the assessment practices of Kurdish tertiary TESOL teachers and the amount of their assessment literacy in AA. The goal was to examine any room for development regarding their AAs and/or their assessment literacy in AA, which is necessary in action research in order to plan professional support for the participants as an intervention (Bachman, 2001; Riel, 2007). Therefore, the aim of the second phase was, if necessary, to increase the participants’ knowledge of AA,
change their beliefs about it, and encourage them to pay greater attention to it. This was also to make them more aware of the demands of conducting AAs, so that they could reflect on the practicality issues of implementing AAs within their professional context in order to improve them. Bailey and Wolf (2012), Troudi et al. (2009), and Brindley (2001) suggest that providing professional development opportunities for TESOL teachers is fundamental to facilitating the acquisition of greater assessment literacy.

The aims of this study were compatible with the main goal of action research, which is generally an understanding of practice and the community in which practice occurs in order to improve them in a process of reform (Carr & Kemmis, 2003; Riel, 2010a; Hopkins, 2008; Ebbutt, 1985, cited in Cohen et al., 2007) (see Section 4.2.1). The aims of this action research, which was divided into three phases, were represented by its five research questions. These phases were based on some elements of Bachman’s (2001) action research spiral, in which “participants gather information, plan actions, observe and evaluate those actions, and then reflect and plan for a new cycle of the spiral”. The phases were also partially based on Riel’s (2007) “progressive problem solving through action research model [that] takes the participant through four steps in each cycle: planning, taking action, collecting evidence, and reflecting” (Mertler, 2013:16).

1.4 Research questions

As shown below, the sequence of the research questions of this action research was based on the three phases of the research design (see Table 1). These phases were based on the “cycles of planning, acting, observing and reflecting”, which constitute the central phases of any action research approach (Carr & Kemmis, 2003:162). More specifically, these phases were based on some elements of the phases of Bachman’s (2001) and Riel’s (2007) models of action research presented in Mertler (2013). The three research questions of the first phase investigated the assessment practices and AAs used by the participants, as well as their knowledge of AA and their beliefs on whether AAs are beneficial. These pre-intervention questions were used as a preparation for the action research intervention. This represented the phases of gathering data and planning actions according to the two models of action research designed by both Bachman (2001) and Riel (2007), in which data should be collected to plan and justify the actions. The second phase had no research question as it was devoted to the action research intervention, which was based on the phase of taking action
and observing according to the two models. The third phase contained two research questions, the first of which focused on the impact of the intervention on the participants. The last question, which explored the participants’ beliefs regarding the practicality issues of implementing AAs, was divided into two sub-questions, one that looked at the feasibility of AAs and the other at the challenges involved. These post-intervention research questions were compatible with the last two phases of reflection on and evaluation of the outcome of actions according to the two adopted models.

With the exception of Research Question 1, all the other four research questions focused on the participants’ cognitions of AA, and thus contained three key terms associated with teacher cognition research: cognition, knowledge, and belief (see Section 3.7.1 for clarification of these terms as were intended and justified to be used for this action research).

This study seeks answers to the following five research questions. The first three questions pertain to the pre-intervention phase of this action research, whereas the latter two are part of the post-intervention phase.

RQ1. What are the current assessment practices and alternative assessments used by in-service Kurdish tertiary TESOL teachers?

RQ2. Do in-service Kurdish tertiary TESOL teachers have knowledge about alternative assessment?

RQ3. Do in-service Kurdish tertiary TESOL teachers believe that alternative assessments are beneficial?

RQ4. Have the in-service Kurdish tertiary TESOL teachers’ cognitions of alternative assessment been changed by the intervention?

RQ5. What do in-service Kurdish tertiary TESOL teachers believe about the practicality issues of alternative assessments?

   a. Do they believe that implementing alternative assessments is feasible?

   b. Do they believe that performing alternative assessments would face challenges?
1.5 Organisation of the thesis

This thesis is divided into six chapters. Following this introduction, Chapter 2 gives an explanation of the general and academic context where this study is conducted; namely, the Kurdistan Region and the English departments of its universities. Chapter 3, the literature review, briefly reviews the contrast between assessment and testing. Then, it reviews the literature on the emergence, evolution, conceptualisation, theoretical background, benefits, feasibility, and challenges of AAs, and reviews many relevant empirical studies. Chapter 4 presents the informing research paradigm, methodology, instruments, and procedures adopted and employed for the conducting of this study. Moreover, it explains the role of the researcher, participants, ethical considerations, and limitations of this study. Chapter 5 presents and discusses the research findings by addressing all the research questions in sequence to check to what extent the research aims were achieved. Finally, Chapter 6 summarises the findings into some conclusions, their implications, and subsequent recommendations. It also clarifies the contribution this study can make to the field of AA and makes some suggestions for future research in the area of AA within the context of this study. Also, it provides some personal reflections of the researcher on the overall process of this action research.

This introductory chapter has elucidated the nature of the problem and rationale for conducting this action research in the English departments of the public universities of the Kurdistan Region. This chapter also presented the aims of this study and their associated research questions as to how to achieve them. It also explained the aspects of the problem, rationale, aims, and research questions of this study. In the context of this action research, these aspects concern the current assessment practices and AAs used by the participants, their knowledge about AAs, their beliefs about whether AAs are beneficial, intervention to change their cognitions of AA, and their beliefs regarding the feasibility and challenges of AAs. Finally, this chapter has clarified the organization of this thesis regarding the sequence of the chapters and their contents. The next chapter will explain the context in which this action research is undertaken.
2.1 The Kurdistan Region

This section provides a brief historical, geographic, and political description of the region where this study was undertaken. The Kurdistan Region is mostly mountainous, but in the south, it falls away to the dry Mesopotamian plain. Its population is around 6 million (House of Commons, Foreign Affairs Committee, 2015), with the majority of the inhabitants being Kurds, although there are also Assyrians, Chaldeans, Turkmen, Armenians, and Arabs, who have lived together in peace for centuries (Kelly, 2009). After the invasion of Iraq in 2003 and its political changes, a new Constitution of Iraq was ratified in 2005, according to which Iraqi Kurdistan became a federal entity of Iraq, and Arabic and Kurdish became the joint official languages of Iraq (Vernez et al., 2014). Currently, the Kurdistan Region is a multi-party democracy with a cabinet drawn from the 111-member Kurdistan National Assembly, elected via a party list system (House of Commons, Foreign Affairs Committee, 2015). As a de jure region, it has its Kurdistan Regional Government (KRG), which is elected under democratic standards (Kelly, 2009). Previously, it existed in a de facto sense, emerging after the uprising of March 1991 in the Kurdish region, whereby the ethnic Kurdish people and Peshmerga forces launched a rebellion against Saddam Hussein’s regime (Kelly, 2009). After that uprising and after the USA, the UK, and France enforced the no-fly zone in that area by citing UN Security Council authority (House of Commons, Foreign Affairs Committee, 2015), the three northern Kurdish governorates gained autonomy (Harb, 2008), and thus the region became independent of Saddam’s regime (Kelly, 2009). After being freed from Saddam’s repressive hand, the Kurds progressed both politically and economically, providing social services, healthcare, and education (Kelly, 2009), which gained prominence since the invasion of Iraq in 2003. Since then, this region has become more secure and stable than the rest of Iraq, and its economy has developed (Sharif, 2013).

2.2 Education in the Kurdistan Region and recent reforms

UNESCO has been supporting the KRG in the development of education since 2007 by offering expertise, experiences, and evaluation to modernise the curriculum, assistance that was greatly needed because one of the problems of the public schools in the Kurdistan Region prior to UNESCO’s involvement was the dearth of teaching methods and materials
The results of the UNESCO cooperation can be evidenced by the report of the UN Assistance Mission for Iraq, entitled “Educational Reform in the Kurdistan Region of Iraq”, which stated that the region has taken some essential steps towards a better educational system to meet international standards. One of the crucial elements is a focus on creative thinking, which is encouraged in the school curriculum (Sharif, 2013). In 2010, upon the request of the KRG, the RAND Corporation assessed the K–12 system and recommended procedures to improve the quality of education in the Kurdistan Region. That was to help the KRG to build on the recently instituted reforms towards a universal, basic education (Vernez et al., 2014).

Furthermore, through support for improving the quality of education in Iraq, the British Council has worked with teachers and school leaders in all types of educational contexts to support and improve the quality of education. The support included technical and vocational education, training in institutions, physical education, and sport in secondary schools in Kurdistan. The most prominent support, however, is the training of over 45,000 teachers, principals, and supervisors in a more child-centred approach to teaching and learning, which is being implemented in more than 3,700 schools (British Council, 2015). These reforms have also improved the ELT materials and methods in the schools. In support of the major developments, several other changes have been encouraged or implemented in teaching methods, student retention, and assessment of students (Vernez et al., 2014). Therefore, emphasis is also placed on assessment, though only 12% of teachers in the Kurdistan Region announced, as their training priority in 2010, their need to be trained in developing student assessments and analysing the results to improve instruction (Vernez et al., 2014).

Regarding the ELT developments in the school curriculum, the overwhelming IT revolution, globalisation, use of the internet, and use of English as an international language (Hagerman, 2009) have given English language teaching a special place in school timetables worldwide, including the lower stages in primary schools (Graddol, 2008). This is also the situation in the Kurdistan Region, where teaching English starts in the first grade of primary school (Sharif, 2013), which can be considered the first remarkable improvement of ELT, because previously it was taught from the fifth year of primary school onwards. That was the norm of the traditional system of the three distinct levels of education (primary: grades 1-6, intermediary: grades 7-9, and secondary: grades 10-12) (Wells, 2012), which has been...
replaced by a new K–12 two-level system consisting of basic education (grades 1–9) and secondary education (grades 10–12) (Vernez et al., 2014).

Providing schools with useful English books, imported from Finland to the Kurdistan Region (Sharif, 2013), can be regarded as another improvement. For incorporating the new materials into teaching and learning, Vernez et al. (2014) recommend that the KRG should establish regional training centres for upgrading the knowledge and expertise of practicing teachers, and develop curriculum maps to provide them with step-by-step guidance on how and what to teach. These maps should combine recommended content, suggestions for teaching methods and classroom exercises, student assessment, monitoring, and teaching plans, to ensure that teachers present the curriculum in a standardised fashion (Vernez et al., 2014). The recommendations of Vernez et al. (2014) also include assessment and they denote that there is a need for further development in the aspects of qualified teachers, a modern curriculum, and professional development for teaching, learning, and assessment processes. In this regard, for supporting English teachers and learners, the British Council run a number of exciting initiatives to improve the teaching, learning, and assessment of English language by working closely and directly with teachers, ministries, and institutions to raise the standards of ELT. The support covers teacher and trainer training in the primary, secondary, vocational, and tertiary education, establishes professional networks like the ELTPN (English Language Teaching Professional Network), and provides the opportunity to take the Cambridge Teaching Knowledge Test (TKT) (British Council, 2015).

Thus, reforming assessment has already begun but it needs further work, especially at the tertiary level.

2.3 Higher education in the Kurdistan Region and ongoing reforms

The commencement of the development of higher education in the Kurdistan Region dates back to the time after the first Gulf War and the imposition of a no-fly zone in 1991, when the region attained a level of intellectual liberation, during which the Universities of Salahaddin and Sulaimani acquired independence from the Ministry of Higher Education in Baghdad. Since then, “the number of universities has soared from one to 22 and educational reform is a major work in progress” (All-Party Parliamentary Group on the Kurdistan Region in Iraq, 2012:14). The Universities of Duhok and Koya were established by the KRG in 1992 and
2004, respectively (Harb, 2008). Universities in the Kurdistan Region have received assistance programmes from U.S. and European governments and universities, which have included academic exchanges, curriculum materials, information services, and teacher training programmes. This has encouraged a robust higher education system that offers a modern education (Harb, 2009). This includes the growing relationship between the UK and the Kurdistan Region, in which education is a central part. By this, “the Kurdistan Region has invested in transforming the old top-down system, which discouraged open thinking and focused on learning by rote” (All-Party Parliamentary Group on the Kurdistan Region in Iraq, 2012:14).

Three aspects clearly show the developments of higher education in the Kurdistan Region. Firstly, the KRG has launched the Human Capacity Development Programme (HCDP) in order to provide top scholars with the opportunity to study overseas, which is regarded as one of the most important governmental investment projects. Secondly, the MHESR/KRG has introduced the new system of the Quality Assurance Programme, which started with the review of the system and the decision-making process in 2010, after which the reforms began with a pilot project in 2011. Subsequently, the teaching and administrative staff of the Kurdistan Region universities accepted the implementation of the programme (Palander, 2013), which included reforming and updating the undergraduate and postgraduate curricula and creating training programmes compatible with the market needs. Moreover, it focused on increasing students’ abilities and supporting them to be more self-directed and critical learners (Palander, 2013). According to the MHESR/KRG, this new quality culture of the higher education system is associated with economic integration because of the countless investors in the Kurdistan Region that necessitate producing well-qualified graduates with considerable proficiency in English to secure employment (Palander, 2013).

Thirdly, all the aforementioned efforts for reforms may help to internationalise higher education in the Kurdistan Region. However, as mentioned by Ahmad (2014), as Kurdistan is a developing region, the KRG spends less on internationalisation, and so its weakness is rooted in the lack of resources and the universities’ dependence on the government. Therefore, the implementation of internationalisation is regulated by policies that lack flexibility in universities (Ahmad, 2014). Further, although the role of teaching staff is essential in the internationalisation of universities, many teachers did not develop the relevant competence (Ahmad, 2014). This might be firstly due to the outdated local system, which
may explain the need for some improvement in the performance of teaching staff in the Kurdistan Region (Schneider, 2003, cited in Ahmad, 2014). Secondly, while curriculum reform occurred rapidly in the region, there was less focus on the teachers (Ahmad, 2014).

I noticed that the reforms through the quality assurance, HCDP, and internationalisation have, to a certain extent, developed the quality of higher education in many aspects including ELT, especially in relation to the empowerment of students who are now able to assess their teachers.

2.4 ELT in the Kurdistan Region and recent developments

The necessity of English proficiency is a result of the political, economic, and academic developments currently occurring in the Kurdistan Region. Consequently, learning English in the region is much more established than it was in the past. At present, there are 18 English departments at the public universities in the region, in addition to those offered by the private universities. The public universities are Salahaddin, Sulaimani, Duhok, Hawler Medical, and Koye universities, along with the very recently established Soran, Zakho, Raniya, Halabja, Kalar, and Charmo universities. The numerous private universities include the American University in Iraq in Sulaimani city, Kurdistan University in the city of Erbil (Quantrell & Khidir, 2011), and the American University in Kurdistan in Duhok city, established in 2016.

In the English departments of the public universities, all teachers hold either an MA or PhD in English language and linguistics or applied linguistics. The majority are Kurdish, and as such there are only a few native English speakers, Europeans, Arabs, or other foreign TESOL teachers. Students in those departments are required to have graduated from secondary school with high average scores, usually 75 percent or above, and higher scores in English. This is after they pass the Ministerial exam, by which graduates are assigned to the university programmes (Vernez et al., 2014). The majority of students are Kurdish, with a few hailing from the other ethnic minorities in the region.

2.5 English departments of the universities of the Kurdistan Region and their assessment system

Regardless of the focus on English language in the higher education system of the Kurdistan Region, the tertiary English departments need further development and reform in their
teaching, learning, and assessment approaches. This is despite the fact that English has slowly become a second language in the region, and is taught from the very beginning of primary school (Sharif, 2013). The reforming of teaching, learning, and assessment requires teacher education programmes for TESOL teachers, as UNESCO (2004) confirms that English language educators need training programmes for interactive teaching methods, classroom management, guidance, and counselling. The MHESR/KRG has concentrated on training local teachers to understand teaching methods better (Sharif, 2013), but this has not yet reached perhaps as many as half of the Kurdish tertiary TESOL teachers (see Appendix 3).

In the Kurdistan Region, some Kurdish tertiary TESOL teachers have attended short and intensive ELT training programmes abroad, in particular in English-speaking countries such as the United States and the United Kingdom, but these programmes were not available for the majority of the individuals (see Appendix 3). There have been mandatory qualification courses in the universities of the Kurdistan Region for all tertiary teachers regarding the teaching methods in all social, natural, and applied sciences, but these did not include any specific ELT subjects (see Appendix 3). Perhaps for this reason, the teaching, learning, and assessment of English language, including AAs, need more development.

Regarding the assessment system of the English departments of the universities of the Kurdistan Region, the implementation of tests constitutes around 90% of the assessment of students. This possibly confines teachers’ AAs because 80-90% of the marks are decided by tests (see Section 5.1.2 and Table 5). Although teachers are given the freedom to perform and mark AAs with only 10-20 marks throughout the academic year, this is not compulsory. Indirectly, this encourages the conducting of tests mostly at the expense of AA, and therefore there is no balance between tests and AAs (see Sections 5.1.1 and 5.1.2). This also limits committing to the two basic purposes of assessment, ‘assessment as learning’ and ‘assessment for learning’, which are essential parts of implementing AAs as mentioned in Earl and Katz (2006), whereas ‘assessment of learning’ is conducted in an organised way by following traditional testing. Therefore, I question whether AA in these English departments is sufficiently integrated into teaching and learning. Overall, based on my experience as a teacher and learner with some international experience, I feel that practice and teacher development regarding assessment and AA in the English departments need to be reviewed and reformed.
2.6 Efforts to reform ELT and assessment in the English departments and the influence of Quality Assurance

Waters (2009) suggests that innovation, as an attempt to achieve beneficial change, has only recently become a characteristic of ELT; since then efforts have been made to resolve many problems in the field of ELT (Silver et al., 2002). Concerning the English departments of the universities of the Kurdistan Region, up-to-date courses and the latest teaching, learning, and assessment approaches need to be adopted. To this end, a number of practices have taken place recently such as some TESOL teachers attending training courses overseas for developing the ELT methodology in the Kurdistan Region. For instance, ten teachers from four universities were invited to participate in a six-week intensive teacher-training programme at Huddersfield Technical College thanks to the support of the British Council (British Council, 2014). This, and maybe some other attempts, have led to some reforms. For example, in the English Department of Duhok University, team teaching and role-play have been introduced in the conversation classes for the first time. Another improvement has been the introduction of a college notice board, where tertiary EFL students can display their articles, comments, and poems. Furthermore, the English Department of Sulaimani University produces a newsletter in English that features contributions from staff and students (British Council, 2014). Indeed, these new activities can be used as a basis from which to develop AAs.

However, despite all the efforts made by the British Council to reform the English departments of the universities of the Kurdistan Region, there remain some shortcomings. This is perhaps due to the need for more resources, materials, professional training courses, and professionally qualified teachers to follow the newly introduced materials and methods (see Section 5.3.2 and Appendix 3). Regarding the performance of AAs, it is also apparent that little has been done to introduce AA to the Kurdish tertiary TESOL teachers (see Sections 5.1 and 5.3). Therefore, I believe that there is still room for further reforms and improvements in this aspect of implementing AAs as part of ELT. In this direction, additional reforms have already commenced by introducing Quality Assurance of the MHESR (see next section), and it is hoped that this current action research can contribute to this reform process by introducing AAs to participants in two English departments.
Regarding ELT and assessment, according to Palander (2013), some of the key goals of Quality Assurance are: (1) to modify the English language testing by following international standards for evaluation; (2) to raise the trust in the courses of ELT; (3) to establish social justice; and (4) to stay connected with the world. This has led to some bold attempts to reform English language assessment. For instance, teachers are now required to provide students with samples of tests prior to exams (MHESR/KRG, 2010b), which perhaps follows the recommendation of TESOL International Association (2010) that TESOL teachers must be sure of their students’ prior experience with the test questions and answer formats. Another aim of Quality Assurance is to collect students’ opinions about many kinds of teachers’ classroom performances (including teachers’ assessment practices), obtained through detailed questionnaires designed to assess teachers’ performances and gauge students’ perceptions (Quality Assurance Programme, 2013). This can be regarded as an institutional trust and respect for students, which complements AA’s implicit focus on students as active and responsible agents, capable of needs analysis, goal setting, and assessment of achievement (Finch, 2002). Presently, these accreditation and assessment processes are put into practice in all universities in the Kurdistan Region. This fulfils the recommendation of Carla Sanderson, the former Provost of Union University in Tennessee, about “a distinctive model of quality assurance in which transparency, democracy and accountability are highlighted” (MHESR/KRG, 2010a). This transformation process tries to develop and empower students through the learning processes or the institutional changes (Palander, 2013), whereas in the past the students’ views were not given any credence.

Within English departments, the empowerment of students and the assessment of teachers’ performances need time to improve English language assessment effectively. A period of time is necessary to estimate the extent to which the Kurdish tertiary TESOL teachers will be influenced by the Quality Assurance Programme, in terms of their methods of English language teaching, learning, and assessment.

In this chapter, many aspects of the context of this study - the English departments of the universities of the Kurdistan Region - were explained, and a brief description was given of the situation of the Kurdistan Region and a general description of its education and higher education. It also covered the recent reforms in education and higher education including the improvement of ELT in the school curriculum and higher education. This chapter also described the situation of the English departments of the universities of the Kurdistan Region.
regarding some aspects of their assessment system and the need for further reform and development, and the recent efforts in that respect. Then, it presented the positive influence of the Quality Assurance programme on reforming assessment in higher education and how it affected the English departments. This chapter also showed that, so far, the full potential of AA has not been achieved, and it stressed the necessity of doing action research in this context in order to exploit a greater potential of AAs. The next chapter will present the literature review on the emergence, evolution, conceptualisation, theoretical background, benefits, feasibility, and challenges of AA in the other contexts worldwide.
Chapter Three: Literature Review

3.1 Introduction

This literature review of AA begins by briefly defining assessment and testing, and comparing and contrasting them. This is presented as a starting point for explaining and analysing the emergence, evolution, conceptualisation, background theories, benefits, feasibility, and challenges of the use of AAs at the tertiary level, covering the myriad of types, methods, practices, activities, and strategies of AA. While it is preferable to investigate each practice alone, dealing with all the varieties of AA is more efficient in the context of this action research. This is because of the situation of the English departments of this context, where the implementation of all AA forms needed improvement through teacher development (see Sections 5.1 and 5.3).

The structure and content of this literature review reflect the topics and issues explored in the phases of the design of this action research (see Table 1). This review starts with the conceptualisation of alternative assessment (AA) and its background theories, and also presents empirical studies on TESOL teachers’ AAs, their knowledge of AA, and their beliefs about whether AAs are beneficial at the tertiary level; this covers various types, methods, practices, activities, and strategies of AA. This is compatible with the aims and research questions of the first phase of this action research, in which the participants’ assessment practices and use of AAs, their knowledge about AA, and their beliefs about whether AAs are beneficial were explored. The seminar contents of the second-phase intervention was also based on the review of the literature, which covers many aspects of AA, including the emergence, evolution, conceptualisation, theoretical background, benefits, feasibility, and challenges of AA (see Section 4.3.3 and Appendix 7). This literature review also elucidates the practicality issues of AAs, and presents numerous empirical studies about the feasibility and challenges of performing AAs at the tertiary level. This reflects the aims and research questions of the third phase of this action research, in which, after the intervention, the study investigated the participants’ beliefs about the feasibility and challenges of conducting AAs at the tertiary level in their professional context.
3.2 Assessment and testing

Testing is a universal facet of social life in that throughout history people were tested to prove their abilities (McNamara, 2000). Educationally speaking, assessment, measurement, testing, and evaluation are similar procedures for collecting information about a taught skill (Bachman & Palmer, 2010), having the same objective of making inferences about students’ proficiencies (Schmitt, 2002). However, assessment is significantly distinct because of the wide range of aspects it considers. Schmitt (2002) explains that assessment is used in a broader sense, which includes both formal measurement tools and other kinds of qualitative assessment modes. Thus, it includes a much wider range of practices (Green, 2014). Additionally, careful observations, various methods and measures, and integration of information have been identified as key features of assessment (Phye, 1996). Assessment also has several phases: planning, development, administration, analysis, feedback, and reflection (Coombe et al., 2007), and three basic purposes. Firstly, ‘assessment for learning’ yields information to modify learning activities, to target instruction and resources, and to give feedback to students. Secondly, ‘assessment as learning’ develops and supports metacognition for learners by focusing on their pivotal roles as critical connectors between assessment and learning through reflection and critical analysis of their learning. Thirdly, ‘assessment of learning’ is the summative outcome used to confirm students’ knowledge and abilities (Earl & Katz, 2006).

AA, which is informed by ethical considerations, critical language testing principles, and TESOL standards, and which is underpinned by many social, educational, and critical theories, embraces most of the aspects of assessment mentioned above. In spite of the positive aspects of AA, including the depth and breadth of assessment, teachers still mostly use traditional testing methods (Shohamy et al., 2008). This is also the situation of the English departments of the universities of the Kurdistan Region, in which around 90% of student assessment is completed through tests (see Section 5.1). This happens despite many concerns about language testing such as dissatisfaction with the kinds of information gathered, the washback effects of high-stakes/standardised tests, and the possibility of generating inaccurate results from teacher-made tests (Tsagari, 2004).
3.3 Emergence and evolution of alternative assessment

AA refers to a movement that has been particularly prevalent in US schools, which have endeavoured to move away from using standardised multiple-choice tests in favour of performance-based assessment modes (McNamara, 2001). The necessity and invention of AA may have started with the problem of some students who performed very well in class but could not get high grades from tests (Ajideh & Nourdad, 2012), because the multiple practices of AA can assess and enhance students’ performances and skills (see Sections 3.7.3 and 5.6). Also, attempts to overcome the limitations of the assessment of teachers as a single assessor, which can trigger potentially biased evaluations, increased interest in AA in the field of education (Matsuno, 2009).

As a result, since the 1990s the focus of assessment development has shifted from traditional assessment to AA (Bahous, 2008). It was a paradigm shift from psychometrics to a wider model of assessment, from a testing culture to an assessment culture (Derakhshan et al., 2011). At the heart of that change was the investigation of learner-centred AA methods (Ross, 2005), which led to the emergence of numerous AA types, methods, and practices that have become widespread (Derakhshan et al., 2011) and have been important elements of teaching and learning processes of foreign languages (Anusienė et al., 2007). Changes to assessment processes and the progress of AA demonstrate the considerable efforts made to reform assessment in general. In my understanding, one of the key factors that contributed to this reform is the increasing significance of the relationship between teaching, learning, and assessment mentioned in Hamayan (1995) in order to improve education. At present, introducing changes in assessment practices in order to reform education is probably most developed in the United States and the United Kingdom (Torrance, 1995).

3.4 Conceptualisation of alternative assessment

The following conceptualisation of AA covers some of its general definitions, elucidating the intricacy and overlapping nature of its conceptualisation, the intermixed nature of its types, methods, activities, practices, and strategies, and some of its notable characteristics.
3.4.1 General definitions of alternative assessment

There is no single definition of AA (Tannenbaum, 1996), but generally such practices can be regarded as alternatives to conventional evaluation methods (Chirimbu, 2013). AAs are ongoing processes that involve teachers and students in making judgements on the latter’s progress by using non-conventional strategies (Hancock, 1994), which could be performed outside the classroom or even the institution (Tsagari, 2004). More broadly, there are a mixture of facts and claims about AAs presented by their proponents: AAs are regarded as ongoing processes that are informal, value-added (Rudner & Schafer, 2002), untimed, long-term, process-oriented, and interactive performances, and they include contextualised communicative tasks, and open-ended and free-response formats (Anusienė et al., 2007). They are less formal, formative in function, and often have low stakes (Alderson & Banerjee, 2001). They provide individualised feedback, and generate washback, criterion-referenced scores, and creative answers. Furthermore, they have the ability to foster intrinsic motivation in students (Anusienė et al., 2007).

3.4.2 Intricacy and overlapping nature of conceptualisation of alternative assessment

The complexity of conceptualising AA is, firstly, related to the different terms used to refer to it (Tsagari, 2004), some of which are actually considered types of AA (Coombe et al., 2007, cited in Azarnoosh, 2013; Chirimbu, 2013; Grabin, 2009). Secondly, many forms of AA are referred to as methods (Bachman & Palmer, 2010; Tsagari, 2004; Grabin, 2009; Hamayan, 1995; Tannenbaum, 1996), most of which are categorised into constructed-response and personal-response types of AA (Brown & Hudson, 1998). Thirdly, many modes of AA are regarded as activities (Hamayan, 1995) or formative assessment strategies of AA (Regier, 2012) (see Appendix 4 for a description of numerous types, methods, activities, practices, and strategies of AA, along with the researchers who described them). This indicates that almost all student activities are AAs if they involve student assessment, regardless of what they are called. All AAs share some basic features, although they have subtle differences in their focus (Hamayan, 1995), and ‘alternative assessment’ as a term is more generic and can incorporate the features of other commonly used terms (Hamayan, 1995; Tsagari, 2004).

As a basis of exploring the forms of AA conducted by the participants, this current action research used a selection of them that are of relevance to the assessment of the four English language skills: oral presentation, direct question and answer, conversation, oral discussion,
immediate feedback, oral description of an event or object, retelling of a story after listening to a passage, oral interviewing, book report, writing assignment, editing a piece of writing, students’ summaries of what is read or listened to, oral reading, self-assessment, and peer-assessment (see Table 3). The selection of these AAs were, firstly, based on my familiarity with the context of this study as a partial insider researcher (see Section 4.4), and secondly, they were approved by six in-service Kurdish tertiary TESOL teachers in three English departments after I piloted the questionnaires with them (see Section 4.5.1.1).

3.4.3 Intermixed nature of the types, methods, practices, activities, and strategies of alternative assessment

The overlapping character of the forms of AA, at the implementation level, can be displayed best by portfolio, which covers several other types, methods, practices, activities, and strategies of AA. This may further complicate the conceptualisation of AA, as these practices are slightly different in their focus. Portfolios allow students to self-assess their knowledge and skills, and to participate actively in the evaluation and development of reflective thinking (Anusienė et al., 2007). The potential of portfolios to include numerous modes of AA is confirmed by Wolf’s definition of portfolios (1989, cited in Bruton, 1999) that they collect the biographies of works, which include a range of students’ works and reflections. Tannenbaum (1996) lists the sorts of materials that can be incorporated in a portfolio, which encompasses many outcomes of various AAs like audiotaped and videotaped recordings of readings or oral presentations, dialogue journal entries, book reports, writing assignments, reading log entries, artwork (e.g., pictures and drawings), graphs, charts, conference notes, interview notes, tests, and quizzes. All these can be considered products of the processes of AA that can be put in a portfolio (Tannenbaum, 1996).

So far, the intertwining and overlapping terminologies and forms of AA have been discussed. There are also a number of facets and features of AA that reveal the potential of such assessment and contribute to conceptualising it further.

3.4.4 Some notable characteristics of alternative assessment

The following sub-sections present some basic characteristics of AA such as its performance and authenticity, its multiple practices and holistic view of language, its integrity with instruction, and its duration as claimed by the proponents of AA.
3.4.4.1 Performance and authenticity

The two terms ‘performance’ and ‘authenticity’, which are used synonymously to describe AA (performance assessment and authentic assessment), illustrate the two core characteristics of these assessment modes. Performance means to produce a directly or indirectly observable response, whereas authentic means that the nature of the task and its context are relevant and represent aspects of real life (Chirimbu, 2013). Bachman and Palmer (1996:23, cited in Bachman, 2002) conceptualise authenticity as “the degree of correspondence of the characteristics of a given language test task to the features of a target language use”. Thus, AA can be characterised by proximity to actual language use, which is based on activities having authentic communicative functions, that is, real performance in authentic contexts (Hamayan, 1995). This is a major goal of AA, as Huerta-Macias makes clear: AAs “gather evidence about how students are approaching, processing, and completing real-life tasks in a particular domain” (Huerta-Macias, 1995:9, cited in Tannenbaum, 1996). As mentioned in Chirimbu (2013), authentic and real-life activities of a process-oriented nature (i.e., performance and authenticity) are the two features that make AAs superior.

3.4.4.2 Multiple practices of alternative assessment and a holistic view of language

AA has an array of measures that can be adapted for different contexts (Tannenbaum, 1996) and can be used for multiple referencing in order to obtain information about learners from many sources by various means (Hamayan, 1995). Hence, one main reason for the popularity of AA in the United States is that it is believed these assessment forms can give those learners who do poorly on selection-type tests the chance to demonstrate their achievements by alternative means (Grabin, 2009), that is, by classroom instructional activities that are close to real-life situations (Ataç, 2012; Tsagari, 2004). Concerning language assessment in general, including that of English language, the multiple practices of AA might create a holistic view of language as mentioned in Hamayan (1995). This is aligned with the belief that the interrelationships among the aspects of language (phonology, grammar, and vocabulary) and the four language skills (listening, speaking, reading, and writing) are considered components of a structurally integrated whole. Through AAs, language could be assessed not so much as a structure but rather as an instrument for communication and self-expression, yet they allow for structural analyses as well. Moreover, because AA takes learners’ social, academic, and physical contexts into consideration and assesses them in a
variety of natural settings, a more holistic assessment is possible (Hamayan, 1995). AA is referred to as a form of holistic assessment already in the work of Chirimbu (2013).

3.4.4.3 Alternative assessment integrity with instruction

AA results, based on learner progress over time, could be used to improve instruction (Tsagari, 2004). Hence, an influential characteristic of AA is its relation to and integrity with instruction, which exhibits many other features of such assessments. As mentioned in Ataç (2012), AAs can be integral components of teaching and learning, in which each student is approached as a unique person and compared to both his/her past performances and aims. In AA, there could be more informative views, the possibility of several perspectives, and more focus on strengths, progress, a culture-fair environment, improving and guiding learning, and collaborative learning, and such assessments also have intrinsic learning potential (Ataç, 2012). Thus, through AAs, instruction and assessment are intrinsically integrated (Ripley, 2012).

In addition, active learning can be integrated into AA and both can be performed simultaneously. Generally, active learning is any kind of activity that involves students in the learning processes (Bell & Kahrhoff, 2006; Faust & Paulson, 1998) by completing a wide variety of strategies either inside or outside the class (Eison, 2010). Active learning includes a continuum of activities from listening practices to writing activities and complex group exercises (Faust & Paulson, 1998), engaging students in some important activities such as critical or creative thinking, conversation with a partner, expressing ideas through writing, and reflecting on their learning processes (Eison, 2010). By these active learning processes, students are engaged actively in building their understanding of facts, ideas, and skills by completing instructors’ directed tasks and activities (Bell & Kahrhoff, 2006), by which students can apply course material to real life situations (Faust & Paulson, 1998).

3.4.4.4 Duration of performing alternative assessments

AAs cannot be carried out in a short period because they should meet the criteria of documenting individual learners’ growth over time rather than comparing the learners with each other (Tannenbaum, 1996). Portfolios are an obvious example as they are collections of students’ work accumulated over time and organised to assess their competencies in a given
standard or objective (Ripley, 2012). Therefore, AAs basically need a long time, maybe one term or even the whole academic year, to be performed successfully.

3.5 Theoretical background of alternative assessment

Generally speaking, “all theories are the product of some practical activity, so all practical activities are guided by some theory” (Carr & Kemmis, 2003:113). Educationally, because social contexts are constantly changing, education should be accordingly reconstructed (Burden & Troudi, 2007). In language education contexts, the progress of different assessment practices is usually influenced by various language learning methods (Dragemark Oscarson, 2009). Also, the development of theories of learning logically triggered the appearance of different assessment and testing methods (Chirimbu, 2013). Behaviourism, cognitivism, constructivism, and multiplicity of intelligence are among the prominent theories that have influenced assessment and resulted in changes in it (Grabin, 2009). Research paradigms like positivism, interpretivism, critical theories, and socio-cultural approaches have also affected assessment methods. Under the critical theory, critical pedagogy, critical applied linguistics (CALx), and critical language testing have emerged and influenced language assessment. Figure 1 displays the types of theories that underpinned, influenced, or interrelated with AA.
Figure 1: Classification of the types of theories underpinning, influencing, or interrelated with alternative assessment
3.5.1 Theories that emphasise the social aspect of learning and learners’ pivotal roles

The transition from behaviourist theories to constructivist and cognitivist theories of learning led to a new concept of learning in which the learner’s pivotal role is recognised (Grabin, 2009). The theories of the socio-cultural approach, communicative approach, constructivism, cognitivism, multiple intelligences, and interpretivism are different in some aspects, but regarding the consideration of learning and assessment as part of the social world, emphasising the social side of learning, and focusing on learners’ pivotal roles, they are, to some extent, similar. These theories have significantly reformed assessment, and may possibly have led to the evolution of AA.

3.5.1.1 Socio-cultural approach

Lev Vygotsky’s socio-cultural approach to learning has impacted assessment considerably. Vygotsky’s works focused on the importance of discourse with others and language mediation to reach a higher level of learning (Dragemark Oscarson, 2009). From the Vygotskian perspective, mediation can be defined as any human activity in which higher mental functions are mediated by another person (Shrestha & Coffin, 2012). AA is underpinned by a socio-cultural approach because within such an approach, learning and assessment are contextualised at any given time through the social interaction between learners and their environments as conceptual tools, physical tools, and people (Scarino, 2013). Here, the Vygotskian notion of the Zone of Proximal Development (ZPD) and mediation are especially central to dynamic/alternative assessment (Shrestha & Coffin, 2012). Dynamic assessment is one of the various terms used to refer to alternative assessment (Tsagari, 2004), and is also regarded as a type of AA that shares some of its basic features, although they have subtle differences in their focus (Hamayan, 1995). Vygotsky (1978:86, cited in Shrestha & Coffin, 2012) defines the ZPD as “the distance between the actual development level as determined by independent problem-solving, and the level of potential development as determined through problem-solving under adult guidance or in collaboration with more capable peers”. Thus, the ZPD means cognitive development takes place through language-mediated activities in interaction with people who have more advanced cognitive abilities such as teachers or smarter peers (Vygotsky, 1978, cited in Ishihara, 2009). The collaboration among peers creates a collective ZPD from which learners can benefit equally (Turuk, 2008). From Vygotsky’s viewpoint, it is essential to know what individual students
might be able to do in the future as well as what they can do currently (Shrestha & Coffin, 2012). The link between the socio-cultural approach and AA is through the ZPD. Dynamic/alternative assessment is a process rather than a product and its process is mostly within the ZPD. It is a development-oriented process that reveals students’ present abilities in order to help them overcome any performance problems, and to realise their potential. Here, mediation is as integral to dynamic/alternative assessment as it is to the ZPD. This integration is that the ZPD is about each learner’s potential development, and mediation provides opportunities for such a development (Shrestha & Coffin, 2012). As such, the socio-cultural approach emphasises the fundamentality of social interaction in learning through the ZPD and mediation, and it stresses the role of learners as effective partners in the interaction process for learning.

3.5.1.2 Constructivism

In constructivism, teachers should adjust their teaching methods to allow students to play a more active role. They have to cultivate a safe and encouraging environment where students feel comfortable to express their ideas and develop their concepts (Grabin, 2009). This is in accordance with AA, which embodies a collaborative approach to assessment and makes teachers and students interact during teaching and learning processes (Tsagari, 2004). In AA, teachers play the role of a partner and collaborator, whereas in traditional testing, teachers have a much narrower role of a test generator and examiner (Chirimbu, 2013). Similarly, constructivism places individual learners at the centre of the learning and assessment processes (Grabin, 2009), and social constructivism puts more responsibility of learning on students and emphasises the fundamentality of students’ active involvement in the learning process. This leads students’ metacognitive awareness and strategic abilities to become impressive features of learning more autonomously (Drągemark Oscaron, 2009).

Constructivism argues that humans cannot learn by absorbing knowledge passively. It regards learners as active participants in the acquisition of knowledge who are capable of developing their understandings, drawing on their prior background knowledge, and formulating mental concepts. Therefore, different students construct different meanings because they draw upon their individual learning experiences and backgrounds. Thus, according to constructivism, the core of learning is individual learners’ interpreting of meanings and their discovering and problem-solving (Grabin, 2009). Overall, the constructivist perspective of learning
presupposes an active role that the learner can play, and that human beings can inherently
construct meaning from experience (Dragemark Oscarson, 2009).

3.5.1.3 Multiple intelligences

The multiple intelligences theory, developed by Howard Gardner in 1983, presents a
theoretical foundation for recognising students’ different abilities and talents. It proposes
eight different aptitudes that deserve to be designated as intelligences. Each individual has
some of those types of intelligences to varying degrees, which include musical, bodily-
kinaesthetic, spatial, interpersonal, intrapersonal, natural, logical mathematic, and linguistic
intelligences (Grabin, 2009). Gardner believes that only the latter two forms of intelligence,
logical mathematics and linguistics, are tested and regularly appreciated in schools. He thinks
that all the eight skills are relevant socially and culturally in society, and as such should be
addressed in classrooms. He justifies his belief in multiple forms of intelligence on the
grounds that people think differently, and not all of them are the same; therefore, one
educational approach cannot possibly serve everybody. He proposes that instead of
questioning the degree of each young child’s intelligence, it is better to consider examining
the methods in which individual children can exhibit their intelligence. This has a substantial
influence on assessment because of the understanding about the necessity of offering students
various opportunities to show their knowledge and skills, and to allow a variety of
approaches to learning (Grabin, 2009). Those opportunities can be provided through the
various methods of AA and this is what makes this theory interrelated with AA.

Moreover, there are some other theories that focus on social aspects of learning and the role
of students, and which possibly underpin AA. Grabin (2009) stated that AA is consistent with
cognitive learning as an approach because it demands an active role from learners. I contend
that cognitive learning cannot possibly be assessed by traditional testing only, and this might
be the reason for the invention and development of AA. Similarly, the communicative
approach to language learning involves more implicit learning, and requires students to
communicate with one another in meaningful situations in various contexts (Dragemark
Oscarson, 2009). In my understanding, the focus on communication skills, authenticity, and
context demands an active role from students. This also matches with the major aim of
learning, which is being capable of applying skills and knowledge to perform meaningful
tasks in authentic contexts (Grabin, 2009). Therefore, there is a need for a new assessment
initiative in ESL/EFL that should acknowledge the influence of context on performance (Hancock, 1994). Here, the focus on authenticity and the context of communication in learning and assessment is quite relevant to the two primary features of AA, performance and authenticity. Also, interpretivism as a research paradigm focuses on the social aspect of language. This probably changes assessment because it is claimed that AA is informed by interpretivism, which considers language proficiency as part of the social world (Lynch & Shaw, 2005). In the same manner, AA prioritises language use that can be best understood as part of the social life (Lynch, 2001). Therefore, judgements about language ability cannot be considered as a measurement task since there is no perfect score waiting to be approximated (Lynch, 2001). Hence, there is an argument for an interpretivist hermeneutic approach that probably responds to many essential qualities of AA (Lynch & Shaw, 2005).

Overall, all the theories mentioned above require a more active role from learners, and allow them to employ their knowledge and skills through various means in a social realm. This has influenced AA and can possibly enhance the ethicality of AA chiefly by the social and democratic interactions between teachers and students. These issues might relate to the ethicality of teaching, learning, and assessment in general, which is one of the major aims of critical theories, in which a special focus is put on ethicality.

3.5.2 Critical theories that emphasise students’ assessment rights and ethicality of assessment

The main aim of critical theory is “critiquing and challenging the status quo” (Ary et al., 2006:469). Self-reflection and self-criticism are included among the essential features of a critical approach (Segerholm, 2010), in order to create change for people’s benefit and a better world (Watson & Watson, 2011). Lynch (2001) argues that AA can largely be aligned with the critical approach as a principal theoretical background (Lynch, 2001). I agree with Lynch because critical approaches have affected AA substantially by dealing with assessment directly through CALx, critical pedagogy, and critical language testing. Firstly, CALx has an interest in everyday aspects of applied linguistics like language learning and assessment, and it has a resistance to the normative (Pennycook, 2001, cited in Dragemark Oscarson, 2009). Shohamy (2001a; 2001b, cited in Dragemark Oscarson, 2009) explains that CALx covers the questions about the agendas behind assessment and the nature of knowledge upon which language assessment is based. Secondly, critical language pedagogy focuses on learning
outcomes that can be assessed in different ways, and even in different ways for different learners (Dragemark Oscarson, 2009). Therefore, multiple assessment forms can provide strong evidence to reveal learners’ knowledge (Shohamy, 2001). Here, AA can meet the necessities of critical language pedagogy of multiple assessment practices. Thirdly, critical language testing starts with the assumption that language testing is not neutral, but rather is a product of cultural, social, political, and educational agendas that can affect teachers and learners’ lives (Dai Quang, 2007). In order to eliminate the negative effects of tests, testers should regard some basic principles of testing that are fundamental in classroom assessments like reliability, validity, reference-points, record-keeping (Earl & Katz, 2006), fairness, and wash-back (The EALTA Executive Committee, 2006). Consideration of these principles is to make teachers’ inferences about students’ learning credible, fair, and free from bias (Earl & Katz, 2006). To this end, critical language testing attempts to challenge the psychometrics of traditional language testing and tries to support interpretive approaches to language assessment (Dai Quang, 2007). This signifies a paradigmatic shift, in which many new criteria for understanding validity are reconsidered, such as consequential, systemic, and interpretive validities (Quang, 2007). More important among these, I believe, is the consideration of the ethical issues that are associated with the effects of tests (consequential validity), because the validity framework must be integrated with ethics (Dragemark Oscarson, 2009).

Ethicality covers the issues of harm, consent, fairness, deception, privacy, confidentiality (Lynch, 1997), validity, absence of bias, access, administration, and social consequences (Bachman, 2005). Generally, a test is said to be fair if it is free from bias or favouritism (Karami, 2013). Fairness also means that assessment should be structured in a way to maximise ethical behaviour and to shift the power relations between the assessor and the assessed. From the critical AA perspective, ethics or fairness means that learners’ opinions are taken into account (Dragemark Oscarson, 2009). Through collaboration, AA can result in shared power and thereby empower learners rather than subjugate them (Dragemark Oscarson, 2009). Thus, AA offers a potentially more ethical, democratic, or equitable approach, but only if the assessment outcomes value individual diversity (May & Hornberger, 2008). Hence, AA might indeed help practitioners to address ethicality goals (Lynch, 1997).

Critical language testing supports AA such as portfolios for language assessment to develop a more democratic assessment method, whereby test-takers and local bodies are more active
(Dai Quang, 2007). Howe (1994, cited in Lynch, 1997) adds that the democratic model of assessment is the most moral approach that provides a viable alternative, by including the voices that have historically been banned from negotiating educational issues (Lynch, 1997). Therefore, AA probably provides what Shohamy called democratic testing, in which assessment ethicality and empowering students are stressed. This is what Freire advocates as an anti-authoritarian, dialogical, and interactive approach (Chandella & Troudi, 2013). To foster such a discourse of ethicality, Shohamy's critical language testing principles about questioning the informing values, agendas, goals, needs, purposes, and actual uses of tests should be taken into account in language assessment (Bachman, 2005). Committing to these, assessment practices must enable the construction of the self as a subject of assessment rather than its object (Dragemark Oscarson, 2009). This matches what critical theories, critical pedagogy, and critical research endeavour to achieve by enlightening, empowering, and emancipating people from oppression (Brown & Jones 2001).

I suggest that CALx, critical pedagogy, and critical language testing have influenced AA in many aspects for the purpose of providing more valid and ethical assessment. These theories also have a strong connection with the other aforementioned theories regarding the social teacher-student interaction, empowering students, and giving them their pivotal roles.

Figure 2 below displays the theories that focus on student-centred learning, the social side of learning, and teacher-student interaction, and those theories that emphasise student empowerment, students’ rights, and the ethicality of teaching, learning, and assessment. Those theories underpin, influence, or are interrelated with AA, and their two principal aspects constitute the core of the theoretical framework of this study, and are interrelated with the features and benefits of AAs. The one-direction arrows from the two groups of theories at the top of page, pointing down to their two main aspects, denote that they create and share these common features. The two-direction arrow between these two main aspects indicates that their features inform and enhance each other. The four two-direction arrows among the two boxes of the main aspects of theories, and the features and benefits of AA, display how they inform or interrelate with one another. The two-direction arrow between the features and benefits of AA indicates that they are interrelated with and enhance each other.
Figure 2: The theoretical framework: Interrelation of the features and benefits of alternative assessment with the two main aspects of its underpinning theories.
3.5.3 Theoretical framework of the study

The aforementioned teaching and learning theories, the socio-cultural approach, the critical theories, and the research paradigms have influenced the emergence, evolution, and/or development of AA. This influence has been through two main potential aspects of those theories and approaches: (1) the social interaction side of learning, whereby students play a pivotal role; and (2) the validity and ethicality of teaching, learning, and assessment.

The socio-cultural theory, constructivism, multiple intelligences theory, communicative approach, cognitivism, and interpretivism share the objective of focusing on the social aspect of learning and establishing democratic interactions between teachers and students in a social environment. This is perhaps related to the practice of student-centred learning, in which students are given their active roles, and it will contribute to empowering students and increasing the validity and ethicality of teaching, learning, and assessment. These are, in turn, the issues that are stressed by the critical theories: CALx, critical pedagogy, and critical language testing.

Critical theories emphasise the democratic teacher-student interactions, empowering students and giving them the right to be more active and involved in the relevant decision making about styles of teaching, learning, and assessment. Thus, critical theories and the above-mentioned theories, approaches, and paradigms have these powerful characteristics in common. These features have influenced assessment substantially, which has possibly led to the evolution, adoption, and progress of AAs. The common characteristics and goals of these underpinning theories of AA have become the primary features and benefits of AA under the influence of those theories, and these constitute the theoretical framework of this study (see Figure 2). These theories advocate AA to develop more democratic assessment methods, in which there would be a high degree of assessment reliability, validity, ethicality, and practicality, and empowering students. Traditional testing alone cannot assess a kind of learning based on the adoption of the above-mentioned theories; hence, the need for AA becomes more urgent and inevitable. This is firstly because AA focuses on performance and authenticity (Chirimbu, 2013; Bachman, 2002; Hamayan, 1995; Huerta-Macias, 1995, cited in Tannenbaum, 1996). Secondly, AA has a wide range of types, methods, practices, activities, and strategies that can be used for a more comprehensive and ethical assessment of

3.6 Benefits and feasibility of applying alternative assessments: Theoretical and empirical outcomes

This section covers the interrelation of benefits and feasibility of AAs and explains several significant benefits of AAs for teaching, learning, and assessment.

3.6.1 Interrelation of the benefits and feasibility of alternative assessments

Generally, testing practicality includes significant factors such as the availability of time, personnel, equipment, space, implementation funding (Malone, 2011), test development and maintenance costs, resources, ease of marking, suitable/trained graders, and administrative logistics (Coombe et al., 2007). These aspects of practicality are also relevant to assessment and AA. I contend that the benefits and feasibility of AA are closely related because the innumerable benefits potentially increase the feasibility of these assessments. For example, learners’ involvement in classroom assessments increases their motivation and achievements, and saves the teachers’ time (Arter, 2001). This demonstrates the interrelated and overlapping benefits and feasibility of AA for teachers and students. Another example is self-assessment, in which students are almost entirely the assessors of their work, during which teachers just monitor and supervise them, which is much easier for the teachers than doing the entire job themselves. This represents a significant aspect of the feasibility of self-assessment for teachers, and self-assessment is highly advantageous for students (Azarnoosh, 2013; Khonbi & Sadeghi, 2013; Naeini & Duvall, 2012).

3.6.2 General benefits of alternative assessments

A general benefit of AA for the tertiary level education is that it sets expectations that are appropriate within the cognitive, social, and academic developments of students and meets their individual needs (Hamayan, 1995). Finch (2002:5) explains that AA is suitable particularly for tertiary students because it encourages them to assume responsibility for their learning and the assessment of their learning by the cycle of intention, action, and reflection, which can be facilitated by AAs. AA gives an institutional trust to students and regards them
as active and socially responsible agents who are capable of needs analysis, goal setting, and assessment (Finch, 2002). This is also confirmed by Chirimbu (2013), who notes that AA methods could be used in training needs analysis, in identifying the best teaching materials, in monitoring progress, and in evaluating the extent to which the course objectives have been reached. Here, I suggest that the feasibility of AA lies in the involvement of students in deciding which methods of assessment are compatible with their needs. In addition to saving time, AA offers teachers the opportunity to follow a student’s evolution over time individually rather than comparing levels and knowledge of various students (Chirimbu, 2013).

3.6.3 The wide spectrum of the types, methods, practices, activities, and strategies of alternative assessment

Teachers believe that language proficiency can only be assessed by utilising multiple assessment forms such as portfolios and performance tasks of AA (Shohamy et al., 2008). Some teachers think that the use of multiple methods for assessing students’ work can ensure fairness by benefitting from a variety of instruments (Troudi et al., 2009). This can be made possible by the multiple tasks of AA, the training of judges for using clear criteria, and triangulating decision-making processes by depending on different sources of data, such as students, families, and teachers (Tsagari, 2004). This is necessary because virtually all language education contexts will likely require a variety of AAs to meet their actual inferential demands (Norris et al., 2002). Students should be allowed to choose suitable ways to display their knowledge, skills, and understanding by offering them multiple tasks as options to be adopted by them, in order to accommodate their individual needs (Grabin, 2009). Thus, the application of AA: (1) is more responsive to the students’ distinct learning styles; (2) values the diversity of their institutions (Hancock, 1994); (3) considers each student’s cultural background; (4) regards their level of knowledge (Chirimbu, 2013); and (5) considers the variations in their needs and interests (Grabin, 2009). I assert that what makes AA more ethical, feasible, and beneficial for learners is their right to choose from a wide range of AAs, which can be valid resources to obtain data on their knowledge and abilities.

3.6.4 Multi-faceted and comprehensive assessment of students’ knowledge and abilities

Another aspect of the advantages and feasibility of AA, which is linked to the use of its multiple forms, is the multi-faceted and almost comprehensive assessment of students’
knowledge and skills. By adopting AAs, a more comprehensive assessment than that achieved with traditional tests is possible. Worthen (1993, cited in Grabin, 2009) explains that traditional assessment merely samples “tiny snippets of student behaviour”, whereas AA looks at the bigger picture of students’ performances. AA tries to provide a general framework through which to understand learners’ achievements (Baker, 2010), by considering multiple sources of information when selecting assessment strategies and interpreting the results (Brown & Hudson, 1998). Shohamy, (1998, cited in Piggin, 2012) stresses the assessment of both academic and social English language skills to provide a clearer picture of students’ English proficiency (Stephenson et al., 2004). I suggest that AAs, which take place at various points in time and in different ways both inside and outside the classroom (Tsagari, 2004), are perhaps the best procedures to assess this academic and social English. Regarding this, many educators have come to recognise that AAs are necessary instruments for obtaining a dynamic picture of student’s academic and linguistic developments (Tannenbaum, 1996). Hence, TESOL International Association (2010) recommends the use of various performance-based assessment tools and techniques because, in AA methods, learners are allowed a long time to generate rather than choose a straightforward answer (Tsagari, 2004). Therefore, it is claimed that multiple-choice questions lack the ability to assess accurately whether students know the material or not (Law & Eckes, 2007).

Another reason that contributes to the comprehensiveness of AA is related to its process-oriented nature, which can assess both the product and process of learning (Ajideh & Nourdad, 2012) because the focus is equally on the two (Hamayan, 1995). In AA, it is not only the final product that can be assessed but also the process of reaching that product, which demonstrates the quality and complexity of learners’ processes of thinking (Chirimbu, 2013).

### 3.6.5 Integration of teaching, learning, and assessment

AA focuses on learners’ emergent or dynamic abilities, whereby assessment is integrated with teaching and learning (Behrooznia, 2014), because AA outcomes are based on the knowledge of learner progress (Tsagari, 2004). Therefore, teachers can use assessment results to reach conclusions not only about learners but also about instruction as a whole (Hamayan, 1995). By this, AA attempts to integrate assessment and learning activities (Grabin, 2009),
and thus assessment procedures should contribute towards learners’ ability to learn rather than undermine learning (Dragemark Oscarson, 2009). In this regard, because AAs are used successfully to evaluate students and the whole instruction process, and allow educators to monitor and modify teaching continually, many teachers call for a closer connection between instruction and assessment (Hamayan, 1995).

Moreover, AA has the advantage of being context-specific, with the assessment tools adapted to reflect best the purposes of learning (Chirimbu, 2013). Educationists claim that by utilising AAs, information can be collected on a number of factors that influence achievements, which could be identified in students’ linguistic, cultural, familial, or educational backgrounds (Tsagari, 2004). Furthermore, regarding the receiving of students’ views on learning, AA can alter and enrich students’ attitudes towards learning and increase their linguistic self-confidence (Chirimbu, 2013). I contend that this integrates assessment and learning much further, because considering students' cognition possibly leads to the democratisation of assessment, in which students can be more interested and involved in learning. I suggest that the advantage of the integration of instruction and assessment is in favour of students’ learning, which shows an obvious side of the benefits as well as the feasibility of AA.

### 3.6.6 Self-assessment and critical thinking

Self-assessment encourages learners to be aware, to obtain confidence, to acquire a view of evaluation, and to see errors as helpful (Alderson & Banerjee, 2001). Self-assessment involves metacognition that can be described as having two functions: self-assessment (assessing one’s cognition) and self-management (managing further cognitive development). The metacognitive function plays an effective role in the construction of new knowledge (Dragemark Oscarson, 2009). This is related to some AAs that are based on learners’ evaluations of their learning, which allow them to reflect on their linguistic development as well (Anusienė et al., 2007). Thereby, students are more confident as they can see their success and understand the meaning of being in charge of their learning, which is a principal advantage of AA (Grabin, 2009). Thus, self-assessment enhances students’ achievements (Rudner & Schafer, 2002), and is associated with critical thinking encouraged by AA in general (Chirimbu, 2013). AA instils in students lifelong skills like critical thinking that constitute a basis for future learning, and enable learners to evaluate what they learn (Hancock, 1994). Through AAs, the teaching and learning processes move to a superior,
richer, and more reflexive level, the level of critical thinking and reflection (Chirimbu, 2013). This great benefit of AA justifies the greater effort that such practices require, because participation in AAs can possibly assist students to become skilled judges of their own strengths and weaknesses, and to set realistic objectives for themselves that lead them to become self-directed and autonomous learners (Tsagari, 2004).

3.6.7 Focus on students’ strengths

AA focuses on students’ strengths more than their weaknesses (Tannenbaum, 1996; Chirimbu, 2013). Thus, one main benefit of AAs is that they measure students’ strengths, assessing what they can do (Delett et al., 2001; Grabin, 2009). Through AAs, students are assessed on what they can integrate and produce (their strengths) rather than what they can recall (Azarnoosh, 2013). Therefore, in the teaching of foreign languages at the tertiary level, using permanent AA methods during the whole academic year is recommended (Chirimbu, 2013). Stressing students’ strengths, I argue, makes AA more feasible and beneficial despite having several practicality problems.

3.6.8 Democratisation of assessment through alternative assessments

The democratisation of assessment constitutes a significant facet of the AA benefits and feasibility. AA can be used to develop a more democratic assessment, in which test-takers and local parties are more involved (Dai Quang, 2007). Piloting tests by eliciting students’ opinions (The EALTA Executive Committee, 2006) is a procedure for developing such a democratic assessment. To this end, Shohamy’s guidelines for making testing more democratic (Broad, 2001), Freire’s anti-authoritarian, dialogical, and interactive approach (Chandella & Troudi, 2013), and Foucault’s consideration of ethics as the practice of freedom can be achieved in portfolio assessment. Here, freedom lies in the students’ abilities to shape their portfolio processes and forms (Lynch & Shaw, 2005). Thus, because portfolio features the best work of a student (mostly selected by students), those assessment procedures become powerfully motivating for each student (Hamayan, 1995). The democratisation of assessment and empowerment of students would increase the feasibility and benefits of AA.
3.7 Empirical research on the benefits and feasibility of alternative assessments at the tertiary level

This section covers the effectiveness of teacher cognition and the empirical research on the tertiary TESOL teachers’ beliefs about the benefits and feasibility of AAs, including additional research findings on the same issues.

3.7.1 Effectiveness of teacher cognition: Empirical studies of teachers’ beliefs on alternative assessments

The reason why teachers’ beliefs are so important is that teachers are seen as one of the most significant personnel in educational systems, as they are on the front line of education, and are heavily involved in a variety of teaching and learning processes (Jia, 2004). Teachers are also the final implementers of educational principles and theories, yet were absent from educational studies for many years, and their voices were seldom heard. However, from the 1990s onwards, many researchers advocate examining teachers’ cognition on issues and including that in educational investigations (Jia, 2004). As evidenced by both the research questions and collected data, this study focuses principally on the teachers’ cognitions of AA in their context. Such cognitions included their knowledge about AA, their beliefs about whether AAs are beneficial as well as their feasibility and challenges.

Teacher cognition involves “the beliefs, knowledge, theories, assumptions and attitudes that teachers hold about all aspects of their work” (Borg, 2015:57), including their self-reflections, beliefs, and knowledge about their teaching, students, and content, as well as their awareness of problem-solving strategies for classroom teaching (Borg, 2006:36). Language teacher cognition is defined as the study of what teachers think, know, and believe (Borg & Al-Busaidi, 2012), in order to focus primarily on the unobservable dimensions of teaching such as teachers’ mental lives (Borg, 2009). The nature of language teacher cognition has a practical orientation, as it is personally defined, often tacit, systematic, and dynamic, as well as being highly context-sensitive. Therefore, “language teaching can be seen as a process which is defined by dynamic interactions among cognition, context and experience” (Borg, 2006: 275). One aspect of teacher cognition not well understood currently is the emotional or affective dimension. According to Borg, “This is an issue which is likely to become more
prominent in continuing research as we seek to understand how cognitive and affective factors interact in shaping what teachers do” (Borg, 2006: 271-273).

Teacher knowledge is part of teacher cognition (Borg, 2015). Grossman and Richert (1988:54, as cited in Ben-Peretz, 2011) define teacher knowledge as a body of professional knowledge encompassing knowledge about general pedagogical principles and skills and the subject to be taught. As such, there are several types of teacher knowledge, such as general pedagogical knowledge, knowledge of learners and their characteristics, and knowledge of educational contexts (Ben-Peretz, 2011). Teacher belief which is also part of teacher cognition (Borg, 2015) needs another kind of knowledge, practical knowledge, which is the “first-hand experience of students’ learning styles, interests, needs, strengths and difficulties and a repertoire of instructional techniques and management skills” (Borg, 2015:44). Practical knowledge is vital as it is generated by teachers from their own experiences and subsequent reflections (Borg, 2015: 57). All these types of teacher knowledge, which are possibly helpful in conceiving teachers’ beliefs, were considered in this study to explore TESOL teachers’ cognitions regarding the implementation of AAs.

Generally, a belief is a mental state in which a proposition is accepted as true by an individual (Borg, 2001). More broadly, Borg (2001, cited in Xu, 2012) defines belief as a proposition that is held by an individual either consciously or unconsciously, which is evaluative because it is accepted as true, imbued with emotive commitment, and serves as a guide to one’s thoughts and behaviours. Teachers’ beliefs, which are personal, unconscious, practical, and systematic (Wang, 2011), are convictions or opinions formed by experience or the intervention of ideas through the learning processes (Borg, 2006). Teachers’ beliefs are a form of personal knowledge, which consists of implicit assumptions about students, their learning, their classrooms, and the subject matter they are taught (Borg, 2006:36). Therefore, it can be argued that teachers’ knowledge and beliefs are interrelated and combine to form part of their cognition.

This study explores TESOL teachers’ assumptions, convictions, opinions, and propositions, which form their beliefs regarding whether AAs are beneficial, and the practicality issues of implementing them. The ultimate aim of this action research is to achieve a wider implementation of AA, and thus this is the major rationale for conducting this study, which covered the three types of teachers’ beliefs: beliefs about learners, about learning, and about
themselves (Xu, 2012). This is important because it is discovered that teachers are immensely affected by their beliefs, which are closely associated with their values, views of the world, and their understanding of their place within it (Xu, 2012). The British educational theorist Pajares (1992, as cited in Xu, 2012) notes that teachers’ beliefs have a greater impact than their knowledge on their methods of planning lessons, their decision-making processes, and their general classroom practice, and therefore teachers’ beliefs are both fundamental in determining their behaviours towards students (Xu, 2012) and influential in shaping their classroom practices (Wang, 2011).

Regarding assessment, systematic research has only recently begun to describe the nature and complexities of language teachers’ assessment practices (Cumming, 2009), and investigations of classroom-based assessment methods within the ESL/EFL contexts also have recently started to appear (Cheng et al., 2004). My exhaustive literature review has revealed that there are only a few studies about tertiary TESOL teachers’ beliefs on AA. Hence, little is known about ESL/EFL teachers’ assessment and evaluation procedures, especially at the tertiary level (Cheng et al., 2004). Currently, AA is used extensively at tertiary level education (Anusienė et al., 2007), so more research is needed on the factors behind tertiary ESL/EFL instructors’ selection of assessment methods, and how these affect learning (Cheng et al., 2008). To this end, this current study deals with some present AA issues from the cognitions of tertiary TESOL teachers in the Kurdistan Region.

3.7.2 Empirical research on the tertiary TESOL teachers’ beliefs about the benefits and feasibility of alternative assessments

Regarding teachers’ general views of AA, many believe that the improvement of foreign language education would be helped by pursuing assessment initiatives that are beyond traditional tests (Shohamy et al., 2008). This echoes students’ perceptions, which are fairly positive towards AA in general, as students tend to perceive that learning by doing and by simulating real practices is the best way of acquiring knowledge and developing abilities and competencies (Chirimbu, 2013). Concerning TESOL teachers’ cognition of the adoption of multiple assessment methods of AA, both Shohamy et al. (2008) and Troudi et al. (2009) concluded that teachers believe in the use of multiple assessment practices like portfolios, performance tasks, and essays as essential to ensuring valid and fair outcomes.
In reviewing the literature of AA, I have placed particular emphasis on tertiary TESOL teachers’ cognitions of AA in the other contexts because this study focused on Kurdish tertiary TESOL teachers’ cognition of AA. To this end, I discuss in detail the studies of Shohamy et al. (2008), Troudi et al. (2009), Estaji (2012), and Ataç (2012), all of whom have explored tertiary TESOL teachers’ cognition of AA, in addition to Cheng et al. (2004) who explored teachers’ assessment practices. Taken together, these indicate the limited number of studies on AA at the tertiary level from TESOL teachers’ cognitions.

Cheng et al. (2004) performed a comparative analysis of instructors’ assessment practices at some universities in Canada, Hong Kong, and China (Beijing). In the first year (2000-2001) of their three-year study, the researchers conducted a survey questionnaire with 267 ESL/EFL teachers. The participants explained the assessment purposes, methods, and procedures in the three contexts. The results reflected some differences related to class size. For example, less structured assessments (AAs) need more time to be marked, and therefore they may not be feasible in large classes that are common in Beijing. There were also differences in the assessment purposes, methods, and reporting procedures across the three contexts. However, the total time proportion spent on assessment in relation to teaching was essentially the same in the three settings. The researchers presented a likely explanation for this similarity, positing that the greater use of AA methods in Hong Kong and Canada needs more marking time, which might be the same as the greater amount of time required to process the objectively scored assessments (tests) in Beijing. Also, they discovered that the assessment methods and procedures used in these three contexts were diverse (Cheng et al., 2004).

The research of Shohamy et al. (2008) was situated within the Centre for Advanced Language Proficiency Education and Research (CALPER) at Pennsylvania State University, USA. The assessment project was one of the CALPER’s initiatives, which considered a variety of aspects of the assessment of Advanced Language Proficiency (ALP). Their study focused on the opinions of EFL/ESL teachers of schools and universities on the assessment of ALP and their assessment practices. The researchers collected data through a survey aimed at exploring the perceptions and practices of 467 participants. Approximately 33% were secondary school teachers while more than 50% were tertiary teachers. Regarding their perceptions of assessment, teacher participants believed that the compound ALP construct could be assessed only by applying multiple assessment procedures of AA. They also believed that assessment is always an ongoing process (Shohamy et al., 2008).
The study of Troudi et al. (2009) was about the assessment roles and philosophies of some EFL teachers in the United Arab Emirates and Kuwait. The participants were 21 tertiary teachers who represented nine colleges and universities. The researchers used an open-ended questionnaire to investigate the participants’ views on the nature of assessment, revealing that teachers are aware that using multiple AA methods are one of the ways to provide fair tests and ensuring fairness for students. Although the teacher participants recognised the role of traditional testing, they believed that students should benefit from various instruments. They stressed the impact of AA for learning purposes, and they called for incorporating these assessments into the curriculum. Troudi et al. (2009) further discovered that teachers’ assessment orientations were affected by internal and external criteria, their knowledge of their students, their educational and cultural contexts, their educational principles, and their pedagogical experience.

Estaji (2012) used a questionnaire and, for further qualitative data, interviewed 36 intermediate and advanced English teachers from two language institutions in Iran. Estaji also conducted a paired sample t-test to compare the group means about the teachers’ perceptions of AAs that reflected the principles of AA. The results showed that the participants’ perceptions of AA have a great influence on their practices. However, regarding the definition of AA, the majority of the teacher participants put little emphasis on the formative aspect of assessment. Moreover, there was a level of disagreement among the participants concerning the scope of assessment, as some teachers viewed assessment as encompassing all activities throughout courses, whereas a few viewed it as a small part of an exam. This implies that not all the instructors had studied or had been exposed to the concept of assessment. Nonetheless, they recognised various forms and purposes of AA. They also viewed assessment for educational purposes, not merely for scores or grades (Estaji, 2012).

Ataç (2012) distributed a questionnaire to 37 randomly selected instructors at Atilim University Preparatory School of English in Turkey. The aim was to explore the opinions and attitudes of English language instructors about the adoption of AA in language teaching classes. The results showed that the participants were strongly inclined towards adopting AA in their classes. Ataç (2012) also concluded that it would be pedagogically useful to attach more significance to AA in the curriculum and educational programmes for language teaching. This inclination of teachers probably demonstrated the feasibility and advantages of AA from their perceptions.
The findings of the studies of Cheng et al. (2004), Shohamy et al. (2008), Troudi et al. (2009), Estaji (2012), and Ataç (2012) concern tertiary TESOL teachers’ AA practices and their cognition of them. The relation of these findings to those of this current action research will be examined in Chapter 5.

3.7.3 Additional research findings of the benefits and feasibility of alternative assessments

Concerning students’ reflective and critical thinking through self-assessment, peer-assessment, and portfolios, Azarnoosh (2013), Sharifi and Hassaskhah (2011), Grami (2010), Bahous (2008), and Lucas (2007) demonstrated that students’ attitudes to portfolio and peer-assessment were positive, and these assessments helped them enhance their reflective and critical thinking. Nezakatgoo (2011), Lynch and Shaw (2005), and Anusienė et al. (2007) reached some findings on the effectiveness of students’ reflections on their learning process by looking at their strengths and weaknesses (Nezakatgoo, 2011). Determining strengths and weaknesses by self-assessment and peer-assessment can foster students’ learning as well (Azarnoosh, 2013), which was also confirmed by Khonbi and Sadeghi (2013), who found that EFL students performed much better on the post-test after applying self-assessment and peer-assessment. This is because these two assessments make students aware of their weaknesses, and encourage them to face those weaknesses (Lim, 2007). By focusing on weaknesses, many researchers such as Azarnoosh (2013), Nezakatgoo (2011), Grami (2010), Jaeverbakhsh (2010), Burkšaitienė and Teresevičienė (2008), Lucas (2007), and Roskams (1999) discovered that several types of AA can lead to autonomous learning, as students can take responsibility for their learning and address their linguistic problems and deficiencies. Regarding dynamic/alternative assessment and self-assessment, many researchers like Naeini and Duvall, (2012), Javaherbakhsh (2010), and Anusienė et al. (2007) concluded that self-assessment and dynamic assessment can help language teachers in gauging students’ levels of understanding and awareness more accurately. As a result, teachers become more familiar with their students’ capabilities, deficiencies, and needs, and they can develop methods for addressing the difficulties subsequently (Anusienė et al., 2007).

Regarding peer-assessment, Azarnoosh (2013), Grami (2010), Cheng and Warren (2005), and Roskams (1999) discovered that students did not find peer-assessment to be difficult or boring (Azarnoosh, 2013). Besides, students discussed and exchanged ideas more freely and
openly in the less formal atmosphere of peer-assessment (Grami, 2010), which encouraged learners to regard assessment as a shared responsibility (Azarnoosh, 2013). Finally, Roskams (1999) tells us that Ng (1993) cited more than 20 studies with findings that demonstrated that peer feedback is sometimes more informative than teacher feedback. This is probably because peer-assessment enhances self-awareness, that is, noticing the gap between one’s own and others’ perceptions (Azarnoosh, 2013). By this, peer feedback helped students gain new skills and improve existing ones as concluded by Grami (2010). That is why EFL students held positive attitudes towards peer-assessment, and the 20 student interviewees of Peng’s (2008) study wished to adopt peer-assessment in the future. Positive attitudes of students’ on peer-assessment indicates, in some way or another, the feasibility of this important method of AA.

Improvement of productive skills (writing and speaking) by AA was found by several researchers. Ghahremani and Azarizad (2013), Javaheerbakhsh (2010), Grami (2010), Wakabayashi (2008), Buršaitienė and Teresevičienė (2008), and Roskams (1999) discovered that using various AA types such as dynamic assessment, self-assessment, peer-assessment, and portfolio improved student writing significantly. Among these, self-assessment is a viable and useful learning task to develop oral performance (Chen, 2008). With regard to the receptive skills (listening and reading), Birjand and Azad (2014) and Ghahremani (2013) revealed that dynamic/alternative assessment mediation showed the difficulties of listening to texts (Birjand & Azad, 2014); hence, learners in dynamic groups outperformed the other groups in their listening abilities (Ghahremani, 2013). Also, the mediatory scaffolding of dynamic assessment could enhance students’ reading comprehension as concluded by Shabani (2012) and Ajideh and Nourdad (2012).


The aforementioned findings about the effectiveness, feasibility, and benefits of AAs will also be addressed in this current study with reference to the in-service Kurdish tertiary TESOL teachers’ beliefs. Overall, regarding the benefits and feasibility of AAs, many of the theoretically assumed advantages have been confirmed by empirical research. However, we should not forget that, despite the feasibility of AA and its numerous benefits, there are undoubtedly some associated challenges.
3.8 Challenges of the implementation of alternative assessments: Theoretical and empirical outcomes

This section explains some general challenges of implementing AAs. Generally, AAs need hard work and more effort than tests in most cases, as Law and Eckes (2007) mention that AAs might require considerable thought, space, logistics, and training. However, this does not mean they are impractical, although teachers and students might face some obstacles while performing them. Some challenges are related to some common AAs, for example Brown and Hudson (1998) state that portfolios, self-assessment, peer-assessment, and conferences are relatively difficult to produce and organise.

3.8.1 Challenges of the integration of learning and assessment

The integration of learning and assessment is an essential feature of AA, yet teachers are possibly neither able to connect AA in support of individual learning, nor to communicate information effectively about students’ developments to stakeholders (May & Hornberger, 2008). This view demonstrates the possible unfeasibility of the most important function of AA, which is the improvement of instruction. This also disputes Hamayan’s (1995) statement that interpreting and understanding information from AA is easy and helpful for all stakeholders.

3.8.2 Ethical challenges of alternative assessments

Ethicality of AA, which constitutes one of its fundamental aspects possibly under the influence of its foundation theories, is also questioned. The marking of AA, which is based on learners’ observed performance, raises equity concerns (Baker, 2010) because it involves subjective scoring (Brown & Hudson, 1998). Raters’ expectations may be affected by student differences such as race and ethnicity. Therefore, some researchers suggest that it is false to assume that AA can ensure equity automatically for diverse populations (Hamayan, 1995). Baker (2010) explains that, paradoxically, the movement of AA faces paralysing challenges of equity, and thus Hamp-Lyons (1997, cited in Tsagari, 2004) recognises the need to conduct further studies on the impact of AA. It cannot be assumed that just because those assessments are based in humanistic concerns, they solely yield good outcomes with no harm; hence, safeguards should be put in place, since the potential bias needs to be assessed (Baker, 2010).
3.8.3 Challenges of class size and limited time

The challenge of class size is that, in large classes, each student’s initiatives are difficult to arrange and assess (Bruton, 1999). High numbers of students in classes, together with the added time needed for AA activities, can create an enormous obstacle to the implementation of AAs (Tsagari, 2004). Similarly, the documentation of AAs is time-consuming to administer and analyse thoughtfully by teachers in order to give accurate feedback to their learners (Tsagari, 2004); hence, some teachers are reluctant and hesitant to adopt AAs properly. In contrast, Stiggins (2002, cited in Grabin, 2009) rejects those ideas by mentioning that AA enables teachers to save time while they develop and utilise those assessments more efficiently, and thus it benefits language teachers.

There are some virtual solutions to the challenges of large classes and limited time by computers, which can virtually replace humans in raising students’ consciousness by directing their attention to the key sections of the text in order to assist them to understand better (Shabani, 2012). This can be accomplished by administering computer-based dynamic assessment (Poehner & Lantolf, 2013), in which dynamic tutor mediation in the form of wikis and email exchanges can be provided, which is an important method to support the development of students’ academic writing (Shrestha & Coffin, 2012). This solution is also confirmed by Teo (2012), who suggested that for EFL college freshmen, the mediation in computer-based dynamic assessment allowed for information internalisation, which helped promote learners’ potential development.

3.9 Empirical research on the challenges of implementing alternative assessments at the tertiary level

Empirically speaking, there are fewer studies about the challenges of AA than those about their benefits and plausibility. This section covers the studies on some challenges of AAs.

3.9.1 Studies on teachers’ cognition of the challenges of alternative assessments

Relatively fewer studies focus on the challenges of AAs from teachers’ perspectives than those from students’ perceptions and practices. Yu (2013) conducted semi-structured interviews to explore the beliefs of 26 Chinese tertiary EFL teachers from several universities in China about the role of peer feedback in L2 writing classes. The two most significant
findings were that, firstly, although teachers practised what they believed, their practices sometimes did not match their beliefs when it came to peer-assessment. Secondly, the study discovered that some EFL teachers might not be aware of the value and potential of peer-assessment for their students’ learning, which prevents students from engaging in and benefiting from peer interactions in L2 writing. Therefore, teachers’ knowledge of and qualification in AA might be serious obstacles to the implementation of those assessments.

Abbas (2012) also investigated the difficulties of implementing AAs in teaching from Iraqi teachers’ perspectives by using a rating scales questionnaire with 30 participants as an instrument of measurement. She conducted her study in the two English departments of the College of Education (AL-Assmaie) and the College of Basic Education at Diyala University, which is geographically near to the Kurdistan Region. She concluded that the challenges of using AAs were: (1) the lack of technologies for engaging student in AAs; (2) the lack of a supportive classroom environment to enable students to feel comfortable with each other; (3) the limited time to implement AAs; (4) the unfamiliarity of most Iraqi students to take responsibility for assessment, as they were used to being taught in a traditional teacher-centred approach; and (5) the general difficulty of incorporating AAs into classroom activities (Abbas, 2012). Most of these challenges are closely related to those of this current study due to their contextual similarities.

Wach (2012) investigated 87 EFL teachers’ beliefs about the place of assessment in their teaching practice, the challenges it poses to them, and their roles as assessors. A general conclusion from the study was that the role of the assessor is probably among the most challenging roles that teachers perform in the contemporary foreign language classroom. The findings indicated that classroom assessment (AA) posed a significant challenge to the teacher participants, despite the fact that they appreciated it. That was chiefly due to the dual roles of instructor and assessor, which requires considerable levels of awareness, organising skills, expertise, and human sensitivity (Wach, 2012).

Yu (2013), Abbas (2012), and Wach (2012) indicated that the implementation of AAs posed a major challenge to the teachers in spite of their positive attitudes towards AA. Thus, even if language assessment practitioners wish to adhere to a code, they might face working conditions that make it difficult to fulfil all the items of that code (Jia, 2009). This demonstrates that teachers’ beliefs, perceptions, and knowledge are not enough to conduct
AAs, since there would be many contextual factors that hinder performing those assessment modes.

3.9.2 Challenges of self-assessment and peer-assessment

Lim (2007) explained that the ELLs were not confident in assessing their own or their peers’ grammatical mistakes and faulty pronunciation, particularly those mistakes made by more proficient students than the assessors. Suzuki (2009) found that L2 writing teachers might need to give instructions to their students in order to make them more confident in their self-assessment and L2 ability. The students were either not confident or very strict as Khonbi and Sadeghi (2013) concluded, since in peer-assessment the students were found to be more willing to assess their peers strictly, thereby underestimating them. A counter finding, also presented by the same two researchers, revealed that in self-assessment students tend to be lax. Furthermore, there are cultural and educational challenges associated with peer-assessment. Some teachers believed that in Asian cultures, peer feedback is constrained by a number of issues, such as the fear of mistakes, politeness norms, and the belief that peer feedback lacks credibility (Roskams, 1999). Concerning credibility, Grami (2010) discovered that students approved teachers’ feedback, but they were apprehensive about peer feedback, because it originated from fellow students whose language proficiencies were lower than that of the teacher. Maybe because of this, Roskams (1999) found that all students were unsure about the fairness of peer feedback. Consequently, students had a less positive attitude towards assessing their peers (Cheng & Warren, 2005). Also in the affective domain, Lim (2007) concluded that the repetition of both self-assessment and peer-assessment could make ELLs bored, possibly because students objected to the amount of work they were supposed to accomplish, as Bahous (2008) revealed. In sum, students’ beliefs and feelings about the credibility of peer feedback, their apprehension about this assessment, and the workload of self-assessment and peer-assessment can be pressing challenges to AA that need to be addressed urgently.

This chapter presented the emergence, evolution, conceptualisation, underpinning theories, benefits, and practicality issues of AA in various dimensions and in various contexts all over the world. It started with a brief comparison between assessment and testing, then how AA emerged and evolved. This chapter conceptualised AA through presenting some of its general definitions, showing the intricacy and overlapping nature of its conceptualisation and the
intermixed nature of its types, methods, practices, activities, and strategies. It further conceptualised AA by explaining some of its notable characteristics, such as its performance and authenticity, its possession of multiple practices, its holistic view of language, its integrity with teaching and learning, and its duration. This chapter also explicated some important underpinning theories of AA, such as the socio-cultural approach, constructivism, multiple intelligences, and critical theories, explaining how their two main aspects constitute the theoretical framework of this action research. Then it reviewed many studies about the benefits, feasibility, and challenges of applying AA at the tertiary TESOL level from theoretical and empirical literature, especially from teachers’ cognitions. Many of the benefits of AA and its practicality issues were also revealed by this current study. The next chapter will present the research paradigm, methodology, methods, and procedures employed to explore these AA issues from the cognitions of Kurdish tertiary TESOL teachers in the English departments.
Chapter Four: Design of the Study

4.1 The informing research paradigm: Critical theory

Generally, the aim of critical paradigm is “critiquing and challenging the status quo” (Ary et al., 2006:469) by attempting to enlighten, empower, and emancipate people from oppression (Brown & Jones 2001). This critical stance is value-laden and leads the inquiry in certain directions to evaluations of any question, including their questions, since self-reflection and self-criticism are essential features of a critical approach (Segerholm, 2010). Overall, this is to create change for the benefit of participants and a better world (Watson & Watson, 2011).

Educationally, the goal of critical educational researchers is not only to understand or describe behaviours in societies, but also to change such performances (Mack, 2010) because critical studies, as transformative research, provide alternative understandings, practices, and social relations (Simon et al., 2012). Critical studies attempt to inform educational appraisal for the improvement of educational actions (Mack, 2010) by identifying quality teaching, and refining and redefining the notions of curricula, pedagogy, literacy, and achievement (Morrell, 2009). To this end, researchers need to be guided by a critical agenda when conducting critical research in language education to raise awareness of the lack of professional development or problematise assessment practices in some contexts, because critical research should be systematic, theoretically justifiable, and credible (Troudi, 2015).

Due to the main features of the critical approach in the creation of positive change, I have followed a critical research paradigm and its action research methodology in conducting this study. These are best suited to achieve the research aims and answer the research questions of this study, which are intended to introduce alternative assessment (AA) to some teacher participants in order to improve their assessment practices in the English departments of the universities of the Kurdistan Region.

4.2 Methodology: Action research

This section provides the definitions and aims of action research and the rationale for choosing action research for this study. It also provides a short explanation of the two action research models adopted for this study.
4.2.1 Definitions and aims of action research

Action research is defined variously by different researchers, but Carr and Kemmis (2003) and Hopkins (2008) display its core features. Carr and Kemmis (2003:162) define action research as a form of self-reflective enquiry, undertaken by participants in a social situation to improve the justice and rationality of their practices, their understanding of them, and the situation in which these practices are implemented. Hopkins (2008:47) defines it as that which “combines a substantive act with a research procedure; it is action disciplined by inquiry, a personal attempt at understanding while engaged in a process of improvement and reform”. Action research emphasises action, change, and empowerment (Troudi, 2015), whereby change processes are the ways in which teachers implement innovation (Griffee, 2012). In professional terms, “action research is the systematic reflective study of one’s actions and the effects of these actions in a workplace context”, involving deep inquiry into an individual’s professional actions (Riel, 2010b:1). It links practices, understandings, and situations to each other, and thus, “it involves discovering correspondences and non-correspondences between understandings and practices, between practices and situations, and between understandings and situations” (Carr & Kemmis, 2003:180). Educationally, action research has been utilised in the curriculum development and improvement programmes of schools, professional development, planning of systems, and policy development (Carr & Kemmis, 2003).

Concerning the improvements and reforms achieved through action research, teachers are challenged continuously by a variety of problems they are required to solve in order to improve their practice. To achieve this, teachers need to engage in action research and become more flexible in their thinking and more receptive to new ideas. Here, identifying a problem, finding solutions, and testing them are the three elements of problem-solving (Johnson, 2008). Therefore, action research, as a means of exploring change, takes place in cycles, and each cycle is an experiment (Riel, 2010a). Kurt Lewin, who coined the phrase ‘action research’, described it “in terms of planning, fact-finding and execution”; hence, the self-reflective cycles of planning, action, observation, and reflection are central to action research (Carr & Kemmis, 2003:162). By this, action research could bridge the gap between theory and practice, as previously many teachers believed that education research is practically irrelevant (Johnson, 2008).
4.2.2 Rationale for choosing action research methodology for this study

Although action research has a long history, its use in ESL studies is relatively recent. It is something new in this field (Crookes, 1993), and it has been influential in TEFL and TESL studies even more recently (Troudi, 2015). This current study has adopted action research in order to improve the implementation of AAs in the English departments of the universities of the Kurdistan Region. To this end, it attempts to follow the central steps of action research: reaching conclusions about research questions, exploring practical significance of research findings, taking action based on the conclusions, and sharing the findings (Dai Quang & Hang, 2008). More specifically, it is based on some elements of Bachman’s (2001) action research spiral and Riel’s (2007) model of the progressive problem-solving action research (see Figure 3 and 4 below).

Figure 3: Bachman’s action research spiral (Mertler, 2013:18)
Bachman’s (2001) model “suggests that participants gather information, plan actions, observe and evaluate those actions, and then reflect and plan for a new cycle of the spiral”. Riel’s (2007) “action research model takes the participant through four steps in each cycle: planning, taking action, collecting evidence, and reflecting” (Mertler, 2013:16). It is also based on the general design of action research that originates from an individual but the change processes are always social, extending to a widening group of stakeholders (Riel, 2010a). Therefore, action research calls for teachers’ direct participation, whereby an impact will be created on their cognitions and skills, contributing to a pedagogical renewal (Pellerin, 2011). For creating the required impact on the participants’ cognition of AA, I presented two seminars to introduce AA to them, which might lead to a better performance of AAs. This can be further enhanced by the function of action research, which helps shape and refine assessments by promoting dialogues among assessment stakeholders, including dialogues between assessors and researchers (Singh, 2006).
three rounds of interviews of this current study. So, action research can fulfil the aims of this study and answer its questions, because its goals are based on creating improvements in AA cognitions and practices of the in-service Kurdish tertiary TESOL teachers.

4.3 Data collection methods

This section explains the stages of data collection during which the first-phase used survey questionnaires and interviews, the second phase was intervention via the presentation of two seminars, and the third-phase employed interviews and focus group interviews. These are based on the research questions and aims of this study, the latter of which are: (1) to explore the amount of AA in the assessment practices of Kurdish tertiary TESOL teachers; (2) to investigate the amount of their assessment literacy in AA; (3) to plan a small professional support for the teacher participants as an action research intervention, with the goal of changing cognitions of AA in order to raise awareness of the demands of AAs, and (4) to encourage teachers to reflect on the practicality issues of implementing AAs within their professional context. Fulfilling the aims of this study will eventually improve the implementation of AAs, which can be achieved through obtaining answers for the following five research questions: (1) What are the current assessment practices and AAs used by the participants? (2) Do they have knowledge about AA? (3) Do they believe that AAs are beneficial? (4) Have their cognitions of AA been changed by the intervention? and (5) What do they believe about the practicality issues of AAs? a. Do they believe that implementing AAs is feasible? b. Do they believe that performing AAs would face challenges?

4.3.1 Stages of data collection

Qualitative techniques such as questionnaires, interviews, and focus group interviews were employed in this study to collect the data needed to answer the research questions and achieve the aims. This action research had a qualitative element by which three rounds of interviews and an open-ended questionnaire question were used. Interviewing is used increasingly in qualitative research, along with open-ended questionnaire questions and observation as significant elements of data collection triangulation (Barnard & Burns, 2012). To this end, this action research follows mixed methods by using 90 questionnaires, 24 interviews, and two focus group interviews. The design of this action research and the mechanism and phases of data collection and analysis are illustrated in Table 1 below, which
represents the phases of the cycle of this action research, based on some elements of the four steps of the cycles of Bachman’s (2001) action research spiral and Riel’s (2007) model of action research (Mertler, 2013).
## Phases of the design of this action research

<table>
<thead>
<tr>
<th>Phases of Bachman’s (2001) and Riel’s (2007) models of action research</th>
<th>Phases of the design of this action research</th>
</tr>
</thead>
<tbody>
<tr>
<td>The first phase: pre-intervention via action research</td>
<td>The first phase: pre-intervention via action research</td>
</tr>
<tr>
<td>1. Bachman’s (2001) step of gathering information and planning actions</td>
<td>1. To obtain data on the current assessment practices of the in-service Kurdish tertiary TESOL teachers, and to explore to what extent their assessment practices include AAs.</td>
</tr>
<tr>
<td>2. Riel’s (2007) step of studying and planning the actions</td>
<td>2. To get data on the participants’ knowledge of AA to determine whether they need to be introduced to AA and promote their assessment literacy in it.</td>
</tr>
<tr>
<td>The second phase: the intervention via action research</td>
<td>The second phase: the intervention via action research</td>
</tr>
<tr>
<td>1. Bachman’s (2001) step of acting on the problem</td>
<td>1. After the analysis of the first-phase collected data, the seminars were designed to introduce AA to the attendants to increase their knowledge of AA, to change their attitudes about it, and to encourage them to do more about AAs.</td>
</tr>
<tr>
<td>2. Riel’s (2007) step of taking actions</td>
<td>2. Some other teachers of the two English departments also attended the seminar but not as participants.</td>
</tr>
<tr>
<td>The third phase: post-intervention via action research</td>
<td>The third phase: post-intervention via action research</td>
</tr>
<tr>
<td>1. Bachman’s (2001) steps of observation, evaluation, and reflection on the actions taken</td>
<td>1. To explore to what extent the participants’ cognitions of AAs had changed. That was to ensure the catalytic validity for this action research by the required changes and improvements achieved in the participants.</td>
</tr>
<tr>
<td>2. Riel’s (2007) steps of collecting and analysing the evidence of the actions and reflecting on them</td>
<td>2. To attain data on the participants’ beliefs about the feasibility and challenges of implementing AAs. That was after their knowledge was increased on the nature of AAs and their requirements, so that they can have a clearer viewpoint about the practicality issues of AAs.</td>
</tr>
</tbody>
</table>

### Table 1: Phases of the design of this action research
4.3.2 First-phase instruments: Survey questionnaires and interviews

As a partial insider researcher, I reflected on the possible areas for development surrounding the teacher participants’ cognitions and implementation of AAs before conceiving the three research questions for the first phase of this action research, during which the development potential should be diagnosed. Therefore, in the first phase, data were collected to discover the nature of the necessity for development by exploring the participants’ assessment practices, their knowledge of AA, and their beliefs about whether AAs are beneficial, and then identifying what could be improved in relation to their knowledge of AA. That was used as a basis for the action research intervention in the second phase.

For the first-phase data collection, I distributed questionnaires as a survey in 18 English departments in 10 public universities across the Kurdistan Region. The recruitment of the survey participants was based on the willingness of the teaching staff of those 18 English departments to volunteer to participate in my study. However, I made sure that, in each department, at least half of the teaching staff who had been teaching ELT subjects filled in the questionnaires. I used questionnaires primarily because they are tools of data collection for survey research designs (Griffie, 2012), by which a survey could provide three types of data: (1) factual information (factual questions to discover the characteristics of teachers and learners); (2) behavioural information (behavioural questions to explore what teachers or students do in their language teaching and learning aspects); and (3) attitudinal information (attitudinal questions to find out the teachers’ or learners’ opinions, beliefs, or interests) (Dornyei, 2003, cited in McKay, 2006).

The survey questionnaires yielded all these kinds of information from the participants. The questionnaire comprised: (1) three items on the qualification and experience of the participants; (2) eight items on the types of test questions participants use in their test papers (in a five-point Likert-scale of always, often, sometimes, rarely, and never); (3) 15 items on AAs that the participants implement (the same five-point Likert scale); and (4) 12 items on the participants’ cognition of AAs (in a five-point Likert-scale rating their situations and beliefs as: strongly agree, agree, unsure, disagree, and strongly disagree). Among these last 12 items, the third one was an open-ended question so that the participants could provide a definition and their understanding of AAs. This also follows Brown and Rodgers’ (2002) explanation that surveys are typically used to collect and describe the characteristics,
attitudes, views, and opinions of students, teachers, and administrators via interviews or questionnaires, or both. Here, the questionnaires that include some broad open-ended questions make respondents capable of providing answers in line with their own perceptions (Barnard & Burns, 2012). Overall, in the first phase, 90 questionnaires were used as instruments for the survey, which were supported by 12 interviews; thus, the first phase follows a language education survey. Brown (2001:2, cited in McKay, 2006) defines this kind of survey as gathering “data on the characteristics and views of informants about the nature of language or language learning through the use of oral interviews or written questionnaires”.

The survey questionnaire was an informative tool for this study, since it could investigate the current assessment practices and AAs of the in-service Kurdish tertiary TESOL teachers and their cognition of AAs in all the 18 English departments of the public universities of the Kurdistan Region (excluding the evening classes, as they have the same staff as the morning classes). The survey participants may represent the population of all those departments, because the distribution of the questionnaires covered around half of the staff in all the English departments. Then, 12 randomly selected teachers in the two selected English departments were interviewed to support the survey data. The first-phase interviews were semi-structured, comprising 17 questions that allowed for some follow-up questions if any new issues were to arise. The questions focused mostly on the participants’ assessment practices and AAs, their knowledge of AA, and their beliefs about whether AAs are beneficial.

The selection of the two English departments at a public university for the interviews and intervention was based on a practical reason, because action research requires intervention and it was not practical for a single researcher to do the intervention in 18 English departments. Therefore, I selected the two departments in which I was a member of the teaching staff for several years and they are at the same university in the same city. For the randomisation of the sample of 12 interview participants (six from each of the two English departments), I depended on the willingness of the participants to be volunteers in the series of the research instruments of my action research in addition to the intervention: the first-phase questionnaires, first-phase interviews, second-phase intervention seminars, third-phase interviews, and third-phase focus group interviews. I made sure of the variety of their
experiences and their degrees: six interviewees have PhDs and six have MAs with differing number of years of the experience of teaching ELT subjects (see Section 4.4).

4.3.3 Second-phase intervention: Presentation of two seminars

More than half of the questionnaire and interview content of the first phase were used to ascertain the amount of participants’ knowledge of AA and their beliefs about whether AAs are beneficial. That was to identify any gaps therein as a preparation for the second-phase action research intervention, during which I presented two seminars to the participants in the two selected English departments. The intention was to increase their assessment literacy in AA, and thus the seminars covered the two areas of assessment literacy: teachers’ assessment knowledge and teachers’ perceptions of their assessment knowledge (Coombe et al., 2012). This can function as a starting point for the participants’ professional development in the area of English language assessment, especially AAs, because professional development activities for teachers range from courses, seminars, and workshops to formal qualification programmes, which are highly influential in developing teachers’ skills, knowledge, and expertise (OECD, 2009). Through seminars we can potentially obtain the three advantages of action research: (1) encouraging teachers to reflect on their practices for the purpose of changing them; (2) empowering teachers by releasing them from the ideas handed down by their past experiences; and (3) providing teachers with access to new ideas from academic researchers and trainers that they can implement in their classrooms (Griffee, 2012).

The seminars covered the subjects and issues surrounding the emergence, evolution, conceptualisation, theoretical background, benefits, feasibility, and challenges of AA, which are presented in the literature review of this study. Because the two seminars were only around half an hour each in duration, I was very selective in including those parts of the literature review that provide the basic information necessary to introduce AA, beginning with some information about assessment and its status as a profession. In order to explain the professionalisation of language assessment, along with the demand for professional development for TESOL teachers in assessment, I depended on studies by the TESOL International Association (2012), Coombe et al. (2012), Bailey and Wolf (2012), Troudi et al. (2009), Taylor (2009), Tsagari (2004), and Brindley (2001). For a general comparison between AA and testing, I relied on Green (2013), Schmitt (2002), McNamara (2000), and Worthen (1993, cited in Grabin, 2009). To facilitate an explanation of AA evolution, I used
the work of Ajideh and Nourdad (2012), Matsuno (2009), and McNamara (2001), and to define AA, I used Tannenbaum (1996) and Hancock (1994). In order to display the characteristics of AA, I depended on Ripley (2012), Derakhshan et al. (2011), and Lynch and Shaw (2005). To present the different terms used for AA, along with an overview of how some of AAs are regarded as methods, practices, or activities of AA by different researchers, and how they are intermixed, I used Chirimbu (2013), Azarnoosh (2013), Coombe et al. (2007), Anusienè et al. (2007), Tsagari (2004), Tannenbaum (1996), Hamayan (1995), and Wolf (1989, cited in Bruton, 1999). In order to explain the general challenges of AA, I used Baker (2010), Law and Eckes (2007), Tsagari (2004), and Brown and Hudson (1998). Finally, and specifically in relation to the Kurdistan Region, I used some of the studies presented in Chapter Two of this thesis to show the KRG’s attempts to reform assessment, including the work of Palander (2013), the Quality Assurance Programme (2013), MHESR/KRG (2010a), and MHESR/KRG (2010b) (see Appendix 7 for the slides).

4.3.4 Third-phase instruments: Interviews and focus group interviews

This action research focuses primarily on interviews, of which there are three rounds in the first and third phases. Kvale (2007:9) shows the significance of the qualitative interview as “a key venue for exploring the ways in which subjects experience and understand their world” by “a unique access to the lived world of the subjects, who in their own words describe their activities, experiences and opinions”. This is necessary because investigating the social world requires another understanding that “allows insight into the hidden meaning behind human action” (Wilson, 2009:261), the “subjective world of human experience” (Cohen et al., 2007:21), and “the complex world of lived experience from the point of view of those who live it” (Mertens, 2010:16). Creswell (2007) confirms that data from those who experience the phenomenon often can be collected by interviewing them. Moreover, interpretivists pursue rich data from interviews to understand what is happening from the perspective of the participants (Radnor, 2002). Indeed, interviews provide access to the essence of individuals’ experience (Ary et al., 2006), probably because interviews are inter-subjective, that is, neither subjective nor objective, by which knowledge is generated among humans; hence, “human interaction for knowledge production” (Cohen et al., 2007:349). Interviewing is one of the most frequently employed techniques to obtain qualitative data about participants’ feelings, opinions, and beliefs regarding situational issues (Ary et al., 2006). Thus, interviews can gather many types of data, including personal perceptions, experiences, opinions,
preferences, and ideas (Wallace, 1998). As a type of interviewing, face-to-face interviews have long been the dominant method in the field of qualitative research (Opdenakker, 2006), providing researchers with a scope for greater exploration than is normally possible (Brace, 2004). So, the three rounds of interviews of this current study would provide abundant data for finding answers to its research questions.

Regarding the third-phase interviews, there were two rounds. Firstly, individual face-to-face, semi-structured interviews were conducted with 12 interviewees, who were previously interviewed in the first phase. The third-phase interview protocol had 14 questions that allowed for some follow-up questions to be added if new issues were to arise. A semi-structured interview is a technique by which the interviewer has several questions in mind to ask the interviewees, but not following a specific and predetermined order (Savenye & Robinson, 1996). Thus, it “combines a certain degree of control with a certain amount of freedom to develop the interview” (Wallace, 1998:147). This is perhaps the most popular method of interviewing, and its benefit is its flexibility that permits the pursuit of unexpected aspects of inquiry. Secondly, two subsequent focus group interviews were held with the same 12 interviewees of the first and third phases (six per focus group interview), in which the same procedures of semi-structured interviewing were followed. That is, the 12 participants of the first phase completed the first-phase questionnaires, were interviewed in the first phase, attended the two seminars, were interviewed again in the third-phase after the seminars, and finally were once again invited to attend the two focus group interviews. The purpose of organising these two focus group interviews was to discuss the feasibility and challenges of performing AAs more thoroughly. This is because “focus groups are a form of group interview that capitalises on communication between research participants in order to generate data” (Kitzinger, 1995:299). The effectiveness of focus group interviews is the question the moderator asks participants, which reflects on the entire discussion, and then participants offer their opinions on the subject of central importance to the investigator (Krueger, 2002).

4.4 The participants

Usually, research participants can be active or passive subjects of the methods of data collection, which may include observation, experiment, auto/biographical reflection, survey, or test (BERA, 2011). This action research had only active participants for its questionnaires,
interviews, and focus group interviews. All the teacher participants of this study were volunteers who signed the consent form for the three research instruments in advance. The number of participants was based on the necessity of the research instruments in the two phases of data collection. The entire research cohort for all the instruments of this action research were in-service Kurdish tertiary TESOL teachers from all the 18 English departments of the public universities of the Kurdistan Region. The survey questionnaire participants totalled 90 individuals (including the 12 interviewees), all of whom completed the first-phase questionnaires in the 18 English departments. Of these, 14 teachers have PhDs and 76 have MAs, 56 teachers have 1-5 years of experience of teaching English at the tertiary level, 17 have 6-10 years, seven have 11-15 years, two have 16-20 years, four have more than 20 years, and four did not declare their years of teaching English. The sampling of the 90 survey questionnaire participants from the 18 English departments possibly approximates the representation of the population of the in-service Kurdish tertiary TESOL teachers. This is because in all those English departments, around half of the teachers who taught ELT subjects completed the questionnaires. Therefore, regarding the participants’ current assessment practices and AAs, their knowledge of AA, and their beliefs about whether AAs are advantageous, generalisations can be made about Kurdish tertiary TESOL teachers, since the process can be considered as probability sampling (Creswell, 2005, cited in Grabin, 2009). When quoting from the survey participants’ responses to the questionnaire’s open-ended question, I referred to them as SP followed by a number (e.g. SP42 denotes ‘survey participant number 42’).

The other 12 participants of the three rounds of interviews, with six coming from each of the two English departments, took part successively in the first-phase questionnaires, the first-phase interviews, the second-phase seminars, the third-phase interviews, and the two focus group interviews of the third phase. Of these 12 interviewees, six have PhDs and six have MAs. All the 12 interviewees are the in-service Kurdish tertiary TESOL teachers who taught ELT subjects in the two English departments where the intervention of this action research was undertaken. Five interviewees, Ala, Goran, Ara, Dara, and Naska, have PhDs and are specialised in English language and linguistics, whereas Lara has a PhD and is specialised in English language and applied linguistics. Ala and Goran had been teaching English for nearly 14 years, Ara 10 years, Dara and Naska eight years, and Lara seven years. The other six interviewees had MAs; four of them, Nazanin, Mardan, Ashty, and Darawan have MAs in English language and applied linguistics, and the other two, Aram and Dildar, have MAs in
TESOL. Nazanin had been teaching English for nearly seven years, Aram six years, Mardan and Dildar five years, and Ashty and Darawan four years. The interview participants were representative of their two English departments, but not necessarily of all those in the region’s universities, although I believe they are similar significantly. When directly quoting from the three rounds of interviews, I used the above-mentioned interviewees’ pseudonyms.

All the questionnaire and interview participants should be afforded anonymity (concealing their identities in all research findings) and confidentiality (ensuring who has the right to access the participants’ collected data) (Norton, 2009); I tried hard to maintain these ethical issues in this study.

4.5 Data collection procedures

This section presents the procedures of the first and third phases of the data collection.

4.5.1 First-phase data collection

4.5.1.1 Piloting and distributing the questionnaires

For the first-phase survey questionnaires, I prepared and distributed a sample to six in-service Kurdish tertiary TESOL teachers in three English departments, two teachers from each. The purpose was to pilot the suitability of the questionnaire items, in order to acquire feedback that could inform the slight amendment of the questionnaire. Then, I tested the internal consistency reliability of the questionnaire items by applying Cronbach’s alpha for inter-reliability correlations (see Table 2 below). Cronbach’s alpha coefficient is an indicator most commonly used for internal consistency (Pallant, 2016), and its value is sensitive to the number of items in a scale: fewer than 10 items will affect its value (Pallant, 2016), and if the number of items increases, the value of α will increase (Field, 2009). Also, the negative wording of the items needs to be reversed before testing the reliability (Pallant, 2016). Based on these, I changed some negative wording in the questionnaire items, which were 31, and this was supposed to be helpful in making the value of Cronbach’s alpha higher. In spite of this, because it is usually difficult to achieve the internal consistency of a scale (Subedi, 2016), in the first run of Cronbach’s alpha to the questionnaire items, the result was below .7 indicating that the degree of correlation among items was not enough but it suggested that item number 8 should be deleted. This is necessary if an item causes a substantial decrease in
α, the researcher should consider dropping that item from the questionnaire (Field, 2009). After deleting item number 8, the Cronbach’s alpha coefficient value reached above .7 indicating that the survey met the internal consistency reliability. The Cronbach’s alpha coefficient value of a scale should ideally be above .7 (DeVellis, 2012, cited in Pallant, 2016), which is regarded as acceptable, although above .8 is preferable (Pallant, 2016).

<table>
<thead>
<tr>
<th>Cronbach’s Alpha Based on Standardized Items</th>
<th>Number of Items</th>
</tr>
</thead>
<tbody>
<tr>
<td>.728</td>
<td>31</td>
</tr>
<tr>
<td>.624</td>
<td></td>
</tr>
</tbody>
</table>

Table 2: SPSS Cronbach’s Alpha test for the internal consistency reliability of the questionnaire items

Following the pilots, I copied the questionnaires and distributed them to approximately 5-10 participants in each of the 18 English departments. The distribution included in-service Kurdish tertiary TESOL teachers who specialised in English language and linguistics or applied linguistics, and who were teaching ELT subjects. Not included were those teachers who specialised in English literature, as they were not TESOL teachers in the sense of teaching English language. In each department, 3-7 participants completed the questionnaires, with a total of 90 returned, and I was available to help if any questions arose about the questionnaire items.

4.5.1.2 Piloting and conducting the interviews

Generally, during the three rounds of interviews, I focused on the credibility and dependability of the interviews as much as possible. First, I considered that a valid argument can be obtained when it is sound, well grounded, justifiable, strong, and convincing, as described by Kvale (2007). Here, validity refers to the truth, correctness, and strength of a statement (ibid). For the reliability of interview data, I also took into account the general concern of interviewing, proposed by Kvale (2007), that subjects may change their answers during an interview, as they might yield different responses to different interviewers. I tried hard to avoid this by making the interviewees feel as relaxed and comfortable as possible, in order to encourage them to express their ideas freely, clearly, and without hesitation, leading to credible findings. In the first-phase interviews, each question was categorised to belong to an aspect of the teachers’ current assessment practices and AA's, their knowledge of AA, and
their beliefs about whether AAs are beneficial. These included some follow-up questions to obtain optimal responses from the participants (Turner, 2010). The questions were open-ended and as neutral as possible in order to enable the researcher to get insight into the cognition of the teachers (Turner, 2010). For the purpose of piloting the first-phase interview questions, I conducted two interviews with two participants from the two English departments. After reducing the number of questions based on the interviewees’ feedback, I conducted semi-structured interviews with 12 participants, six from each of the two English departments of the two faculties within a university. This phase was followed by the second phase, which was devoted to the intervention and was not a phase for data collection.

4.5.2 Third-phase data collection

Because of the participant-led nature of the design of this action research, a modification was made in the third-phase instruments. According to the initial design, in the third phase the interviews and questionnaires should have taken place. However, during the third-phase interview discussions about the challenges of AAs, I realised that the use of questionnaires does not guarantee good insight into the teachers’ cognitions of those challenges, the factors behind them, and how to solve them. Therefore, I decided to replace the questionnaires with another method that better fit the purpose of dealing more deeply with the challenges of AAs. After reading about focus group interviewing, I made the decision to replace the third-phase questionnaires with focus group interviews, which provided the opportunity for the participants to discuss the challenges of AAs with each other and attempt to find solutions.

One advantage of that flexible approach to data collection and to teacher development by the two focus group interviews was that it allowed for an in-depth discussion of the AA challenges and the possible solutions. Another benefit was creating a wider teacher development opportunity, as I responded to broader teacher interest in AA by inviting other members of the teaching staff to join the focus group interviews as observers, which enabled them to become more familiar with AAs and their associated challenges. However, a possible disadvantage arose in the form of a potential ethical implication related to the participants’ anonymity, since the attendance of other teachers meant that the participants could not be granted full anonymity, because everyone knew who said what in those two focus groups. Nonetheless, I protected their recordings on my PC and used pseudonyms in the presentation of the findings.
4.5.2.1 Piloting and conducting the interviews

As with the first phase, each third-phase interview question was open-ended and as neutral as possible, and intended to belong to a side of the teachers’ cognitions of AA including some follow-up questions. Moreover, the same procedures of piloting the first-phase interviews were applied to the piloting of the third-phase interviews. The teachers interviewed in the third phase were the same 12 who were interviewed in the first phase. By interviewing the same teachers before and after the intervention, it was possible to explore the potential increase in their knowledge of AA, the changes in their beliefs about AA, and their cognitions of the feasibility and challenges of performing AAs in their context. This enabled an analysis of the catalytic validity of this action research to see what improvements were achieved after the intervention seminars.

4.5.2.2 Holding the two focus group interviews

After I had discussed the challenges of the performance of AAs with the 12 interviewees in the third phase, I realised that these interviewees needed to discuss these challenges with each other to provide more clarifications. To this end, I held two focus group interviews in the two English departments. The participants were the 12 previous interviewees and some other teachers as listeners. Thus, I was able to avoid overly small or large groups, which risk being less productive or difficult to manage, respectively (Morgan, 1997).

4.6 Data analysis procedures

The majority of the findings of this empirical study provide insights into the participants’ AA practices and their cognition of their performance of those AAs. I organised the reporting of the findings through sub-sections, sections, and main sections by presenting, interpreting, and synthesising sub-categories to reach categories and synthesising categories to reach some notable themes. In this section, I explain the counting and processing of the responses of the questionnaire items, the coding and thematic data analysis of interview transcriptions, and the responses to the open-ended questionnaire question.
4.6.1 Counting and processing the responses of the first-phase questionnaire items

The answers to the open-ended question of the questionnaire were analysed qualitatively in the same way as the interviews. In contrast, the rating responses to the items about the assessment practices and AAs, and the cognition of the teacher participants on AA were calculated and analysed quantitatively. I added the responses manually and used a calculator to ascertain the percentages. I mentioned the number of a few participants who did not answer certain items or the open-ended question in the last column of each table of the statistical results of the questionnaires (see Appendix 3). The statistical percentages of the questionnaire items were displayed in tables and presented in the relevant sections of Chapter Five to support the qualitative findings of the interviews.

4.6.2 Coding, categorisation, and thematisation of interview transcriptions and responses to the open-ended questionnaire question

For coding, categorising, and thematising the interviews, focus group interviews, and the responses to the open-ended questionnaire question, I benefited from the following procedures and criteria for data analysis. I generally followed the seven stages of interview investigations by Kvale: designing, interviewing, transcribing, analysing, thematising, verifying, and reporting (Knox & Burkard 2009). Coding, which is fundamental to the stages of analysing, thematising, and verifying, is described by Kvale (2007:105) as it “involves attaching one or more keywords to a text segment in order to permit later identification of a statement”. Thus, words and short phrases can be used in coding (Saldaña, 2012); this is after segmenting the data into crucial analytical components (Grabin, 2009), by which the qualitative richness of the phenomenon can be captured (Fereday & Muir-Cochrane, 2006). Also for coding and categorising, I took benefit from Strauss and Corbin (1990:61, cited in Kvale, 2007), who explain that open coding refers to “the process of breaking down, examining, comparing, conceptualizing and categorizing data”. Kvale (2007) adds that the coding of a text’s meaning into categories makes it possible to quantify how often specific themes were addressed in a text, and the frequency of themes could be compared and correlated with other measures. Hence, the interview findings can be “compared, contrasted and even converted into statistics” (Grix 2004:127-128).

After coding and categorisation, thematisation is the most important stage among Kvale’s seven stages of interview investigations. Boyatzis (1998:161, cited in Fereday & Muir-
Cochrane, 2006) defines a theme as “a pattern in the information that at minimum describes and organises the possible observations and at maximum interprets aspects of the phenomenon”. Thematisation involves thematic content analysis, which is a descriptive presentation of qualitative data that may be interview transcripts or other relevant identified texts (Anderson, 2007). Therefore, thematic analysis is a search for themes that arise as effective aspects for describing the phenomenon, and this involves the identification of themes by reading and re-reading the data carefully. Thus, it is pattern recognition where emerging themes within the data become the categories for analysis (Fereday & Muir-Cochrane, 2006). Additionally, I took advantage of the 15 steps developed by Anderson (2007) for the thematic content analysis of interviews and open-ended questionnaire questions. However, thematic content analysis cannot suffice as a complete analysis as it is only descriptive. In this regard, I tried to follow phenomenological research procedures that require researchers to situate meaning units in relationship to context and structure (Anderson, 2007).

Moreover, I could take benefit from the ideas illustrated in Saldaña (2012) that, for the quantities of qualities, most qualitative studies in education research generate 80-100 codes that can be organised into 15–20 categories. Then, they will be synthesised eventually into 5-7 main concepts. The final number of the major themes must be held to a minimum in order to retain the coherence of the analysis (Saldaña, 2012). Committing to this, my research contains five main themes, each of which features several categories and sub-categories. Moreover, I benefited from the points listed by Alhojailan (2012), who notes that several points should be considered for data conclusions: (1) the notation of themes and the relevance of any similar or contradicting statements; (2) the grouping or establishment of categories of data that go together; (3) identification of interrelations among factors and variables; and (4) the building of conceptual coherence and consistency. Ultimately, these should be used to investigate the validity of the findings to fit the theoretical framework of the research (Alhojailan, 2012).

For the purpose of coding, categorisation, and thematisation, all the interviews and focus group interviews were transcribed fully, and the responses to the open-ended questionnaire question were typed in a summary. After I added up the qualitative data from the transcribed interviews and focus group interviews, and the typed responses to the open-ended question, I started coding what the participants stated by using short phrases (see Appendix 14). During
the coding process, I strove to arrive at findings deductively and inductively by making the categorisation and thematic analysis more flexible in order to allow the appearance of new sub-categories and categories if they arose. However, I kept in mind the five themes represented by the five research questions of this study, which influenced my search for sub-categories and constituting categories and themes.

After completing the coding process in this way, I came up with around 400 codes. In order to decrease the number of codes, I attempted to combine those that were similar by collapsing them into amalgamated structures, and also deleted many that were less relevant to the topics of the three research instruments. As such, the codes were reduced to 64, at which point I clustered the closely related codes to form 13 categories. After that, I clustered the closely related categories to form five themes, which I sequenced according to the design and the five research questions of this action research (see Appendix 1 for a table of all themes, along with their categories and sub-categories from the whole collected data, and Appendix 4 for a table that presents all themes, categories, and sub-categories from the answers to the open-ended questionnaire question). I allocated each theme a main section, within which I organised their categories and sub-categories in sections and sub-sections, respectively. Based on this coding, categorisation, and thematisation, I created main sections, sections, and sub-sections in NVivo for the organised themes, categories, and sub-categories, respectively. The purpose of this step was to put the relevant codes of each sub-category into an NVivo sub-section under a specific section and main section. By using NVivo, I was able to organise the coded data in a way that each theme included its categories, and each category included its sub-categories.

After I copied out the coded data distributed to the main sections, sections, and sub-sections in NVivo, I started synthesising what the participants stated about each sub-category, with adequate interpretation about it and how it was related to the other sub-categories in a category. Then, I provided enough interpretation about the interrelation of categories in a specific theme, and how the five themes were associated with one another. The interpretations were based on my critical thinking, understanding, reflection, and analysis of the synthesised data as a partial insider researcher.

It is noteworthy that, after data analysis, I used some direct quotations from the open-ended questions and the two kinds of interviews in order to report the findings. However, I first
modified them slightly by removing repetitions, hedges, digressions, and fillers that did not contribute to the meaning, thus preserving authenticity.

4.7 Role of the researcher

Prior to the start of this research, approximately two years before the data collection, I was a member of the academic context under study. Among the 18 English departments, I was involved in three of them for several years in a number of roles, including undergraduate student, master’s student, TESOL teacher, visiting TESOL teacher, coordinator of quality assurance, Deputy Head of Department, and Head of Department. Although I was not part of the other 15 English departments, I nevertheless attended several meetings with representatives from all these departments in order to determine the teaching subjects of the undergraduate studies. Therefore, when I started data collection, I was very familiar with many aspects of ELT in the context of this study. As such, my experience of the processes of data collection and analysis was as a partial insider researcher. My lengthy experience in the context was very helpful in facilitating this action research, including planning how to explore AA and the participants’ cognitions of it, and my insider status may have yielded more multi-dimensional and in-depth data than an outside researcher could have achieved.

Regarding my acquaintance with the participants, I did not personally know the majority of the survey participants of 15 English departments, with the exception of a very few individuals whom I had met at general meetings. However, in the three English departments where I worked, the participants were my former colleagues. Concerning the 12 interviewees in the two selected English departments where I did the interviews and intervention, all of them were my previous colleagues for several years. My long experience in the context and my familiarity with the interviewees enabled me to have a deeper critical understanding and analysis of the situation of AA in the English departments. Furthermore, I was able to understand the interviewees’ concerns about all aspects of AA more deeply than an outsider researcher, which might have contributed to the credibility of the findings. This is important because it is usual in a qualitative study for the researcher to be employed as a major data-collection instrument, because he or she can construct a composite and holistic picture, and analyse words before reporting the comprehensive views of respondents (Grabin, 2009). Therefore, in qualitative research, the researcher can assume a prominent role, since conclusions are not reached from the statistical data but rather from the understandings and
sensitive analyses made by researchers (ibid). That is why I contend that I had the leading role in the entire process of data collection, as well as the data analysis procedures and the reporting of the findings and conclusions. In order to eliminate my subjectivity as a researcher, I endeavoured to be as objective as possible to avoid the major criticism against qualitative research, which is the lack of objectivity (Walker & Evers, 1999, cited in Grabin, 2009); this also may add to the credibility of the findings.

4.8 Credibility and dependability of the findings

I attempted to achieve the reliability and validity of quantitative data findings, firstly by piloting the questionnaires through considering six participants’ opinions to improve the questionnaires. Secondly, I used Cronbach’s alpha as an additional piloting method to test the internal consistency of the questionnaire items, subsequently deleting one item and thus ensuring an internal consistency of more than .7. Then, for obtaining the credibility and dependability of qualitative data findings of the three rounds of interviews, I piloted the first- and third-phase interviews with four interviewees. Their feedback focused on the number of questions of the interview protocols rather than the content of those questions. Also, I could further reinforce the reliability and validity of quantitative data findings and the credibility and dependability of qualitative data findings by the triangulation of data collection, which was fulfilled by employing three research instruments: survey questionnaires, interviews, and focus group interviews. Triangulating the data collection, piloting the interviews and questionnaires, and using Cronbach’s alpha to test the questionnaire items may have contributed to more robust action research findings. Therefore, I believe that the results, findings, conclusions, implications, recommendations, and suggestions of this action research can be credible and dependable.

More importantly, the catalytic validity in critical research is the most influential criterion that an action study should suggest an agenda, give rise to specific actions, and strive to ensure that the research results in those actions. This is the research characteristic that tries to help participants to understand their worlds in order to transform them (Cohen et al., 2007). Also, the historical situatedness of critical research is another criterion that takes account of the “reality-shaping values” (Conrad & Serlin, 2006:409) of people’s situations, in this study, teachers’ situations. To commit to these criteria, the second-phase intervention by the two seminars was based on the findings and recommendations of the first phase, as it is highly
important in action research to prove that there is a need for intervention in a situation. It is stated by Troudi and Riley (1996) that an action research intervention plan must be tested after reviewing the collected data. After the intervention, the third-phase interviews and focus group interviews were used to check the changes in the participants’ cognitions of AA and their increasing willingness to pay greater attention to it. I suggest that the catalytic validity of this action research has been achieved satisfactorily and that the findings are credible enough to be depended upon by future studies.

4.9 Ethical considerations

I placed specific focus on the ethicality issues of the three rounds of interviews and the open-ended questionnaire question because the vast majority of these dealt with the participants’ cognitions of AA, which also constituted the major part of the qualitative data collected for this study. The emphasis on the ethicality issues of interviews was necessary because an interview study is a ‘moral enterprise’ in its means and aims, since the produced knowledge can influence our understanding of human situations (Kvale, 2007). Thus, interview inquiries are “saturated with moral and ethical issues” (Kvale, 2007:23), including the need to ensure the confidentiality of subjects, which means that their private data must not be revealed, or if there is any information that can be easily identifiable (Kvale, 2007), or it would be released, the participants should agree prior to this happening (Kvale, 1996). I also seriously considered the following ideas about all participants’ rights. For example, the protection of participants from harm in research might be psychological, in that the research might influence the participants’ self-esteem and academic confidence (Norton, 2009). Therefore, research participants have to be informed about the investigation purposes and the main design features (Kvale, 1996). Researchers should inform the participants about any expected detriment arising from the research process or research findings (BERA, 2011), and the risk of harming someone must be the least possibility (Kvale, 1996). To this end, ethical issues such as informed consent, confidentiality, and anticipated consequences must be taken into account with any qualitative interview (ibid).

For the quantitative data from the responses to the questionnaire Likert scale items, I applied the same ethical procedures and considerations of the qualitative data. That included the informed consent and confidentiality protection, which are the key issues in the ethical treatment of participants in social research (Singer, 2008). There is almost consensus about
the basic principles of ethical research, which are informed consent, confidentiality, and protection from harm (Norton, 2009). The protection of the participants of surveys and interviews can be possible by making them free from risk by informing them about the general nature of the research and their role as participants, taking informed consent from them, giving them the right to refuse to participate or withdraw without penalty, and protecting their anonymity and confidentiality of their data (Zeni, 1998). This protection is important because there are various threats to the confidentiality of survey data, for example, carelessness in removing the participants’ names, addresses, or telephone numbers from questionnaires, not locking the cabinets in which questionnaires are kept, or not encrypting files in order to secure electronic files physically (just as one does for paper copies) (Singer, 2008).

Committing to the above-mentioned ethical issues, first I gave the information sheets about this action research to all the participants so that they could read them before deciding whether to participate. Secondly, I prepared the consent forms for them to sign before starting their participation. The ethical approval form of the University of Exeter had been previously issued to me before commencing the data collection processes. All the participants gave consent to volunteer in all the research instruments of this action research in one consent form (see Appendix 11), in which they admitted that they have read the information sheet which included a brief explanation of all research instruments in the first and third phases, in addition to mentioning the second-phase intervention seminars (see Appendix 10). The questionnaires were mentioned in the information sheet and the consent form as well, and all the survey participants read and signed them prior to their participation (see Appendices 10 and 11).

4.10 Limitations of this study

Regarding the data collection in the two phases of this action research, the use of the three research instruments required a lot of effort. That being said, the most difficult, time-consum ing, and costly element was the distribution of the first-phase survey questionnaires to 18 English departments of 10 universities in 11 cities and towns across the Kurdistan Region. As a result, because of this large workload and the restricted time I had for the three phases of data collection and intervention, there are some limitations to this study.
The first limitation is that a few survey participants avoided filling in a few items on the questionnaires, with eight participants failing to answer the open-ended question (see Appendix 3). A second limitation is that I could only attempt to change the participants’ cognitions of AA through the use of two seminars, designed to increase their knowledge of AA and change their beliefs about it. As such, I was not able to organise workshops to train them in how best to perform AAs, especially their marking techniques and how to document progress. A third limitation I perceived is that I could not investigate the affective challenges related to students’ attitudes towards AAs.

A fourth limitation of this study can be linked to the use of English language in the three rounds of interviews. All the participants of this study are bilingual TESOL teachers, since Kurdish is their native language, yet they are also proficient in the English language. An individual can be regarded as bilingual if he or she has a native-like competence in two languages or a minimal proficiency in a second language (Hamers & Blanc, 2000). In this study, I could have used the Kurdish language or have adopted a bilingual approach to data collection in the interviews, which might have yielded more in-depth and broader responses. This is because the degree of access a bilingual individual has to linguistic codes will vary along a number of dimensions, including the linguistic dimension (Hamers, 1981). Perhaps Merino (1997) supports this, observing that non-native English-speaking teachers sometimes feel unsafe when they use English in their teaching. Feeling unsafe is probably related to the fear of making mistakes due to the extent of their access to the linguistic codes of ESL or EFL. The participants of this study use Kurdish rather than English to communicate with each other in most day-to-day conversations, even when talking about ELT issues. The use of both languages in the focus group interviews did occur at times, as in a very few cases Kurdish was used by some participants. I transcribed those Kurdish-used parts and translated them into English, mentioning that they were translated from Kurdish using the acronym ‘TFK’. However, for the majority of the three rounds of interviews and focus group interviews, the participants were happy to use English. The decision to use English was made because of the terminology and precision of the concepts that were going to be discussed. Generally, the interviewees were confident in using English, and other than the few instances noted, it rarely presented as a barrier to expressing their ideas.

A fifth limitation is related to the Likert scales presented in Tables 11 and 12, in which the former ranged from very challenging to unchallenging (see Section 5.9.1), and the latter...
ranged from very practical to impractical (see Section 5.9.2). These Likert scales, along with some other scales in the questionnaire that are unconventional and non-ordinal, were used to explore the survey participants’ beliefs about the practicality or other issues of AA. Whilst these Likert scales were generally able to display the beliefs of the 90 survey participants about the feasibility and challenges of AAs as well as their benefits, they perhaps only did so partially rather than providing a deep insight into the participants’ beliefs, and thus they can be regarded as a limitation. This is because biases may intervene in surveys, such as central tendency bias (participants avoiding extreme response categories), acquiescence bias (participants agreeing with certain statements to please the researcher), and social desirability bias (participants portraying themselves in a more socially favourable light rather than being honest) (Subedi, 2016). Nonetheless, in this study the three rounds of interviews after the survey questionnaires might have decreased the influence of these biases considerably.

This chapter explained the informing research paradigm of this study, which was critical theory and its methodology, action research. After accounting for critical theory briefly, it provided some definitions of action research including its aims and the rationale for adopting it for this study, as a suitable methodology for achieving the aims of this study. This chapter explained data collection methods and procedures employed in this study such as survey questionnaires, interviews, and focus group interviews. It also clarified the stages of data collection and analysis according to the three phases of the design of this action research, including the data analysis procedures, coding, categorisation, and thematisation processes. This chapter also provided brief biographic information about the participants, the ethical considerations, and the role of the researcher as a partial insider. Finally, it presented the rigorousness of the findings including their credibility and dependability, and the limitation of this action research. The next chapter will present all the findings, providing multi-dimensional interpretations for them, and discussing them with the prior literature and the relevant theories.
Chapter Five: Findings and Discussions

In order to find answers to the five research questions of this action research, the data were collected in two phases using questionnaires, interviews, and focus group interviews. These research instruments produced abundant data, especially the three rounds of interviews that provided a deep insight into the participants’ cognitions of alternative assessment (AA). Thus, the findings reached in this study were derived, firstly, from the quantitative data of the responses to the questionnaire items that presented statistical measures, secondly from the qualitative data of the answers to the open-ended questionnaire question, and thirdly from the lengthy discussions of the interviews and focus group interviews, which comprised the majority of the data. In seeking answers to a specific research question, I searched the entire collected data, though I usually found the most relevant data in a certain research phase or a certain research instrument.

The findings of this action research are reported with reference to its five research questions that are, in turn, guided by its main goals. These are based on completing some elements of Bachman’s (2001) action research spiral and Riel’s (2007) model of action research (Mertler, 2013) (see Table 1). Depending on the five research questions, the findings constitute five themes, which are presented in five main sections in this chapter. The themes are about the participants’ assessment practices and AAs, their knowledge of AA, their beliefs about whether AAs are beneficial, the changes of their cognitions of AA following the intervention, and finally, their beliefs about the feasibility and challenges of the implementation of AAs. Each of these themes has several categories and sub-categories (see Appendix 1 for a list of themes, categories, and sub-categories), and each category is allocated a section in order to report it. In reporting the findings, quotations from interviews, focus group interviews, and answers to the open-ended questionnaire question are used, in addition to presenting tables from questionnaires that show relevant statistical data (see Appendix 3 for the tables of the questionnaire). Each main section of a theme is followed by another main section to provide in-depth discussions about its categories, by integrating them with prior empirical research findings and conclusions, and by linking them to the relevant theories.

Based on some elements of Bachman’s (2001) action research spiral and Riel’s (2007) model of action research (Mertler, 2013), the investigation of the current assessment practices and AAs of the participants, their knowledge of AA, and their beliefs about whether AAs are
beneficial constituted the first phase of the cycle of this action research. The purpose was to discover the extent to which the participants’ assessment practices included AAs, and to identify any areas for development. In this phase, I ascertained that there was ample room for the participants to develop their knowledge and skills of AA and its technical, duration, frequency, and marking requirements. Subsequently, I reflected on how to address this opportunity and I decided to plan some actions as a preparation for the second phase, in which I presented two seminars in order to introduce AA to the participants and increase their knowledge about its demands and procedures. After that came the third phase, in which I explored and evaluated the outcomes of the intervention to discover the extent to which the participants’ knowledge of AA had increased, whether their beliefs on AA had been altered, and whether their willingness to do more about AA was enhanced. In that phase, I also investigated thoroughly the participants’ beliefs about the feasibility and challenges of implementing AAs within their professional context. That was after the participants were more familiar with the requirements of AAs, and thus they became more able to give more realistic opinions about the practicality issues of AAs (see Appendix 2 for a brief presentation of the topics of all the major findings, conclusions, intervention, implications, and recommendations according to the three phases of this action research).

### 5.1 The assessment practices and alternative assessments of Kurdish tertiary TESOL teachers

This main section addresses the first research question, which explored the current assessment practices and AAs of in-service Kurdish tertiary TESOL teachers. The quantitative data from the questionnaires were mainly used to address this question, although some other relevant issues were taken from the interviews as well. The theme of this main section is divided into two categories. The first concerns the AAs and types of test questions used by the participants, including the time and frequency of AAs, whereas the second is about the marking of AAs and tests.

#### 5.1.1 Alternative assessments and test questions

Table 3 below presents the participants’ use of a variety of AAs, with the data acquired from the 90 survey questionnaires.
Table 3: Alternative assessments used by the survey participants

Because I did not do classroom observations to see what exactly the participants used to do regarding AAs, I had to rely on their reflections on their own practices, as expressed in the questionnaires and interviews. Table 3 shows that almost all the survey participants conducted various AAs with varying frequency. Among these AAs, their implementation of oral activities such as oral presentations, conversations, oral discussions, and direct questions and answers constituted the highest usage. Writing assignments were also used more frequently than other AAs. Most of these practices of the participants were more akin to active learning practices with an element of AA because they did not completely integrate AAs into the active learning practices during the lecture time, since they allocated 5-25 minutes for conducting AAs usually at the end of lectures (see below). AAs could not be properly performed in a short time, separated from teaching and learning times, and the continuous assessment of active learning practices should have been carried out by AAs simultaneously (see Section 3.4.4.3). This indicated that the participants needed support, firstly to increase their knowledge about AA and its technical, duration, frequency, and marking requirements, and secondly to develop their skills as to how to implement them.
However, some of their AAs contained positive assessment elements because they were performed in a longer class time and involved marking out of ten marks (e.g., presentations and writing assignments were assessed with 5-10 marks) (see Section 5.1.2). Nonetheless, many other AAs were marked with only 0.5 to 1.5 marks for each activity, which might indicate room for development in terms of assessment accuracy and meeting the three basic purposes of assessment for improving teaching and learning (see Sections 5.1.2 and 5.2). The predominant factor behind this was the allocation of nearly 80-90% of the marks for traditional exams as a top-down technical rule, by which the major part of student assessment was carried out (see Section 5.1.2).

Regarding the length of class time for the performance of AAs during the 45-50 minutes of lectures, only two participants (Lara and Naska) (2 of 12) stated that they engaged their students in AAs throughout the class time. As such, they could probably integrate AAs with their teaching and their students’ learning. However, the other 10 interviewees (10 of 12) devoted 5-25 minutes to AAs usually at the end of lectures but they admitted that this was not enough given the large number of students in their classrooms. Because of the limited lecture time and the high number of students, four interviewees (4 of 12) stated that they performed some AAs infrequently: once a week, once every 10 days, or, for some AAs such as presentations, up to three times a year. Dara confirmed this saying: “The first course, I will ask them to prepare one seminar, and the second course also the second presentation.” The other participants perhaps had the same case because of two determining factors, large classrooms and limited time, though they did not talk about it. The infrequent use of some AAs possibly decreased the continuous monitoring of each student’s language growth over time and slowed down the progress of AAs based on the development of students’ learning, and thus it may not have committed to the socio-cultural ZPD and mediation (see Section 5.2). Therefore, as some of these AAs were infrequently performed, it was highly likely that their general effectiveness in enhancing students’ learning would have been compromised (see Section 5.2). These issues in the context of this study indicated that there was room for developing and improving the use of AAs, in order to achieve better the desired results of AA. The reported frequency and extent of using AAs also suggested that there was greater potential for the use of AAs in the English departments of the universities of the Kurdistan Region.
Regarding tests, the statistical results of the questionnaires pointed to a variety of test questions and items that the survey participants used in their test papers, as presented in Table 4 below.

Table 4: Types of test questions and items used by the survey participants

<table>
<thead>
<tr>
<th>Questions and items</th>
<th>Always</th>
<th>Often</th>
<th>Sometimes</th>
<th>Rarely</th>
<th>Never</th>
<th>Not answered</th>
</tr>
</thead>
<tbody>
<tr>
<td>Multiple choices</td>
<td>8=9%</td>
<td>34=38%</td>
<td>27=30%</td>
<td>14=16%</td>
<td>7=8%</td>
<td>0=0%</td>
</tr>
<tr>
<td>Fill-ins</td>
<td>8=9%</td>
<td>26=29%</td>
<td>33=37%</td>
<td>12=13%</td>
<td>6=7%</td>
<td>5=5.55%</td>
</tr>
<tr>
<td>Matching items in two columns</td>
<td>3=3%</td>
<td>14=16%</td>
<td>23=26%</td>
<td>31=34%</td>
<td>19=21%</td>
<td>0=0%</td>
</tr>
<tr>
<td>True–false items</td>
<td>7=8%</td>
<td>30=33%</td>
<td>21=23%</td>
<td>19=21%</td>
<td>13=14%</td>
<td>0=0%</td>
</tr>
<tr>
<td>Short answers</td>
<td>26=29%</td>
<td>28=31%</td>
<td>27=30%</td>
<td>7=8%</td>
<td>2=2%</td>
<td>0=0%</td>
</tr>
<tr>
<td>Sentence-completion items</td>
<td>8=9%</td>
<td>28=31%</td>
<td>32=36%</td>
<td>16=18%</td>
<td>5=6%</td>
<td>1=1.11%</td>
</tr>
<tr>
<td>Writing paragraphs</td>
<td>32=36%</td>
<td>26=29%</td>
<td>14=16%</td>
<td>11=12%</td>
<td>5=6%</td>
<td>2=2.22%</td>
</tr>
</tbody>
</table>

Table 4 shows that most participants revealed a high frequency of using most of the types of test questions. Among these, short answers and writing paragraphs constituted the highest usage. The high frequency of most of the test questions could be attributed to the heavy reliance on traditional testing as the primary assessment of students under the requirements of the assessment system of the participants’ English departments, and thus they depended far less on AAs (see next section).

5.1.2 The marking of alternative assessments and tests

Table 5 shows the marking allocation of AAs and tests by the participants.
As shown in Table 5, only 10-20 marks were allocated for AAs from 100 marks of the whole academic year; meanwhile, 80-90 marks were assigned to traditional ‘pen and paper’ exams. This distribution was mostly based on traditional testing methods, which were apparent in the series of tests of the first and second terms, followed by a final year exam. The questionnaire results were indicative of the limited number of marks devoted to AAs because even from the 10-20 marks that the system allows teachers to allocate for AAs, half of the survey participants (45: 50%) devoted only 10 marks to them, and less than a quarter (13: 14%) devoted fewer than 10 marks to AAs (see Table 5). One of the interviewees highlighted the marking imbalance saying:

> Usually, here in our universities, a few tests are conducted, and usually these tests are the main tests that count heavily against the final score of the student... I am allowed only five marks or ten marks [to allocate for AAs] for the entire year, no more than this (Goran, Int1).

Because only such a small proportion of marks are set aside for AAs, teachers cannot award more than 0.5-1.5 marks for each AA activity, as noted by five of the interviewees (5 of 12). Among the 12 interviewees, only four (Dara, Dildar, Darawan, and Ashty) could use a technique by which they were able to assign nearly 10 marks for each AA activity, before adding all the marks together and dividing them into the number of performed activities to be reduced to 10-20 marks allowed to be assigned to AAs. Allocating only a few marks to AAs
indicated that the English departments studied did not value such activities highly, especially in comparison with traditional tests that were given more weight.

The marking of AAs needed improvement, since half of the interviewees (6 of 12) stated that they did not use the marking techniques for AAs such as rubrics and checklists, and a quarter of them (3 of 12) did not reveal whether they used them or not. Rather, most interviewees (9 of 12) referred to a simple list prepared for ticking (marking) students’ participation in AAs. For example, Dara said: “We have lists of participation, whether they [students] participate, and I will give them marks on that participation”. Lara clarified that “ticking their [students’] names in order to know that this [student] has participated today”, with Aram confirming that this technique “is not a rubric in its sense but it is like the name of the students in a list and then checking the activities”. There was some doubt about the feasibility of such ticking (marking) lists due to the high number of students, as Ara explained that “the challenge is that we have 45 students per class, [and so] if all of them raise their hands, a ridiculous situation comes up, and the teacher should tick for 45 students, and you cannot”. Perhaps more problematic than this was that three of the interviewees (3 of 12) marked students’ AA performances by memory. Marking AAs with few marks without using checklists and rubrics was related to the limited marks devoted to AAs within the assessment system of the English departments, in which assessment of students was mostly accomplished through tests (see Section 5.1.1).

If AAs were marked in a more rigorous way, using rubrics or checklists, then it was likely that their accuracy, validity, and ethicality would have been improved. I contend that most interviewees’ (9 of 12) marking of AAs by tick lists was more akin to giving rewarding marks than evaluative marks, for example, Mardan said, “The students participate more, will get more [marks]”. Dara added: “We have lists of participation of the students whether they participate and I’ll give them marks on that participation”. Lara confirmed: “I have a sample here, indicating or ticking their [students’] names in order to know that this has participated today”. Goran clarified further:

In my absentees sheet, the names list of students, I am ticking usually on a daily basis,... So, right next to the students name there is a tick of daily participation on that day, [indicating] that student took part and the name of the activity is usually given above (Goran, Int1).
Nevertheless, Ala and Darawan stated that they used a simple rubric, and Dildar used “a checklist for spelling mistakes, grammatical mistakes, and punctuation mistakes”.

Marking students’ participations in AAs was not always carried out. Seven interviewees (7 of 12) admitted that they were unable to mark all AAs, including Darawan, saying: “Maybe sometimes I give them grades, sometimes, not all the time”. Naska confirmed that “in the majority of the cases, especially in my class, it is always oral, and you cannot have the recording of all the things and correct [mark] them”. Regarding student-marked assessments like self-assessment and peer-assessment, almost all the interviewees (11 of 12) stated that they did not depend on or count the marks of those assessments because they believed that their students’ low level and subjectivity in these assessment forms might cause biases in favour of themselves or in favour of or against certain students (see Section 5.9.3). For this, Lara noted that, “it [assessment by students] is not reliable”. Nazanin added, “when they have peer-assessment..., there is no mark from the side of the teacher”. Naska confirmed, “for matters of grading and things like that, I don’t depend on their [students’] assessments”.

These limitations to AAs in the context of this study revealed teachers’ need for support in terms of their implementation and marking of AAs.

5.2 Discussion of the assessment practices and alternative assessments of Kurdish tertiary TESOL teachers

The participants’ use of various AAs turned out to be mainly active learning practices with an element of AA, which means that they need to develop their usage of AAs mentioned in the literature by Grabin (2009), Tsagari (2004), Brown and Hudson (1998), Tannenbaum (1996), and Hamayan (1995). However, regarding diversity, their reported AAs are in line with the recommendation of TESOL International Association (2010), which advocates the use of various performance-based assessment tools and techniques. Cheng et al. (2004) also discovered that the assessment practices of tertiary TESOL teachers in China, Canada, and Hong Kong were diverse. Nonetheless, the participants of this current study need to take greater benefit from more attention to integrating AAs with teaching and learning and establishing marking criteria for AAs. This is because their AAs usually took place at the end of lectures with a tick list to record student participation, arguably a more simple way of conducting AA, and therefore there is room to develop their implementation and marking of AAs. Comparatively, studies from Canada and Hong Kong (Cheng et al., 2008) suggested
that AAs could be more integrated into teaching and learning throughout the duration of a class, because teachers in those two contexts mostly depended on coursework (that includes numerous AAs) rather than formal tests. Nevertheless, the participants of this study indicated an inclination to use AAs without perhaps using the full potential of AA. This is similar to the teacher participants of Ataç (2012), who were strongly inclined towards adopting AA.

Within the context of this present study, the limited class time had an effect on the ability to conduct AAs frequently, along with the high number of students. This supports the findings of Tsagari (2004), who discovered that large class sizes and insufficient time to perform AAs can be enormous obstacles to their successful implementation. The infrequent performance of some AAs, by which the process of learning cannot be assessed continuously, does not adhere to the modern assessment culture (mostly AA), which has the feature of the evaluation of both the process and product of assessment tasks (Lynch, 2001). AAs require time to implement adequately, in order to meet the criteria for documenting individual learners’ growth over time rather than comparing learners with each other (Tannenbaum, 1996). For example, portfolios can be used to include all recordings of AAs over a period of time, collating students’ work to be organised to assess their competencies in a given standard or objective (Ripley, 2012).

In the context of this study, another potential development in the usage of AA is related to the adequate allocation of marks. The participants of this research were only able to devote 0.5-1.5 marks to each AA activity, and these were mostly given as rewards rather than serving an assessment purpose. In contrast, traditional testing made up nearly 90% of the marks. This is similar to what Shohamy et al.’s (2008) teacher participants reported, as the highest percentage of the grades they gave were for quizzes and tests. The sole or heavy reliance upon traditional testing might be a reaction to the large classes. This is also the case in Cheng et al.’s study of tertiary EFL teachers in China, who used selection item standardised tests and translation exercises frequently. The heavier focus on tests than coursework was due to the large class sizes in the country, since tests can be more easily marked than AAs, and thus are less labour intensive (Cheng et al., 2008). This confirms Ataç’s (2012) assertion that, until recently, assessments were based mostly on standardised ‘pen and paper’ tests, which still cover a significant portion of assessment. In fact, this indicates the practicality of traditional testing in large classes in many contexts.
Devoting more marks for AAs, using rubrics and checklists, and establishing continuous marking would all help to develop the presence and use of AAs in the English departments. Further, such developments would lead to the ability to use AAs when collecting data on students’ progress, as Bachman and Palmer (2010) describe a broad range of AAs through which teachers can gather information on students. The reported AAs of the participants’ of this study, although containing an element of assessment, needed to be developed in terms of the two fundamental purposes of assessment, assessment as learning and assessment for learning (Earl & Katz, 2006), and teachers should no longer rely heavily on tests as their commitment to the requirements of their departments. This situation of the English departments of the universities of the Kurdistan Region might be akin to some English departments in China, where teachers focused more heavily on tests than on coursework (which incorporates many AAs) (Cheng et al., 2008). However, English departments in Canada and Hong Kong can be comparatively used as examples of how teachers can assign greater credence to coursework than formal tests (Cheng et al., 2008). Here again, however, the high number of students per class is a significant determiner.

Concerning the student-marked assessments, the interviewees needed to be encouraged to view self-assessment and peer-assessment more favourably and not place quite so much emphasis on the subjectivity of students, which according to the participants leads to overestimation or underestimation of their own or their peers’ work. The interviewees of this current study are similar to those of Yu’s (2013) study, which found that some EFL teachers may not be aware of the value and potential of peer-assessment for their students’ learning. However, overestimation and underestimation in self-assessment and peer-assessment was also observed by other researchers, including Khonbi and Sadeghi (2013), Chen (2008), and Suzuki (2009), which indicates that it is a pressing challenge in many contexts because subjectivity in such assessments would compromise their advantages for learning.

Additionally, in the context of this study, the infrequent performance of AAs could potentially affect their multi-dimensional benefits for enhancing students’ learning. As such, the most important feature of AA, which is promoting students’ learning, could be compromised (see Section 5.5.2). Black and Wiliam (1998, cited in Earl & Katz, 2006) synthesised more than 250 studies, revealing that the intentional use of AAs to promote learning improved students’ achievements. Therefore, concerning one of the basic purposes of assessment, which is ‘assessment as learning’, the AAs used by the participants of this
study need to be improved. ‘Assessment as learning’ is to develop and support metacognition for students through assessment (Earl & Katz, 2006).

The limited implementation of AAs in the context of this study can potentially be addressed yet are currently hindered by class size, limited time, infrequency of performing AAs, and not marking all of them. Solving these issues could fulfil the goals of the underpinning theories of AA such as socio-cultural theory, constructivism, multiple intelligence theory, communicative approach, and interpretivism. These theories have the common objective of focusing on the social aspect of learning and student-centred learning, whereby students are given an active role. Also, these challenges would not contribute to empowering students and increasing the validity and ethicality of teaching, learning, and assessment, which are emphasised by the underpinning critical theories of AA like CALx, critical pedagogy, and critical language testing.

5.3 Assessment literacy of Kurdish tertiary TESOL teachers in alternative assessment

This main section addresses the second research question, which explores whether in-service Kurdish tertiary TESOL teachers have knowledge about AA. The questionnaires and interviews of the first phase were used to answer this question. The theme of this main section has two categories: the participants’ knowledge of AA, and the areas for enhancing their knowledge of AA.

5.3.1 Participants’ knowledge of alternative assessment

As displayed in Table 6 below, the majority of the survey participants (78: 87%) said that they had read about AA, except 12 (13%), and the vast majority (86: 95%) said they possessed varying knowledge about AA, except 2 (2%).

<table>
<thead>
<tr>
<th>Have you read about AAs?</th>
<th>Read a lot</th>
<th>Read</th>
<th>Read a little</th>
<th>Have not read</th>
<th>Not answered</th>
</tr>
</thead>
<tbody>
<tr>
<td>14=16%</td>
<td>34=38%</td>
<td>30=33%</td>
<td>12=13%</td>
<td>0=0%</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Do you have enough knowledge about AAs?</th>
<th>Have a lot</th>
<th>Have</th>
<th>Have a little</th>
<th>Do not have</th>
<th>Not answered</th>
</tr>
</thead>
<tbody>
<tr>
<td>12=13%</td>
<td>46=51%</td>
<td>28=31%</td>
<td>2=2%</td>
<td>2=2.22%</td>
<td></td>
</tr>
</tbody>
</table>

Table 6: Survey participants’ readings and knowledge of alternative assessment
However, in their answers to the open-ended questionnaire question aimed at ascertaining what they know about AA, the majority of the survey participants (71: 79%) talked about the general usefulness of AA for their students’ learning, but they did not define it conceptually. This implies that they possibly found it difficult to conceptualise AA despite being familiar with its benefits. Other than the 71 participants, three listed some AA activities while eight did not write anything (or anything of relevance). While the survey participants showed that they used a range of AAs (see Table 3), from the open-ended question (see Appendix 3 and 4) I understood that some were not aware of the conceptual and theoretical frameworks associated with AAs. Therefore, the participants needed support to possess more conceptual knowledge about AA, and how it is underpinned by several teaching, learning, assessment, critical, and socio-cultural theories. However, being aware of the benefits of AA probably indicates that they had an idea and a general knowledge about AA, which they possibly obtained from their experience of performing AAs with their students. It is worth mentioning that eight of 90 survey participants could provide a definition of AA. For example, SP42 defined AA as “AAs encompass formative evaluations done by the teacher as a quick reminder, and as a tool to touchstone students’ performances and progress inside a classroom”, and SP51 defined it as “AA is the ability to multi-evaluate the capability or feedback of students”.

The answers of almost all the interviewees (11 of 12) to the questions about their familiarity with AA were short explanations and did not include conceptualisation of AA. In fact, two thirds of the interviewees (8 of 12) admitted that their knowledge of AA was inadequate. For example, Mardan said, “I think most of our teachers follow the strategies of conducting AA, but they may not have an idea about the concept itself”. It is noteworthy that among the 12 interviewees, Aram provided a concise definition: “AA includes different ways of giving assessment, not having conventional ones, only testing on paper and marking. It could be a continuous process of daily assessment, either by conversation, talking or probably sometimes that involves quizzes”. Almost all the interviewees (11 of 12) were unable to provide a clear definition of AA, and many of them (7 of 12) possessed some misconceptions about it. For instance, Dara linked the performance of AA to production skills only, and Dildar, Nazanin, Darawan, Aram, Lara, and Ala all conflated AA with speaking skills only. This revealed the potential room for development in relation to their knowledge about how all language skills can be assessed by AAs.
Regarding the differences between traditional testing and AA, two thirds of the interviewees (8 of 12) did not differentiate between traditional testing and AA conceptually but they distinguished them technically and psychologically. Psychological differentiation included participants’ beliefs that testing is stressful or causes fear, whereas AAs give freedom to students. However, two of the 12 interviewees, Lara and Ashty, clarified that AAs can sometimes cause shyness and reticence. Nonetheless, four interviewees (Aram, Nazanin, Ara, and Naska) distinguished testing from AA conceptually.

Concerning the purposes of assessment or AA, more than two thirds of the interviewees (9 of 12) talked about ‘assessment of learning’ as their major purpose. I asked them whether AAs could have any influence on their teaching styles and materials, and their students’ learning, which prompted further discussion. In particular, three interviewees demonstrated some knowledge about this, including Lara who talked about assessment for learning: “another purpose will be how, for example, am I successful in my teaching? If there is a problem..., I will try to improve myself... my style of teaching” (Lara, Int1). Aram and Ara talked about assessment as learning saying: “[AA is] mostly for improving their [students’] performance” (Aram, Int1), and “assessment in the class is really important because it is a way of encouragement” (Ara, Int1).

The knowledge about AA that the participants demonstrated in the survey and in the interviews provided a mixed picture, with a minority being able to provide definitions of varied complexity, while others had limited knowledge about AA. Nonetheless, the participants were aware of the long time needed for a large number of students to participate in certain kinds of AAs and the benefits of those assessments for learning mentioned in the literature (see Section 5.5.2), but they were apparently in need of appropriate support in knowledge and training. As I observed, it was clear that they would benefit from greater conceptual and theoretical understanding of AA and more knowledge about: (1) the continuous nature of AAs throughout the class time when integrated into teaching and learning; (2) the necessity of the frequency of AAs throughout the academic year to observe the students’ growth over time; and (3) the criteria of implementing and marking AAs with rubrics and checklists. The participants needed to be informed of these issues in order to be able to reflect on the feasibility and challenging aspects of using AAs. These issues became part of the intervention by the seminars in the second phase (see Appendix 7), since this study
had the aim of familiarising teacher participants with conceptual and theoretical knowledge of AA to strengthen their understanding of it, and encourage wider use of these assessments.

5.3.2 Factors behind the participants’ limited assessment literacy in alternative assessment

Several factors were identified as areas to improve in order to increase the participants’ knowledge of AA. Firstly, AA was not part of the MA and PhD syllabi of teachers who were not specialised in applied linguistics or related fields. Dara, Naska, Ashty, Goran, and Ara mentioned that they were not specialised in applied linguistics, and they had not studied AA. For example, Naska said: “I think teachers should be given more courses, especially teachers like our specialty because we didn’t have any classes [about AA]”. Goran added: “No, I haven’t [got any knowledge about AA] because I am not a graduate of a school of education”. I suggest that this could apply to all teachers who were not specialised in applied linguistics, and contend that they constitute the majority of the teaching staff of the English departments of the universities of the Kurdistan Region. Even those who were specialised in applied linguistics had not studied AA, as indicated by Mardan, Nazanin, Lara, Dildar, and Darawan, who were either specialised in this field or in a related one but who had not studied AA. Nazanin, for example, said: “No, I haven’t studied it in this way, since I’m not specialised in teaching methods”. Among the interviewees, only Aram, who obtained a Master’s degree in TESOL from an institution in the United Kingdom, said “I took those items as part of my MA courses”. This might indicate that the majority of the participants had not been introduced to the concept of AA in their previous courses, which meant that there was the opportunity for them to develop this skill and use it in their practice.

Secondly, another factor that could lead to improvement in AA knowledge and usage is the provision of more resources on AA guidelines and standards. Dissatisfaction with the provision of resources in the English departments was expressed by almost all of the interviewees (11 of 12). Further, three interviewees (Goran, Darawan, Aram) clarified that they had only rules for technical regulations of the distribution of marks for the two main term tests and the final examination, and a dean of one of the faculties confirmed that. As such, I was able to infer that the participants’ approaches towards the use of AA stemmed from their experience and their efforts rather than their previous training or their departments’ guidelines and rules. That was despite the fact that two thirds of the interviewees (8 of 12)
demonstrated their awareness of the importance and necessity of the guidelines for implementing AAs, as Ashty indicated: “I think it is a necessity”. Aram confirmed, “It is very highly recommended, [and] the more guidelines you have, the better teachers perform, [and thus] the better ideas they have about what to do... It is very important”. However, Lara and Naska had a different idea about guidelines, perceiving them as a kind of restriction to teachers. In particular, Naska believed that there should be general guidelines but not restrictive ones. Thus, standards and guidelines were regarded as influential by most interviewees, but in reality their practices possibly were not influenced by any standards, guidelines, or principles of AA.

Thirdly, the limited number of training courses in assessment and/or AA was another factor, as displayed in Table 7 below.

<table>
<thead>
<tr>
<th>Attendance of any training course(s) on teaching and/or assessment</th>
</tr>
</thead>
<tbody>
<tr>
<td>The courses</td>
</tr>
<tr>
<td>Number and percentage of teachers that attended</td>
</tr>
</tbody>
</table>

Table 7: Training courses that the survey participants attended

Table 7 shows that a few of the survey participants (7: 8%) had attended courses on assessment, whereas the majority (83: 92%) attended some courses, which were not on assessment, AA, and/or testing. The 12 interviewees completed the questionnaires too, and thus they had the same situation as the survey participants. Expressing his concern at this situation, Goran said: “I think the department has to involve its own staff in some brief, short training courses, workshops to familiarise them with these concepts, and then expect them to apply, to implement them in class”.

Fourthly, the participants could benefit from more knowledge about the underpinning theories of AA, which might be achieved by the establishment of a modern curriculum in the English departments. Such a curriculum would help to ensure that the participants have adequate knowledge of AA and conduct it appropriately. A modern curriculum is supposed to provide the standards and guidelines of English language teaching, learning, and assessment, informed by newly developed educational theories; therein, some focus can inevitably be put
on AA. This also signifies that the English department administrators could develop more of an awareness of the recent approaches of learning, which could lead to greater use of AA. They could at least provide their teachers with some resources, general guidelines, seminars, or workshops on AA.

Overall, although the participants were free to perform AAs, albeit with very limited marks allocated to them (see Section 5.1.2), the aforementioned factors contributed considerably to limit their knowledge of and skill in performing AA. This shows the participants apparently needed appropriate training and support and revealed areas for development.

5.4 Discussion of the assessment literacy of Kurdish tertiary TESOL teachers in alternative assessment

The importance of the interrelationship between theory and practice is well explained by Carr and Kemmis (2003:113):

Teachers could not even begin to ‘practice’ without some knowledge of the situation in which they are operating and some idea of what it is that needs to be done. In this sense those engaged in the ‘practice’ of education must already possess some ‘theory’ of education which structures their activities and guides their decisions. A ‘practice’, then, is not some kind of thoughtless behaviour which exists separately from ‘theory’ and to which theory can be applied.

(Carr & Kemmis, 2003:113).

Carr and Kemmis’s (2003) idea generally denotes the importance of teachers’ knowledge about teaching, learning, and assessment theories. The unprecedented expansion of language testing during the first part of the twenty-first century necessitated the consideration of assessment literacy (Fulcher, 2012). Due to the need for knowledge of assessment in general, Mahboob and Tilakaratna (2012) recommend that it is urgent to equip teachers with the knowledge of assessment so that they can empower students and measure their developments properly. Bailey and Wolf (2012), Troudi et al. (2009), and Brindley (2001) stress the necessity of providing teacher-training programmes as professional development opportunities for TESOL teachers, in order to help them acquire greater assessment literacy. This is because such literacy is regarded as an influential aspect of TESOL teachers’ professional development (Newfields, 2006). Therefore, increasing TESOL teachers’ assessment literacy for implementing AA is essential and inevitable.
Assessment literacy is complex and multi-dimensional, and has been defined by numerous prominent assessment experts in various ways (Coombe et al., 2012). Despite all the definitions, Fulcher (2012) says that there is still slight agreement on what might constitute assessment literacy. However, the definitions provided in Coombe et al. (2012), Malone (2013), Inbar-Lourie (2008), Popham (2011), and Stiggins (1991, 1997, cited in Fulcher, 2012) have some elements in common. They describe assessment literacy as the necessary expertise, understanding, familiarity with, and knowledge of a myriad of skills, procedures, conceptualisations, quality standards, and principles of sound assessments, which are likely to affect educational decisions, including the application of such knowledge and decisions in classroom assessment practices. Also important is that there are two areas of assessment literacy: teachers’ assessment knowledge and teachers’ perceptions on their assessment knowledge (Coombe et al., 2012). More broadly, Davies (2008, cited in Mosquera et al., 2015:305) describes language assessment literacy “as being composed of skills (the how-to or basic testing expertise), knowledge (information about measurement and about language), and principles (concepts underlying testing such as validity, reliability, and ethics)”.

ESL/EFL teaching is a career that requires a specialised knowledge base (Richards, 2009), and successfully playing the dual roles of instructor and assessor requires a high level of awareness, organising skills, expertise, and human sensitivity (Wach, 2012). This is directly relevant to AAs, as they require lots of thought, space, logistics, and training (Law & Eckes, 2007). Hence, Rea-Dickins and Rixon (1999, cited in Shim, 2008) argue that if teachers are to be recognised as good assessors, they need to have more knowledge about assessment and the theory of assessment. That is why it was essential for this current study to explore whether the participants have knowledge about AA and the relevant theories, in order to estimate the need for intervention. The findings showed that the knowledge of the participants of this study about AA was limited, and they would greatly benefit from more knowledge on AA. In particular, a clearer understanding of the conceptual definitions of AA provided in the literature (e.g., Chirimbu, 2013; Ripley, 2012; Anusienė et al., 2007; Tsagari, 2004; Rudner & Schafer, 2002) would be useful. Most participants (71: 79%) talked about the general usefulness of AA instead of defining it conceptually, which is the opposite of Estaji (2012), who discovered that, in defining AA, the majority of the participants placed little emphasis on the formative aspect of assessment, which is mostly related to AA benefits. However, Shim’s (2008) study on classroom AAs revealed that the teacher participants had adequate knowledge of general teaching and assessment theory, although most of them felt
that they needed more professional knowledge and skills in language assessment. These findings generally support those of Coombe et al. (2012), who argue that almost all teachers carry out assessments but without having learned the principles of sound assessment. This, I suggest, can be deemed a constraining concern.

Knowledge about the differences between testing and assessment or AA is also essential because language assessment literacy is about teachers’ familiarity with both testing and assessment conceptualisations (Malone, 2013). It was found that the interviewees’ knowledge of AA would greatly improve if they were able to identify the differences between testing and assessment more broadly. In addition, awareness of the fact that assessment (mostly AA) covers both formal measurement tools and other kinds of qualitative assessments (Schmitt, 2002) would be beneficial. Likewise, mindfulness of other aspects of assessment that make it more comprehensive, such as its features, principles, phases, and goals (Coombe et al., 2007; Phye, 1996), would be also advantageous in increasing assessment literacy, particularly in AA. The interviewees’ limited knowledge about the differences between testing and assessment goes partially with the misunderstanding of some of Estaji’s (2012) participants, who viewed assessment as a small part of an exam.

Concerning participants’ knowledge of assessment purposes, one of the issues that TESOL teachers should be aware of is the purposes of assessment (TESOL International Association, 2010), which includes assessment for learning, assessment as learning, and assessment of learning (Earl & Katz, 2006). AA addresses those purposes, and therefore teachers’ knowledge about them is relevant to their assessment literacy in AA. In this regard, Estaji’s (2012) participants recognised various purposes of AA, and they viewed assessment as having educational purposes rather than merely being useful for scores or grades. However, the interviewees of this current study needed more knowledge about the three purposes of assessment. Therefore, they would benefit from enhanced knowledge about these purposes in order to prevent over reliance upon one purpose of assessment, which was assessment of learning. More knowledge about these assessment purposes would improve teacher literacy in general, and Stiggins (1995) suggests that, according to the standards, teachers should be aware of assessments’ original functions and what clear purposes they serve. Assessment must also derive from and reflect appropriate achievement targets, because assessment-literate teachers should know what they are assessing and for what reason (Stiggins, 1995).
That is why sufficient knowledge about these fundamental purposes of assessment is crucial to increase assessment literacy, particularly in AA.

Several factors could lead to the improvement of the participants’ assessment knowledge of AA. Firstly, AA should be covered by the MA and PhD syllabi of teachers, thus providing them with an introduction. This is in line with Green (2014:6), who mentions “in language education, teacher training programmes tend to give assessment rather little attention. Assessment is often the topic of just a few sessions: apparently an issue of little direct concern to the teacher, perhaps better left to expert specialists”. Estaji (2012) also discovered that some instructors had not studied or had not been exposed to the concepts of assessment. These findings also support Coombe et al. (2012) who state that almost all teachers have not learned the principles of sound assessment. Covering AA in the syllabi of MA and PhD courses in English departments would prevent the critical impact of limited assessment literacy of teachers on their professional development.

Secondly, providing resources on the standards and guidelines of AA in English departments can break down barriers to assessment literacy. Coombe et al. (2012) explain that the lack of resources on assessment can hinder the development of assessment literacy. Educational institutions have introduced various assessment guidelines to make decisions on selection, certification, and achievement (Brindley, 2001). This means that the English departments of the universities of the Kurdistan Region need to provide resources on AA due to its value and necessity. Assessment and AA resources like standards and guidelines were regarded as imperative by two thirds of the interviewees (8 of 12) that took part in this study, thereby concurring with the extant literature. For instance, Coombe et al. (2012) state that assessment literacy means teachers should be familiar with how to meet specific quality standards. At present, there are 11 TESOL standards that fall under five conceptual domain frameworks: language, culture, planning, assessment, and professionalism (Thibeault et al., 2011). More knowledge about assessment standards and actual use of them could strengthen English language teaching and learning because TESOL standards have almost reached universal status (Cumming, 2009), and assessment in language teaching is especially influenced by the standards movement that thrust assessment into a more prominent position in the learning process (Grabin, 2009).
Thirdly, more training courses for the teacher participants of this study would contribute to increasing their knowledge of AAs and enhancing their skills of implementing them. According to Taylor (2009), teachers perform assessments but often without assessment training to equip them adequately for this role. This is meaningful because the availability of suitable and trained graders is part of assessment practicality (Coombe et al., 2007).

Fourthly, the participants’ potential development of assessment literacy in AA is related to enhancing their familiarity with the recent teaching, learning, socio-cultural, and critical theories that underpin AA. This is because the existence of learning theories logically triggers the appearance of different assessment and testing methods (Chirimbu, 2013), or the progress of different assessment practices is usually influenced by various language learning methods (Dragemark Oscarson, 2009). Enhanced knowledge might also include the AA consideration of assessment use, its societal values, its consequences, and its ethicality (Bachman, 2007), as well as Shohamy’s critical language testing principles, such as questioning the values, agendas, and needs of tests (Bachman, 2005). If the participants know more about those assessment theories and principles, they possibly could develop their understanding and knowledge of AA.

More broadly, limited knowledge about socio-cultural theory, constructivism, multiple intelligences, communicative approach, interpretivism, CALx, critical pedagogy, and critical language testing would lead to limited knowledge about the common goals of those theories. These shared goals include focusing on the social aspect of learning, the democratic interactions between teachers and students, student-centred learning, students’ active roles, and the validity and ethicality of assessment. These are also the essential features of AA that are possibly under the influence of those theories, and greater knowledge about them can lead to more valid and ethical teaching, learning, and assessment. Sufficient knowledge of the relevant theories of any practice is of vital importance because “those engaged in the ‘practice’ of education must already possess some ‘theory’ of education which structures their activities and guides their decisions” (Carr & Kemmis, 2003:113). More specifically, familiarity with the concepts and theories of AA is crucial, because this meets a basic requirement of language assessment literacy, which is language teachers’ familiarity with the testing and assessment conceptualisations, and applying them to classroom assessment practices (Malone, 2013). Additionally, teachers’ literacy in assessment, especially in AA, is a strength in their teaching skills because the evaluation of students’ progress is considered a
major aspect of a teacher’s job (Brumfit & Johnson, 1979). This is because of the inevitability of assessment as an integral part of education, and because without enough teacher assessment literacy, students cannot attain high levels of academic achievement (Coombe et al., 2012).

5.5 Beliefs of Kurdish tertiary TESOL teachers on whether alternative assessments are beneficial

This main section will address the third research question, which explored whether the in-service Kurdish tertiary TESOL teachers believe that AAs are beneficial. In order to find an answer to this question, I chiefly took advantage of the data from the first-phase questionnaires and interviews. The participants of this study held a variety of beliefs concerning the value and benefits of AA. However, their beliefs stemmed mostly from their conducting of active learning practices with an element of AA, which is generally good and contributes significantly to increasing their cognition of AA. The theme of the benefits of AA from the beliefs of the participants has two categories: the benefits of AAs for teaching and the advantages of AAs for learning.

5.5.1 Benefits of alternative assessments to teaching

According to an item of the questionnaires about whether AA improves their teaching (see Table 8), the vast majority of the survey participants (88: 98%) agreed that they could use the outcomes or observations of AAs to improve their teaching methods.

<table>
<thead>
<tr>
<th>I can use the results of AAs to improve my teaching.</th>
<th>Strongly agree</th>
<th>Agree</th>
<th>Unsure</th>
<th>Disagree</th>
<th>Strongly disagree</th>
<th>Not answered</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>34=38%</td>
<td>46=51%</td>
<td>8=9%</td>
<td>0</td>
<td>0</td>
<td>2=2.22</td>
</tr>
</tbody>
</table>

Table 8: Survey participants’ ratings of the benefits of alternative assessments to teaching

Then, in answering the open-ended questionnaire question, a few survey participants (6: 7%) talked about the influence of AAs on their teaching. For instance, SP24 said: “It makes me correct the way of teaching that I follow”, with SP57 adding: “I think AAs encourage teachers to adjust their teaching styles accordingly”. SP48 confirmed: “AA gives a full picture of students’ performance and this helps teachers to improve their own teaching”.

Also, the majority of interviewees (10 of 12) stated that they used AAs for improving their teaching by changing their methods and strategies, mainly to accommodate different students. Goran noted that “AA improves my way of teaching, a kind of regular monitoring of myself and the teaching process... to tailor my teaching methods to the needs of the students”. Dildar added: “If I see the results [of AA] are not good, I have to change my strategy, my approach, my method”. Naska confirmed that AA “will help you to find the best way to approach the students”. However, I contend that it was perhaps difficult for them to manipulate AA outcomes adequately in order to improve their teaching methods and materials, primarily because they needed to develop their simultaneous combination of AAs with teaching and learning (see Section 5.1.1). Teaching improvement can be more successful if it is based on continuous observation of students’ learning progress by AAs (see Section 5.6).

5.5.2 Benefits of alternative assessments to learning

The second category, which is about the benefits of AAs for learning, has three subcategories: the improvement of students’ learning by AAs, the utilising of a variety of AAs, and the students as assessors, critical, reflective, and problem-solving thinkers and autonomous learners.

5.5.2.1 The improvement of students’ learning by alternative assessments

Based on the answers to a questionnaire item about the general benefits of AAs for students’ learning, the majority of the survey participants (66: 73%) regard it as very or quite beneficial (see Table 9).

<table>
<thead>
<tr>
<th>AAs will be beneficial for my EFL students more than testing.</th>
<th>Very beneficial</th>
<th>Quite beneficial</th>
<th>Unsure</th>
<th>Slightly beneficial</th>
<th>Not beneficial</th>
<th>Not answered</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>27=30%</td>
<td>39=43%</td>
<td>20=22%</td>
<td>3=3%</td>
<td>1=1%</td>
<td>0=0%</td>
</tr>
</tbody>
</table>

Table 9: Survey participants’ ratings of the benefits of alternative assessments to learning

In the responses to the open-ended questionnaire question, in their attempts to define AA, the majority of the survey participants (71: 79%) talked about the general usefulness of AAs in increasing students’ learning, by motivating them to practise the English language and master
the required skills. For example, SP22 said: “They [AAs] are practical, motivating, exciting, build self-confidence, and all in all, create the English environment for students in classes”. SP34 added: “They [AAs] help the students improve their language abilities and performance”, with SP68 confirming that “doing AAs improves the students’ language skills”. Additionally, all interviewees (12 of 12) perceived, experienced, or observed several benefits of AAs, as all of them believed that AAs improved students’ learning of English language skills by encouraging them to be more active and more involved in practising language through active engagement, participation, and interaction in the classroom. For instance, Nazanin observed that “because it [AA] encourages them [students] to practise the language. They [AAs] motivate them to be involved inside the class with the topics. It makes them more active”. Ara added: “If you are applying assessment [AA], you are involved with the students. You get feedback at the same time, [and] you motivate them to study, to learn”.

Regarding the advantages of AAs for students’ performances, half of the interview participants (6 of 12) spoke about the necessity of AAs for improving students’ performance, although in doing so they mostly mentioned improving students’ speaking skill through classroom communications. They regarded communication as vital for students attempting to learn English language, as Ala said, “Actually, learning a language means communication, taking into consideration that we are non-native speakers [of English]”.

Concerning whether to focus more on students’ strengths or weaknesses in order to improve their learning by AAs, two thirds of the interviewees (8 of 12) believed that the focus should be on both strengths and weaknesses. However, they admitted that, previously, they focused more on providing feedback on students’ weaknesses rather than focusing on their strengths, as Goran said: “Sometimes, when I find that they are discouraged, I change the strategy and switch to focusing on the strengths, but, by nature, I am more focused on weaknesses”. Dildar confirmed that students’ “strengths are also important but, in our classes, in EFL classes, we have more weaknesses, unfortunately, than strengths... They have to get benefit from their weaknesses”. When giving feedback, those eight interviewees emphasised the need to provide encouragement rather than offering direct or embarrassing feedback.

Overall, the interviewees’ implementation of AAs improved their students’ learning but it might be more improved if AAs are more frequently performed, because this provides
continuous adjustment of AAs to fit the progress of their students’ language growth, which is usually in the form of a socio-cultural ZPD and mediation (see Section 5.6).

5.5.2.2 Using various alternative assessments to improve learning and assessment accuracy

More than two thirds of the interviewees (9 of 12) believed that enhancing students’ learning of English language needs various AAs to provide fair opportunities for them so that they can choose and perform suitable AA activities frequently. They also believed that students have the right to perform multiple AAs to present their knowledge and abilities, as part of providing them with a fair assessment, rather than just depending on a few tests. For example, Ala explained: “Learning a language should not depend on only tests, it should have activities, presentation, speaking, dialogue, listening”. That is why more than half of the interviewees (7 of 12) believed that the systematic allocation of more marks for AAs would increase the value of these assessments in their English departments. This would encourage students to view AAs as valuable, which could lead to improvements in their learning of English, and give them more opportunities to show their knowledge and skills. The beliefs of those seven interviewees could be related to the possible situation of their students as being test- and certificate-oriented and thus looking more at passing than learning. Consequently, students prepare for tests rather than AAs because around 90% of the marks are devoted to tests (see Sections 5.1.2 and Table 5). I suggest that this situation could be the reason why some participants asked for more marks for AA, including Aram: “There is a strong relationship [between marks and AAs] because as long as there are marks to get, they study more and they learn more. It is a kind of incentive”. Ara agreed with Aram, noting that if more marks are allocated for AA, then such practices “give values to the students’ activities day by day... Everything is counted... It is a way of encouragement to them”. Goran confirmed: “The students are more focused on these tests than AAs. Sometimes, they don’t care about such assessments [AAs] in class because they say [there are no marks for them]”.

Ascribing more value to AAs by devoting more time and marks will make students more interested in them, which in turn can boost their learning because more than two thirds of the interviewees (9 of 12) believed that testing causes stress for students, while AAs make them involved in their learning and assessment and even enjoy it. Ara explained how her students enjoyed peer-assessment: “I told them to do this exercise, ‘give me the answers’ and I exchanged the papers to the other group. ‘OK, correct them’, and it was really playful, and
they were laughing”. In addition to providing more time and marks for AAs, students’ interest in performing them might be reinforced further by linking them to real world contexts, as all the interviewees (12 of 12) believed that their AAs were related to and thus can be applied in real life situations. For instance, Aram said: “I was teaching something practical, practical in the sense that they use it in real life. I try to involve them with the topics that can be used in those areas”. These 12 interviewees believed that this would improve students’ learning, as Darawan stated: “The class activity is connected to their realities, it helps them more”.

If AAs are awarded more weight in the assessment system of the English departments by allocating more marks, more time, and more frequent opportunities for usage, this will likely lead to a more comprehensive and accurate assessment of students. This is because almost all the survey participants (87: 97%) agreed that AAs allow their EFL students to show what they know, and the majority (80: 89%) agreed that AAs allow their students to exhibit what they can perform (see Table 10). This indicates that almost all of them agreed that AAs can be used to assess students’ English language proficiency accurately.

<table>
<thead>
<tr>
<th>AAs like oral presentation, doing conversations, oral discussion, writing assignments, self-assessment, peer-assessment, etc. allow my EFL students to show what they know.</th>
<th>Strongly agree</th>
<th>Agree</th>
<th>Unsure</th>
<th>Disagree</th>
<th>Strongly disagree</th>
<th>Not answered</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>44=48.88%</td>
<td>43=47.77%</td>
<td>3=3.33%</td>
<td>0=0%</td>
<td>0=0%</td>
<td>0=0%</td>
</tr>
<tr>
<td>AAs allow my EFL students to show what they can perform.</td>
<td>Strongly agree</td>
<td>Agree</td>
<td>Unsure</td>
<td>Disagree</td>
<td>Strongly disagree</td>
<td>Not answered</td>
</tr>
<tr>
<td></td>
<td>27=30%</td>
<td>53=59%</td>
<td>9=10%</td>
<td>1=1%</td>
<td>0=0%</td>
<td>0=0%</td>
</tr>
</tbody>
</table>

Table 10: Survey participants’ beliefs on the accuracy of alternative assessments

Also, two thirds of the interviewees (8 of 12) believed that AA is more direct and immediate, and its assessment would be clearer and more accurate because AA prevents students from relying on memorisation and instead assesses their current performance, which has priority in language learning. For example, Dara said: “They use their language exactly, and you can hear... You can assess easily”. Nazanin added: “you can hear [while performing AA] but, in testing, you just find the [answers]”. Mardan confirmed: “I think AA [is more accurate] because, sometimes, they memorise things, and they can write down what they have [remembered]”. Aram clarified further:
When they talk about a certain topic, from their talk, you can realise to what extent they are aware of the language... In order to measure student’s achievement, I would give priority to AA... I think as long as we are talking about language learning, performance is prior to the knowledge on paper, because knowledge on paper probably cannot be reflected in real life, probably cannot be performed (Aram, Int1).

In addition, two thirds of the interview participants (8 of 12) believed that more ways for students to show their knowledge and performance lead to more accurate assessments, as Aram said: “I always think that the more ways that you have for assessing somebody, the more accurate results that you can get”. Dara supported Aram’s idea that “multiple resources are better instead of depending only on the test papers”. They believed that assessment decisions are better and more accurate when based on multiple assessment results. They related that to some ethical issues, such as certain students potentially having psychological or physical problems that affect their ability to attend all tests. For this reason, Ala said: “I cannot assess, [or] decide the destiny of a student for one assessment”. Ashty explained more, saying: “The multiple one is fairer because even if the students didn’t make it in the test, they can make it in their performance, they can make it in their assignments, they can make it in their presentations”. Interestingly, some survey participants (14: 16%), in their attempts to define AA, demonstrated their understanding that, through AAs, a more comprehensive assessment is possible, since students have diverse ways to be assessed, and thus should not be limited to just one exam.

However, the majority of the participants had a say about the importance of traditional testing too, and how it should be combined with AA. For example, two interviewees, Dildar and Ala, believed that tests could also accurately show students’ knowledge. Further, the majority of the interviewees (10 of 12) believed that test results can be valuable, beneficial, and can improve learning. Moreover, almost all of the survey participants (88: 98%) agreed that the results of tests could be used for advancing their teaching. So, the majority of the survey and interview participants agreed that traditional testing also has significant impacts in terms of the three essential purposes of assessment. For this reason, the majority of the interviewees (10 of 12) called for an integration of AA and testing due to their importance in assessing the process and product of learning, as parts of a complementary process of assessment. Regarding this, Ala favoured both AA and testing: “I cannot depend on one of them unless the topic demands [it]... That is why we need both of them”.
5.5.2.3 Students as assessors, critical thinkers, and autonomous learners

Concerning student-marked assessments, all the interview participants (12 of 12) believed that generally teachers are not the only assessors of students. However, one interviewee, Lara, said that in Kurdistan it is possible for this role to be solely occupied by teachers. All the interviewees (12 of 12) believed, to differing degrees, that students could assess themselves or each other, as clarified by Naska, who said: “Especially in classes like writing, they can assess each other”. Nazanin added: “They can find out the errors, not only in vocabulary but also in grammar. So, this is also a way in order to enhance their ability”. The 12 interviewees believed that peer-assessment benefited students’ learning by enabling them to perceive their mistakes through those of their peers, as Ara explained: “they have learned not only from their mistakes [but also] they have learned from the others’ mistakes”. Ala added: “[If] they are aware of the other people’s mistakes, they will be more aware of their [own] mistakes”. However, it was apparent from the interviewees’ statements that they rarely allowed their students to be assessors of themselves by performing self-assessment, while they allowed students to assess each other by performing peer-assessment but they rarely counted the markings done by students to add marks to their scores because they did not trust and did not depend on students’ assessments. In this respect, Nazanin said: “When they have peer-assessment..., there is no mark from the side of the teacher”. Ara confirmed: “they are testers of themselves but not for the marks”. This is perhaps because of the subjectivity and low level of some of the students, as perceived by most interviewees (see Section 5.9.3). It could also stem from some interviewees’ belief that self-assessment and peer-assessment cannot produce the desired results in English departments because of some cultural and social factors, as Goran clarified: “Maybe due to some cultural problems in Kurdistan, we are different from the other cultures, this peer-assessment can be possible in some western communities, but I don’t think they will produce the desired results in Kurdistan”.

Regarding students’ subjectivity, more than half of the interviewees (7 of 12) believed that some students can determine their strengths and weaknesses by self-assessment and peer-assessment, but they do not reveal that to their teachers or peers. Those interviewees also believed that determining strengths and weaknesses requires critical and reflective thinking, and although students might possess this ability, they do not apply it to finding and displaying their own or their peers’ mistakes. According to Darawan:
They can [find strengths and weaknesses] but they will not tell. They know that, for example, a
student has a difficulty in speaking or writing, but he is not going to mention that for the
teacher or for the other students... They will not tell/reveal their weaknesses, [so as] not to
lower their marks (Darawan, Int1).

This could be also because, in their context, issues were taken personally, as Aram suggested:
“If your peers tell you that you’ve made a mistake, probably you don’t take it in an easy way
as your teacher tells you. So, it is kind of psychological”. Additionally, this is probably
because those interviewees believed that some students were not fair in their assessments,
with Ashty pointing out that “actually, no, they are biased”. Therefore, some of the
interviewees (e.g., Dara, Dildar, Goran, and Nazanin) believed that teachers’ must review
students’ peer-assessments afterwards, as Dara said: “I think that the teacher must take a
glance at least hurriedly on that piece of information [students’ peer-assessment]”. That is
why Goran recommended the following:

We need to create an environment of trust in class. The teacher has to make sure that all the
students, the assessments that they are giving, are valid, objective, and trusted. That when
student A is evaluated, is assessed by student B, he/she takes into account the points that are
against him or her (Goran, Int1).

This is despite the fact that more than two thirds of the interviewees (9 of 12) believed that
the adoption of AAs, especially self-assessment and peer-assessment, might help their
students to become more confident, critical thinkers, and autonomous learners, as Ashty said:
“They will be responsible for what they are doing; they will be free... It leads them to be
autonomous”. Ara added: “You give them a kind of confidence that they can be learners by
their own reading and by their own working”. However, Lara, Ala, Darawan, and Goran had
doubts about students’ abilities to be autonomous, with Darawan noting: “No, they cannot be
[autonomous]. They have to have a teacher because, if they are not instructed to do
something, they will not do that... So, they cannot be autonomous”. Goran agreed with
Darawan that students needed “to be more autonomous! Autonomy is a very important point
here, something that we are lacking in our classes”.

In sum, the participants believed in many of the benefits of AAs mentioned in the literature
(see Sections 3.6 and 5.6). However, increased knowledge of AA, improved AA skills, more
time, and fewer students per class would enable the participants to take more advantage of the
benefits of AA, as would continual adjustment of AAs upon which to base further teaching and learning developments.

5.6 Discussion of the beliefs of Kurdish tertiary TESOL teachers on whether alternative assessments are beneficial

Borg (2006:87) briefly shows the effectiveness of teacher cognition in forming teachers’ classroom practices, saying: “Diverse, too, is the range of concepts which have been invoked in accounting for the cognitions which shape in-service language teachers’ classroom practices” According to Borg, teacher cognitions include their instructional concerns or considerations, principles or maxims they try to implement, considering different levels of context, their pedagogical knowledge and personal practical knowledge, and their beliefs (Borg, 2006:87). Teacher cognition, which is the study of what teachers think, know, and believe (Borg & Al-Busaidi, 2012), is very powerful in developing educational practice, as it is explained above in Borg (2006). The nature of language teacher cognition has a practical orientation, and it is highly context-sensitive (ibid). Therefore, “language teaching can be seen as a process, which is defined by dynamic interactions among cognition, context, and experience” (ibid: 275). Teachers’ beliefs and attitudes are influential in shaping their classroom practices (Wang, 2011) and have an enormous influence on their assessment and evaluation practices (Estaji, 2012). This is because the teacher is “the agent of the assessment process” (Harlen, 1996:129; Rea-Dickens, 2004:251, cited in Estaji, 2012) and their beliefs can impact upon the whole learning environment, particularly the methods by which students’ success and failure are assessed and interpreted in the classroom (Dragemark Oscarson, 2009). More specifically, Shim (2008:63) mentions many researchers (e.g., Breen et al., 1997; Arkoudis & O’Loughlin, 2004; Davison, 2004; Edelenbos & Kubanek-German, 2004) who argue, “teachers’ beliefs and attitudes can have a significant influence on the assessment practices, which indicates that their assumptions might shape and constrain their practices in the assessment process”. Despite teachers’ major roles in assessment and the fact that they are the ultimate practitioners of educational principles and theories, they were absent from educational studies for numerous years, and their voices were seldom heard (Jia, 2004). Therefore, this study focused especially on the teacher participants’ AAs, their knowledge of them, their beliefs whether AAs are beneficial, and the feasibility and challenges of implementing them.
The attempts of the survey and interview participants of this study to employ AA outcomes for improving their styles of teaching commit to the status of assessment as an integral part of the process of teaching and learning (Coombe et al., 2012). Additionally, this could be part of the efforts of using assessment results to reach conclusions not only about learners but also about instruction as a whole, by which they use assessment to guide instruction (Hamayan, 1995). Empirically, this supports Shim’s (2008) finding that 39 teacher participants seemed to be interested in AA to reflect on their teaching issues, in order to modify what to teach and how to teach. The focus of most interviewees of this current study on students’ weaknesses to develop their teaching techniques denotes that they had diagnostic competence, which is defined by Edelenbos and Kubanek-German (2004, cited in Shim 2008) as “the ability to interpret students’ foreign language growth... and to provide students with appropriate help”. Nonetheless, the participants’ focus on students’ weaknesses might be at the expense of their strengths, which indicates the participants’ need to commit more to the focus of AA on learners’ emergent or dynamic abilities, whereby assessment is integrated into teaching and learning (Behrooznia, 2014). This is to avoid ignoring the first part of the ZPD of socio-cultural approach in AA, which is a development-oriented process that reveals students’ present abilities (Shrestha & Coffin, 2012). However, the participants’ focus on students’ weaknesses concurred with the second part of the ZPD, in which assistance is provided to overcome any problems of students’ performances, which can be accomplished by mediation (Shrestha & Coffin, 2012).

The participants used AAs to promote their teaching, but enhanced knowledge of AA and developing skills of implementing them could lead to further integration of AA into their teaching in order to advance learning. Such improvements would help the participants of my study to employ AAs to collect more valuable information on some factors, which can be traced in students’ linguistic, cultural, familial, or educational backgrounds that can influence their achievements (Tsagari, 2004). Additionally, addressing the participants’ busy workload, limited time, and the large number of students in their classes would enable them to improve further learning by adjusting teaching methods and material through AAs. This would be similar to the TESOL teachers in Canada, who conducted their AAs as an integral part of their teaching (Cheng et al., 2008).

The participants’ belief in the benefits of AAs for enhancing learning, which they possibly obtained from their experience of implementing AAs, is quite convincing because specialised
knowledge on ESL/EFL teaching can be acquired by practical experience as well (Richards, 2009). Their belief in the benefits of AAs agrees with the function of assessment results in general, which give crucial implications for developing the quality of education (Wolf et al., 2008). This belief is also in line with AAs in which there could be more focus on strengths, progress, improving and guiding learning, and collaborative learning (Ataç, 2012). This can be in the framework of ‘assessment as learning’, which is to develop and support metacognition for students by focusing on their roles as critical connectors between assessment and learning (Earl & Katz, 2006). Consequently, learners’ involvement in classroom assessments increases their motivation and achievements (Arter, 2001). In this regard, several benefits of AAs for enhancing students’ learning of English language were perceived, experienced, or observed by the vast majority of the survey and interview participants of this study (see Section 5.5.2). This is in line with the findings of several researchers such as Ghahremani (2013), Ross (2005), Dunn and Mulvenon (2009), in addition to Black and Wiliam (1998, cited in Earl & Katz, 2006), who synthesised more than 250 studies whose conclusions indicated that the AAs designed to promote learning improved students’ achievements. In the same dimension, the belief of the participants’ of this study that AA classroom communications help students to master English language skills supports the findings of several studies, including those of Ghahremani and Azarizad (2013), Javaherbakhsh (2010), Grami (2010), Wakabayashi (2008), Burkšaitienė and Teresevičienė (2008), and Roskams (1999), who all examined the enhancement of productive skills by AAs. It also matches Birjandi and Mahnaz (2014), Ghahremani (2013), Shabani (2012), and Ajideh and Nourdad (2012), all of whom investigated the improvement of the receptive skills through the performance of AAs.

The interviewees’ belief that English language learning needs the performance of a variety of AAs to reinforce students’ learning concurs with one of the most important features of AA mentioned in the literature: diverse practices to accommodate different students (see Section 3.6.3). This also adheres to the recommendation of TESOL International Association (2010), which advocates the use of a range of performance-based assessment techniques. This goes partially with the recommendations of Troudi et al. (2009) and Shohamy et al. (2008), who discovered that teachers believed in the use of multiple assessment practices as fundamental to ensure valid and fair outcomes.
The interviewees’ belief that students enjoy performing AAs as it gets them involved might be related to linking various AAs to real world contexts. This is essential because a major goal of AAs, as Huerta-Macias (1995:9, cited in Tannenbaum, 1996) explains, is that they “gather evidence about how students are approaching, processing, and completing real-life tasks in a particular domain”. Authentic and real-life activities with the process-oriented nature of AAs are the two features that make such assessments superior (Chirimbu, 2013). More broadly, this commits to the major aim of learning, which is being capable of applying skills and knowledge to perform meaningful tasks in authentic contexts (Grabin, 2009). Simulated real-life activities of AAs might also contribute to making students more interested, because across the many studies reviewed by Raymond (2008), the performance of interesting activities was one of the common prescriptions for sustaining students’ attention and interest.

Conducting a variety of AAs may also lead to a fairer, more comprehensive, accurate, and valid assessment of students’ knowledge and performance (see Section 3.6.4). Most interviewees of this study believed that the accuracy of assessment could be achieved by using various AAs on which to base assessment decisions. This can be accomplished better by training teachers to use clear criteria, and through the triangulation of decision-making processes from those different sources of data (Tsagari, 2004). Regarding this, the participants of this study agreed that AAs allow their EFL students to show what they know and what they can perform. The participants’ belief that AA might be a more direct and clear method to display students’ knowledge and performances is also mentioned in Tsagari (2004), who asserts that AAs are valid in themselves and of themselves because of the direct nature of these assessments. The beliefs of the participants of this study are also in agreement with the finding of Shohamy et al. (2008), which suggests teacher participants believed that the construct of the complex advanced language proficiency (ALP) could be assessed only by using multiple assessment procedures of AA. Overall, this follows AA that features multiple referencing, that is, it obtains information about learners from a number of sources through various means (Hamayan, 1995), which are fundamental because virtually all language education contexts require a variety of AAs to meet their inferential demands (Norris et al., 2002). Such a variety of AAs includes self-assessment and peer-assessment, which are indispensable when demonstrating students’ knowledge and performance more accurately. This was discovered by Naeini and Duvall (2012), Javaherbakhsh (2010), and Anusienė et al. (2007). Matsuno (2009) adds that peer-raters have the potential to make significant
contributions to the whole assessment process. Despite the value and vitality of these two types of AA, the participants of this current study need to do more to embrace them rather than emphasising student subjectivity and low level as challenges to self-assessment and peer-assessment. This also might be because some EFL teachers may not be aware of the value and potential of peer-assessment for their students’ learning, as found by Yu (2013).

Regarding the ethicality of assessment through the performance of multiple AAs, the interviewees believed that students have the right to perform many kinds of AA in order to demonstrate their knowledge and abilities, and subsequently to achieve a fair assessment. This supports the studies of both Troudi et al. (2009) and Shohamy et al. (2008), who found that teachers believed that applying various assessment practices is fundamental to guaranteeing valid and fair outcomes. In fact, this also concurs with critical language pedagogy that focuses on learning outcomes that can be assessed in different ways, and even in different ways for different learners (Dragemark Oscarson, 2009). The fairness might be through the ability of students or people to have more than a single opportunity to show their capacities (Hancock, 1994). This also commits to multiple intelligences theory, according to which students need to be offered a range of opportunities to exhibit their knowledge and skills (Grabin, 2009).

With regard to traditional testing, the participants of this study believed that testing could be generally important, beneficial, accurate, and improve learning. As such, they agreed that traditional testing has good impacts in terms of the three core purposes of assessment: ‘assessment for learning’, ‘assessment as learning’, and ‘assessment of learning’ (Earl & Katz, 2006). Hence, the interviewees called for an integration of AA and testing as parts of a complementary process of assessment, which concurs with the scope of assessment that Schmitt (2002) presented. Schmitt explains that assessment is used in a broader sense, which includes both formal measurement tools and other kinds of qualitative assessments. This also supports the findings of Troudi et al. (2009), who observed that teacher participants recognised the role of traditional testing but believed students should benefit from various assessment instruments.

There are some benefits of AA related to students’ roles as assessors too. As a starting point, the attempts to overcome the limitations of teachers’ assessments as single assessors increased interest in AA in the field of education (Matsuno, 2009). That probably led to
Finch’s (2002) clarification: AA is particularly suitable for tertiary students because it encourages them to assume responsibility for their learning and the assessment of their learning, “as the cycle of intention, action and reflection becomes a personal endeavour, facilitated by: portfolios, projects, self- and peer-assessment, learning conversations, and reflective journals” (Finch, 2002:5).

Regarding this, the interviewees of this study believed that teachers are not the only assessors, and students can be assessors of themselves, but they seldom let their students do self-assessment. The participants need to improve this because self-assessment is highly advantageous for students, as concluded by several researchers such as Azarnoosh (2013), Khonbi and Sadeghi (2013), Lim (2007), Naeini and Duvall (2012), Javaherbakhsh (2010), and Anusiené et al. (2007). The interviewees did permit their students to perform peer-assessment but infrequently, which is perhaps because students may have a less positive attitude towards assessing their peers, as shown by Cheng and Warren (2005). More contextually, it might be because most interviewees of this study believed that students’ subjectivity, their low level, their personal relationships, or their propensity to take issues personally might prevent them from performing self-assessment and peer-assessment objectively and appropriately. Roskams (1999) unveiled some teacher-perceived obstacles of peer-assessment in Asian cultures, finding that peer feedback is viewed as constrained by the fear of mistakes, politeness norms, and a perceived lack of credibility. These challenges potentially can limit the function of socio-cultural ZPD, which is the cognitive development through language-mediated activities in interaction with people who have more advanced cognitive abilities, such as teachers or more able peers (Vygotsky, 1978, cited in Ishihara, 2009). As such, collaboration among peers creates a collective ZPD so that learners can take benefit equally (Turuk, 2008), which is significant because Vygotsky’s works focused more on the effectiveness of discourse with others and language mediation to reach a higher level of learning (Dragemark Oscarson, 2009).

Nevertheless, the interviewees believed that the adoption of AAs, especially self-assessment and peer-assessment, might help their students to become more confident, critical thinkers, and autonomous learners, which supports the findings of many researchers like Azarnoosh (2013), Sharifi and Hassaskhah (2011), Grami (2010), Bahous (2008), and Lucas (2007). These researchers discovered that students’ attitudes to portfolio and peer-assessment were positive as they helped them enhance their reflective and critical thinking. Furthermore, some
researchers such as Azarnoosh (2013), Nezakatgoo (2011), Grami (2010), Javaheerbakhsh (2010), Burksaitienë and Teresevičienë (2008), Lucas (2007), and Roskams (1999) revealed that several types of AAs lead to students’ autonomous learning, as they encourage students to take responsibility for their learning and address their linguistic areas in need of development. In addition, the recent reforms of the MHESR/KRG Quality Assurance addressed how to increase students’ abilities and support them to become more critical and autonomous learners (Palander, 2013).

In broad terms, the advantages of AAs for learning that the participants of this study believed in are in line with the two main features of the underpinning theories of AA. These are, firstly, the focus on the social aspect of learning in a student-centred type of learning, whereby students are given their active roles and are empowered, and, secondly, the increasing of the accuracy, validity, and ethicality of teaching, learning, and assessment. The validity and fairness of AAs that the participants believed in were mainly related to the rights of students to choose and perform a variety of AAs. This commits to the aspirations of CALx, critical pedagogy, and critical language testing for giving rights to students to encourage them to become more active and involved in the decision-making process. The most important and also the initial point of the reforms in assessments was possibly the focus on students’ roles, which was the transition from the behaviourist theories to the constructivist and cognitivist theories of learning, and which led to a new concept of learning in which the learner’s pivotal role is recognised (Grabin, 2009). The beliefs of the participants of this research about the benefits of AA for giving students an active role were concerned mostly with the wide range of AAs that students can perform, especially those that can be completed by students themselves, like self-assessment and peer-assessment. This particularly commits to sociocultural theory, constructivism, and multiple intelligences theory, in which students are active by doing various activities (see Section 3.5.3).

As a consequence of the first-phase findings, by which the action research plan for intervention can be tried out (Mertler, 2013; Troudi & Riley, 1996), I realised that the teacher participants implemented AAs in the first place, although some aspects needed more development. The participants were aware of the benefits of AAs and they believed in them, but their knowledge of AA needed to be increased. This indicated that an introduction for teachers to AA would help them increase their knowledge about the necessary requirements
of performing such assessment. I believe that this would be a crucial first step to improving the usage of AAs.

All the benefits of AAs that the participants of this study believed in can increase their feasibility, but the numerous challenges will inevitably make such assessments impractical if not addressed. I understood that it would be much more credible to discuss the practicality issues of AAs with the interviewees after the intervention, by which I could inform them about what exactly is AA and its core demands, criteria, time duration, and continual adjustments. That was the end of the first phase, and the findings were used in preparation for a second intervention phase, in which I presented the two seminars in the two selected English departments.

5.7 The intervention outcome: Introducing participants to alternative assessment

This main section is devoted to explicating the intervention outcome, and explores whether the two seminars succeeded in introducing the participants to AA, in increasing their knowledge of AA, in changing their beliefs on AA, and, subsequently, whether the seminars could enhance the participants’ willingness to pay greater attention to AAs in the future. The aim was to provide answers to the fourth research question, which belonged to the post-intervention phase of this action research, and which explored the degree to which the in-service Kurdish tertiary TESOL teachers’ cognitions of AA changed after the intervention.

Regarding the nature of the intervention, the first-phase questionnaires and interviews can be considered part of the intervention because some participants heard about ‘alternative assessment’ as a term for the first time, after which they began searching for it, as Mardan said: “When I heard this concept, I started reading about AA just to elaborate my knowledge”. Goran added: “Right now, I am working and reading some literature on AA and assessment in general”. However, for the evaluation of the catalytic validity of this action research, I examined only the influence of the two seminars on the teacher participants’ cognitions of AA.

During the seminars (see Appendix 6), I attempted to briefly explain the numerous points about AA, which included: (1) definitions of AA; (2) its relevant concepts; (3) its underpinning theories; (4) its core requirements, such as criteria in the form of a rubric or a
checklist; (5) its continuity to assess students’ language growth over time, when integrated into teaching and learning; (6) its frequency; (7) its need for technology to facilitate the implementation of AAs; (8) its need for objectivity to increase the validity and ethicality of AAs; and (9) its need for more marks, for which I suggested a 50-50 divide between AAs and tests. The content of the seminars was based on the data analysis of the first phase, which showed that the participants needed more knowledge about the above issues. Therefore, I planned to inform them about these AA issues to enable them to have a clearer picture of the assessment system of their English departments in order to critique it and discuss the practicality issues of AAs after this intervention phase. These practicality issues concerned to what extent AAs are feasible, and what challenges they expect the implementation of such assessments will face. The post-intervention discussions were designed to find solutions and give recommendations to tackle the challenges, in order to enhance the performance of AAs.

Following the presentation of the two seminars in the two English departments, reports from all 12 attendants indicated that there was an increase of their knowledge of AA, and that their beliefs regarding the conducting of such assessment changed, indicating a willingness to pay greater attention to AAs in the future. In the answers to the question of whether the attendants’ knowledge about AA increased, all the 12 participants stated that their knowledge of AA had been raised to a good extent, as Dildar acknowledged: “I think it increased my knowledge about assessment and testing, actually, and encouraged me to look for more information about these topics”. Mardan agreed with Dildar, saying: “I have learnt some new concepts and the idea of presenting the seminar was something good”. Ala added: “It gave me some small details that I do them naturally, and I didn’t know there are actually studies [about them]... It increased my knowledge about this issue [of AA]”. Concerning the implementation of AAs, Ashty said that the seminar was beneficial “in terms of what are the AAs, kinds of AAs, and how can we use it in class”. Naska agreed with Ashty, saying: “The seminar helped me in how to apply it [AA] in a more active and better way”. Aram had a broader view about the seminar:

Through the seminar, I realised that we lack that AA in the application. I think, partly concerning the references, the people you mentioned in the theory part, as well as some of the content, actually, it was more of very interesting or useful (Aram, Int3).

Additionally, the participants’ knowledge of the differences between AA and testing, and the three purposes of AA, were ultimately higher than before the intervention. However, Ara
believed that the seminar gave [her] “some kind of knowledge about assessment”, but “not too deep, it was just a rough introduction”.

Concerning whether the participants’ beliefs about AA changed following the intervention, almost all the attendants (11 of 12) confirmed that their beliefs were changed after they attended the two seminars. After the intervention, they believed that they should use more AAs by devoting more effort and allocating more time to them. They were eager to pay more attention to AA, and expressed their willingness to incorporate more AAs into their classroom teaching and learning. Nine of those 12 participants called for some changes in the assessment system of their English departments in order to allow more room to implement AAs. They also called for addressing the challenges of AAs, especially the number of students per class and the limited time.

Regarding the influence of the intervention on changing the beliefs of the participants, Goran’s statement can be taken as typical:

I think [AA is] not only important but it is essential. It should be an integral part of our assessment system. Not only I, but I advise my colleagues also to follow suit; reading about it is not enough, applying, implementing these assessments or procedures in the next course, or maybe the next academic year is a must. Whether the department will hold training courses for the teachers, or maybe a series of seminars or forums... The teachers must be instructed about this, and should be shown the importance of the application of these assessments. It is a flaw, it is a weakness, it is a serious weakness (Goran, Int3).

Ashty also demonstrated her eagerness to follow AA to achieve more valid, ethical, and democratic assessment:

I wish that all the teachers were applying this system here because actually it is a necessity. We don’t have to rely on just paper-based [tests]. We have to rely on their [students’] performances, their attitudes, their personality, their knowledge..., everything, communication, how well they are. So here, we have to depend on all these tools [AAs] (Ashty, Int3).

Aram also showed his enthusiasm to implement AA and its methods of marking:

Now, I think we should do more in order to increase it... Generally speaking, we need to make more effort; we need to read more about it; we need to learn the standards of doing it; we need to learn how to design rubrics in order to do it; and we need to push our students to get involved with learning. I had an idea about it, but now I think we should do more in order to
increase it, because I think the measurement of the performance, as you mentioned in the seminar, is very important (Aram, Int3).

Additionally, Dara admitted that his beliefs had changed after the seminar, and revealed his readiness to perform AA more widely:

Actually, my view toward that point [AA], and also toward the whole process of learning and teaching in the English department has changed. I will do my best to change the style of my teaching and using this method [AA] (Dara, Int3).

It is noteworthy that the heads of the two English departments were among the 12 interviewees, and they were most eager and promised to develop the implementation of AAs. Throughout the study period, I was able to meet almost all the heads of the 18 English departments in which I distributed the survey questionnaires, who were happy to discuss with me their support for the use of AAs. This indicates that the impact of the intervention might lead to real developments in the English departments of the universities of the Kurdistan Region, especially after the recommendations of this study are disseminated.

Although increasing the participants’ knowledge of AA was successful to a good extent, as a researcher, an interviewer, and a seminar presenter, I perceived that my intervention was more successful in changing the attendants’ beliefs about performing AAs, as I observed the eagerness of most of them after the seminars. Besides, I contend that their opinions were quite justifiable regarding the necessity of taking further actions to tackle some of the contextual factors (see Sections 5.9.2, 5.9.3, and 5.9.4), because these challenging factors presented serious obstacles to the implementation of AAs.

However, one participant, Lara, stated that her beliefs had not changed after the seminar, saying: “I don’t think that your presentation can have a great effect on my presentation inside the class, because I have a time, which is specified, a limited time, so I can do whatever I do”. She seemed to be disappointed about the limited time, seeing it as a pressing challenge that does not allow any changes and improvements.

5.8 Discussion of the intervention outcome

The attempt to integrate both attitudinal and functional developments of teacher participants (Evans, 2008), through the presentation of two intervention seminars, was one of the basic
aims of this study, since teacher participants needed to be introduced to AA and acquire more knowledge about it. That was based on the nature of action research, which attempts to address particular groups with distinct needs (Troudi, 2006), and which concentrates on what works best with particular learners in particular contexts (Griffee, 2012). Essentially, doing action research on AA is quite contextual, and thus it might lead to positive results. Research shows that greater implementation of AAs would benefit students (see Section 3.6), thus teachers need support to enable this. Presenting the two seminars as a starting point for professional development and as an action research intervention was designed to increase the participants’ knowledge of AA to enable ultimately a better implementation of AAs. That was also to achieve the catalytic validity, which is a major criterion to ensure that an action research should suggest an agenda, give rise to some specific actions, and result in those actions (Cohen et al., 2007). In this regard, after the intervention, there was an increase in the participants’ knowledge of AA and substantial changes in their cognitions of the performance of AAs. The participants’ knowledge of the differences between AA and testing increased, which is important because assessment literacy involves teachers’ familiarity with both testing and assessment conceptualisations (Malone, 2013). Also, their knowledge of the three purposes of AA increased after the intervention, which is crucial as well, since TESOL teachers should know about the purposes of assessment (TESOL International Association, 2010).

Changing the participants’ beliefs enhanced their willingness to do more about AA in the future, as the intervention stimulated reflections on AAs, so that the participants can explore the feasibility and challenges of performing such practices in their professional context. This might have strengthened my action research intervention to integrate both attitudinal and functional developments, as parts of the teachers’ professional development (Evans, 2008). That was achieved by giving the participants access to new ideas (about AA) from academic researchers and trainers to implement in their classrooms (Griffee, 2012), because ultimately action research can help shape and refine assessment (Singh, 2006). Refining and reforming assessment could also be fulfilled by the discussion of the challenges of AA with the participants (after they were more familiar with AA and its requirements following the intervention) in order to find solutions for them and to give recommendations about them. Such recommendations would be effective because if those challenges are resolved, it would be a great step towards a more successful implementation of AA. So, after I disseminate the recommendations of this study to the 18 English departments, the outcome of the intervention
would be more influential, because most of the Heads of Departments that I met promised to work on developing AA. That is why I suggest that my intervention, to a good extent, achieved what I planned regarding the intended outcomes. As such, equipping some Kurdish tertiary TESOL teachers with introductory information on AA gives this study a pedagogical significance about the implementation of AAs.

Expanding teachers’ knowledge about AAs by presenting two seminars had a special place in the design of this action research. As an endeavour for professional development, this action concurs with the suggestions of Bailey and Wolf (2012), Troudi et al. (2009), and Brindley (2001), who argue that it is urgent to equip TESOL teachers with assessment literacy in order to broaden and improve their English language assessment practices. The great need for acquiring assessment knowledge by teachers is clarified by Coombe et al. (2012), who say that almost all teachers perform assessments but have not learned the principles of sound assessment. Taylor (2009) also suggests that many people are involved in assessment processes, but they have not got a background in assessment. Therefore, the efforts of this study to promote the assessment literacy of the participants can be potentially influential for the development of their AAs in particular, and the English language teaching, learning, and assessment in their English departments in general. Further, the intervention of this study might be a contribution to Waters’ (2009) statement that innovation, as an attempt to achieve beneficial change, has recently become a characteristic of ELT. Also, this intervention can be interpreted as a small part of the enormous efforts to resolve many problems in the field of ELT (Silver et al., 2002).

Nonetheless, time is a crucial factor in the development of assessment literacy, as teachers need enough time to obtain more knowledge and skills, and to implement meaningful and sustainable changes in classroom procedures (Lukin et al., 2004). However, as a starting point, I was able to help the teacher participants to get on the right track of reading about AA, developing their implementation of AAs, and taking advantage of them. Increasing the participants’ willingness to reflect on AA in order to use it more efficiently is critical because mastery of knowledge and skills might not comprise the whole professional development. This is because it may ignore the dynamics of the social construction of professionalism, because teachers’ sense of self, their beliefs, and their knowledge could be influenced by their previous life experiences (Shin, 2012).
During the seminars, I also briefly mentioned the underpinning theories of AA, including socio-cultural theory, constructivism, multiple intelligences, communicative approach, cognitivism, interpretivism, CALx, critical pedagogy, and critical language testing. That possibly encouraged the participants to search for those theories, which might help them to restructure their teaching and assessment practices. This is quite necessary because, as Carr and Kemmis (2003) state, teachers and educators must already have knowledge of some theories of education to guide them in restructuring their activities and making decisions.

### 5.9 Beliefs of Kurdish tertiary TESOL teachers on the feasibility and challenges of alternative assessments

Investigating teachers’ beliefs, which stem from their contextual experiences and practices in their classrooms, is a reasonable way to explore the practicality issues of AAs. This main section is designed to answer the fifth research question, which belongs to the post-intervention of this action research, and it explores the beliefs of in-service Kurdish tertiary TESOL teachers on the practicality issues of AAs, which include the feasibility and challenges of these assessments. For this question, data from all the research instruments was used, but the vast majority of it was from the third-phase 12 interviews and the two focus group interviews, during which some systemic, teacher-related, managerial, and technical challenges of AAs were discussed in detail. The theme of this main section is divided into four categories. The first category is about one feasibility side of AAs, which is related to decreasing the time of marking tests at home, while the other three categories concern the challenges of AAs that pertain to the assessment system of the English departments and their context, the teachers’ subjectivity, and the managerial and technical facilities.

#### 5.9.1 Feasibility of implementing alternative assessments

In the first-phase questionnaires, most survey participants (58: 64%) believed that AAs are quite or very practical (see Table 11).

<table>
<thead>
<tr>
<th>AAs will be practical for my EFL students more than testing.</th>
<th>Very practical</th>
<th>Quite practical</th>
<th>Unsure</th>
<th>Slightly practical</th>
<th>Impractical</th>
<th>Not answered</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>19=21%</td>
<td>39=43%</td>
<td>17=19%</td>
<td>13=14%</td>
<td>0</td>
<td>2=2.22</td>
</tr>
</tbody>
</table>

Table 11: Survey participants’ beliefs on the feasibility of alternative assessments
In the first phase, the beliefs of the participants about the practicality issues of AAs were about performing AAs but probably not about their standards and requirements, such as criteria of marking. That was because the participants needed to be informed about the standards, criteria, and requirements of AA (see Section 5.3.1). Consequently, I did not discuss these issues with the first-phase interviewees because I thought they should first be familiar with the demands of AA, such as rubrics and checklists, in order for them to acquire a clearer picture of their situations and the assessment system of their English departments, and within which to consider and discuss the practicality issues of AAs. That is why in the third phase, after the intervention, their beliefs were related to AAs including the consideration of their requirements. Therefore, they had concerns about the challenges of AA, in which only one aspect of feasibility was mentioned by them, whereas all other aspects were seen as challenges to AAs (see Sections 5.9.2, 5.9.3, and 5.9.4).

After the intervention, only decreasing the time of marking tests at home was identified by more than two thirds of the interviewees (9 of 12) as a feasibility aspect of AAs in the context of this study, because doing a lot of marking at home was a concern. For instance, Goran marked 80 test papers in five days, Dara marked 126 tests in more than a week, Lara marked 78 in 10 days, and Naska needed 50 days to correct all the mistakes. As such, more class time and marks need to be devoted to AAs to relieve teachers from the pressure of marking traditional tests at home. For this, Aram said: “The more activities you have in the class, the less time you spend grading and giving marks [of test papers]. ...It will be not only saving time, it is gonna be very beneficial”. Ara added: “It is going to be easier for the teacher, not all of the marks for the [test] questions”. Nonetheless, Naska, Ara, and Nazanin believed that written AA assignments also take too long to assess at home, as Naska clarified:

It depends on the type of the assessment that you are doing. In the majority of the cases, assessment [AA] increases my job at home, because I usually ask for a report or class reading, open-ended questions. I usually receive the answer in the form of a paper, and I have to take it home and read it (Naska, Int3).

5.9.2 Challenge (1): The assessment system of the English departments and their context

Because of teachers’ major role in teaching, learning, and assessment, their cognitions of the challenges of performing AAs are effective in diagnosing and addressing them. To this end, during the 12 interviews and the two focus group interviews of the third phase, the
interviewees discussed in detail several challenges that they had encountered, and will possibly face again in the future if they implement AAs more widely. Therefore, the challenges of AAs will possibly make them impractical, despite having numerous benefits that will increase their feasibility in some aspects. The participants’ beliefs about the advantages of AAs signified a hopeful future for a more successful implementation of them, because they believed in the benefits of AAs for teaching, for learning, and for increasing the accuracy of assessing students’ English language proficiency. However, regarding the challenges of performing AAs, I sensed an air of disappointment. The participants believed that some challenges related to their assessment system, their subjectivity, management issues, and technology facilities cannot be easily solved. For example, Ara expressed her disappointment:

The most important thing about AA is that you give freedom, you give space, and you give values to the students’ activities day by day. ...This is generally true... but this is an imaginary one or a dreamy one... When you come to reality, you face so many difficulties, and if these problems are solved, these challenges are met, I think it is a good step for a teacher to practice [AA]... The idea is excellent but I think it is far reaching because you cannot control all the challenges (Ara, Int3).

Ara’s general concern echoed the majority of the participants’ concerns about AAs. Even in the first phase, most survey participants (54: 60%) believed that performing AAs would be quite or very challenging (see Table 12).

<table>
<thead>
<tr>
<th>AAs will be more challenging for my EFL students than testing.</th>
<th>Very challenging</th>
<th>Quite challenging</th>
<th>Unsure</th>
<th>Slightly challenging</th>
<th>Unchallenging</th>
<th>Not answered</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>16=18%</td>
<td>38=42%</td>
<td>13=14%</td>
<td>21=23%</td>
<td>2=2%</td>
<td>0=0%</td>
</tr>
</tbody>
</table>

Table 12: Survey participants’ beliefs on the challenges of alternative assessments

More than two thirds of the interviewees (9 of 12) mentioned several contextual factors that can make AAs more difficult in the context of their English departments. One of the factors was their assessment system, which was principally based on the traditional teaching, learning, and assessment methods. The participants’ statements generally denoted that this system might have prevented the improvement of teaching, learning, and assessment procedures in their English departments. More specifically, several participants directly or indirectly blamed their assessment system for the limited performance of AAs, such as
Mardan who said: “The problem is not with the teachers. We have something called system”. Ara added: “Sometimes, it is not you who can do that, it is the overall system”. Ala confirmed: “I feel that the system will not let us [do AA more widely]”. Therefore, the first challenge that hindered the successful performance of AAs could be the assessment system of the English departments of the universities of the Kurdistan Region.

This system follows a traditional system of lecturing style and testing that usually encourages memorisation and spoon-feeding, and thus leads students to be test-oriented because the major part of assessment is completed through tests with around 90% of the marks (see Section 5.1.2). Consequently, students do not focus on skills, but rather they focus on knowledge, and so do not need any daily studying and practices. Instead, they can study prior to exams, as Nazanin observed: “I think, here, no one cares about these assessments nowadays, because we only depend on tests”. Goran agreed with Nazanin, saying: “Multiple assessment results are more accurate to make decisions; although, currently this is not the case. Currently, the decisions mainly are out of the two main tests”.

The assessment system of these English departments is determined chiefly by top-down technical guidelines imposed on their teachers about the frequency, intervals and duration of exams, number of questions, and number of marks for each test. Naska explained: “For testing, the general purpose is to grade the student, and this is something a requirement, it is kind of imposed by the department and the policy of the university”. Teachers predominantly relied on traditional testing because of the prioritisation of testing over AA in their departments. Having first and second term tests with 20-40% of marks and the final exam with 60% of marks encourages focusing on knowledge of English language rather than its four language skills (see Section 5.1.2). As such, those departments are possibly test-oriented that have made teachers and students become test-oriented and certificate-oriented, because of the inequality between testing and AA. The participants perceived this situation as a result of allocating a few marks to AAs, which are not as compulsory as tests (see Section 5.1.2). This is regardless of the fact that the majority of the interviewees (10 of 12) believed in the significance of both AAs and tests, and they called for an integration of them by focusing on both equally in order to achieve a complementary student assessment process.

The majority of the interviewees (10 of 12) were willing to change the system of teaching, learning, and assessment from a teacher-centred to student-centred approach. To achieve this,
they wanted to implement more AAs in order to provide more valid and fairer assessment for their students; this is in case teachers and students are not subjective anymore, as Ashty expressed in her willingness: “I wanna change the system of teaching to be the student-centred, not the teacher-centred”. Mardan further clarified:

If I have authority, I will change the whole grading system in English department because we have to work on four language skills. We have to focus to assess the students’ listening, speaking, reading, and writing... but because of this system, listening and reading are ignored (Mardan, Int3).

This might be because of teachers’ feeling that they cannot be fair by doing most of the assessment through tests, as Lara explained:

Sometimes, I have a student who is very, very good in a daily activity, in assessment, but when it comes to the paper exam, they have problems. So, the teacher will be in a conflict how to assign the mark to that student in daily... Nowadays, we have no equal marks for [tests and AAs], but if these apply, like this, I will prefer (Lara, Int3).

5.9.3 Challenge (2): Teachers and students’ subjectivity in marking alternative assessments

The Kurdish tertiary TESOL teachers’ acceptance of the prescribed rules of testing from their English departments is probably because they find them suitable to eliminate some contextual factors related to their own and their students’ subjectivity in marking AAs. Almost all interviewees (11 of 12) believed that students were subjective in doing self-assessment and peer-assessment because of their social relationships and the tendency not to show their own and their friends’ mistakes, in order to avoid losing marks. “If they find that it is their friend..., they are not going to assess that, they are just saying it is correct” (Darawan, Int3), because “they don’t wanna spoil the relationship with their peers” (Ashty, Int3). Concerning teachers’ subjectivity, all interviewees (12 of 12) believed that teachers were subjective in marking AAs due to some psychological and emotional reasons, such as teachers’ feelings during assessment, including sympathising with or being intolerant of some students. They considered that the validity and ethicality of marking AAs were sometimes dependent on the teachers’ mood, and maybe sometimes on the nature of the relationships between a teacher and an individual student. Some teachers, such as Darawan, Aram, and Naska, frankly stated that they tried hard not to be subjective, but they still struggled to remain objective. Perhaps
that is why many students complained about the marks given to them for AAs. The head of one of the two English departments attempted to eliminate teacher subjectivity in marking students’ speaking skill (which is completely through AA) by forming a committee of three teachers instead of one teacher, as it had previously been. It is likely that because of teachers’ subjectivity in AA, half of the interviewees (6 of 12) believed that testing was more objective and fairer to students, by which subjectivity can be avoided, since test papers have everything recorded. For this, Darawan said: “It [AA] is not an objective thing because it depends on your mood at that time... but if it is a testing thing, for example, there are blanks or just ticking”. Naska confirmed: “I am always trying to avoid that sort of assessment [AA]. I try to depend on papers... because of matters of subjectivity”.

Concerning the factors behind teachers’ subjectivity, Aram, Ashty, and Goran related it to the unavailability of criteria for AAs, which is possibly due to not using the marking techniques of AAs such as checklists and rubrics that can be employed as criteria. Aram explained: “If you do not have some standards, some rubrics or some criteria, usually personal feeling would involve”. Most interviewees admitted that they had only a simple list for ticking students’ participations in AAs (see Section 5.1.2), although Aram, Naska, and Goran believed that their subjectivity was also influenced by some cultural factors. In this regard, Aram said: “I would say this is partly related to culture. Culturally, we are not brought up in an objective way. That is, we mix our feelings with the things that we do”. Naska agreed with Aram, saying: “I guess, right now in our culture, that [AA] will not be that easily achieved because of matters of subjectivity”. These beliefs indicate that in their culture, people probably mix their relationships, emotions, and feelings with whatever they do, which might lead to subjectivity.

During the focus group interviews, after I had suggested a redistribution of marks to 50 for traditional tests and 50 for AAs, there were strong concerns about the potential for subjective marking, and therefore, the majority of the attendants disagreed with me. Naska justified her concern, saying:

No matter how hard you try, still teachers are subjective... I think changing right now [to 50-50 redistribution of marks] will not be a good idea. I think teachers should be acquainted with the notion of objectivity... Unless that is achieved, I do not recommend any change. Some of them [AAs], especially oral ones, teachers cannot avoid being sometimes subjective. Personally, I do not claim that I am always objective in grading students. We, as teachers, do not have a very
comprehensive understanding of the term ‘being objective.’ We need a lot of workshops; we need a lot of practices, and we need assistants to supervisors... to be sure that our marking is totally objective. Unless that is fulfilled, there is no need to do that [awarding AA 50 marks]. If the department does not check how I mark the students, and how I put my rubrics, and how I do the checklist, then I can do it the way I like (Naska, Foc3).

As I noticed, most of the focus group interviewees agreed with Naska’s rejection of my suggestion. They believed that they need much more professional development and supervision to become objective in using AA, largely because of their need to be more familiar with the concept of ‘being objective’. It is likely that this stemmed from cultural factors: “Culturally we are not brought up in an objective way” (Aram, Int3).

5.9.4 Challenge (3): Management and technology

The managerial and technical challenges that the interviewees believed they were hindering obstacles to AAs were the large number of students per class, the limited time for teaching, and the costs of providing and maintaining classroom technological facilities. Perhaps because of those three challenges, the interviewees considered AAs to be very difficult. Concerning the number of students and the limited lecture time, all interviewees (12 of 12) believed that a large number of students in their classrooms made the implementation of AAs considerably challenging, because they did not have enough time during their lectures to give an opportunity to all students to perform AAs individually. In this regard, Dildar expressed his concern: “We have lots of students in class, [and so] how can you assess all of them?” Goran explained: “It is time-consuming. Listening to so many students, talking and answering questions, anything, debate, conversation... Time management will be a problem”. Ara clarified more:

There is a large number of students in the class and you are limited with a period of time, that is, 45 minutes... You can’t manage it and it is very difficult... One of this is time, time management... You can’t give the role to every single student to have the assessment because of the large number of students (Ara, Int3).

In the two focus group interviews, I observed that most of the attendants nodded their heads to show agreement with Naska and Ara’s opinions that “because students apply only for the Kurdistan Region universities, we will have crowded classes continuously” (Ara, Foc3, TFK), and that “this is part of our reality, we have to deal with crowded classes” (Naska,
The challenge of the high number of students is mostly associated with the nature of the practice-oriented lectures/classes that constitute most sessions in the English departments. In the focus group interviews, I recommended dividing large classes to have fewer students in each, but that was deemed impractical because of the limited number of teachers and classrooms.

Regarding the limited time, some interviewees (4 of 12) believed that the holidays and skipping classes by students made their time much more limited. Many focus group interviewees agreed with Ala and Naska when they presented their concerns about these time issues. Ala revealed:

First-year students start in January [losing nearly three months already]... The issue is that because we lose a lot of lectures because of the holidays, weekends, [and] the students sometimes don’t show. So, this will minimise the time that we meet the students. We need to have systematic weekends and holidays... You have a certain syllabus, [and] suddenly, comes a week of holidays and all my system just crashed... (Ala, Foc3).

Regarding the provision of technology for implementing AAs, all interviewees (12 of 12) believed that providing and maintaining technological facilities for teaching, learning, and assessment required improvement (e.g., internet, computers, data shows, projectors, copy machines, and TVs), since in some classrooms, no technology instruments were provided or they did not work. In this regard, Nazanin said, “If the time allows us, [and] if we have tools, we can do that [AA], but unfortunately we do not have any tools in order to help us to do that”. Ashty added: “the facilities that we don’t have, unfortunately.” Also, these interviewees believed that providing and maintaining such equipment is costly, which may constitute a constraining challenge due to the current economic crisis in the Kurdistan Region. The provision of classroom technologies, their maintenance, and training teachers on how to utilise them are quite urgent for conducting AAs because some of these technologies are quite practical for large classes (see Section 5.10).

Given these systemic, affective, managerial, technical, and financial challenges in the context of this study, a successful implementation of AAs would involve more effort than in other contexts. In fact, all interview participants (12 of 12) believed that AAs required much work, effort, and energy from them, since AA included understanding and evaluating many aspects of students through various means. They also believed that performing AAs was unlike
testing, which was much less frequent and easier. That is why some participants admitted that they could not always mark students’ AAs (see Section 5.1.2), and Darawan, Mardan, and Naska confirmed that daily markings of such assessments were very difficult, while tests were easy. Naska said: “Actually, testing is much easier because it is done only once or twice a year..., which is to do everything at that time. Actually, assessment is time-consuming; it is a kind of burden for the teacher”. Dildar agreed with Naska: “More time and more effort [is needed] especially in EFL classes... We don’t have enough time to do marking in the classroom”. Aram added:

One challenge might be the effort; probably it takes a lot of time getting prepared. Honestly, it is not like a traditional one; you go just put some questions there, and then students go and answer. No, [for AA] you need to have a continuous plan (Aram, Int3).

I contend that the above-mentioned challenges of AAs are the most pressing ones. However, there are possibly some other hindering factors related to teachers, students, management, or technology that can be further investigated in future studies.

5.10 Discussion of the beliefs of Kurdish tertiary TESOL teachers on the feasibility and challenges of alternative assessments

Because teachers play a major role in assessment, their cognitions are highly influential in diagnosing the obstacles to AAs and in finding feasible ways to perform them. Feasibility includes some determining factors such as the availability of time, personnel, equipment, space, implementation funding (Malone, 2011), maintenance costs, resources, ease of marking, suitable/trained graders, and administrative logistics (Coombe et al., 2007). After the intervention, most interviewees believed in only one positive aspect of the feasibility of AAs, which was decreasing the time of marking tests at home, since they can do half of the assessment in the classroom with AAs. The interviewees’ understanding of that seems to be right because the length of time for doing and marking tests and AAs might be approximately the same. That was discovered by Cheng et al. (2004), who showed that the greater use of assessment methods (AAs) in Hong Kong and Canada leads to more marking time, which might be the same as the greater amount of time required to process the objectively marked assessments (tests) in Beijing (Cheng et al., 2004). However, this aspect of feasibility was rejected by three interviewees of this current study, who believed that written AA assignments also took a long time to mark at home. In contrast, Stiggins (2002, cited in
Grabin, 2009) argues that AAs allow teachers to save time when they develop and utilise those assessments more efficiently.

Regarding the general challenges of AAs, Law and Eckes (2007) mention that such assessments might require a considerable amount of thought, space, logistics, and training. Maybe for this reason, most participants of this study believed that the performance of AAs would be generally challenging, which might also concur with the findings of Yu (2013) and Wach (2012), who observed that teachers’ attitudes to AAs were positive and thus they appreciated them, but implementation nevertheless posed a major challenge. The challenges perceived or experienced by the participants of this study are contextual, and if efforts are made to resolve them, they can be a basis for improving ELT in the English departments. This is because “language teaching can be seen as a process which is defined by dynamic interactions among cognition, context and experience” (Borg, 2006:275).

The majority of the interviewees believed that the assessment system of their English departments could benefit from development, as testing is currently prioritised over AA and students therefore tend to be test-oriented. This situation might be global, with Ataç (2012) suggesting that until recently, assessment processes were based primarily on standardised tests, which still cover a significant portion of assessments. Regarding the reliance on tests, the situation of the Kurdistan Region’s English departments is perhaps related to the large number of students per class, which is similar to the case observed by Cheng et al. (2008) in China, as this probably leads to the use of selection assessment methods (tests) because they can be easily marked. Alternatively, it might be similar to the ‘product orientation’ of the English departments in some Asian contexts such as Vietnam and Thailand (Saengbon, 2006, cited in Fitzpatrick, 2011), where communicative language teaching does not always function well. Those English departments that are product-oriented cannot implement communicative approaches successfully, and thus AAs cannot be performed well during the language communication exercises, since they only need tests to assess the product of students’ learning rather than its process. As such, even if language assessment practitioners wish to adhere to a code, they might face working conditions (as contextual challenges) that could prevent following the items of that code (Jia, 2009).

However, most survey and interview participants believed in the significance of both AAs and tests, and thus most interviewees called for an integration of them by incorporating more
AAs into their teaching methods, in order to achieve a student-centred education. Incorporating more AAs means that they attempt to eliminate the contextual challenging factors of teaching, learning, and assessment because AA takes learners’ social, academic, and physical contexts into consideration, and assesses learners in a variety of natural settings (Hamayan, 1995). For this, AA has an array of measures that can be adapted for different contexts (Tannenbaum, 1996), and also it has the great advantage of being context-specific, whereby assessment tools are adapted to best reflect the purposes of learning (Chirimbu, 2013).

Another contextual challenge that the interviewees believed in is teachers’ and students’ subjectivity in doing and marking AAs, which usually results in overestimation or underestimation. Students’ overestimation and underestimation in self-assessment and peer-assessment as a result of their subjectivity were also discovered in some other studies, including those of Khonbi and Sadeghi (2013), Chen (2008), Suzuki (2009), and Boud and Falchikov (1989, cited in Chen, 2008). This can lead teachers to distrust their students, which in turn challenges the essence of AA’s validity and ethicality. AA offers a potentially more ethical, democratic, and equitable approach if the assessment outcomes value individual diversity (May & Hornberger, 2008), which might help practitioners to address ethicality goals (Lynch, 1997). However, this might be under question because all interviewees of this study expressed the belief that teachers need to be more fair when marking AAs due to their subjectivity. This supports some findings in the literature about the ethicality shortcomings of AA. For instance, the marking of AAs, which is based on learners’ observed performance, raises equity concerns (Baker, 2010) because it involves subjective scoring (Brown & Hudson, 1998). Subjectivity leads to an unfair assessment of students, and fairness is crucial in teaching. Raymond’s (2008) teacher participants from the science and English faculties situated fairness as the third most prominent effective teaching characteristic. Subjectivity also compromises the quality of teaching; Raymond (2008) reviewed 10 studies that discussed the significance of objectivity to excellent teaching methods. In the focus group interviews of my study, teachers insisted that they must learn how to be objective in order to eliminate their subjectivity and to avoid unfair marking of AAs. Some interviewees related their subjectivity to the unavailability of criteria for AAs. The marking techniques of AAs like checklists and rubrics can be used as multi-dimensional criteria in order to eliminate teachers’ subjectivity in marking AAs. Raymond (2008) explains that when students complain, they should know that they are assessed by a higher standard of evaluation, which
is set by an internal system based on a rubric that cannot be adjusted through external pressure.

In the context of this current study, the teachers’ desire to be more objective in conducting AAs, in order to ensure validity and ethicality, is also a cultural issue, as three interviewees believed that their culture influences them. They believed that in their culture, people probably mix their emotions with work, which creates subjectivity; hence, teachers possibly mix their emotions with assessment. I believe that this situation of AA will lead us to reconsider the cultural challenges of AA more thoroughly. In some cultural contexts, teachers intending to use AAs may need intensive teacher training courses to ensure valid and ethical implementation, in order to avoid the cultural influence that leads to subjectivity. I could find only a few studies that mention the subjectivity of teachers in AAs (e.g., Baker, 2010; Brown & Hudson, 1998; Raymond, 2008). Further, a wide search of the extant literature failed to yield any study that deals with teacher subjectivity culturally, though this issue is essential for valid, ethical, and successful implementation of AA.

Perhaps because of the teachers’ tendency to be subjective in AA, half of the interviewees believed that traditional testing is more objective and fairer to students, despite their awareness of the importance of AA and the drawbacks of testing. I contend that considering some basic principles of testing that are fundamental in classroom assessments is necessary for eliminating subjectivity in assessment, AA, and testing. These principles include reliability, validity, reference-points, record-keeping (Earl & Katz, 2006), fairness, and washback (The EALTA Executive Committee, 2006), which might enable assessors to employ a variety of assessment measures with minimal bias (Newfields, 2006). To avoid further bias, safeguards should be put in place, since the potential bias needs to be assessed (Baker, 2010).

Regarding the high number of students per class, all the interviewees of this study deemed it to be a big challenge to AA, primarily because their class time is too limited to give all students the opportunity to perform AAs. This supports Tsagari’s (2004) finding that high numbers of students in classes and limited time pose huge obstacles to the implementation of AAs. This may be because in large classes, each student’s initiatives are difficult to arrange and assess (Bruton, 1999). In the English department of Diyala University, which is very near to the Kurdistan Region, Abbas (2012) also concluded that the limited classroom time to
implement AAs was very challenging. That is why the interviewees of this current study found tests more practical, which is similar to the teachers’ assessment practices in the English departments of the universities of Beijing that Cheng et al. (2004) studied, where less structured assessments (AAs) need more time to mark; therefore, as the authors found, they may not be feasible in large classes. Thus, even if TESOL teachers try to implement AA in their large classes, it might be demotivating to them because of the huge workload that AA requires in such classes. Zafar Khan (2011) discovered that the increasing workload and demand on teachers’ time would be demotivating to them. The situation of the Kurdistan Region is confirmed by Vernez et al. (2014:XVI), who found that teachers should be trained on lecturing as a way of teaching “rather than the still too poorly defined student-centred methods and should focus on the most practical techniques for large classrooms”.

Additionally, the holidays and skipping classes by students decrease the number of sessions, reduce the time of AAs, and consequently reduce the opportunities for large numbers of students to perform AAs. Skipping classes by students was also observed by Zafar Khan (2011) in Oman, where bonus holidays made massive changes to the teaching and examination dates because of the lack of a strict academic calendar, which was especially demotivating to the expatriate teachers.

Because of the above-mentioned challenges to AAs, all the interviewees believed that doing and marking AAs require a lot of work, effort, and energy, in contrary to the easiness and infrequency of traditional tests, chiefly because of the large number of students. This is relevant because ease of marking and suitable/trained graders are parts of assessment practicality (Coombe et al., 2007).

There are some virtual solutions to large classes, including the use of computers to replace humans in raising students’ consciousness on topics (Shabani, 2012), which can be possible by administering computer-based dynamic assessment (Poehner & Lantolf, 2013). However, as demonstrated by all the interviewees of this study, this might not be easily achieved in the classrooms of their English departments because of the need for more provision of costly technological instruments (e.g., computers). Abbas (2012) also concluded that the need for more use of technology to engage students in AAs in the classrooms of the English department of Diyala University was a pressing challenge. This would decrease the feasibility of AAs because assessment practicality also includes the maintenance costs and administrative logistics (Coombe et al., 2007). Nonetheless, the interviewees of this study are
interested in and regard classroom technologies as effective, and they are concerned about them. This is promising because teachers’ opinions on the use of technology in language classrooms are found to be a major determiner for integrating technology into the curriculum and its subsequent success (Al-Mahrooqi & Troudi, 2014).

The existence of all the aforementioned challenges to AAs as areas for development possibly signifies a need for greater effort in the English departments of the universities of the Kurdistan Region. This is necessary to concur with the current rapid and sustainable development in the area of assessment, as its issues have recently come to prominence in the agendas of educational authorities all over the world (Brindley, 2001). The need for greater development and modernisation of the methods of teaching, learning, and assessment in those English departments would lead to the adoption of the recently emerged theories and approaches to teaching and learning, such as socio-cultural theory, cognitivism, constructivism, multiple intelligences, communicative approach, CALx, critical pedagogy, and critical language testing. This is particularly pertinent because those theories influence the processes of teaching, learning, and assessment considerably, by which a special focus is put on the social aspect of learning, student-centred learning, empowering students, and increasing the validity and ethicality of teaching, learning, and assessment. Regarding assessment, these theories have encouraged the use of AA, which embodies their features and objectives (see the Sub-Sections of 3.5 especially 3.5.3, and Figure 2).

This chapter presented in detail all the findings this action research discovered in its first and third phases with adequate interpretations and in-depth discussion with literature and the relevant theories. For the first phase, it started with explicating and discussing the findings concerning the assessment practices and AAs of the participants, including the marking of AAs and tests. The chapter then presented and discussed the assessment literacy of the participants in AA, covering the factors behind their limited assessment literacy in AA as areas for improvement. It also presented and discussed the beliefs of the participants about whether AAs are beneficial to teaching, to learning, and to assessment. The second phase was the intervention with no data collection and analysis. For the third phase, this chapter presented and discussed the outcome of the intervention regarding its impact on introducing the participants to AA, and increasing their knowledge and changing their beliefs about it. Finally, it presented and discussed the findings related to the participants’ beliefs on the feasibility and challenges of implementing AAs, including one aspect of feasibility of AA
related to time, and some main aspects of its challenges related to the assessment system of the English departments, teachers and students’ subjectivity in marking AAs, management, and technology. After presenting and discussing all the findings, the next chapter will present the conclusions of these findings, and other concluding issues.
Chapter Six: Conclusion

This chapter is divided into six sections: (1) conclusions; (2) implications; (3) recommendations; (4) the contribution of this study to the field of AA; (5) suggestions for further research in the field of AA in the English departments of the universities of the Kurdistan Region; and (6) my reflections on the journey of this action research project. An explanation of the conclusions is followed by mentioning their potential implications on several dimensions, such as pedagogy or teacher education. Then, based on the synthesised conclusions and implications, a list of recommendations will be offered for improving the assessment system of the English departments.

The following conclusions, firstly, concern the assessment practices and alternative assessments (AAs) of the Kurdish tertiary TESOL teachers. AA is immensely important because of its emphasis on validity and ethicality. AAs are ongoing processes that involve teachers and students in making judgements on students’ progress, through the application of non-conventional strategies (Hancock, 1994). Worthen (1993:445, cited in Grabin, 2009) explains that traditional assessment merely samples “tiny snippets of student behaviour”, but AA looks at the bigger picture of students’ performance. Secondly, some conclusions relate to the necessity of assessment literacy in AA, which covers the knowledge of its concepts, requirements, criteria, features, and benefits as mentioned in the literature. Knowledge about these aspects of AA requires a deeper knowledge of the underpinning theories of AA, especially the two most important aspects of those theories, which are interrelated with the characteristics and benefits of AA (see Figure 2). Adequate knowledge of the underpinning theories of any practice is necessary because “those engaged in the ‘practice’ of education must already possess some ‘theory’ of education which structures their activities and guides their decisions” (Carr & Kemmis, 2003:113). After it was evident that the participants’ knowledge about the aspects of AA needed to be increased, I carried out an intervention to help provide further information. Hence, several conclusions were reached about increasing the participants’ knowledge of AA and changing their beliefs about it. Afterwards, it was predicted that they would have a broader understanding of the feasibility and challenges of using AAs in their contexts. Therefore, the final conclusions relate to the practicality issues of AAs, particularly the challenges, since addressing them is tremendously influential in finding
solutions and, subsequently, providing appropriate recommendations. This was a fundamental part of the last step of the cycle of this action research.

6.1 Conclusions

Almost all the findings of this study display some congruence with previous empirical studies, especially concerning teachers’ cognitions of contextual factors. After synthesising all the findings, I arrived at some conclusions. Regarding the assessment practices and AAs of the in-service Kurdish tertiary TESOL teachers, I could conclude that their AAs were more akin to active learning practices with an element of AA. Also, it was clear that the survey participants need to develop their implementation of AAs, because only around 10% of the total marks are devoted to them, along with limited class time allocated to AAs in large classrooms. The performance of some AAs is therefore infrequent, which would weaken their general effectiveness for students’ learning. When marking AAs, the participants tend to use a list for ticking (marking) each student’s participation without referring to any criteria in the form of rubrics or checklists. Therefore, the participants’ marking procedures of AAs are mostly akin to rewarding participation rather than evaluating performances. However, their marking of some AAs such as writing assignments and presentations is more accurate. Reserving only 10% of the marks for AAs without providing any guidelines to regulate the implementation of them indicates that the assessment system of the English departments of the universities of the Kurdistan Region is mostly based on traditional testing methods. Hence, AA is given low value in comparison with the weight given to testing, whereby exams are regulated by rules and are allotted around 90% of the total marks. The participants’ acceptance of this assessment system is probably because of their concern about teachers’ subjectivity in marking AAs. Overall, in terms of the three basic purposes of assessment: ‘assessment for learning’, ‘assessment as learning’, and ‘assessment of learning’ (Earl & Katz, 2006), the AAs of the in-service Kurdish tertiary TESOL teachers require some development.

Concerning the assessment literacy of the participants in AA, the results of the questionnaires demonstrated that most of the survey participants had some confidence in relation to their knowledge about AA. However, their responses to the open-ended question revealed room for improvement, since they did not talk about AA conceptually, and none of them mentioned the concepts and theoretical assumptions of AA and its underpinning theories. Also, it was
identified that they would benefit from additional knowledge about the purposes of assessment and AA, and how to distinguish them from testing. Nevertheless, the survey participants including the 12 interviewees had greater awareness and belief in the benefits of AAs for their students’ learning, which showed their general familiarity with and knowledge of AAs. Overall, the participants’ knowledge of AA needs further development, and there are four factors behind this: (1) AA was not part of their MA and/or PhD syllabi; (2) there is a limited number of resources on AA guidelines and standards; (3) there is limited training courses for implementing AAs; and (4) the participants need to be more familiar with the recent teaching, learning, socio-cultural, and critical theories that underpin AA.

The participants were generally aware of the multi-dimensional benefits of AA. They agreed that they can use the AA outcomes for adjusting and improving their styles of teaching, but their improvements were mostly based on students’ weaknesses. Thus, focusing on students’ strengths, a significant feature of AA, needs to be developed. Regarding the benefits of AA for students’ learning, the participants perceived, experienced, or observed how AA improves students’ learning, and they believed in these benefits. Nevertheless, their ability to take advantage of the AA benefits requires improvement, especially with regard to integrating it into teaching and learning more broadly in a longer class time. This is in order to enable teachers to identify continually students’ strengths to focus on, and diagnose and address their weaknesses. Nonetheless, they believed that students have the right to perform a variety of AAs, since they also believed that the more ways students have to display their knowledge and performance, the more accurate assessment will be for making assessment decisions, rather than depending on a few exams on certain days alone. The participants also believed in both traditional testing and AA for achieving a complementary process of assessment, and they called for an integration of them due to their importance in assessing both the process and product of learning. Regarding the benefits of AA in allowing students to be critical, reflective, and problem-solving thinkers, as well as assessors and autonomous learners, the participants had several opinions. For instance, they believed that teachers are not the only assessors, as students can be assessors and AAs might help them to be confident, critical thinkers, and autonomous learners. In contrast, they infrequently allowed their students to perform self-assessment and peer-assessment because of their belief about student subjectivity and their low level.
After analysing the data collected from the first phase, I realised that the participants were in need of increasing their knowledge of AA and its requirements, as a starting point to implementing such assessment more robustly. To this end, in the second intervention phase of this action research, I gave priority to increasing their knowledge of AA by presenting two seminars in two English departments.

Following the intervention, I can conclude that the two seminars have influenced the attendants both attitudinally and functionally (Evans, 2008). All the attendants agreed that there was a significant increase of their knowledge of AA, and significant changes in their beliefs about AA, which has affectively enhanced their willingness to pay greater attention to AA. However, the intervention seminars will not have an immediate effect because, as the literature makes clear, time is a crucial factor in the development of assessment literacy. Nevertheless, I believe that the two seminars have put the participants on the right track of doing more about AA, primarily by improving the participants’ cognition of AA, which will possibly lead to stronger pedagogical implications in the future. The interviewees showed their willingness to devote more effort to AA, allocate more time for it, incorporate more AAs into their classes, and they called for adjustments in their assessment system to address the challenges of implementing AA. Additionally, because the participants have already started to read about and reflect on AA, the intervention can be potentially influential for the development of ELT in the English departments. Similarly, when the recommendations of this study are disseminated to the 18 English departments, the intervention impact will be broader and more influential. Overall, I can conclude that this action research has achieved its catalytic validity, which is an important criterion to ensure that an action research should suggest an agenda, give rise to some specific actions, and result in those actions (Cohen et al., 2007).

Concerning the participants’ beliefs on the feasibility and challenges of AAs, I conclude that despite the survey participants’ belief in AAs as quite practical, this belief would probably change if they acquired more knowledge about the practicality requirements of AA. This can be seen by the numerous hindering challenges as believed by the third-phase interviewees after they became more aware of the requirements of AA in terms of time, frequency, and criteria of marking. Consequently, after the intervention, the only positive aspect of feasibility that most third-phase interviewees believed in was decreasing the marking time of
tests at home. This is in case the marks are redistributed for AAs and tests to have equal portions, whereby teachers can do half of the assessment in the classroom through AAs.

Regarding the difficulties, the participants believed that one of the main challenges to the implementation of AA was the current assessment system of the English departments of the universities of the Kurdistan Region, which is principally based on traditional testing methods, and in which the testing is prioritised over AA. Thereby, 90% of the assessment of students is carried out through tests, leading to teachers and students becoming test- and certificate-oriented. This is despite the interviewees’ willingness to change their departments’ system of teaching, learning, and assessment from a teacher-centred to student-centred approach, whereby traditional testing must not be ignored but should rather be integrated with AA in equal proportions and significance. However, currently the participants were hesitant to give more marks for AA because of their belief about teachers’ subjectivity in marking AAs, and also students’ subjectivity in self-assessment and peer-assessment. They believed that teachers must first learn how to be objective before being allowed to mark AAs with 50% of the total marks. They were worried about that teachers’ subjectivity will possibly make AA invalid and unfair, thus violating its basic elements. Three interviewees believed that their subjectivity stemmed from some cultural assumptions in which emotions have a role, which can create bias on some issues. There were several managerial and technical challenges to AAs such as the large number of students per class, limited time of sessions, limited number of sessions during an academic year, and the costs and maintenance of technological facilities. The interviewees were somewhat concerned with these challenges, as they believed that they cannot be easily addressed, especially because of the current economic crisis. Their main concerns were the large number of students and their limited lecture times. This disappointment matches what was concluded by Tsagari (2004), who found that high numbers of students in classes and the limited time to perform AAs create huge obstacles to the implementation of such assessments.

The existence of all the aforementioned challenges to AA indicates the need for greater efforts from the English departments of the universities of the Kurdistan Region to address them. Endeavours to develop and modernise the methods of teaching, learning, and assessment in those English departments are needed, which would facilitate the adoption of the recently developed theories and approaches to education, including socio-cultural theory, constructivism, multiple intelligences, communicative approach, cognitivism, and critical
theories such as CALx, critical pedagogy, and critical language testing. These influence the processes of teaching, learning, and assessment, and they underpin and encourage the use of AAs. Paying little attention to those theories is incongruent with the current rapid and sustainable development in the area of assessment, since its issues have recently come to prominence in the agendas of educational authorities worldwide (Brindley, 2001).

Finally, I have realised that the use of AAs in the English departments requires development in terms of their benefits for teaching and learning, and their effectiveness in showing students’ knowledge and performance accurately, validly, and ethically. Moreover, because of the current existing challenges, the implementation of AA would be mostly impractical. Therefore, performing AAs in the English departments can only partially fulfil the essential goals of AA that stem from their underpinning theories. These goals relate to focusing on the social aspect of learning, student-centred learning, empowering students, students’ active roles, and increasing the accuracy, validity, and ethicality of assessment.

6.2 Implications: Pedagogy and teacher development

After disseminating the conclusions of this action research to the English departments of the universities of the Kurdistan Region, there would be a number of pedagogical and teacher education implications.

1. The senior administrators of the English departments would be informed about the need to develop their assessment system to allow for a wider implementation of AA. Their potential for development relates to the validity and fairness of the assessment of their EFL students, and improvements may have a positive effect on students’ learning.

2. The English departments’ administrators would be informed about the potential for developing their teachers’ knowledge of AA, and their need to learn about how to implement AAs more widely. This might lead to some actions by these departments to solve this issue and provide proper training.

3. The conclusions about the participants’ beliefs on the benefits of various AAs to teaching, learning, and assessment in their context would be pedagogically significant because the English department administrators can use these benefits to plan how to develop AA in order to increase their advantages even more.
4. Because of the success of the intervention seminars, another pedagogical implication is that the English departments can similarly organise more seminars on AA for their teachers, especially to inform them about how to incorporate AAs into teaching and learning more robustly.

5. The most important pedagogical implications can arise from the several challenges to AAs discovered by this study. These challenges, whether related to the whole assessment system of the English departments, subjectivity of teachers and students, or the management and technology, can lead the senior administrators of these departments to find solutions that may lead to improvements in teaching, learning, and assessment.

6.3 Recommendations

The conclusions reached above lead me to recommend the following actions to improve the implementation of AA. The conclusions and recommendations of this study will be disseminated to all the English departments of the universities of the Kurdistan Region. The recommendations can be modified, based on updated needs analysis of teachers and students of the departments. The recommendations can also be a foundation for several action research studies to work on AA, especially in order to tackle its challenges in a department or a group of departments together. At present, I recommend that the English departments need to:

1. Change and improve the assessment system by focusing equally on tests and AAs;

2. Organise some seminars to increase teachers’ knowledge of AA, focusing on how AAs are important, influential, and beneficial as mentioned in the literature, and how they are underpinned and interrelated with many recently developed teaching, learning, socio-cultural, and critical theories;

3. Provide some training courses or workshops on how to implement, mark, and document AAs, and how to benefit from them;

4. Focus on educating teachers and students about how to be objective in assessment and AA through the seminars, training courses, and workshops;

5. Incorporate some subjects about AA into MA and PhD courses;
6. Provide resources on assessment, especially guidelines and standards of AA;

7. Allocate more marks for AAs so that they can be equally valued as tests;

8. Decrease the number of students per class;

9. Devote more time to AAs through incorporating them into the teaching and learning processes simultaneously; and

10. Install and maintain more technology instruments in classrooms, particularly computers.

If the above recommendations are met, I further suggest that English departments need to:

1. Elicit the help of experts and specialists to monitor teachers’ implementation of AA; and

2. Reform teacher evaluation processes to include AAs as an important part of teachers’ assessment of their students, as well as a main point to evaluate teachers’ success.

### 6.4 Contribution of the study to the field of alternative assessment: Research and pedagogy

Research on TESOL teachers’ assessment practices is somehow limited, and it is only recently that investigations into classroom-based AA within the ESL/EFL contexts have started to appear. Therefore, not much is known about TESOL teachers’ assessment and evaluation procedures at the tertiary level (Cheng et al., 2004). Hence, more research into the factors behind tertiary TESOL teachers’ selection of assessment methods, and how these affect their students’ learning, is imperative (Cheng et al., 2008). Moreover, because AA is used widely at the tertiary level nowadays (Anusienė et al., 2007), it requires much further research. For that reason, this current action research can make a contribution to the area of AA at the tertiary level, especially because of its context, since this study can be a starting point to conducting more studies in the field of AA in the English departments of the universities of the Kurdistan Region. The contribution can be influential as it deals with a number of non-researched issues in those English departments, such as (1) the Kurdish tertiary TESOL teachers’ current assessment practices and AAs; (2) their knowledge of AA; (3) their beliefs about whether AAs are beneficial; (4) the impact of the action research
intervention on their cognitions of AA; and (5) their beliefs on the feasibility and challenges of implementing AAs.

Regarding implementation, increasing the participants’ knowledge of AA and its standards and requirements via two seminars’ gave this study a pedagogical implication and significance. This is because the intervention was followed by examining the participants’ beliefs about the feasibility and challenges of AAs in their professional context. I believe that the major part of the contribution of this action research is the diagnosing of some contextual challenges in the context of the English departments of the universities of the Kurdistan Region that hinder a successful implementation of AA. This will add to studies that explore the factors behind the obstacles to AA in order to tackle them. In a general sense, this will contribute to the development and improvement of the methods of ESL/EFL teaching, learning, and assessment.

More specifically, I have discovered some issues about AA that have surprised me and that I have not come across in the extant literature. I believe that the following findings are unique to this study: (1) the practice of giving marks to tertiary EFL students as a reward for their participation in AAs in place of evaluating them (in some cases); (2) the participants’ awareness of and belief in the numerous benefits of AAs for teaching, learning, and assessment through their experience of conducting them in spite of their limited knowledge about them; and 3) the participants’ worries about being unable to be adequately objective when implementing AAs is probably a cultural issue. Three interviewees were aware of their cultural influence, which can create bias towards individuals, groups of people, or specific situations, including based on gender (Ismael, 2013). The participants’ cultural assumptions perhaps stem from emotional beliefs that allow them to give or withhold marks based on ideas of helping, rewarding, or punishing students, leading to the perpetuation of teacher subjectivity. In this respect, I have found only a few studies that mention the subjectivity of teachers (e.g., Baker, 2010; Brown & Hudson, 1998; Raymond, 2008). However, to the best of my knowledge, there is no study that deals with teacher subjectivity culturally in the field of AA. I hope that this finding will raise the question of the cultural challenges of AA to attract more attention to them, in order to be investigated further in the future. This will diagnose and thus address specific cultural assumptions of teachers that are obstacles to AA in specific contexts.
6.5 Suggestions for further research

Because of the lack of research of this nature in the English departments of the universities of the Kurdistan Region, many areas of English language assessment merit many other studies to explore them. However, for the current situation, I will call for further research into the following areas in the field of AA:

1. Investigating each of the contextual challenges uncovered by this study in order to find more factors behind them and to examine solutions for them;

2. Exploring the feasibility and challenges of AA from the perspectives of the students and administrators;

3. Studying the external cultural, social, and economic factors that might intervene in the implementation of AA successfully;

4. Examining each of the types, methods, practices, activities, and/or strategies of AA separately in order to address them more in-depth; and

5. Exploring each of the underpinning theories of AA separately in terms of their interrelation with AA, or their influence on it, in order to develop further the connection between them and thus achieve a more successful implementation of AAs.

The above suggestions for further investigations into AA can be part of the increasing interest in assessment methods (Ataç, 2012), and are in line with the processes of identifying and addressing certain contextual factors that are responsible for facilitating or obstructing implementation, which is quite necessary (Bullock, 2011). The suggestions can also contribute to the rapid and sustainable development in the area of assessment (Brindley, 2001).

6.6 Personal reflections on the study

Before starting this action research as a partial insider researcher, I expected that Kurdish tertiary TESOL teachers might need more assessment literacy in the complexity and intermixed nature of AA and its standards and requirements. My expectation turned out to be right; therefore, after the intervention, I felt satisfied that I had managed to introduce the
participants to AA and the possibilities it could offer. In addition, I mentioned in my previous study (Ismael, 2013) that some Kurdish tertiary TESOL teachers were perhaps subjective in AAs and tests, which resulted mostly from the typically large number of students per class and gender discrimination. However, during the implementation of this current action research, I realised that a researcher cannot presume or hypothesise about many significant issues of TESOL teachers, even if he or she has been teaching in the same context for a long time. This is because during and after the data analysis of this study, I understood and analysed numerous essential aspects of teaching, learning, and assessment in these English departments, particularly their assessment and testing procedures that I had not anticipated. Also, the grave concern of the participants about their subjectivity in AA, stemming from cultural or other factors, was an issue I did not foresee. Moreover, I noticed their slight disappointment about some systemic, affective, managerial, and technical challenges of AA that are believed to be difficult to tackle.

However, the interaction with the participants and heads or deputy heads of the 18 English departments I visited for data collection was incredibly rewarding. I enjoyed the way we discussed the potential developments of AA in their departments after they filled in the questionnaires. Visiting the 18 English departments to distribute the questionnaires was very demanding, but the volunteer participants’ helpfulness was a daily encouragement for me as a researcher. It was a fruitful experience in terms of understanding their opinions about the subject, as most of them showed an interest in using AAs. It was also inspiring that most of the heads of those English departments were very helpful and friendly. They appreciated what I was doing, and some of them happily and proudly explained to me that they had done something about many sorts of learning activities, and that teachers usually give some daily marks for AAs. Furthermore, after I reflected on the challenges of AA, in order to draw up some necessary recommendations as to how to address them, I realised that my action research can direct some road map solutions to eliminate the obstacles to the implementation of AA. It is encouraging and rewarding to me that my efforts will, in some way or other, be valuable in the future.
### Appendix (1): List of themes, categories, and sub-categories from the whole collected data

<table>
<thead>
<tr>
<th>Theme</th>
<th>Category</th>
<th>Sub-category</th>
<th>Explanation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Assessment practices and AAs used by the participants</td>
<td>AAs and test questions</td>
<td>AAs</td>
<td>Almost all the survey participants conducted various AAs with varying frequency. Among these AAs, their implementation of oral activities such as oral presentations, conversations, oral discussions, and direct questions and answers constituted the highest usage. Writing assignments were also used more frequently than other AAs (see Section 5.1.1 and Table 3).</td>
</tr>
<tr>
<td>Duration and frequency of AAs</td>
<td></td>
<td>Two interviewees (2 of 12) engaged their students in AAs throughout the class time. However, the other 10 interviewees (10 of 12) devoted 5-25 minutes to AAs usually at the end of lectures. Because of the limited lecture time and the high number of students, four interviewees stated that they performed some AAs infrequently (see Section 5.1.1).</td>
<td></td>
</tr>
<tr>
<td>The nature of AAs</td>
<td></td>
<td>Most of these practices of the participants were more akin to active learning practices with an element of AA, because they did not completely integrate AAs into the active learning practices during the lecture time, since they allocated 5-25 minutes for conducting AAs usually at the end of lectures (see Section 5.1.1).</td>
<td></td>
</tr>
<tr>
<td>Types of test questions</td>
<td></td>
<td>Most participants revealed a high frequency of using most of the types of test questions. Among these, short answers and writing paragraphs constituted the highest usage (see Section 5.1.1 and Table 4).</td>
<td></td>
</tr>
<tr>
<td>The marking of AAs and tests</td>
<td>Marks allocated for AAs</td>
<td></td>
<td>Only 10-20 marks were allocated for AAs from 100 marks of the whole academic year; meanwhile, 80-90 marks were assigned to traditional ‘pen and paper’ exams. The questionnaire results were indicative of the limited number of marks devoted to AAs because even from the 10-20 marks that the system allows teachers to allocate for AAs, half of the survey participants (45: 50%) devoted only 10 marks to them, and less than a quarter (13: 14%) devoted fewer than 10 marks to AAs (see Section 5.1.2 and Table 5).</td>
</tr>
</tbody>
</table>
Techniques of marking AAs

Half of the interviewees (6 of 12) stated that they did not use the marking techniques for AAs such as rubrics and checklists, and a quarter of them (3 of 12) did not reveal whether they used them or not. Rather, most interviewees (9 of 12) referred to a simple list prepared for ticking (marking) students’ participations in AAs (see Section 5.1.2).

The nature of the participants’ marking of AAs

Most interviewees’ (9 of 12) marking of AAs by tick lists were more akin to giving rewarding marks than evaluative marks (see Section 5.1.2).

Marks allocated for tests

Regarding tests, 80-90% of the marks were assigned to ‘pen and paper’ exams. This distribution was mostly based on traditional testing methods, which were apparent in the series of tests of the first and second terms, followed by a final year exam (see Section 5.1.2 and Table 5).

Assessment literacy of the participants in AA

The participants’ knowledge of AA

The participants’ conceptual knowledge of AA

The majority of the survey participants (78: 87%) said that they had read about AA, except for 12 (13%), and the vast majority (86: 95%) said they possessed varying knowledge about AA, except for 2 (2%). However, in their answers to the open-ended questionnaire question aimed at ascertaining what they know about AA, the majority of the survey participants (71: 79%) talked about the general usefulness of AA for their students’ learning, but they did not define it conceptually (see Section 5.3.1 and Table 6).

Knowledge of the differences between testing and AA

Two thirds of the interviewees (8 of 12) did not differentiate between traditional testing and AA conceptually, although they distinguished them technically and psychologically. Psychological differentiation included the participants’ beliefs that testing is stressful or causes fear, whereas AAs give freedom to students. Two of the 12 interviewees clarified that AAs can sometimes cause shyness and reticence. Nonetheless, four interviewees distinguished testing from AA conceptually (see Section 5.3.1).

Knowledge of the purposes of AA

More than two thirds of the interviewees (9 of 12) talked about ‘assessment of learning’ as their major purpose. One interviewee demonstrated some
knowledge about ‘assessment for learning’, and two interviewees talked about ‘assessment as learning’ (see Section 5.3.1).

| Knowledge about some important requirements of AAs | It was clear that the participants would benefit from more knowledge about: (1) the conceptual and theoretical aspects of AA; (2) the continuous nature of AAs throughout the class time when integrated into teaching and learning; (3) the necessity of the frequency of AAs throughout the academic year to observe the students’ growth over time; and (4) the criteria of implementing and marking AAs with rubrics and checklists (see Section 5.3.1). |
| Factors identified as areas to improve for increasing the participants’ assessment literacy in AA | MA and PhD syllabi | AA was not part of the MA and PhD syllabi of teachers who were not specialised in applied linguistics or related fields, and even those who were specialised in applied linguistics had not studied AA (see Section 5.3.2). |
| Resources on AA | Dissatisfaction with the provision of resources in the English departments was expressed by almost all of the interviewees (11 of 12), and further, three interviewees clarified that they had only rules for technical regulations of the distribution of marks for the two main term tests and the final examination, and a dean of one of the faculties confirmed that. That was despite the fact that two thirds of the interviewees (8 of 12) demonstrated their awareness of the importance and necessity of the guidelines for implementing AAs (see Section 5.3.2). |
| Training courses in AA | Table 7 shows that a few of the survey participants (7: 8%) had attended courses on assessment, whereas the majority (83: 92%) attended some courses, which were not on assessment, AA, and/or testing. The 12 interviewees completed the questionnaires too, and thus they had the same situation as the survey participants (see Section 5.3.2 and Table 7). |
| Knowledge of the underpinning theories of AA | The participants could benefit from more knowledge about the underpinning theories of AA, which might be achieved by the establishment of a modern curriculum in the English departments. A modern curriculum is |
supposed to provide the standards and guidelines of English language teaching, learning, and assessment, informed by newly developed educational theories; therein, some focus can inevitably be put on AA (see Section 5.3.2).

<table>
<thead>
<tr>
<th>Beliefs of the participants on whether AAs are beneficial</th>
<th>Benefits of AAs to teaching</th>
<th>Use of AAs to improve teaching methods</th>
</tr>
</thead>
<tbody>
<tr>
<td>Use of AAs to improve teaching methods</td>
<td>The vast majority of the survey participants (88; 98%) agreed that they could use the outcomes or observations of AAs to improve their teaching methods. In answering the open-ended questionnaire question, several survey participants also talked about the influence of AAs on their teaching. The majority of interviewees (10 of 12) stated that they used AAs for improving their teaching by changing their methods and strategies, mainly to accommodate different students (see Section 5.5.1).</td>
<td></td>
</tr>
<tr>
<td>AA integration into teaching and learning</td>
<td>It was perhaps difficult for them to manipulate AA outcomes adequately in order to improve their teaching methods and materials, primarily because they needed to develop their simultaneous combination of AAs with teaching and learning (see Section 5.5.1).</td>
<td></td>
</tr>
<tr>
<td>Benefits of AAs to learning</td>
<td>The improvement of students’ learning by AAs</td>
<td>The majority of the survey participants (66; 73%) regard AAs as very or quite beneficial (see Table 9). Also, in the responses to the open-ended questionnaire question, in their attempts to define AA, the majority of the survey participants (71; 79%) talked about the general usefulness of AAs in increasing students’ learning. Additionally, all interviewees (12 of 12) perceived, experienced, or observed several benefits of AAs, as all of them believed that AAs improved students’ learning of English language skills (see Section 5.5.2.1).</td>
</tr>
<tr>
<td>Focus on the strengths and weaknesses of students’ learning</td>
<td>Two thirds of the interviewees (8 of 12) believed that the focus should be on both strengths and weaknesses. However, they admitted that previously they had focused more on providing feedback on students’ weaknesses rather than focusing on their strengths. When giving feedback, those eight interviewees emphasised the need to provide encouragement rather than offering direct or embarrassing feedback (see Section 5.5.2.1).</td>
<td></td>
</tr>
<tr>
<td>Using a variety of AAs for improving learning</td>
<td>More than two thirds of the interviewees (9 of 12) believed that enhancing students’ learning of English language needs various AAs to provide fair opportunities for them so that they can choose and perform suitable AA activities frequently. Also, more than half of the interviewees (7 of 12) believed that the systematic allocation of more marks for AAs would increase the value of these assessments in their English departments (see Section 5.5.2.2).</td>
<td></td>
</tr>
<tr>
<td>Affective aspect of AAs</td>
<td>Two thirds of the interviewees (9 of 12) believed that testing causes stress for students, while AAs make them involved in their learning and assessment and even enjoy them (see Section 5.5.2.2).</td>
<td></td>
</tr>
<tr>
<td>Authenticity of AAs</td>
<td>All the interviewees (12 of 12) believed that their AAs were related to and thus can be applied in real life situations, and they believed that this would improve students’ learning (see Section 5.5.2.2).</td>
<td></td>
</tr>
<tr>
<td>Students’ right to do AAs</td>
<td>More than two thirds of the interviewees (9 of 12) believed that students have the right to perform multiple AAs to present their knowledge and abilities, as part of providing them a fair assessment, rather than just depending on a few tests (see Section 5.5.2.2).</td>
<td></td>
</tr>
<tr>
<td>AAs for accurate assessment of knowledge and abilities</td>
<td>Almost all the survey participants (87: 97%) agreed that AAs allow their EFL students to show what they know, and the majority (80: 89%) agreed that AAs allow their students to exhibit what they can perform (see Table 10). Two thirds of the interviewees (8 of 12) also believed that AA is more direct and immediate, and its assessment would be clearer and more accurate because AA prevents students from relying on memorisation and instead assesses their current performance, which has priority in language learning (see Section 5.5.2.2).</td>
<td></td>
</tr>
<tr>
<td>Using various AAs for assessment accuracy and ethicality</td>
<td>Two thirds of the interviewees (8 of 12) believed that more ways for students to show their knowledge and performance lead to more accurate assessments. They believed that assessment decisions are better and more accurate when based on multiple assessment results. Interestingly, some survey participants (14: 16%), in</td>
<td></td>
</tr>
<tr>
<td>The combination of AAs and tests</td>
<td>The majority of the interviewees (10 of 12) believed that test results can be valuable, beneficial, and can improve learning, and two interviewees believed that tests could also accurately show students’ knowledge. Moreover, almost all of the survey participants (88: 98%) agreed that test results could be used for advancing their teaching (see Table 8). For this reason, the majority of the interviewees (10 of 12) called for an integration of AA and testing due to their importance in assessing the process and product of learning, as parts of a complementary process of assessment (see Section 5.5.2.2).</td>
<td></td>
</tr>
<tr>
<td>Students as assessors</td>
<td>All the interview participants (12 of 12) believed that generally teachers are not the only assessors of students. However, one interviewee, Lara, said that in Kurdistan it is possible for this role to be solely occupied by teachers. Also, all the interviewees (12 of 12) believed, to differing degrees, that students could assess themselves or each other (see Section 5.5.2.3).</td>
<td></td>
</tr>
<tr>
<td>Students’ awareness of their own and their peers’ mistakes</td>
<td>All the interview participants (12 of 12) believed that peer-assessment benefited students’ learning by enabling them to perceive their mistakes through those of their peers (see Section 5.5.2.3).</td>
<td></td>
</tr>
<tr>
<td>Implementing self-assessment and peer-assessment</td>
<td>It was apparent from the 12 interviewees’ statements that they rarely allowed their students to be assessors of themselves by performing self-assessment. However, while they allowed students to assess each other by performing peer-assessment, they rarely counted the markings done by students to add marks to their scores because they did not trust and did not depend on students’ assessments (see Section 5.5.2.3).</td>
<td></td>
</tr>
<tr>
<td>Critical thinking through self-assessment and peer-assessment</td>
<td>More than half of the interviewees (7 of 12) believed that determining strengths and weaknesses requires critical and reflective thinking, and although students...</td>
<td></td>
</tr>
<tr>
<td>assessment</td>
<td>might possess this ability, they do not apply it to finding and displaying their own or their peers’ mistakes due to their subjectivity (see Section 5.5.2.3).</td>
<td></td>
</tr>
<tr>
<td>Students’ subjectivity in self-assessment and peer-assessment</td>
<td>More than half of the interviewees (7 of 12) believed that some students can determine their strengths and weaknesses by self-assessment and peer-assessment, but they do not reveal that to their teachers or peers due to their subjectivity (see Section 5.5.2.3).</td>
<td></td>
</tr>
<tr>
<td>AA and autonomous learning</td>
<td>More than two thirds of the interviewees (9 of 12) believed that the adoption of AAs, especially self-assessment and peer-assessment, might help their students to become more confident, critical thinkers, and autonomous learners (see Section 5.5.2.3).</td>
<td></td>
</tr>
</tbody>
</table>

### The intervention outcome: introducing the participants to AA

| Knowledge of AA | Increasing the participants’ knowledge of AA | All 12 attendants’ stated that their knowledge of AA had been raised by the intervention to a good extent (see Section 5.7). |
| The participants’ knowledge about the differences between AA and testing | The participants’ knowledge of the differences between AA and testing were ultimately higher than before the intervention (see Section 5.7). |
| The participants’ knowledge about AA purposes | The participants’ knowledge of the three purposes of AA were more increased than before the intervention (see Section 5.7). |

### Changing the participants’ beliefs

| Changing the participants’ beliefs about AA | Changing the participants’ beliefs about AA | Almost all the attendants (11 of 12) confirmed that their beliefs were changed after they attended the two seminars. After the intervention, they believed that they should use more AAs by devoting more effort and allocating more time to them (see Section 5.7). |
| Increasing the participants’ willingness to implement more AAs | Increasing the participants’ willingness to implement more AAs | Almost all the attendants (11 of 12) were eager to pay more attention to AA, and expressed their willingness to incorporate more AAs into their classroom teaching and learning (see Section 5.7). |
| The participants’ calling for some improvements to allow a wider implementation of AAs | The participants’ calling for some improvements to allow a wider implementation of AAs | Nine of the 12 attendants called for some changes in the assessment system of their English departments in order to allow more room to implement AAs. They also called for addressing the challenges of AAs, especially the number of students per class and the limited time (see Section 5.7). |

<p>| Beliefs of the Feasibility of | The participants’ | In the first-phase questionnaires, most survey |</p>
<table>
<thead>
<tr>
<th>participants on the feasibility and challenges of AAs</th>
<th>implementing AAs</th>
<th>general beliefs about the feasibility of AAs</th>
<th>participants (58: 64%) believed that generally AAs are quite or very practical (see Table 11). In the first phase, the beliefs of the participants about the practicality issues of AAs were about performing AAs but probably not about their standards and requirements such as criteria of marking (see Section 5.9.1).</th>
</tr>
</thead>
<tbody>
<tr>
<td>One feasible aspect of AA</td>
<td></td>
<td>Decreasing the time of marking tests at home was identified by more than two thirds of the interviewees (9 of 12) as a feasibility aspect of AAs in the context of this study, because doing a lot of marking at home was a concern (see Section 5.9.1).</td>
<td></td>
</tr>
<tr>
<td>Challenge (1): Assessment system of the English departments and their context</td>
<td>General beliefs about the challenges of AA</td>
<td>Most survey participants (54: 60%) believed that performing AAs would be quite or very challenging (see Section 5.9.2 and Table 12).</td>
<td></td>
</tr>
<tr>
<td></td>
<td>The assessment system of the English departments</td>
<td>More than two thirds of the interviewees (9 of 12) mentioned several contextual factors that can make AAs more difficult in the context of their English departments. One of the factors was their assessment system, which was principally based on the traditional teaching, learning, and assessment methods (see Section 5.9.2).</td>
<td></td>
</tr>
<tr>
<td></td>
<td>The assessment system as a challenge to improving education</td>
<td>The statements of more than two thirds of the interviewees (9 of 12) generally denoted that their assessment system might have prevented the improvement of teaching, learning, and assessment procedures in their English departments (see Section 5.9.2).</td>
<td></td>
</tr>
<tr>
<td></td>
<td>The assessment system causing the limited performance of AAs</td>
<td>More than two thirds of the interviewees (9 of 12) directly or indirectly blamed their assessment system for the limited performance of AAs (see Section 5.9.2).</td>
<td></td>
</tr>
<tr>
<td></td>
<td>The assessment system’s encouragement of a traditional approach</td>
<td>The assessment system of the English departments follows a traditional system of lecturing style and testing that usually encourages memorisation and spoon-feeding, and thus leads students to be test-oriented because the major part of assessment is completed through tests, around 90% of the marks (see Section 5.9.2 and 5.1.2).</td>
<td></td>
</tr>
<tr>
<td></td>
<td>The assessment system’s focus on</td>
<td>The assessment system of these English departments is determined chiefly by top-down technical guidelines</td>
<td></td>
</tr>
</tbody>
</table>
knowledge rather than performance imposed on their teachers about the frequency, intervals, and duration of exams, number of questions, and number of marks for each test. Teachers predominantly relied on traditional testing because of the prioritisation of testing over AA in their departments. Having first and second term tests with 20-40% of marks and the final exam with 60% of marks encourages a focus on knowledge of English language rather than its four language skills (see Sections 5.9.2 and 5.1.2).

The participants’ willingness to change their assessment system The majority of the interviewees (10 of 12) were willing to change the system of teaching, learning, and assessment from a teacher-centred to student-centred approach. To achieve this, they wanted to implement more AAs in order to provide more valid and fairer assessment for their students; this is in case teachers and students are not subjective anymore (see Section 5.9.2).

Challenge (2): teachers and students’ subjectivity in marking AAs Students’ subjectivity in self-assessment and peer-assessment Almost all interviewees (11 of 12) believed that students were subjective in doing self-assessment and peer-assessment because of their social relationships and the tendency not to show their own and their friends’ mistakes, in order to avoid losing marks (see Section 5.9.3).

Teachers’ subjectivity in marking AAs All interviewees (12 of 12) believed that teachers were subjective in marking AAs due to some psychological and emotional reasons, such as teachers’ feelings during assessment, including sympathising with or being intolerant of some students (see Section 5.9.3).

The validity and fairness of marking AAs All interviewees (12 of 12) considered that the validity and fairness of marking AAs were sometimes dependent on the teachers’ mood, and maybe sometimes on the nature of the relationship between a teacher and an individual student (see Section 5.9.3).

Testing is believed to be more objective Half of the interviewees (6 of 12) believed that testing was more objective and fairer to students, by which subjectivity can be avoided, since test papers have everything recorded (see Section 5.9.3).

Factors behind Three interviewees (3 of 12) related teachers’
<table>
<thead>
<tr>
<th>Challenge (3): management and technology</th>
<th>Large classes</th>
<th>Rejection of allocating 50% of the marks for AAs</th>
<th>Factors behind teachers’ subjectivity: cultural influence</th>
</tr>
</thead>
<tbody>
<tr>
<td>teachers’ subjectivity: unavailability of criteria</td>
<td>subjectivity to the unavailability of criteria for AAs, which is possibly due to not using the marking techniques of AAs such as checklists and rubrics that can be employed as criteria (see Section 5.9.3).</td>
<td>Three interviewees (3 of 12) believed that their subjectivity was also influenced by some cultural factors. For example, in their culture, people probably mix their relationships, emotions, and feelings with whatever they do, which might lead to subjectivity (see Section 5.9.3).</td>
<td>Three interviewees (3 of 12) believed that their subjectivity was also influenced by some cultural factors. For example, in their culture, people probably mix their relationships, emotions, and feelings with whatever they do, which might lead to subjectivity (see Section 5.9.3).</td>
</tr>
<tr>
<td>The need for learning the concept of objectivity</td>
<td>The majority of the attendants of the focus group interviews believed that they need much more professional development and supervision to become objective in using AA, largely because of their need to be more familiar with the concept of ‘being objective’ (see Section 5.9.3).</td>
<td>During the focus group interviews, after I had suggested a redistribution of marks to 50 for traditional tests and 50 for AAs, there were strong concerns about the potential for subjective marking. Therefore, the majority of the attendants disagreed with me (see Section 5.9.3).</td>
<td>During the focus group interviews, after I had suggested a redistribution of marks to 50 for traditional tests and 50 for AAs, there were strong concerns about the potential for subjective marking. Therefore, the majority of the attendants disagreed with me (see Section 5.9.3).</td>
</tr>
<tr>
<td>Challenge (3): management and technology</td>
<td>Limited class time</td>
<td>Limited classes</td>
<td>Limited number of technological facilities in classrooms</td>
</tr>
<tr>
<td>Large classes</td>
<td>All interviewees (12 of 12) believed that a large number of students in their classrooms made the implementation of AAs considerably challenging, because they did not have enough time during their lectures to give an opportunity to all students to perform AAs individually (see Section 5.9.4).</td>
<td>All interviewees (12 of 12) believed that their class time is too limited to practice AAs with such high numbers of students (see Section 5.9.4).</td>
<td>All interviewees (12 of 12) believed that providing and maintaining technological facilities for teaching, learning, and assessment required improvement (e.g., internet, computers, data shows, projectors, copy machines, and TVs), since in some classrooms, no</td>
</tr>
<tr>
<td>Limited class time</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Limited classes</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Limited number of technological facilities in classrooms</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Limited number of technological facilities in classrooms</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cost of technological facilities</td>
<td>All interviewees (12 of 12) believed that providing and maintaining such equipment is costly, which may constitute a constraining challenge due to the current economic crisis in the Kurdistan Region (see Section 5.9.4).</td>
<td></td>
<td></td>
</tr>
<tr>
<td>----------------------------------</td>
<td>---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------</td>
<td></td>
<td></td>
</tr>
<tr>
<td>The difficulty of AA and the easiness of testing</td>
<td>All interviewees (12 of 12) believed that AAs required much work, effort, and energy from them, since AA included understanding and evaluating many aspects of students through various means (see Section 5.9.4).</td>
<td></td>
<td></td>
</tr>
<tr>
<td>The difficulty of marking AAs</td>
<td>Some interviewees admitted that they could not always mark students’ AAs (see Section 5.1.2), and three interviewees confirmed that daily markings of such assessments were very difficult, while tests were easy (see Section 5.9.4).</td>
<td></td>
<td></td>
</tr>
<tr>
<td>The easiness and infrequency of tests</td>
<td>All interview participants (12 of 12) believed that performing AAs was unlike testing, which was much less frequent and easier (see Section 5.9.4).</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Appendix (2): The findings, intervention, implications, and recommendations according to the phases of this action research

First phase

- Assessment practices and AAs of Kurdish tertiary TESOL teachers
  - The participants’ alternative assessments and test questions
  - The participants’ marking of alternative assessments and tests
- Assessment literacy of Kurdish tertiary TESOL teachers in alternative assessment
  - The participants’ knowledge of alternative assessment
  - The factors behind the participants’ limited assessment literacy in AA as areas for improvement
- Kurdish tertiary TESOL teachers’ beliefs about whether alternative assessments are beneficial
  - The participants’ beliefs about the benefits of alternative assessments to teaching
  - The participants’ beliefs about the improvement of students’ learning by alternative assessment
  - The participants’ beliefs about utilising a variety of alternative assessments for improving learning and assessment accuracy
  - The participants’ beliefs about the benefits of alternative assessments for allowing students to be assessors, critical thinkers, and autonomous learners

Second phase

- The action research intervention via two seminars
  - Increasing the participants’ knowledge of alternative assessment
  - Changing the participants’ beliefs about alternative assessment
  - Enhancing the participants’ willingness to do more about alternative assessment in the future
Beliefs of Kurdish tertiary TESOL teachers on the feasibility and challenges of alternative assessments

Third phase

The intervention outcome of introducing the participants to alternative assessment

- Examining whether the participants’ knowledge of alternative assessment was increased
- Examining whether the participants’ beliefs about alternative assessment were changed
- Examining whether the participants’ willingness to do more about alternative assessment was enhanced

Beliefs of Kurdish tertiary TESOL teachers on the feasibility and challenges of alternative assessments

- Feasibility: performing more alternative assessments in the classroom decreases the time of marking tests at home
- Challenge (1): The assessment system of the English departments
- Challenge (2): Teachers and students’ subjectivity in conducting and marking alternative assessments
- Challenge (3): Managerial and technological obstacles such as the high number of students per class, limited time, and costs and provision of technology

Implications and recommendations

- Implications: Dissemination of the findings to all 18 English departments would encourage their teachers and administrators to take further actions
- Recommendations: The solutions to the challenges of alternative assessments will be presented to all 18 English departments to take more actions
Appendix (3): Statistical results of the first-phase questionnaire items

Qualification and experience

<table>
<thead>
<tr>
<th>Highest qualification</th>
<th>MA, MSc., etc.</th>
<th>PhD, EdD, etc.</th>
</tr>
</thead>
<tbody>
<tr>
<td>76</td>
<td>14</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Years of teaching English at the tertiary level</th>
<th>1-5 years</th>
<th>6-10 years</th>
<th>11-15 years</th>
<th>16-20 years</th>
<th>More than 20 years</th>
<th>Not mentioned</th>
</tr>
</thead>
<tbody>
<tr>
<td>56</td>
<td>17</td>
<td>7</td>
<td>2</td>
<td>4</td>
<td>4</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Attendance of any training course(s) on teaching and/or assessment</th>
<th>1 general teaching course</th>
<th>2 ELT training courses</th>
<th>3 ELT training courses or more</th>
<th>1 or more assessment and testing course(s)</th>
<th>No course</th>
</tr>
</thead>
<tbody>
<tr>
<td>65=72.22%</td>
<td>27=30%</td>
<td>13=14.44%</td>
<td>7=7.77%</td>
<td>7=7.77%</td>
<td>4=4.44%</td>
</tr>
</tbody>
</table>

Types of test questions or items you use in the test papers

<table>
<thead>
<tr>
<th>No</th>
<th>Questions and items</th>
<th>Always</th>
<th>Often</th>
<th>Sometimes</th>
<th>Rarely</th>
<th>Never</th>
<th>Not answered</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Multiple choices</td>
<td>8 = 8.88%</td>
<td>34 = 37.77%</td>
<td>27 = 30%</td>
<td>14 = 15.55%</td>
<td>7 = 7.77%</td>
<td>0 = 0%</td>
</tr>
<tr>
<td>2</td>
<td>Fill-ins</td>
<td>8 = 8.88%</td>
<td>26 = 28.88%</td>
<td>33 = 36.66%</td>
<td>12 = 13.33%</td>
<td>6 = 6.66%</td>
<td>5 = 5.55%</td>
</tr>
<tr>
<td>3</td>
<td>Matching items in two columns</td>
<td>3 = 3.33%</td>
<td>14 = 15.55%</td>
<td>23 = 25.55%</td>
<td>31 = 34.44%</td>
<td>19 = 21.11%</td>
<td>0 = 0%</td>
</tr>
<tr>
<td>4</td>
<td>True–false items</td>
<td>7 = 7.77%</td>
<td>30 = 33.33%</td>
<td>21 = 23.33%</td>
<td>19 = 21.11%</td>
<td>13 = 14.44%</td>
<td>0 = 0%</td>
</tr>
<tr>
<td>5</td>
<td>Short answers</td>
<td>26 = 28.88%</td>
<td>28 = 31.11%</td>
<td>27 = 30%</td>
<td>7 = 7.77%</td>
<td>2 = 2.22%</td>
<td>0 = 0%</td>
</tr>
<tr>
<td>6</td>
<td>Sentence-completion items</td>
<td>8 = 8.88%</td>
<td>28 = 31.11%</td>
<td>32 = 35.55%</td>
<td>16 = 17.77%</td>
<td>5 = 5.55%</td>
<td>1 = 1.11%</td>
</tr>
<tr>
<td>7</td>
<td>Writing paragraphs</td>
<td>32 = 35.55%</td>
<td>26 = 28.88%</td>
<td>14 = 15.55%</td>
<td>11 = 12.22%</td>
<td>5 = 5.55%</td>
<td>2 = 2.22%</td>
</tr>
</tbody>
</table>

Classroom assessment activities you do with your EFL students

<table>
<thead>
<tr>
<th>No</th>
<th>The type of classroom assessment activity</th>
<th>Always</th>
<th>Often</th>
<th>Sometimes</th>
<th>Rarely</th>
<th>Never</th>
<th>Not answered</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Oral presentation of the students</td>
<td>29 = 32.22%</td>
<td>42 = 46.66%</td>
<td>16 = 17.77%</td>
<td>3 = 3.33%</td>
<td>0 = 0%</td>
<td>0 = 0%</td>
</tr>
<tr>
<td>2</td>
<td>Direct question and answer</td>
<td>54 = 60%</td>
<td>29 = 32.22%</td>
<td>5 = 5.55%</td>
<td>2 = 2.22%</td>
<td>0 = 0%</td>
<td>0 = 0%</td>
</tr>
<tr>
<td>3</td>
<td>Doing conversations</td>
<td>38 = 42.22%</td>
<td>26 = 28.88%</td>
<td>21 = 23.33%</td>
<td>4 = 4.44%</td>
<td>1 = 1.11%</td>
<td>0 = 0%</td>
</tr>
</tbody>
</table>
4 Oral discussions 39 = 43.33% 38 = 42.22% 12 = 13.33% 1 = 1.11% 0 = 0% 0 = 0%
5 Giving immediate feedback to the students 16 = 17.77% 37 = 41.11% 24 = 26.66% 13 = 14.44% 0 = 0% 0 = 0%
6 Provide an oral description of an event or object 16 = 17.77% 37 = 41.11% 27 = 30% 8 = 8.88% 1 = 1.11% 1 = 1.11%
7 Retell a story after listening to a passage 11 = 12.22% 22 = 24.44% 34 = 37.77% 10 = 11.11% 12 = 13.33% 1 = 1.11%
8 Oral interviewing 16 = 17.77% 16 = 17.77% 26 = 28.88% 24 = 26.66% 8 = 8.88% 0 = 0%
9 Oral readings 16 = 17.77% 27 = 30% 24 = 26.66% 18 = 20% 5 = 5.55% 0 = 0%
10 Book reports 6 = 6.66% 12 = 13.33% 25 = 27.77% 22 = 24.44% 24 = 26.66% 1 = 1.11%
11 Writing assignments 30 = 33.33% 27 = 30% 21 = 23.33% 9 = 10% 1 = 1.11% 2 = 2.22%
12 Students’ summaries of what is read or listened to 8 = 8.88% 25 = 27.77% 30 = 33.33% 12 = 13.33% 3 = 3.33% 12 = 13.33%
13 Editing a piece of writing 8 = 8.88% 26 = 28.88% 23 = 25.55% 17 = 18.88% 15 = 16.66% 1 = 1.11%
14 Self-assessment 8 = 8.88% 22 = 24.44% 32 = 35.55% 19 = 21.11% 9 = 10% 0 = 0%
15 Peer-assessment 15 = 16.66% 17 = 18.88% 29 = 32.22% 17 = 18.88% 11 = 12.22% 1 = 1.11%

The distribution of the marks for AAs and testing

<table>
<thead>
<tr>
<th>Number of teachers</th>
<th>percentage</th>
<th>Marks of testing</th>
<th>Marks of AAs</th>
</tr>
</thead>
<tbody>
<tr>
<td>45</td>
<td>50%</td>
<td>90</td>
<td>10</td>
</tr>
<tr>
<td>14</td>
<td>15.55%</td>
<td>80</td>
<td>20</td>
</tr>
<tr>
<td>7</td>
<td>7.77%</td>
<td>92</td>
<td>8</td>
</tr>
<tr>
<td>7</td>
<td>7.77%</td>
<td>85</td>
<td>15</td>
</tr>
<tr>
<td>4</td>
<td>4.44%</td>
<td>95</td>
<td>5</td>
</tr>
<tr>
<td>3</td>
<td>3.33%</td>
<td>75</td>
<td>25</td>
</tr>
<tr>
<td>2</td>
<td>2.22%</td>
<td>87</td>
<td>13</td>
</tr>
<tr>
<td>2</td>
<td>2.22%</td>
<td>50</td>
<td>50</td>
</tr>
<tr>
<td>2</td>
<td>2.22%</td>
<td>60</td>
<td>40</td>
</tr>
<tr>
<td>1</td>
<td>1.11%</td>
<td>98</td>
<td>2</td>
</tr>
<tr>
<td>1</td>
<td>1.11%</td>
<td>94</td>
<td>6</td>
</tr>
<tr>
<td>1</td>
<td>1.11%</td>
<td>86</td>
<td>14</td>
</tr>
<tr>
<td>1</td>
<td>1.11%</td>
<td>76</td>
<td>24</td>
</tr>
</tbody>
</table>
Your cognition of AAs

<table>
<thead>
<tr>
<th>No</th>
<th>Items</th>
<th>Read a lot</th>
<th>Read</th>
<th>Read a little</th>
<th>Have not read</th>
<th>Not answered</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Have you read about AAs?</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>14 = 15.55%</td>
<td>34</td>
<td>30</td>
<td>12</td>
<td>0=0%</td>
</tr>
<tr>
<td>2</td>
<td>Do you have enough knowledge about AAs?</td>
<td>Have a lot</td>
<td>Have</td>
<td>Have a little</td>
<td>Do not have</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>12 = 13.33%</td>
<td>46</td>
<td>28</td>
<td>2</td>
<td>2=2.22%</td>
</tr>
<tr>
<td>3</td>
<td>What do you know about AAs in general?</td>
<td>An open-ended question</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>The test questions that I use like multiple-choices, fill-ins, matching, true-false, etc. allow my EFL students to show what they know.</td>
<td>Strongly agree</td>
<td>Agree</td>
<td>Unsure</td>
<td>Disagree</td>
<td>Strongly disagree</td>
</tr>
<tr>
<td></td>
<td></td>
<td>20 = 22.22%</td>
<td>63</td>
<td>4</td>
<td>3</td>
<td>0=0%</td>
</tr>
<tr>
<td>5</td>
<td>I can use the results of the questions to improve my teaching.</td>
<td>Strongly agree</td>
<td>Agree</td>
<td>Unsure</td>
<td>Disagree</td>
<td>Strongly disagree</td>
</tr>
<tr>
<td></td>
<td></td>
<td>35 = 38.88%</td>
<td>53</td>
<td>1</td>
<td>1</td>
<td>0=0%</td>
</tr>
<tr>
<td>6</td>
<td>AAs like oral presentation, doing conversations, oral discussion, writing assignments, self-assessment, peer-assessment, etc. allow my EFL students to show what they know.</td>
<td>Strongly agree</td>
<td>Agree</td>
<td>Unsure</td>
<td>Disagree</td>
<td>Strongly disagree</td>
</tr>
<tr>
<td></td>
<td></td>
<td>44 = 48.88%</td>
<td>43</td>
<td>3</td>
<td>0</td>
<td>0=0%</td>
</tr>
<tr>
<td>7</td>
<td>AAs allow my EFL students to show what they can perform.</td>
<td>Strongly agree</td>
<td>Agree</td>
<td>Unsure</td>
<td>Disagree</td>
<td>Strongly disagree</td>
</tr>
<tr>
<td></td>
<td></td>
<td>27 = 30%</td>
<td>53</td>
<td>9</td>
<td>1</td>
<td>0=0%</td>
</tr>
<tr>
<td>8</td>
<td>I can use the results of AAs to improve my teaching.</td>
<td>Strongly agree</td>
<td>Agree</td>
<td>Unsure</td>
<td>Disagree</td>
<td>Strongly disagree</td>
</tr>
<tr>
<td></td>
<td></td>
<td>34 = 37.77%</td>
<td>46</td>
<td>8</td>
<td>0</td>
<td>0=0%</td>
</tr>
<tr>
<td>9</td>
<td>AAs will be beneficial for my EFL students more than testing.</td>
<td>Very beneficial</td>
<td>Quite beneficial</td>
<td>Unsure</td>
<td>Slightly beneficial</td>
<td>Not beneficial</td>
</tr>
<tr>
<td></td>
<td></td>
<td>27 = 30%</td>
<td>39</td>
<td>20</td>
<td>3</td>
<td>1=1.11%</td>
</tr>
<tr>
<td>10</td>
<td>AAs will be practical for my EFL students more than testing.</td>
<td>Very practical</td>
<td>Quite practical</td>
<td>Unsure</td>
<td>Slightly practical</td>
<td>Impractical</td>
</tr>
<tr>
<td></td>
<td></td>
<td>19 = 21.11%</td>
<td>39</td>
<td>17</td>
<td>13</td>
<td>0=0%</td>
</tr>
<tr>
<td>11</td>
<td>AAs will be challenging for</td>
<td>Very</td>
<td>Quite</td>
<td>Unsure</td>
<td>Slightly</td>
<td>Unchallenging</td>
</tr>
<tr>
<td></td>
<td>challenging</td>
<td>challenging</td>
<td>challenging</td>
<td>g</td>
<td></td>
<td></td>
</tr>
<tr>
<td>-----------------------------</td>
<td>-------------</td>
<td>-------------</td>
<td>-------------</td>
<td>---</td>
<td>--------------------------</td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>16 = 17.77%</td>
<td>38 = 42.22%</td>
<td>13 = 14.44%</td>
<td>21 = 23.33%</td>
<td>2 = 2.22%</td>
<td>0 = 0%</td>
</tr>
<tr>
<td>I need a course, some sessions or workshops on conducting AAs.</td>
<td>Strongly agree</td>
<td>Agree</td>
<td>Unsure</td>
<td>Disagree</td>
<td>Strongly disagree</td>
<td></td>
</tr>
<tr>
<td></td>
<td>27 = 30%</td>
<td>51 = 56.66%</td>
<td>9 = 10%</td>
<td>1 = 1.11%</td>
<td>0 = 0%</td>
<td>2 = 2.22%</td>
</tr>
</tbody>
</table>
### Appendix (4): The themes, categories, and sub-categories of the answers to the open-ended questionnaire question

<table>
<thead>
<tr>
<th>Theme</th>
<th>Category</th>
<th>Sub-category</th>
<th>Explanation</th>
</tr>
</thead>
<tbody>
<tr>
<td>The participants’ assessment literacy in AA</td>
<td>The participants’ knowledge of AA</td>
<td>Conceptual knowledge of AA</td>
<td>The majority of the survey participants (71: 79%) did not define AA conceptually, and they did not mention the teaching, learning, critical, socio-cultural, or assessment theories that underpinned AA (see Section 5.3.1).</td>
</tr>
<tr>
<td>Beliefs of the participants on whether AAs are beneficial</td>
<td>Benefits of AA to teaching</td>
<td>Improving teaching methods</td>
<td>A few survey participants (6: 7%) talked about the influence of AAs on their teaching as they give a full picture of students’ performances, which help teachers improve their teaching methods (see Section 5.5.1).</td>
</tr>
<tr>
<td></td>
<td>Benefits of AA to learning</td>
<td>Improving students’ learning</td>
<td>The majority of the survey participants (71: 79%) talked about the general usefulness of AA for their students’ learning (see Section 5.5.2.1).</td>
</tr>
<tr>
<td></td>
<td>The assessment accuracy of AA</td>
<td>Some survey participants (14: 16%), in their attempts to define AA, demonstrated their understanding that, through AAs, a more comprehensive assessment is possible, since students have diverse ways to be assessed, and thus should not be limited to just one exam (see Section 5.5.2.2).</td>
<td></td>
</tr>
</tbody>
</table>
### Types, methods, activities, practices, and strategies of alternative assessment and its different terms in literature

<table>
<thead>
<tr>
<th>Types, methods, activities, practices, strategies, and/or different terms of alternative assessment</th>
<th>Categorisations, conceptualisations and definitions of alternative assessments</th>
<th>The researchers who have categorised, conceptualised, and/or defined alternative assessments</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Authentic assessment</strong></td>
<td>“Various types of assessment procedures for evaluating test takers’ achievement or performance using test tasks that resemble real-life language use as closely as possible” (Richards and Schmidt, 2010:42). “Engaging and worthy problems or questions of importance, in which students must use knowledge to fashion performances effectively and creatively. The tasks are either replicas of or analogous to the kinds of problems faced by adult citizens and consumers or professionals in the field” (Wiggins, 1993:229).</td>
<td>Richards &amp; Schmidt (2010) Wiggins (1993)</td>
</tr>
<tr>
<td><strong>Performance assessment</strong></td>
<td>“This involves teachers thinking about learners’ classroom performance to assess how well learners communicate during specific tasks by checking learners’ performance against criteria. Teachers can see if learners have achieved the purpose of the task by using the criteria” (Cambridge English Language Assessment, 2015:4). “An approach to assessment that seeks to measure student learning based on how well the learner can perform on a practical real-world task such as the ability to write an essay or carry out a short conversation. This approach is thought to be a better measure of learning than performance on traditional tests such as multiple-choice tests” (Richards and Schmidt, 2010:428). “Any tests that are designed to elicit performances of the specific language behaviors that the testers wish to assess. Examples of performance assessments that are not necessarily task-based assessments are composition tasks, oral interview tasks, and so forth. They are designed to elicit students’ abilities to write or speak, but they are typically scored in terms of the linguistic characteristics of the writing or speaking performances that the test designer feels are important for theoretical and/or pedagogical reasons” (Brown, 2004:92).</td>
<td>Cambridge English Language Assessment (2015) Richards and Schmidt (2010) Brown (2004)</td>
</tr>
<tr>
<td><strong>Dynamic assessment</strong></td>
<td>“Dynamic Assessment (DA) is a type of formative assessment which is viewed as an instrument to evaluate the present skill and then to predict future ability of the learners” (Caffrey, 2006 cited in Shakki et al., 2016:143). “DA is administrated by a helper who scaffolds and prepares social support for learning, so DA measures both independent performance and the mediated one. Independent performance shows what a learner has achieved alone whereas the mediated performance reflects what a learner obtained with the help of the mediator” (Poehner, 2005 cited</td>
<td>Shakki, Derakhshan and Ziabari (2016) Naeini and Duvall (2012) Shrestha and Coffin (2012) Lidz and Gindis (2003) Poehner (2007)</td>
</tr>
</tbody>
</table>
“Dynamic assessment concentrates on the learner’s errors which are studied in terms of the individual’s ongoing development and learning using mediations to promote growth. Lev Vygotsky (1978) is considered the originator of the theoretical framework upon which dynamic assessment is based; it emphasizes the interdependence of learning that leads development and thus the interlocking of instruction and development” (Naeini & Duvall, 2012:22).


<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>“When a teacher uses formal and informal assessment and information on learners’ progress during a course to give learners feedback on their learning or to change their teaching” (Cambridge English Language Assessment, 2015:4).</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>“Integrating instruction and assessment, formative assessment is a continuous cycle that entails gathering evidence of and judging student learning; providing feedback to students about their learning; and using assessment data to adjust subsequent instruction as needed” (Alvarez et al., 2014:2).</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>“A test that is given during a course of instruction and that informs both the student and the teacher how well the student is doing. A formative test includes only topics that have been taught, and shows whether the student needs extra work or attention.” (Richards and Schmidt, 2010:227-228).</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>“This is used for formative assessment and also continuous assessment. It consists of a collection of learners’ work done over a course or a year which shows development of their language and skills” (Cambridge English Language Assessment, 2015:4).</td>
<td></td>
<td>Bailey (1998)</td>
<td>Arter and Spandel (1992)</td>
</tr>
<tr>
<td>“An approach that makes use of portfolios as a basis of assessment” (Richards and Schmidt, 2010:444). [Portfolio is] a purposeful collection of work that provides information about someone’s efforts, progress or achievement in a given area. It is a learning as well as assessment tool. As applied to language learners, its characteristics are:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>a. the learner is involved in deciding what to include in the portfolio</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>b. the learner may revise material in the portfolio after feedback from the teacher or others</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>c. the learner is required to assess or reflect on the work in the portfolio, thus becoming aware of personal development</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>d. there is evidence of mastery of knowledge</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>e. it may include various forms of work, such as written work, audio recording, video recording, etc.” (Richards and Schmidt, 2010:443-444).</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>“A portfolio is a purposeful collection of students’ work over time that contains samples of their language performance at different stages of completion, as well as the students’ own observations</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Table</td>
<td>Definition</td>
<td>References</td>
<td></td>
</tr>
<tr>
<td>-------</td>
<td>------------</td>
<td>------------</td>
<td></td>
</tr>
<tr>
<td></td>
<td>“Checking one’s own performance on a language learning task after it has been completed or checking one’s own success in using a language” (Richards a&amp; Schmidt, 2010:517).</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>“Self-assessment is a process of formative assessment during which students reflect on and evaluate the quality of their work and their learning, judge the degree to which they reflect explicitly stated goals or criteria, identify strengths and weaknesses in their work, and revise accordingly (Goodrich, 1996; Gregory et al., 2000; Hanrahan &amp; Isaacs, 2001; Paris &amp; Paris, 2001; Andrade &amp; Boulay, 2003). With few exceptions, this definition of self-assessment excludes self-grading. Rather, student self-assessment is a process in which students collect information about their own performance and see how it matches their goals and/or the criteria for their work.” (Andrade and Du, 2007:160).</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>“Activities in which learners assess each other’s performance” (Richards &amp; Schmidt, 2010:425).</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>“An arrangement in which individuals consider the amount, level, value, worth, quality, or success of the products or outcomes of learning of peers of similar status” (Topping, 1998:250).</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Conference</strong></td>
<td>“Conferences, a special type of purposeful conversation or discussion between teachers and learners, can be regarded as a new form of evaluating students’ achievement in different educational settings. Genesee and Upshur (1996) argue that conferences involve both teachers and learners visiting each other in an office or classroom to discuss the students’ performance in their learning process. They stress that during a conference the focus of the instructor should be on the learners and their needs in the learning process they are experiencing” (Baleghizadeh and Zarghami, 2012:132).</td>
<td>Baleghizadeh and Zarghami (2012) Regier (2012) Richards and Schmidt (2010) Brown (2004) Brown and Hudson (1998)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>“Student understanding can be formatively assessed using one-on-one conferences with each student in your class or with select students for whom you want to further assess their learning. Determine your target questions ahead of time to ensure you are gathering information related to your goal or outcome. Take notes during the conferences to refer to later when planning instruction” (Regier, 2012:9).</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>“In teaching, a semi-structured face-to-face conversation between a teacher and a student or a small group of students in which work being undertaken is discussed. For example in a writing class a student may present a collection of his or her writing in a portfolio and discuss the selection in the portfolio, difficulties encountered, and strengths and weaknesses. The teacher gives feedback on progress, suggested improvements, etc.” (Richards and Schmidt, 2010:115).</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Conversation</strong></td>
<td>“Assessment conversations, or dialogic interactions or exchanges, which continuously happen in the classroom and that are at the center of informal formative assessment. It is argued that assessment conversations make students’ thinking explicit in an unobtrusive manner, and when students’ thinking is explicit, it can be examined, questioned, and shaped as an active object of constructive learning” (Ruiz-Primo, 2011:15).</td>
<td>Ruiz-Primo (2011)</td>
<td></td>
</tr>
<tr>
<td><strong>Debate</strong></td>
<td>“Debate refers to the process of considering multiple viewpoints and arriving at a judgment, and its application ranges from an individual using debate to make a decision in his or her own mind to an individual or group using debate to convince others to agree with them” (Freeley &amp; Steinberg, 2005 cited in Kennedy, 2007:183). “A game in which two opposing teams make speeches to support their arguments and disagree with those of the other team” (Krieger, 2005:1).</td>
<td>Freeley and Steinberg (2005) Kennedy (2007) Krieger (2005)</td>
<td></td>
</tr>
<tr>
<td><strong>Interview</strong></td>
<td>“An oral proficiency interview can be viewed as a type of adaptive testing in the sense that an interviewer (i.e. tester) adjusts the difficulty level of language on the basis of an evolving assessment of the interviewee’s (i.e. test taker’s) language ability” (Richards &amp; Schmidt, 2010:10). “This term is intended to denote a context in which a teacher interviews a students for a designated assessment purpose. Interviews may have one or more of several possible goals, in which the teacher assesses the students’ oral production, ascertains a student’s needs before designing a course or curriculum, seeks to discover a student’s learning styles and preferences, asks a student to assess his or her own performance, and requests an evaluation of a course” (Brown, 2004:265-266).</td>
<td>Richards and Schmidt (2010) Brown (2004)</td>
<td></td>
</tr>
<tr>
<td><strong>Discussion</strong></td>
<td>“Having a class discussion part way through a unit of study can provide you with valuable information regarding what your students know about the subject. Focus the discussions on higher level thinking skills and give students a few minutes to reflect on their learning before beginning the discussion. Encourage students to share what they have learned and how that knowledge may have an impact on their daily lives. Brainstorm ways that the knowledge could be transferred to other subject areas or situations the students may come across Listening carefully to the responses given by students will provide useful information for planning future instruction” (Regier, 2012:9-10). “An approach to teaching which consists of a goal-focused group conversation involving either groups of students or the whole class, and which usually involves interaction about subject matter between a teacher and students. Four common types of discussion procedures are used, which differ according to the degree of teacher control. 1. recitation: a teacher directed and highly structured discussion in which the teacher checks to see if students have learned certain facts 2. guided discussion: a less structured discussion in which the teacher seeks to promote understanding of important concepts 3. reflective discussion: the least structured form of discussion in which students engage in critical and creative thinking, solve problems, explore issues, etc. 4. small group discussion: the class is divided into small groups, with students assuming responsibility for the discussion” (Richards and Schmidt, 2010:177-178).</td>
<td>Regier (2012) Richards and Schmidt (2010)</td>
<td></td>
</tr>
<tr>
<td><strong>Open-Ended Question</strong></td>
<td>“A task or question that does not have a right or wrong answer, but which allows learners to offer their own opinions and ideas or to respond creatively”. (Cambridge English Language Assessment, 2015:30). “A test item which allows the person taking the test to answer in his or her own way, in contrast</td>
<td>Cambridge English Language Assessment (2015) Richards and Schmidt (2010) Çakır, and Cengiz (2016)</td>
<td></td>
</tr>
</tbody>
</table>
to questions with limited multiple-choice possibilities” (Richards & Schmidt, 2010:410).

“An open ended question is defined as a question to which a number of different answers would be acceptable, such as Why is Paris the most memorable place you have ever visited?” (Çakır, & Cengiz, 2016:62).

Think-Pair-Share

“The think-pair-share strategy is a great way to gather information about the level of understanding of your students. It is a quick and easy strategy that can be used a number of times throughout a unit of study. Ask students questions such as, “What did you learn during today’s lesson?” or “What connections can you make to your own life based on what you have learned so far?” Give students a few minutes to think about these questions. Pair students up with partners. Students share their thoughts with each other and then join a larger group or the whole class. Randomly call on students to share their ideas. By going through this process, students are able to solidify and refine their thinking before having to share their answers. Circulate throughout the class as students are sharing their thoughts and ideas to assess the overall depth of understanding” (Regier, 2012:17).

“Think-pair-share strategy is a strategy designed to provide students to think a given topic by enabling them to formulate individual ideas and share these ideas with another student. This strategy is a learning strategy developed by Lyman to encourage student classroom participation. The think-pair-share strategy is a cooperative discussion strategy to help students work in group. In applying this strategy, the lecturer poses a question, preferable one demanding analysis, evaluation, or synthesis, and gives students about a minute to think through an appropriate response” (Lyman, 1987 as cited in Usman, 2015:39).

Project

“An activity which focuses on completing an extended task or tasks on a specific topic. Learners may work in groups to create something such as a class magazine. Learners sometimes do some work by themselves, sometimes outside the classroom” (Cambridge English Language Assessment, 2015:35).

“(In teaching) an activity which centres around the completion of a task, and which usually requires an extended amount of independent work either by an individual student or by a group of students. Much of this work takes place outside the classroom. Project work often involves three stages:

1. Classroom planning. The students and teacher discuss the content and scope of the project, and their needs.

2. Carrying out the project. The students move out of the classroom to complete their planned tasks (e.g. conducting interviews, collecting information).

3. Reviewing and monitoring. This includes discussions and feedback sessions by the teacher and participants, both during and after the project. In language teaching, project work is thought to be an activity which promotes co-operative learning, reflects the principles of studentcentred teaching, and promotes language learning through using the language for authentic communicative purposes” (Richards & Schmidt, 2010:467-468).

“Projects are defined as assignments that incorporate student input, with content deriving from real second language use through extensive contact with either native speakers or native texts, integrating language skills, and extending over several weeks or more.” (Eyring, 1997:1).

Essay writing

“The essay is defined as a short literary composition on a subject expressing a personal view. The
An essay consists of the following parts:

- The introduction.
- The body.
- The conclusion (Ibnian, 2011:263).

“A subjective test in which a person is required to write an extended piece of text on a set topic” (Richards & Schmidt, 2010:203).

**Editing**

“The practices in second language writing classes of engaging students in activities that require correction of discrete language errors in their writing, such as errors in grammar, vocabulary, sentence structure, spelling, etc.” (Richards & Schmidt, 2010:189).

“To shorten or change or correct the words or content of some parts of a written text to make it clearer or easier to understand; e.g. learners might edit the first draft of a text they have written to correct the mistakes. See process writing.” (Cambridge English Language Assessment, 2015:14).

**Summarising**

“The ability to produce summaries is sometimes referred to as summary skills and is a focus of instruction in the teaching of reading”. [Summary is] “brief statements of the main ideas in a text or passage, often produced while or after reading something” (Richards and Schmidt, 2010:573).

“To take out the main points of a long text, and rewrite or retell them in a short, clear way. Learners preparing to study at university need to practise summarising skills.” (Cambridge English Language Assessment, 2015:44).

**Diary/journal/Learning journal/Learning log**

“Learning logs are notes students make during a unit of study. Time is set aside at the beginning or end of class for students to write about what they have learned, list any questions about the topic they may have, or make connections between the topic and their own lives. Learning logs provide you with valuable information about what students are learning and possible directions for future instruction. Using learning logs as a formative assessment strategy provides you with information about student learning and what information or skills students may still need to reach the goals and outcomes of the unit” (Regier, 2012:12-13).

“[It is] also learner journal”, “in language teaching, a record prepared by a learner of a student’s learning experiences and describing what activities they have done in class, the progress they have made, and any problems they may have” (Richards & Schmidt, 2010:327).

“[It is] also journal, learning journal”, “the use of a notebook or book in which students write about experiences both in and out of school or record responses and reactions to learning and to learning activities. Learning logs provide students with an opportunity to reflect on learning, and are usually shared with the teacher on a regular basis but not graded. In this way, the teacher may be able to find out how the student is progressing and the students gain additional opportunities to practise writing. In writing classes learning logs may be used as a prewriting activity (see composing processes) and also as a way of encouraging students to develop fluency in writing through writing regularly on topics of their own choice. When learning logs are used as a way of establishing a dialogue between teacher and student (through comments, questions and reactions), they are sometimes referred to as dialogue journals or diaries.” (Richards and Schmidt, 2010:329).

“A journal is a log (or “account”) of one’s thoughts, feelings, reactions, assessments, ideas, or progress toward goals, usually written with little attention to structure, form, or correctness.
Learners can articulate their thoughts without the threat of those thoughts being judged later (usually by the teacher)” (Brown, 2004:260).

<table>
<thead>
<tr>
<th>The categorisation of AA into two basic types</th>
<th>Brown and Hudson (1998) categorised AA into two basic types: constructed-response (e.g. short answers and performance assessments), and personal-response (e.g. portfolio, self-assessment, peer-assessment and conference).</th>
</tr>
</thead>
<tbody>
<tr>
<td>The categorisation of many AA's as types of AA</td>
<td>Self-assessment, portfolio assessment, student-designed tests, learner-centred assessment, projects, presentations (Coome et al. 2007, cited in Azarnoosh, 2013), holistic assessment, integrative assessment (Chirimbu, 2013), performance assessment, peer-assessment (Azarnoosh, 2013) and direct assessment (Grabin, 2009) are regarded as types of AA.</td>
</tr>
<tr>
<td>The categorisation of many AA's as methods of AA</td>
<td>Conferences, debates, demonstrations, diaries/journals, dramatisations, exhibitions, games, observations, peer-assessment (Tsagari, 2004), group assessment (Grabin, 2009), self-assessment, projects, story retelling, thinking-aloud (Tsagari, 2004), inventories (Hamayan, 1995), oral reports, role plays, describing, explaining, summarising, retelling, paraphrasing stories and paraphrasing text materials (Tannenbaum, 1996), observation, self-report, questionnaire, interview and elicitation (Bachman and Palmer (2010) are all seen as methods of AA.</td>
</tr>
<tr>
<td>The categorisation of many AA's as activities of AA</td>
<td>Creative writing, correspondence essays, writing in response to prompts, learning logs, learning journals, classroom projects, interviews, and think-aloud (Hamayan, 1995) are considered as activities of AA.</td>
</tr>
</tbody>
</table>
Appendix (6): First-phase interview questions

1. Are you familiar with the phrase alternative assessments (AA)?
2. What do you believe is the difference between AA and traditional testing?
3. Are there any principles and standards for conducting AAs in your English department?
4. What do you believe is the purpose of your testing and assessment of EFL students?
5. Which of the two, testing or AA, do you think is more accurate to show what your EFL students know?
6. Which of the two, testing or AA, do you think is more accurate to show what your EFL students can perform?
7. Which one do you think is more beneficial for your EFL students, AA or traditional testing?
8. Which one do you think is fairer for your EFL students, AA or traditional testing?
9. How do you incorporate testing into the classroom teaching and learning?
10. How do you incorporate classroom assessments into the classroom teaching and learning?
11. Do you believe that teacher must be the single assessor or the single tester of students?
12. Do you believe that your EFL students can assess themselves or each other?
13. Do you think that your EFL students should discuss ideas openly and informally with each other and with their teachers about how to test and assess them?
14. Which one do you think is more advantageous: assessing students on what they can perform or on what they can recall?
15. Which one do you think is more advantageous: to focus on students’ strengths or weaknesses?
16. Which of the two, testing or AA, do you think will improve your teaching?
17. Generally, do you think that using various AA types can improve your EFL students’ learning?
Appendix (7): Slides used in the seminars

The 2nd phase Action Research Intervention: The Seminar

The conduction of alternative assessment by TESOL teachers at the tertiary level

Slide 2: Professional development of language assessment

Language assessment has been increasingly professionalized. Quality standards, ethical codes and good testing practice guidelines have been generated (Taylor, 2009).

Testing and assessment

Testing is a universal facet of social life that throughout history people were tested to prove their abilities (McNamara 2000:3).

Traditional assessment merely samples “tiny snippets of student behaviour” but alternative assessments look at a bigger picture of students’ performance (Worthen, 1993:445 as cited in Grabin, 2009).

Schmitt (2002:267) distinguishes assessment from testing that assessment is used in a broader sense.

Assessment includes a much broader cycle of practices (Green, 2013).

The situation of the English departments

Following the traditional ‘pen and paper’ tests in the English departments of the Kurdistan Region’s universities

Almost all teachers carry out assessments but without having learned the principles of sound assessment (Coombe’s et al., 2012:20).
<table>
<thead>
<tr>
<th>Unawareness of alternative assessments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Neglecting or being unaware of the recently developed alternative assessments might constitute an enormous problem.</td>
</tr>
<tr>
<td>The reforms of quality assurance of MHESR/KRG has developed higher education and empowered students but it is not quite adequate.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Having testing systems rather than assessment systems</th>
</tr>
</thead>
<tbody>
<tr>
<td>The English departments of the Kurdistan Region universities possibly have testing systems rather than assessment systems.</td>
</tr>
<tr>
<td>This is because alternative assessments are not carried out properly; whereas, what is mostly carried out and scored is testing.</td>
</tr>
</tbody>
</table>
Special skills and assessment literacy for doing alternative assessments

Special skills that teachers need to have in order to implement alternative assessments, is a major concern (Tsagari, 2004).

These skills can be acquired by assessment literacy and training.

The emergence of alternative assessments

Alternative assessment refers to a movement particularly in the USA school contexts that tried to be away from using standardized tests in favour of performance assessments (McNamara, 2001).

Why do some students perform very well but cannot get high grades from the tests (Ajideh and Nourdad, 2012)?

The limitations of teacher assessments as a single assessor increased the interest in alternative assessments (Matsuno, 2009).
The conceptualization of alternative assessments

There is no single definition of alternative assessments (Tannenbaum, 1996).

Generally, alternative assessments are ongoing processes that involve students and teachers in making judgments on students’ progress by using non-conventional strategies (Hancock, 1994).

Some significant facets and characteristics of alternative assessments

The core characteristics of alternative assessments are performance and authenticity.

Multiple resources of alternative assessments can be used for getting data on students’ different abilities.

A nearly comprehensive assessment of student’s knowledge and abilities

Instruction and assessment are intrinsically integrated (Ripley, 2012).

Revealing students’ underlying thinking processes (Lynch and Shaw, 2005).

Requiring higher level thinking and problem solving (Derakhshan et al., 2011).
Alternative assessments are intermixed

There are different terms for the alternative assessment types, methods and practices.

Almost all alternative assessment types, methods and practices are intermixed together and some of them take place at the same time.

Portfolio, which is a type of alternative assessment, covers approximately all sorts, methods and procedures of alternative assessments in addition to traditional assessment methods.

Peer-assessment enhances self-awareness, that is, noticing the gap between one's and others' perception (Azarnoosh, 2013).

Self-assessment and peer-assessment, which are methods of alternative assessments, can be applied to the written assignments, oral presentations, etc., which are procedures of alternative assessments; and all these can be included in a portfolio.

The innovative types, methods and practices of alternative assessments and their different terminology

It is not easy to determine which one is a type, a method, a procedure, a practice or an activity of alternative assessment because they are classified and labelled differently by different scholars and researchers.

Some types of alternative assessments or just different terms with slight differences:

The innovative types, methods and practices of alternative assessments and their different terminology

Some methods or practices of alternative assessments: conferences, debates, diaries, observations, peer-assessment, group assessment, projects, self-assessment, story retelling, think-alouds, oral reports, role plays, creative writing, correspondence essays, writings in response to prompts, interviews, open-ended questions, oral proficiency interview, task-based test, home exam (14 days open books), poster exam (individual or group), presentation (individual or group), reflective journal, research project (mini), short answer questions, performance task (individual or group)

These types, methods and practices of alternative assessments share some basic features, though they have subtle differences in their focus (Hamayan, 1995).

Ways of recording alternative assessments: anecdotal records, checklists, audio and video recordings, rating scales, rubrics.

Practical challenges of conducting alternative assessments

Law and Eckes (2007) obviously mention that alternative assessments might require lots of thought, space, logistics and training.

The subjective scoring of alternative assessment (Brown and Hudson, 1998) might raise equity concerns (Baker, 2010).

The huge obstacle of large numbers of students and the time of doing alternative assessments (Tsagari, 2004)

However, there are some virtual solutions for doing alternative assessments in large classes.
Appendix (8): Third-phase interview questions

1. Has the seminar on alternative assessments (AAs) increased your knowledge on them?
2. Has the seminar on AAs changed your views on them?
3. Now, after the seminar, what do you think of the importance of AA for ELT?
4. Now, what do you believe is the difference between AA and traditional testing?
5. What do you believe is the purpose of your testing and classroom assessment of your EFL students?
6. What do you think of the effort of doing AAs in large classes including their scoring and documentation?
7. What do you think of the time of doing AAs in large classes?
8. What do you think of the cost of doing AAs?
9. What do you think of the fairness of AAs?
10. If the number of students is reduced to 25 students per class, if the number of teaching hours is also increased, and if there are enough teachers, do you think that this can be a good solution for doing AAs and thus enhancing students’ learning of English?
11. Do not you believe that if you do many assessments in the classroom, there will be less test papers and less assessment will remain to do at home?
12. Finally, what is your general view of AAs now?
13. Do you want to learn more about AA practices and train yourself to do them more successfully?
14. Do you want to devote more time for doing more AAs in the classroom?
Appendix (9): Issues discussed in the focus group interviews

1. How to incorporate more alternative assessments in order to achieve a sound assessment system.
2. Redistributing students into more groups to have 25-30 students per class when they have exercise sessions.
3. Redistributing the marks equally for tests and alternative assessments (AAs).
4. Providing teachers with resources on AA.
5. Organising some workshops for training teachers on conducting AA.
Title of Research Project

The Assessment Practices of In-service Kurdish Tertiary TESOL Teachers and their Cognitions of Alternative Assessment

Details of Research Project

This research project is divided into 3 phases. The first-phase aims at exploring the current assessment practices of the in-service Kurdish tertiary TESOL teachers and their knowledge of alternative assessments (AAs) by doing survey questionnaires and interviews. This is to know and understand their assessment practices and to what extent these include AA practices, and to know the gaps of their knowledge of AAs. This understanding of the first-phase is going to be a basis for subsequent intervention in the second-phase in which I present two seminars on AAs. The seminars are for filling in the gaps of their knowledge of these assessments. Eventually, by doing interviews and focus group interviews, the third-phase investigates the effect of the intervention on the teachers’ knowledge and views and then their cognition of the benefits, the feasibility and the challenges of conducting AAs in their professional context. Additionally, the findings of this study can possibly provide a clear understanding of how and to what degree the teacher participants are willing to do more about AAs in the future.
Contact Details

For further information about the research or your data, please contact the researcher:

Mobile: UK: 00 44 (0) 77 8494 7676, Kurdistan Region/Iraq: 00 964 (0) 750 834 2447

Email: daii201@exeter.ac.uk

Graduate School of Education, College of Social Sciences and International Studies, Exeter University, Devon, UK.

If you have concerns/questions about the research you would like to discuss with someone else at the University of Exeter, please contact:

The first supervisor: Dr Salah Troudi

Email: S.Troudi@exeter.ac.uk

or the second supervisor: Dr Susan Riley

Email: S.M.Riley@exeter.ac.uk

The Teachers’ Participation

To participate in this study, you should be an in-service Kurdish tertiary TESOL teacher. Your participation in this study is voluntary. You can stop participating whenever you want by contacting the researcher. Also, you can ask for deleting your data at any time. The methods of this study keep the participants’ anonymity and would not reveal their identity.

The researcher’s requests to the teachers are according to the 3 phases of this study. In the first-phase, the researcher requests the in-service Kurdish tertiary TESOL teachers to supply data on their current assessment practices and their knowledge of AAs by survey questionnaires and interviews.

In the second-phase, the researcher requests the teacher participants to attend a seminar on AAs in each of the two selected English departments.
For the third-phase, the researcher interviews the same 12 interviewees of the first-phase who also attended the seminars. Then the researcher holds two focus group interviews for the same 12 interviews. This is to understand whether the seminars increased the participants’ knowledge of AA, and improved their views about AA and to explore the participants’ views on the benefits, the feasibility and the challenges of conducting AAAs in their professional context.

**Confidentiality and Anonymity**

All the questionnaire responses and the tapes and transcripts of the interviews and focus group interviews are to be held in confidence. They will be accessed and used only by the researcher for the purposes his study, mentioned above. The third parties can be allowed only, if it is required by the law. Then, as a participant if you would like, the researcher provides your given data to you so that you can comment on it and edit it as you see fit (please give your email below). Your data will be held confidentially and indefinitely on an anonymous basis in accordance with the Data Protection Act. So, there is no mention of your name but we will refer to the group of which you are a member.
Appendix (11): Consent form

GRADUATE SCHOOL OF EDUCATION

Statement of Consent

I will participate in the research instruments of this study, entitled “Assessment Practices of the In-service Kurdish Tertiary TESOL Teachers and their Cognition of Alternative Assessment” by the EdD researcher Dler Abdullah Ismael. I have voluntarily agreed to participate in this study and allowed the researcher to use my data for the purposes of his study.

I have been fully informed about the aims and purposes of the research project.

I am aware of and understand that:

1. 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 and may also request that my data be destroyed,

2. I have the right to refuse permission for the publication of any information about me,

3. any information which I give will be used solely for the purposes of this research project, which may include publications or academic conference or seminar presentations,

4. all information I give will be treated as confidential and

5. the researcher will make every effort to preserve my anonymity.

This informed consent form was reviewed and approved by the researchers’ supervisors and the Ethical Approval Committee of the Graduate School of Education/SSIS at the University of Exeter in 18/12/2014 and it will expire in 30/5/2015.
Signature of the participant

...........................................

Date:

If you have any inquiries, you can contact the first supervisor: Dr Salah Troudi, email: S.Troudi@exeter.ac.uk, the second supervisor: Dr Susan Riley, email: S.M.Riley@exeter.ac.uk or the researcher: Dler Abdullah Ismael, email: daii201@exeter.ac.uk.

2 copies to be signed by both the participant and the researcher, one kept by each

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 (12): Application for a Certificate of Ethical Approval

Graduate School of Education

Certificate of Ethical Approval

EdD Thesis

To activate this certificate you need to first sign it yourself, and then have it signed by your supervisor and finally by the Chair of the School’s Ethics Committee.

For further information on ethical educational research access the guidelines on the BERA web site: http://www.bera.ac.uk/publications and view the School’s Policy online.

Your name: Dler Abdullah Ismael

Your student number: 590064924

Return address for this certificate: Exeter Mansion One, Block C, Flat 11, Exeter, EX4 5FH

Degree/Programme of Study: EdD TESOL

Project Supervisor(s): Dr Salah Troudi, Dr Susan Riley

Your email address: daii201@exeter.ac.uk

Tel: UK: 00 44 (0) 77 8494 7676, Kurdistan Region/Iraq: 00 964 (0) 750 834 2447

I hereby certify that I will abide by the details given overleaf and that I undertake in my thesis to respect the dignity and privacy of those participating in this research.

I confirm that if my research should change radically, I will complete a further form.

Signed: Dler Abdullah Ismael date: 25/11/201
Title of Your Research Project:

Assessment Practices of the In-service Kurdish Tertiary TESOL Teachers and Their Cognition of Alternative Assessment

1. Brief description of your research project:

This study aims at exploring the current assessment practices of the in-service Kurdish tertiary TESOL teachers and their cognition in alternative assessments (AAs). This is to know and understand their assessment practices and the gaps of their knowledge of AAs. This understanding is going to be a basis for two subsequent seminars on AAs. The two seminars are for filling in the gaps of their knowledge of these assessments. Finally, this study investigates the effect of the seminars on the teachers’ knowledge and cognition of conducting AAs. The findings of this study can possibly provide a clear understanding of how and to what degree the teacher participants are willing to carry out AAs more successfully in the future.

2. Give details of the participants in the research:

All the participants are divided into the stages of this study as follows:

The participants of the first-phase survey questionnaires include the majority of the in-service Kurdish tertiary TESOL teachers of all the 18 English departments of the Kurdistan Region’s universities (except evening classes English departments as they have the same staff as the morning classes). Also, the first-phase interviews are 12 in-service Kurdish tertiary TESOL teachers from the two selected English departments.

The second-phase participants of the seminars are the 12 interviewees of the first phase plus some other teacher attendants.

The third-phase interview participants are the same 12 in-service Kurdish tertiary TESOL teachers of the first phase that are chosen from the two selected English departments. Finally,
the third-phase focus group interview attendants are the same 12 teacher interviewees who were previously interviewed in the first and third phases.

**Given details regarding the ethical issues of:**

3. **Informed consent:** Copy(ies) of your consent form(s) you will be using must accompany this document

All the questionnaire survey, interview and focus group interview participants were requested to sign the prepared consent form of this study to take part in its research methods. The consent form of this study accompanies a brief explanation on all the research instruments for the participants to be aware of those methods before signing their consent form.

4. **Anonymity and confidentiality**

   It is mentioned in the consent form that the participants’ identity will not be revealed to anybody (except if it is required by the law). Regarding confidentiality, the participants are informed that their data will be used only by the researcher, and it will not be shared with the third parties.

5. **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:**

The first-phase survey questionnaires will be distributed to the in-service Kurdish tertiary TESOL teachers of all the 18 English departments of the Kurdistan Region’s universities. The researcher just requests them to supply data on their current assessment practices and their knowledge of AAs voluntarily. This does not cause any personal harm to them, in case their anonymity and their confidentiality is kept properly. For the first-phase interviews, 12 in-service Kurdish tertiary TESOL teachers from the two selected English departments are asked to take part in the interviews voluntarily on the same above-mentioned issues. Here also, their identity will be kept secret and their confidentiality will be guaranteed, so they will not possibly be harmed.

The second-phase participants of the seminars are 12 interviewees of the first phase and some other in-service Kurdish tertiary TESOL teachers in the two English departments. Their
attendance is voluntary and they will certainly get some benefit from the seminars on AA with no harm.

Finally, the participants of the third-phase interviews and focus group interviews who are the same 12 interviewees of the first phase who also attended the seminars as well, were requested to participate in the interviews again voluntarily with keeping their anonymity and confidentiality appropriately. The anonymity and confidentiality of all the participants are guaranteed by the researcher. In addition, the researcher will not oblige any participant or use his personal relationships with them to force them to take part in the research methods of his study. Thus, all the participations are voluntary.

6. Give details of any other ethical issues which may arise from this project - e.g. secure storage of videos/recorded interviews/photos/completed questionnaires, etc.

The storage of the hard copy completed questionnaires of the first phase will be kept in my personal drawer. The scanned copies of all the questionnaires will be kept in my personal laptop. The interview and focus group interview recordings and transcripts of the first and third phases will be kept in my personal laptop as well. The recording of the seminar will be kept in my personal laptop too. My personal drawer is usually locked and nobody opens it without my permission and my personal laptop has a password to open and the folder of my EdD research has another password. So, the whole collected data will be preserved and secured properly.

7. Special arrangements made for participants with special needs etc.

Not applicable

8. 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):

Not applicable

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:** 18/12/2014        **until:** 30/5/2015

By supervisor:

Dr Salah Troudi:        date: 18/12/2014

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:

Signed:                date: 18/12/2014

Chair of the School’s Ethics Committee
Appendix (13): Certificate of Ethical Approval

CERTIFICATE OF ETHICAL APPROVAL

Title of Project: Assessment Practices of the In-service Kurdish Tertiary TESOL Teachers and Their Cognition of Alternative Assessment

Researcher(s) name: Dler Ismael

Supervisor(s): Dr Salah Troudi
Dr Susan Riley

This project has been approved for the period

From: 18/12/2014
To: 30/05/2015

Ethics Committee approval reference:

D/15/16/36

Signature:  
Date: 18/12/2014

(Dr Philip Durrant, Chair, Graduate School of Education Ethics Committee)
Appendix (14): Transcript of one interview showing how it was coded

An interview

DLER: Thank you so much for participating.
DILDAR: Not at all, you’re welcome.
DLER: We are going to discuss alternative assessments in your classrooms.
DILDAR: Right.
DLER: First of all, are you familiar with the term alternative assessment?
DILDAR: Actually, I’m familiar with different kinds of assessments inside class; it is based on the topic and it is based on the lecture. (Limited knowledge of AA)
DLER: Oh, yeah. Of course. So do you think that you have read about, do you remember you have read about AA, articles, books?
DILDAR: I haven’t read this kind of assessment, but I know there are several kinds of assessment inside the class. the teacher can conduct, for example, we may have, for example, objective, subjective, different kinds of assessment. (Limited knowledge of AA)
DLER: OK, so do you remember you have studied it anywhere?
DILDAR: I studied different kinds of assessment in general.
DLER: Not alternative?
DILDAR: Not alternative specifically but I studied several kinds, and how do you design an assessment for the students. (Not part of their MA or PhD courses)
DLER: OK. So what do you think of the difference between assessment and testing, AA and testing?
DILDAR: AA and testing, testing is not totally different from assessment, they can be seen as synonyms, and they can be seen as alternatives, but when we talk about, for example, testing; we may have several kinds of testing or tests, but AA, teachers, for example, may have several kinds of assessments for the students. This assessment, for example, is different from that assessment. (Limited knowledge of the differences between AA and testing)
DLER: OK. So just I want to I want to remind you that when I say testing, I mean paper and pencil testing, and when I say AA, I mean every assessment that you do in the class,
DILDAR: Right.
DLER: their performance, students’ performance.
DILDAR: Right, right.
DLER: So, what do you think is the difference?
DILDAR: The difference, as you said before, we may have, for example, paper-based testing, we may have internet-based testing, or we may have different kinds of assessment inside the class. For example, I may design, for example, a test just for oral assessment; I want to know how, for example, the students pronounce different words; this is a kind of oral assessment. (Limited knowledge of the differences between AA and testing)

DLER: Right, of course. So er, er do you think, now you have differentiated for me, the differences between AA and testing, er do you think that your assessment literacy, your knowledge about assessment is adequate?

DILDAR: It is not, I can say, it is not completely adequate about assessment; as I said before, I haven’t read much about, for example, these two... things, but I know how do you, for example, design a test, and how do you... how many kinds of assessments do we have, and how many kinds of tests, do we have in English. (Acknowledging their limited knowledge of AA) When I say for example, when we talk about, for example, testing, we may have achievement tests, we may have proficiency tests, different kinds of assessments... tests, sorry. But when we say assessment, the teacher, for example, may employ different ways in order to assess the students inside the class, for example, you may have oral, you may have written, or any kinds; how do you assess your students? (Limited knowledge of AA)

DLER: OK, how, yeah, how many, which of these assessments you do in your classroom?

DILDAR: For example, er I may have oral assessment, I may have, for example, writing assessment... (Limited knowledge of AA) (Performing different kinds of AAs and test)

DLER: Conversations in the class...?

DILDAR: Conversations inside the class, it is based on the topic and the lecture, for example...

DLER: Of course, we are talking about it generally, in all topics and all lectures.

DILDAR: Right now I’m teaching different subjects, for example, I’m teaching grammar and essay writing; it is sometimes based on your experience how do you, for example, assess your students, you may have, for example, er some, for example, points in your mind, you want to design tests, and you want to assess their performances through your points and your experiences, and have this assessment for today, I may have a different assessment for tomorrow... (Limited knowledge of AA) (Their conducting of AAs based on their experience) (Performing different kinds of AAs and test)
DLER: So, which one also do you remember, written, you said, written assignments and you said the one was conversations in the classroom, what else do you remember? Your students do presentations?

DILDAR: Yes, they usually do presentations based on the topics, for example, I usually give them different topics I ask them to prepare, for example, assignments, after they prepare their own assignments, they have to do presentations differently, and each one has its own or assessment, for example, four students may prepare one presentation, I try to assess each student in the group, individual assessment and group assessment. *(Performing different kinds of AAs and test)*

DLER: Wow, you do group assessment?

DILDAR: Yes, group assessment.

DLER: You assess the whole group?

DILDAR: The whole group in general, for example, do they perform well or not.

DLER: Oh, that is very good.

DILDAR: Does she or does he perform well or not?

DLER: So how many? I’m asking about the frequency of these assessments, how many times or how often your students do presentations or participating in conversation throughout the year, academic year?

DILDAR: Twice or three times because we don’t have enough time, we don’t have much time, you know, to perform, the students to give them enough time to present different subjects, you know, they need enough time, yes, they have to have enough time, yes?

*(Limited time for AA) (Infrequency of AA throughout the academic year)*

DLER: Of course, yeah. So if we come back to the topic of the differentiation between AA and testing; I just want to ask you one... I want your opinion about this because this is related to this differentiation. What about that sometimes some students are very good in their performance whether speaking, writing, listening, anything, but they are very bad in the results of their tests, pen and paper tests? What do you think of this?

DILDAR: I think it is different some students, for example, it based on language skills, as you know we have four skills we have reading and writing, listening and speaking...

DLER: Right.

DILDAR: When you, for example, when I design a test for writing, the students may not have, for example, enough vocabulary; I ask them to write an essay, but you see they can’t write well essay, well-performed, well-organised essay.
DLER: So, you mean the level...?
DILDAR: The level, yes.
DLER: Of the students is a problem?
DILDAR: It means that, for example, yes, the level of the student, and for example, the knowledge about the topic sometimes.
DLER: So do you think if he or she is bad in writing, for example, in the classroom; so do you think that in testing maybe she or he can be good as well?
DILDAR: In assessment not in testing, as I said before... She is not good in testing, but she is good in assessment; she has an oral assessment, she can speak or he can speak English very well, he can, but he can’t or she can’t write. She is good in this skill, but she is not in that skill. (Testing and AA misunderstood)
Dler: OK, OK, that is very good. Em is there any principles or standards, or guidelines that the department gives you to em to follow AA or your testing?
DILDAR: Actually, there are not any guidelines from the department... (No guidelines or standards of AA from the department)
DLER: Do you think they are important?
DILDAR: It is, yes, it is important to know how you, for example, design your assessments, and how do you, for example, demonstrate your assessments inside the class, and how do you design your tests; it is important, but as you know, we have several books in applied linguistics... (Understanding of the importance of assessment guidelines and standards)
DLER: Yeah, but here, I mean the department (administrators), whether they give you or not, whether they guide you or not?
DILDAR: They haven’t given us any guidelines for our assessments or tests, but they usually, you know, advise us or they tell us to do quizzes or assessments, presentations, we have to encourage the students...
DLER: It is just saying, no guidelines?
DILDAR: Yes saying, no guidelines. (Understanding of the importance of assessment guidelines and standards)
DLER: So em what is, generally, what is the purpose of your testing and AA? What is the purpose behind it?
DILDAR: The purpose behind, for example, my testing is to know how my students understand my lecture; how they, for example, can they understand the topics? Or did they understand the topics that I have come across or not, but for the assessment, er the purpose
behind my assessment is to know do they understand each point separately in my lecture. For example, I have taught several points and several, for example, headings in my lecture, if I teach, for example, present simple, how many, for example, the structure, I just want to assess each point separately; I have individual and group assessment... *(Knowledge of the purpose of AA: the amount of students’ learning)*

DLER: In the classroom?

DILDAR: In the classroom, if I teach, for example, essay writing, I have group assessment; if I teach, for example, grammar, I have individual assessment.

DLER: So how how do you group assess?

DILDAR: Group assess, for example, I give them a topic, it is task-based; I give them a topic, I write down the title and some vocabularies on the whiteboard, and I always right down some questions, after we demonstrate the questions, they answer the questions, and they know the meaning of vocabularies, and they know about the topic, I give them a model, after they look at the model for a while, they have to return to the model, and they, each group should try one essay. *(Performing different kinds of AAs and test)*

DLER: OK, when you assess in the classroom, do you use rubrics and checklists?

DILDAR: Sorry? *(No knowledge about checklists and rubrics of AAs)*

DLER: We have rubrics and checklists; we have the names of the students, and we have some er areas of learning, for example, when you teach writing, we have grammar, we have spelling, we have punctuation, cohesion, coherence, so we have these items, and we say that student is very good in that, but is not good in this.

DILDAR: yes,

DLER: Do you use checklists with your students?

DILDAR: Yes, I usually use checklists, especially in essay writing; in grammar as well.

DLER: So you have checklists?

DILDAR: Yes, I have a checklist for, for example, spelling mistakes, grammatical mistakes, punctuation mistakes... *(Using a simple checklist)*

DLER: So you mentioned that, for example, a student, you can say (to one) ‘your spelling is good, and this is your result, your grammar is good but your punctuation is bad’?

DILDAR: Is bad, yes.

DLER: Do you have these lists for students?

DILDAR: Yes, I have a list.

DLER: That is very good. OK,
DILDAR: And I usually write my feedback on the results. *(Feedback on the results)*
DLER: Oh, yeah. Yeah, and you give your feedback to the students?
DILDAR: To the students, and I have, for example, abbreviation, for example, I write down ‘G’ it means grammatical mistakes, for example, spelling mistake ‘SPM’. *(Using abbreviations for giving feedback on written texts)*
DLER: Wow, that is very good, that is very good. Now we come back to the comparison between testing and AA. Which of these you think is more accurate to show students’ level in performance and in knowing the subjects? Which one?
DILDAR: I think they both are good, both are good for the students; the teacher should conduct both of them; there are not great differences especially for EFL learners, you know *(Considering both testing and AA as accurate and good)*, but I think testing is better than assessment because we have not had enough time to assess individual students, as you know, we have lots of students inside the class, we may have 50 students. *(Limited time for AA)* *(High number of students per class)*
DLER: but what about when...?
DILDAR: Do I have enough time to assess each student inside the class?
DLER: Yeah, yeah, that is right.
DILDAR: It is difficult.
DLER: Yeah, this is quite right.
DILDAR: But when you design a test, all students can participate in the test, and then you see the results, their performance. *(Practicality of testing)*
DLER: Of course, yes, but what do you think of, for example, a student is very good in performance when he or she speaks while in testing he is not good, so which one is more accurate to show the level of this student?
DILDAR: I think both of them are accurate to show the level of the students, *(Considering both testing and AA as accurate and good)* but as the teacher we have a coursebook, I usually explain everything when I start, for example, my first lecture with the students I have a coursebook, I show them each point separately, what are the objects? What is the, for example, an overview of my lecture, and then we move to explain, for example...
DLER: The testing?
DILDAR: The assessments... we have daily assessment, we have monthly assessment, for example, daily assessment is out of five or ten *(Telling simple things to students about test and assessment designs)*
DLER: Yeah, yeah, actually we come to this point at the end, the scoring of assessment. For now which one of these two, testing and AA, do you think is fairer, is more ethical?

DILDAR: I think there is not fairness here, as I said before, *(No fairness of assessment in participants’ context)*

DLER: Because... Why why there is no fairness?

DILDAR: Because, as I said, we don't have enough time to assess or to test each student...

DLER: So if you have time? *(No fairness of assessment in participants’ context due to the limited time for AA)* *(No fairness of assessment in participants’ context due to the high number of students per class)*

DILDAR: Completely or appropriately, if I have time, yes.

DLER: Which one is fairer, if you have time?

DILDAR: I think they both [are fair]. *(Considering both testing and AA as fair)*

DLER: Both are...

DILDAR: Yes, yes, we can't say this one is better than, completely, but as I said before, testing is, I think, better. *(Considering testing is better because of its practicality)*

DLER: Yeah. So we come to that, how do you, how do you incorporate your testing and your classroom assessment into your teaching and learning, that is, how you linked them?

DILDAR: Testing and assessment?

DLER: Both of them how you linked them to your teaching and students’ learning?

DILDAR: You know, I have to have my procedures, I usually have special procedures to link, to connect them together inside the class...

DLER: How?

DILDAR: For example, as I said before, I have monthly assessment and I have daily, for example, I have monthly testing, and er er this test is out of this, and I have oral assessment; I usually link them like this. *(Weak integration of AA into teaching and learning)*

DLER: Yeah, this is assessment together; how do you, how do you link these assessments to your teaching, do you think that your teaching is influenced by these, the results of the assessments?

DILDAR: Yes, it is influenced, for example, if I encourage my students, for example, to do assessment inside the class, if they have, if they know they have assessment, they have oral assessment, they have other kind of assessment, they have...

DLER: So what do, what do they do?
DILDAR: They encourage, they study more and more. Yes, they try to prepare... (Benefits of AA for learning)
DLER: Now, yeah, I think, I think you reached the point actually, yeah. And how you link these assessments, assessment activities, and testing with the real world situations? Do you have this kind that your assessment resembles a real world situation?
DILDAR: It should be like this. (Connecting assessment to real world situations)
DLER: But you have not done that?
DILDAR: I have I have done that, I usually try to make a link, to make connection between my lecture and the real world, if I teach, for example, a topic, I try to find for example, something in our society and try to link this to my lecture? (Connecting topics of study to real world situations)
DLER: So, you you teach English language, and you relate it to some situations?
DILDAR: You, I I try...
DLER: Do you remember any situation you linked to that?
DILDAR: For example, I teach work and business, for example, museums, why do people visit museums?
DLER: And they should talk about that?
DILDAR: Yes, yes, why do people... and they say, for example, my brother and I, for example, visited museums, we saw things outside, you can also visit the museum, there is a museum, this is the real world, they can do this, and they can write about this. (Connecting activities and assessment to real world situations)
DLER: Yeah, that is very good.
DILDAR: Practical and theoretical.
DLER: Yeah, yeah, of course, yeah, it is total linking actually. Do you believe that the, do you believe that the teacher is the only assessor and tester of the students?
DILDAR: Not the teacher is the only assessor and the tester for the students, they themselves...
DLER: They can be testers of themselves?
DILDAR: Yes, if you have strategies for assessment, if you have strategies for testing...
DLER: That is very good. You think you think your students can assess themselves?
DILDAR: Yes, they can assess themselves, for example, as I said before, if I teach essay writing, a group of students four students, for example, write an essay ,
DLER: Aha, OK,
DILDAR: I will not revise their... you know, their own tasks; this student, for example, wrote this essay, another student can revise it (Understanding of that students can be assessers)

DLER: Actually, peer-assessment...

DILDAR: Peer-assessment.

DLER: Do you think that they can do peer-assessment and self-assessment confidently, validly?

DILDAR: Yes, why, sometimes you know there are some students feel shy, some others you know, may not be confident?

DLER: Why not confident?

DILDAR: Because they think that they have so many mistakes, but they have to do it? (Students as assessers: feeling shy because of having many mistakes)

DLER: So what about when they do their peers’, when they check their peers’ writings or conversation, presentation; do you think they are ethical in that; they are valid, their assessment is valid?

DILDAR: Their assessment is valid, but the teacher should also revise the same, for example, essay, the same test in order to show them finally, to show them their own mistakes. (Students’ assessments can be valid) (Students as assessers: teachers should revise their assessments too)

DLER: Right, what about you now, do you think they are good, they are good in testing, in self-assessment and peer assessment?

DILDAR: I think they are good, especially third and fourth year students. (Students as assessers: level of students)

DLER: What about first and second?

DILDAR: First and second, I think they don’t know each other well, but I think when they become second year students, they can do this. They will be more confident.

DLER: But what is the relation of knowing, being acquainted with each other, and their...?

DILDAR: Because, you know, they don’t know each other well; they are not very very friendly, I think that.

DLER: OK, but do you think this is related to...?

DILDAR: Yes, I think sometimes, it related to do because, for example, one student is not friendly with this student.

DLER: So he cannot assess?
DILDAR: He is not friendly with the student, so he can or she cannot, or if she assesses, for example, this student, he or she, you know, may feel shy. (Students as assessors: socialbility of AA, if they are not friendly and friends they cannot assess each other)

DLER: What about, what about, when they are close friends?

DILDAR: If they are close friends, I think it is normal. (Students as assessors: close friendship does not affect students’ AAs)

DLER: So, it is OK but do you think that they don’t cheat, for example, because he is or she is his friend?

DILDAR: They don’t do cheating because they have to read, for example, the same task orally (self-assessment), if they do, for example, the teacher can er correct the same mistakes. (Less cheating in self-assessment)

DLER: So you think that personal relationships between students don’t affect their assessment, their assessment of each other, peer assessment?

DILDAR: I think that it (personal relationships) may affect their peer assessment, as I said before... (Students as assessors: personal relationships affect peer-assessment)

DLER: But not self-assessment. So, what about self-assessment when a student er for example, just overmark himself or herself?

DILDAR: They can’t do overmark...

DLER: Do you think they do that?

DILDAR: They do that, sometimes, they do that (overmark), but the teacher can do this instead of self-assessment; the teacher, I think, should let them to do peer-assessment, not self-assessment because if students, as I said, they may feel shy, they don’t want to show their own mistakes, the teacher show them their own mistakes. (Preferring peer-assessment over self-assessment due to students shyness to show mistakes)

DLER: Oh, yeah, yeah. Self-assessment and peer assessment; do you think these two assessments make students more autonomous learners?

DILDAR: yes, yes, they...

DLER: OK, how how they can be autonomous learners by self-assessment and peer assessment?

DILDAR: They can be autonomous (by self-assessment and peer assessment ) actually, but I think they feel you do different things, you do different, you use different strategies, it is a kind of demonstration for their own assessments; they will be busy with their own papers, they know their own mistakes.
DLER: So, it makes them autonomous?
DILDAR: Yes. (Students can be autonomous by self-assessment and peer assessment)
DLER: Do you think this peer-assessment and self-assessment; do you think they make students more critical thinkers?
DILDAR: Yes.
DLER: More reflective thinkers?
DILDAR: More reflective and more critical.
DLER: Why?
DILDAR: Because, you know, if they have their own, if they show their own mistakes, for example, one student can criticise another student, this is not right, the right one is this, they can do this... (Critical thinking of AA misunderstood)
DLER: So they need to, they need to...
DILDAR: And the teacher can show them how, for example, how they participate academically, you can say, for example, you can advise them, you can teach them, especially, in conversation if you want to criticise your friend, you have to say, I think that this is not right, the right one is this. (Critical thinking of AA misunderstood)
DLER: Yeah, quite interesting. Do you think that self-assessment and peer assessment by these two assessments, students can determine their weaknesses and their strengths?
DILDAR: Yes, they can determine. (By self-assessment and peer assessment, student can determine their weaknesses and strengths)
DLER: All of them?
DILDAR: Not all of them,
DLER: Which one cannot, maybe; we are talking about your students... which kind of student cannot determine his weaknesses or her weaknesses and strengths?
DILDAR: Especially those they feel shy those they don’t want to show their own mistakes; those they are passive, they are not active. (Student cannot determine their weaknesses and strengths if they are ashamed of their mistakes)
DLER: What about their level, their level of English?
DILDAR: The level, yes, if they think that their level of English is not good, they can’t participate.
DLER: What about if their level is not good, do you think they can? For example, they can show, find their mistakes?
DILDAR: I think they can participate, they can demonstrate, they can show their own mistakes properly.
DLER: Even if their level is low?
DILDAR: Yes, if their level is low, they can do this properly like active students; I have seen those students several times, but the teacher should encourage them more than the active students, to demonstrate, to participate, to show their own mistakes. (Low level student can determine their weaknesses and strengths if they are encouraged)
DLER: Do you believe that even the lower level students can find their mistakes?
DILDAR: Yes, they can...
DLER: First and second year?
DILDAR: Yes, they can find their own... (Low level student can determine their weaknesses and strengths if they are encouraged)
DLER: All their mistakes?
DILDAR: Not all their mistakes, specially, you know, we have major and minor mistakes; they can find, for example, minor mistakes, but they cannot find major mistakes, but when they become, for example, third year students, they can do more than this, and when they become fourth year students; they can do more than this, etc. (Students can show minor mistakes but not major mistakes)
DLER: That is very good. Now we come to another aspect of assessment; this is related to your students,
DILDAR: Right.
DLER: Actually it is related to the rights of students; do you think that the students have the right to discuss with you freely the design of testing, the way of their assessments, the way you follow to assess them; do you think that students have this right?
DILDAR: They have these rights, specially, when they do their term exams, because they have two exams, they usually, you know, I ask them, they usually discuss with the teachers: how are the tests, how many questions, the types of questions...
DLER: Students discuss with the teacher, with you?
DILDAR: yes, with me or with other teachers I know about...
DLER: Now we are talking about you, so they discuss with your?
DILDAR: yes, yes, I usually, you know, discuss... (Students have the right to discuss designs of term tests)
DLER: And what about assessment, do you discuss how to do the other assessments in the classroom?

DILDAR: yes, yes, specially, when they have, for example, presentation, they usually ask me, 'Please teacher tell us how we present, for example, this presentations?' (Students have the right to discuss designs of assessments)

DLER: Right, do you consider their preferences?

DILDAR: yes, why not.

DLER: All of them?

DILDAR: Not all of them, especially those that are academic.

DLER: So they don't know about it, you mean?

DILDAR: yes, they don't know; you have to teach them; I have to teach them. (Not considering academic issues preferred by students)

DLER: How is your relations with your students generally?

DILDAR: In General, you know, my relations is good with all students.

DLER: Do you befriend them?

DILDAR: No, you can’t be friend with the students, but you can for example, you can treat, you can behave with them friendly. (Friendly teacher-student relationship)

DLER: Aha, OK that is very good. Do you think that students have the right to have different kinds of assessments: testing, test questions, different types of test questions, different types of classroom assessment, AA? Do you think that students have this rights or?

DILDAR: Yes, they have these rights. They have, to come across all these kinds of assessment and tests, they have to know what is, for example, assessments, what is testing, what are the kinds of assessment; this is very good, especially for our students because... (Students’ rights to do all assessments)

DLER: Do you think you have given them these rights?

DILDAR: Yes, I have given them these rights, if I have enough time, as I said time. (Time restriction to do all assessments)

DLER: Yes, yes, of course, enough time, number of students...

DILDAR: Number of students

DLER: Yeah, yeah, actually, yeah, this is a challenge.

DILDAR: And we have long holidays,

DLER: Yeah, of course but I think in other countries also we have long holidays.
DILDAR: Every day is holiday, it is a new joke. (Many holidays do not allow doing all assessments)

DLER: Hhhh hhhhh OK, now we come to another side. Which of these processes you consider more advantageous for the students? Assessing students on what they can perform or assessing students on what they can remember when answering, for example, multiple choices, fill-ins, etc. in testing? Which of these do you think is more accurate... sorry, more advantageous?

DILDAR: The process I think the first one is more accurate than the second one...

DLER: Which one?

DILDAR: Because they can, what they can perform

DLER: Is better?

DILDAR: Is better, I think, because if you have this strategy, if you have this procedure, you can for example, go along with the students step by step, (Assessing the process is better than product) and finally you can learn what you want to teach, what they want to learn. But during this, you ha, you ca, you may have, as I said before, as you discussed with me, you may have several, for example, kinds of assessment, several kinds of testing,

DLER: Right, but which one, this is for the students,

DILDAR: For the students, yes.

DLER: So you say performance assessment is better, more advantageous?

DILDAR: I think the first one, yes. (Assessing students’ performances is more advantageous)

DLER: What about testing, is not advantageous? For their learning, generally?

DILDAR: not for the learning, I think assessment is more advantageous than testing for the learning, (Assessment is better than testing for students’ learning)

DLER: Yeah, why?

DILDAR: Because, you know, they know their performance in speaking, but if they do testing, they just for example, memorise... but if you see their performance in a presentation,... (Assessment is better than testing for students’ learning)

DLER: Yeah, which of these, focusing on the students’ strengths or focusing the students’ weaknesses, which of these two is more important and more advantageous?

DILDAR: The weaknesses.

DLER: Weaknesses?
DILDAR: Because, yes, I have to, you know, to indicate to know, what are their weaknesses?

DLER: What about their strengths?

DILDAR: Their strengths are also important, but in our classes, in EFL classes, we have more weaknesses unfortunately, than strengths... they have to get benefit from the weaknesses. (*Focus on students’ weaknesses more than weaknesses*)

DLER: Right, yeah, yeah, quite right.

DILDAR: In order to improve their own performances, and they have to also, they have to encourage their strengths. (*Encouraging students’ strengths as well*)

DLER: Don’t you think that, don’t you think that when you focus a student’s weaknesses, and you say you are wrong in this, this, your level in this is not good. Don’t you think that this discourages that student?

DILDAR: No, actually, as you know, in TESOL or in applied linguistics, we have direct and indirect feedback, how do you, for example, give feedback to the students? Do you give, for example, do I give my own feedback to the students directly or indirectly.

DLER: Indirectly, so when you give it indirectly, you mean...?

DILDAR: Indirectly, yes, it is better than directly...

DLER: How is indirect?

DILDAR: Indirectly, I think you may for example, you may correct the same mistakes, but the students do not know, do not, for example, get embarrassed. (*Indirect feedback to show students’ weaknesses*)

DLER: Ha, not feel shy.

DILDAR: I don’t say, ‘this is a mistake, don’t repeat this mistake’ for example, like this.

DLER: Aha, you write it on a on a paper?

DILDAR: I write on the whiteboard, I show them, you have to write it like this, that is all.

DLER: so what about, don’t you think that if you focus on their strengths, if you say ‘you’re very good in this,’ don’t you think this is also good to encourage them?

DILDAR: yes, you have to have different kinds of incentives. (*Using incentives to encourage students’ strengths*)

DLER: Have you, have you done that with your students, encouraged them on their good sides?

DILDAR: Yes, specially, you know, the students feel that the teacher should give them different words inside the class different incentives, for example, I don’t always say, 'very
good, very good' repeat the same word, you have to have different words: very good, well done, good job, something like this...

DLER: When they have a good performance in the class?

DILDAR: Good performance inside the class. (Using incentives to encourage students’ strengths)

DLER: That is very good. So in your opinion, for your assessment decisions, which one is better, depending on just single set of data, which is testing, or depending on several assessments that you do in the classroom, for example, during a course over time, so which one do you think is better for your decisions about students?

DILDAR: I think er that, for example, the teacher should have several kinds of assessment, several assessments during the academic year,

DLER: Right.

DILDAR: I think this one is better than the first one. (Assessment decisions based on several assessments)

DLER: Right, yeah. Do you think genuine performances that students do in the classroom, this is AA, or just sometimes just choosing an answer in a test, which one do you think motivates students to be more reflective thinkers, to be more successful students?

DILDAR: You said genuine classroom?

DLER: Yeah, we have genuine performances in the classroom, AAs,

DILDAR: Right,

DLER: And we have just choosing an answer in a test paper.

DILDAR: I think the first one, genuine er classroom activities,

DLER: Are better to make...

DILDAR: Yes, are better...

DLER: more critical thinkers,

DILDAR: yes.

DLER: In which way?

DILDAR: In different ways, I think so, we have several kinds: achievement, for example, proficiency any kinds of assessment, you have presentation, it is more critical than the first one, the students, you know they see, for example, one student for example can see another student, he or she can perform well...

DLER: yeah, yeah, that is very good.
DILDAR: he or she can criticise, for example, other students, other friends. (*Critical thinking in AA misunderstood*)

DLER: yeah, that is very good.

DILDAR: It is more encouraging.

DLER: yeah, of course, yeah. So assessing which one is more necessary, the product of learning, which is at the end of a course, or a month; or the process of learning, which is you assess the process of learning of students every day; so which one do you think is more necessary?

DILDAR: The process of learning or assessment?

DLER: The process of learning, every day you assess the process of learning of the students.

DILDAR: I think it is better for the teachers to have their daily assessments; I think daily assessments are more important than the product of learning. (*The importance of assessing the process of learning*)

DLER: yeah, why why daily assessment is more important?

DILDAR: I think daily assessments, for example, are more important than the second one, what is the second? the product of learning, because if you have daily assessments for your students, you see the products, you see the products; but if you just see the results of the students...

DLER: At the end of the course, I think it may not be good for the students.

DILDAR: So you mean you mean every day you see the products of the students by their process of learning?

DILDAR: yes, everyday yes if I have daily assessment, if I have, for example, quizzes every day, I see for example, daily assessment, I see my, I see the assessments every day. (*Everyday assessment of the products of the process of learning*)

DLER: Right, do you think this changes your teaching?

DILDAR: It, yes, it changes my teaching.

DLER: It improves your teaching?

DILDAR: yes, improves our teaching (*Assessing the process of learning improves teachers’ teaching methods*)

DLER: How, how?

DILDAR: Because the teacher should think about different strategies for their own assessment, should demonstrate, should, for example, read more and more.
DLER: OK, but I mean the results of your assessment with the students every day; how these change or improve your teaching?
DILDAR: the result or the assessment?
DLER: The result of daily assessment, when you assess them...
DILDAR: If I see the results, for example, are not good or are good, I have to have my own decision now, for I see, for my teaching, if I see the results are not good, I have to change my strategy, my approach, my method. (Assessing the process of learning improves teachers’ teaching methods)
DLER: yeah, that is very good. Yes, of course, quite good point. How long classroom time usually you allocate for your daily assessments, AA, the time in the classroom from maybe 45 minutes?
DILDAR: Five to ten minutes, (Short time allocated for AAs)
DLER: five to ten minutes, every lecture?
DILDAR: Every lecture, yes, each group, for example, can present a topic, but as I said as I said before, we don’t have enough time in order to assess all the students, all groups, because they need enough time, you have many lectures. (Short time allocated for AAs due to lecture limited time)
DLER: Of course, of course, yeah. And how many scores you give your students for each activity? For example, if a student have a presentation, if a student, for example, have an interview with you, how many marks, scores you give for each activity?
DILDAR: For each activity, it is based on the activity itself, it is based on the topic itself, and it is also based on their own performance, on his or her performance.
DLER: yeah, of course, but how many marks?
DILDAR: I have to consider, it is usually out of, for example, five. (Marking AA with few marks)
DLER: OK, but I mean this is five for one course, I mean for one activity?
DILDAR: I may, for example, give him or give her five for this, and five for the second one, five for the third one, and then I plus them, I collect them and divide it by four, it will be the same mark. (A strategy followed to score AAs by few marks)
DLER: So each, each activity, each assessment activity, er is given how many marks? when you divide...
DILDAR: Five, five marks, or three, or three marks. For each individual student or for the group?
DLER: For each individual student, when for example, she or he has an interview with you, how many marks is given to this activity?

DILDAR: No, actually it is based on the total mark, for example, I usually advice the students, I tell them the procedures of the presentation; this presentation, for example, is out of ten, for example, is out of five, but beside this, they have quizzes, and they have daily activities, they have classroom participation, they have quizzes, I have, you know, to plus, to collect these marks, and divide it, for example, by four and five, this is your, this is your average for your classroom assessment, and so the testing is out of this [mark] or the assessment is out of this. (Performing different kinds of AAs and test) (A strategy followed to score AAs by few marks)

DLER: So when they present, how many marks you give for that presentation?

DILDAR: Five marks.

DLER: And you give five to each student, er aha five marks.

DILDAR: Not to each student, it is based on the individual, for example,

DLER: Sometimes you give four, sometimes, you have three.

DILDAR: Three, yes, It is based on their assessment.

DLER: OK, that is very good.

DILDAR: For example, you get five, you get three, you get two, I have to see their own, for example, each student separately.

DLER: OK, when they have interview with you, you will also base it on five marks?

DILDAR: Five marks, yes.

DLER: And the other five marks, then you collect and divide?

DILDAR: I collect them, yes, this group, for example, got ten out of twenty, or five out of, for example, ten.

DLER: OK, yeah, that is very good. Generally, do you think, these activities, these AA activities, do you think that they improve students’ learning?

DILDAR: yes, they improve students’ learning. (AA improves students’ learning)

DLER: How?

DILDAR: Properly, you see the students always, you know, try to learn, try to find information about different topics. If I encourage them, to prepare presentation about different topics for my lecture, for example, they read more and more, for the topics. (AA improves students’ learning by encouraging them)
DLER: Yeah, yeah, you are very right. Mr. Farhad, thank you so much for your participation.
Appendix (15): A sample showing the summarising and coding responses to the open-ended questionnaire question

TEACHER35 talks about the general usefulness of alternative assessment (AA) for learning (benefits of AA for learning). TEACHER35 thinks that AA provides a realistic way for showing students’ knowledge and abilities. (accuracy of AA)

TEACHER36 talks about the general usefulness of AA for learning (benefits of AA for learning). TEACHER36 said: “AAs help students to express themselves freely. In certain cases, the students are given a bigger platform to display their potentials and skills.” (accuracy of AA)

TEACHER37 talks about the general usefulness of AA for learning (benefits of AA for learning). TEACHER37 could provide a simple definition of AA “AAs encompass any graded assessments done apart from paper-pencil examinations.” (simple definition of AA) “I believe these types of assessment are better than the fixed paper-pencil exams, because they give a better understanding about the students’ background knowledge.” (accuracy of AA)

TEACHER38 talks about the general usefulness of AA for learning (benefits of AA for learning). TEACHER38 said: “I think some of these assessments are more applicable than the others.” (practicality of AA) TEACHER38 thinks that some AAs are more useful and students have less difficulty doing them like writing assignments, editing a piece of writing and direct question and answer. (benefits of AA for learning) (practicality of AA)
Appendix (16): A piloting SPSS Cronbach’s Alpha test for the internal consistency reliability of the questionnaire items

<table>
<thead>
<tr>
<th>Cronbach’s Alpha Based on Standardized Items</th>
<th>N of Items</th>
</tr>
</thead>
<tbody>
<tr>
<td>.728</td>
<td>.624</td>
</tr>
<tr>
<td></td>
<td>31</td>
</tr>
</tbody>
</table>
References


TESOL International Association (2010). Standards for the recognition of initial TESOL programs in P-12 ESL teacher education. Prepared and Developed by the TESOL/NCATE P-12 ESL Teacher Education Program Standards Team and TESOL’s ESL Program Coordinator for P–12 Teacher Education Programs, Alexandria, USA.


