

T. R.
GAZIANTEP UNIVERSITY
GRADUATE SCHOOL OF EDUCATIONAL SCIENCES
DEPARTMENT OF FOREIGN LANGUAGES TEACHING
ENGLISH LANGUAGE TEACHING PROGRAM

NORTHERN IRAQI EFL TEACHERS' ASSESSMENT LITERACY

Master's of Arts Thesis

FRYAD H N. MUHAMMAD

Gaziantep
July 2017

T. R.
GAZIANTEP UNIVERSITY
GRADUATE SCHOOL OF EDUCATIONAL SCIENCES
DEPARTMENT OF FOREIGN LANGUAGES TEACHING
ENGLISH LANGUAGE TEACHING PROGRAM

**NORTHERN IRAQI EFL TEACHERS' ASSESSMENT
LITERACY**

Master's of Arts Thesis

FRYAD H N. MUHAMMAD

Supervisor: Assist. Prof. Dr. Mehmet BARDAKÇI

Gaziantep
July 2017

APPROVAL OF THE JURY

Student : Fryad Hama Najib Muhammad
University : Gaziantep University
Graduate School : Graduate School of Educational Sciences
Department : Department of foreign languages teaching
Thesis Title : Northern Iraqi EFL Teachers' Assessment Literacy
Thesis Date : July 3, 2017

I certify that this thesis satisfies all the requirements as a thesis for the degree of Master of Arts.

(Title, Name and Signature)
Head of Department

This is to certify that I have read this thesis and that in my opinion it is fully adequate, in scope and quality, as a thesis for the degree of Master of Arts.

(Title, Name and Signature)
Supervisor

This is to certify that we have read this thesis and that in our opinion it is fully adequate, in scope and quality, as a thesis for the degree of Master's of Art.

Examining Committee Members	Signature
(Title, Name and SURNAME)
(Title, Name and SURNAME)
(Title, Name and SURNAME)

It is approved that this thesis has been written in compliance with the formatting rules laid down by the Graduate School of Educational Sciences.

(Title, Name and Signature)
Director

RESEARCH ETHICS DECLARATION

The information contained here is, to the best of my knowledge and belief, accurate. I have read the University's current research ethics guidelines, and accept responsibility for the conduct of the procedures set out in the attached application in accordance with these guidelines, the University's policy on conflict of interest and any other condition laid down by the Gaziantep University Research Ethics Committee or its Sub-Committees. I have attempted to identify all the risks related to this research that may arise in conducting this research, and acknowledge my obligations and the rights of the participants.

I have declared any affiliation or financial interest in this research or its outcomes or any other circumstances, which might present a perceived, potential or actual conflict of interest, in accordance with Gaziantep University policy on Conflicts of Interest.

Signature: _____

Name: Fryad Hama Najib Muhammad

Student ID Number: 5028502700150466

Date: _____

Dedication

Dedicated to

My parents, my family and friends, my lovely kids and my love of life.



ACKNOWLEDGEMENTS

I thank God who guide me to the Right Path. Then I would like to acknowledge my special gratitude to my supervisor Assist. Prof. Dr. Mehmet Bardakçı. Without his supervision, this thesis would not have been possible. I also would like to thank the teaching staff in ELT department in Gaziantep University.

I would like to thank all my teachers who have made effort to introduce me to reading, writing, and following academic life. I would like to thank Dr. Craig A. Mertler, Arizona State University, USA. For his permission to use CALI in the study. I would also like to thank all the participant EFL teachers who have given their time to provide data for this study, without their contribution the survey could not have been successfully conducted.

I would like to express appreciation to my family for their supportive efforts and especially from the deep of my heart, I appreciate the role of my love who have always supported me and took the whole responsibility of the family in my absence.

In addition, I would like to appreciate the assistance of my friends and my colleagues.

ÖZET

IRAKLI İNGİLİZCE ÖĞRETMENLERİNİN ÖLÇME/DEĞERLENDİRME KÜLTÜRÜ: ALGILAR VE UYGULAMALAR

Fryad, H N. Muhammad

Yüksel Lisans Tez, İngilizce Öğretmenlik Programı

Danışman: Yrd. Doç. Dr. Mehmet Bardakçı

Haziran-2017, 67 sayfa

Son zamanlarda, öğrenme sürecinde oynadığı kritik rolden ötürü ölçme değerlendirme konusu, daha çok ilgi görmeye başladı. Ölçme ve değerlendirmenin amacı ve işlevleri genişlemiştir; bu nedenle, öğretmenlerin çağdaş değerlendirme ilkeleri hakkında bilgi sahibi olmaları gereklidir. Bu durumun bir sonucu da eğitim literatüründeki güncel akımlardan birinin de öğretmenlerin ölçme ve değerlendirme çalışmaları olmasıdır. Betimsel nicel yöntemin kullanıldığı bu çalışmada, özel sektörde ve devlette çalışan Iraklı İngilizce öğretmenlerinin, ölçme ve değerlendirme kültürü düzeyleri incelenmiştir. Çalışmada, öğrencilerin eğitsel değerlendirilmeleri için, öğretmenlerin yedi tane ölçme değerlendirme becerisine dayanan ve Mertler (2005) tarafından geliştirilen *Sınıf Ölçme Değerlendirme Anketi* uygulanmıştır (AFT, NCME, & NEA, 1990). Veriler, Suleimani ve Erbil valiliklerindeki ortaokullardan ve hazırlık okullarından toplanmıştır. Sonuçlara göre, Iraklı İngilizce öğretmenleri, 35 soru içinde, 15'den daha azına doğru cevap vermiştir. Çalışma bulguları, Iraklı öğretmenlerin ölçme değerlendirme kültürünün, istenilen düzeyden daha az olduğunu ortaya koymuştur. Öğretmenlerin akademik kazanımları ve meslekte geçirdikleri yılları arasında da istatistiksel yönden anlamlı farklılıklara rastlanmamıştır. Öğretmenlerin en düşük performansları, etik ve yasal olmayan durumların, uygun olmayan ölçme ve değerlendirme bilgisi kullanımlarının farkında olunmasını ifade eden 7. standartta görülmüştür. En yüksek performans ise eğitsel karar alırken ölçme değerlendirme sonuçlarından faydalanmayla ilgili olan 4. standartta görülmüştür. Bu sonuçlara dayanarak, öğretmen hazırlık programları ve mesleki eğitim kurslarıyla, öğretmenlerin bu alandaki niteliklerinin artırılması önemle tavsiye edilmektedir.

Anahtar Kelimeler: değerlendirme kültürü, öğretmenlerin değerlendirme kültürü, eğitsel değerlendirme, öğretmen algısı, sınıf pratiği/uygulaması.

ABSTRACT

IRAQI EFL TEACHERS' ASSESSMENT LITERACY: PERCEPTIONS AND PRACTICES

Fryad, Muhammad

MA Thesis, English Language Teaching Program

Supervisor: Assist. Prof. Dr. Mehmet Bardakçı

June-2017, 67 pages

Assessment has gained more considerations since its critical impact on learning became clearer. Assessment purpose and functions have extended and teachers are required to be aware of the principles of modern assessment, as a result, one of the current trends in literature of education is teachers' assessment. This research applied a descriptive quantitative method to investigate the assessment literacy levels of Iraqi English language teachers in both private and public schools. The survey was conducted by using Classroom Assessment Literacy Inventory designed by Mertler (2005) based on the seven standards of teachers' assessment competence for educational assessment of students (AFT, NCME, & NEA, 1990). The data was collected from secondary and preparatory schools of Suleimani and Erbil governorate. The results showed that Iraqi English language teachers responded less than 15 correct responses out of 35 questions. The findings showed that Iraqi teachers' assessment literacy are much less than satisfactory. Statistically significant differences were not found for teachers' academic attainment and years of experience. Teachers' lowest performance was found in standard seven, which is recognizing unethical, illegal, and otherwise inappropriate assessment methods and uses of assessment information. The highest performance was in standard 4 which is about using assessment results when making educational decisions. Revision of teachers' quality through teacher preparation program and professional training courses is definitely needed.

Keywords: assessment literacy, teachers' assessment literacy, educational assessment, teachers' perception, classroom practice.

TABLE OF CONTENTS

APPROVAL OF THE JURY	IV
RESEARCH ETHICS DECLARATION	V
DEDICATION	VI
ACKNOWLEDGEMENTS	VII
ÖZET	VIII
ABSTRACT	IX
TABLE OF CONTENTS	X
LIST OF TABLES	XII
LIST OF FIGURES	XIII
LIST OF APPENDICES	XIV
LIST OF ABBREVIATIONS	XV

CHAPTER I INTRODUCTION

1.1. Presentation	1
1.2. Statement of the Problem	1
1.3. Aim of the Study	2
1.4. Significance of the Study	3
1.5. Assumptions	4
1.6. Limitations	4
1.7. Definitions	5

CHAPTER II LITERATURE REVIEW

2.1. Presentation.....	6
2.2. Assessment and related terms.....	6
2.3. Assessment types	8
2.3.1. Summative assessment	9
2.3.2. formative assessment	10
2.4. The role of Classroom assessment	13
2.5. Assessment reform	14
2.6. the impact of assessment on learning	17
2.7. Assessment literacy.....	18
2.8. Teachers' assessment types	22

CHAPTER III METHODOLOGY

3.1. Presentation	26
3.2. Research design.....	26
3.3. Context of the study	27
3.4. Samling	28
3.5. Procedure	29
3.6. Instrument	29
3.7. Data analysis	31

CHAPTER IV FINDINGS

4.1. Presentation	33
4.2. Analysis of background information	33
4.3. Analysis of teachers assessment literacy	37

CHAPTER V DISCUSSIONS

5.1. Presentation	42
5.2. Iraqi EFL teachers' assessment literacy level	42
5.3. Strong and weak aspects of Iraqi EFL teachers' assessmnet literacy	43
5.4. Iraqi EFL teachers' conception on assessment	44

CHAPTER VI CONCLUSION AND RECOMMENDATIONS

6.1. Conclusion	46
6.2. Suggestions for authorities	47
6.3. Suggestions for teachers	48
6.4. Suggestios for furthur studies	48
REFERENCES	50

APPENDICES	57
Appendix I Classroom Assessment Literacy Inventory (CALI).....	57
Appendix II Answers to (CALI)	65
Appendix III authors permission	66
VITAE	67



LIST OF TABLES

Table 1.The alignment of Stiggins’ competence with the seven standards by AFT, NCME and NEA (1090)	21
Table 2.Alignment of teacher competence (1990) with CALI (2003)	30
Table 3.Particants’ teaching level frequency	34
Table 4.frequency of academic attainment	34
Table 5.teachers’ classroom experience	35
Table 6.standalone course in classroom assessment	36
Table 7. Perception of undergraduate preparation for classroom teaching	36
Table 8. Perception of undergraduate preparation for classroom assessment	37
Table 9. Descriptive statistics of teachers’ assessment literacy level	38
Table 10.statistical analysis of teachers’ performance for each standard	39
Table 11.years of classroom experience with total correct answers.....	40
Table 12.academic attainment with total correct answers	40

LIST OF FIGURES

Figure 1.the relation of teaching, testing and assessment7
Figure 2.the relation among, test, measurement, assessment and evaluation.....8

LIST OF APPENDICES

Appendix I: Classroom Assessment Literacy Inventory56
Appendix II: Title answer to Classroom Assessment Literacy Inventory..... 63
Appendix III: Title permission from the author64



LIST OF ABBREVIATIONS

CALI: Classroom Assessment Literacy Inventory

EFL: English as a foreign language

AFT: American Federation of Teachers

NCME: National Council on Measurement in Education

NEA: National Education Association

ARG: Assessment Reform Group

TALQ: Teacher Assessment Literacy Questionnaire



CHAPTER I

INTRODUCTION

1.1 Presentation

This chapter is an overview about the statement of the problem, aim of the study with research questions and the significance of the study. It is also a description of research assumptions and possible limitations. A brief definition of the most used jargons is presented.

1.2 Statement of the Problem

The quality of teaching is inextricably intertwined with the quality of assessment. Assessment plays a vital role in the whole education process; therefore, it is teachers' responsibility to be assessment literate in order to shape their teaching plans and their educational decisions. Considering literature in this field most of the teachers are not well prepared to face classroom assessment challenge; in some cases it is not a part of teacher preparation courses, or professional development training.

Sometimes there are training courses for teachers about teaching specific course book, when making any change, but teachers' assessment literacy is not updated alongside. Assessment literacy seems to be undervalued or even neglected, which can cause a serious consequence. Any improvement and innovation in the curriculum and course books, without accurate restructure in the assessment policy, will take the full effect.

In recent years language education reform has been done by the ministry of education in Northern Iraq, the reform includes changing the curriculum and pedagogy. Massive investment has been done on training teachers to new teaching methods, but assessment has almost been neglected as if it is not related to teachers' profession. Education policy makers changed assessment policy from one single end-term summative test to 40 % classroom assessments and 60% end-term summative

assessment. Unfortunately, the reform movement has not changed the classroom practices but it has been taken over by the conventional conception of assessment. Teachers continue using end-unit quizzes and end-month paper-pencil summative tests to grade students' classroom performance. It clearly and surely can be claimed that the desired results have still not been approved.

Students study English as a foreign language for twelve years until they graduate high school; however, they cannot give a satisfactory level of language performance. For addressing this educational decisive issue, reform community must act further and shift their attention to deal with unsound assessment, the absence of balance between "assessment for learning" with "assessment of learning" due to inadequate level of teachers assessment literacy. Teachers' assessment literacy needs to be examined and improved. The view of assessment, as judgmental decisions about students' achievement, have negative effects on the reform act, schooling system and learning standards. Teachers are strictly tied to the traditional concept of assessment and they use the same previous methods of gathering information about students' level of performance. The assessment function has been urging students memorize facts and recall in a definite time once or twice a year. They have not been given opportunity to observe their learning progression.

Teachers' insufficient assessment literacy level may have a challenging impact on students learning, since assessment view may shape the teaching prospect and consequently might affect learning. Additionally, teachers' misunderstanding of principles and aims of an act of modification in assessment might reduce the effect of the reform. Teachers' concept, knowledge and practice of assessment in this context need a radical change for the sake of better-quality assessment, improved education and facilitated learning with less anxiety and more self-confidence in the chance for further success.

1.3 Aim of the Study

The aim of this research is to explore the quality of teachers of English as a foreign language; regarding their assessment literacy level by using Classroom Assessment Literacy Inventory (CALI) as a measurement scale, which was developed by Mertler (2003). The present study examines whether teachers are prepared to assess

students' performance and tries to discover teachers' perception about their level of assessment literacy. This research tries to reveal teachers' views about the courses they received in their undergraduate preparation program as teacher candidates.

The current research also tries to find out the weak and strong aspects of teachers knowledge about classroom assessment; what kind of skill, ability they have or they tend to develop. The study inspects whether the level of academic achievement or professional experience affects teachers' information about assessment standards and classroom practice.

The research reviews the level of assessment literacy among Iraqi teachers of English as a foreign language who serve in schools belong to the ministry of education in Northern Iraq regional government. The survey relies on the seven Standards for Teacher Competence in the Educational Assessment of Students (AFT, NCME, & NEA, 1990). To reach the purpose of the study the following research questions are proposed to be investigated:

1. What are the assessment literacy levels of Iraqi EFL teachers?
2. What are the weak and strong aspects of Iraqi EFL teachers' assessment literacy according to the standards of teacher competence?
3. Does year of classroom experience make difference in teachers' assessment literacy level?
4. Does academic attainment make difference in teachers' assessment literacy level?

1.4 Significance of the Study

This study is generally a part of the efforts that have started to raise awareness about the importance of the assessment for learning, or formative assessment, in education process and learning achievement. Similarly, the research is an attempt to shade some lights on teachers' professional development necessities in assessment field.

The importance of this research may be seen in presenting the need of in-service language teachers to improve and the types of the training courses they require. Based on findings about the quality of teachers regarding assessment literacy, this paper provides empirical evidence and suggestions for policy makers, reformists and

teachers. It shows an aspect of restrictions related to teachers that prevent reforms to take effect. The research determines a phase of teachers' quality and the way to enhance to serve student learning. It helps education reformists have evidence in hand to claim for the need of change in teachers' assessment quality and assessment reform.

This study might have influence on promoting other research that may be done in the context of language teaching and assessment. The findings of the thesis, if translated into action, would raise teachers' awareness of assessment as an effective tool that has a great impact on students' motivation, self- confidence and learning as a result. It paves the way on how to promote teachers quality to reach the sound level.

Reforms should not only cover curriculum and pedagogy, but teachers' quality, especially in the aspect of assessment. Assessment needs to be equally handled in any reform attempt to have the desired effects on schooling system and overall society to save time, money and effort.

1.5 Assumptions

This research based on the assumption that Assessment literacy levels of Iraqi English language teachers can be adequately investigated by using Classroom Assessment Literacy Inventory CALI. Another assumption is that teachers understood all the items and responded sincerely according to their own knowledge without being supported by any source. It is also assumed that there are not a significant gap among EFL teachers background knowledge and the samples are the reasonable representatives of the population.

1.6 Limitations

One of the limitations of this study may be the nature of the survey, which depends only on multiple choices; teachers have to choose the right answer among the given prompts.

Another restriction is that the research was applied only in Northern Iraq and covers only teachers of English language. The number of participants can also be another limitation. This study is also limited to the single stage data collection method, which is rather summative-like.

1.7 Definitions

Assessment: is “the process of obtaining information that is used to make educational decisions about students, to give feedback to the student about his or her progress, strengths, and weaknesses, to judge instructional effectiveness and curricular adequacy, and to inform policy.” (AFT, NCME, & NEA, 1990)

“Assessment is the process of collecting information about something that we are interested in, according to procedures that are systematic and substantively grounded” (Bachman, 2004, p. 6-7)

Assessment literacy: “Assessment Literacy is defined as the knowledge of means for assessing what students know and can do, how to interpret the results from these assessments, and how to apply these results to improve student learning and program effectiveness” (Webb, 2002, p.1)

Classroom assessment: Popham (2009) defines classroom assessment as the whole methods and procedures that teachers use to collect accurate information about their students’ knowledge and ability.

Formative assessment: has been referred to as “all those activities undertaken by teachers, and by their students in assessing themselves, which provide information to be used as feedback to modify the teaching and learning activities in which they are engaged. (Black & Wiliam, 1998a, p. 2)

“Formative assessment is an active and intentional learning process that partners the teacher and the students to continuously and systematically gather evidence of learning with the express goal of improving student achievement” (Moss, & Brookhart, 2009, p. 6)

Summative assessment: Defined as “assessments that come at the end of a process or activity, when it is difficult to alter or rectify what has already occurred, are called summative assessments (Airsian, 1994, p.136)

CHAPTER II

LITERATURE REVIEW

2.1 Presentation

This chapter covered a broad overview of assessment and the related terminology. Then assessment types both summative assessment “assessment of learning” and formative assessment “assessment for learning” was addressed in literature. After that the important role of assessment as general, classroom assessment specifically and the impact of assessment on learning were discussed. Finally, the focus was changed to a narrower field of assessment literacy then teachers’ assessment literacy in literature. The aim was to revise the previous works about the field and related issues.

2.2 Assessment and related terms

Assessment has been receiving increasing attention recently and assessment reform has become a counterpart of education reform. Through reviewing related literature, it can be claimed the meaning of assessment has expanded from the measurement of learning to a broader range and more connected to the learning process. Scholars in the field of assessment have distinguished among assessment and tests (Brown, 2004; Bachman, 2004; Popham, 2009; Rogier, 2014).

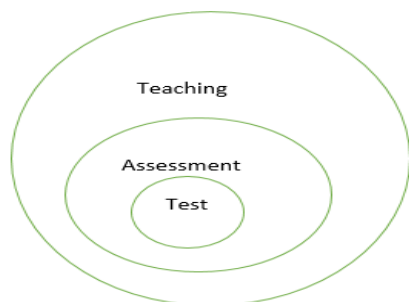
Tests have been defined as subsections of assessment (Brown, 2004). On the other hand, assessment is known as a continuing process that includes a much more expanded range. Assessment includes various methods, activities and mechanisms used by teachers for gathering information during teaching process in order to support and motivate student learning (Brown, 2004; Popham, 2009). The aim of assessment

is to collect information to have suggestions based on evidence so as to adjust the teaching/learning process (Black & Wiliam, 1998a).

One of the distinguishing features of assessment is that assessment observed and the result is used by both teachers and students for instructional decisions; thus, assessment is a student-involved process (Black & Wiliam, 1998a; Rogier, 2014; Stiggins, 2008). There are two main purposes of assessment; first, providing information for educational decisions; second, facilitating and encouraging learning (Black & Wiliam, 1998a; Stiggins, 2006, 2008). The purposes of assessment can be associated with the different types of assessment as summative and formative (Boud & Falchikove, 2006), and each sort of assessment is as important as the purpose that is designed for (Valincia, 2002).

Assessment and testing are seeking to measure (Alderson & Bachman, 2004) students' current level of performance so as to be able to decide about the next necessary step of instruction. Types and purposes of assessment deal with the question of progression.

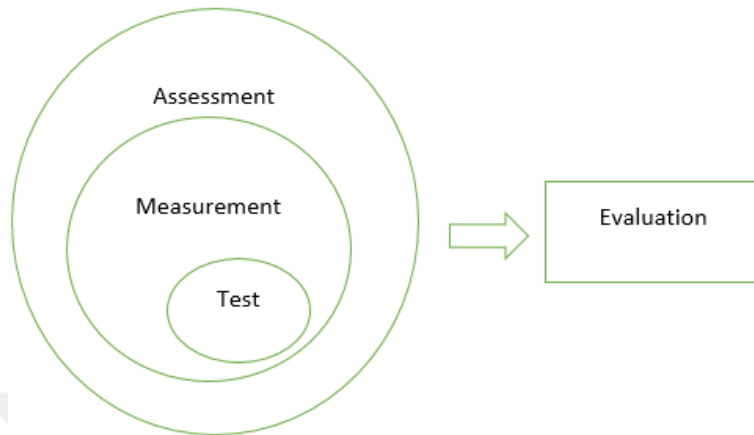
Additionally, Bachman (1990, p. 18) defines measurement as “the process of quantifying the characteristics of an object of interest according to explicit rules and procedures”. She claims that measurement covers a broader meaning than tests and less expanded than assessment in a way that assessment can be both quantifying and non-quantifying while measurement covers only quantifying features of an object. Furthermore, evaluation is known as the possible ways of using assessment results for decisions and depiction of qualities of individuals (Bachman, 2004). The figures 1 and 2 show the relation among teaching, assessment, testing and evaluation.



Brown (2004, p.5)

Figure1. The relation of teaching testing and assessment

Assessment is a part of teaching process and tests are only small scales of assessment (Brown, 2004).



Bachman (2004, p. 10)

Figure 2. The relation among test, measurement, assessment and evaluation.

Tests are a type of measurement the scales, assessment refer to a more general concept it is not only about numbers and scales and the whole purpose is for the aim of evaluation. Evaluation has the broadest meaning in terminology (Bachman, 2004).

Assessment with all the types and purposes are to define and support learners to achieve their goals. It is a professional responsibility for instructors to be aware of the types and purpose of assessment and use the types of assessment properly.

2.3 Assessment types

Assessment has several different types according to the purpose and procedures. Assessment is used for different reasons in various contexts, the use ranges from supporting students' learning to observing national standards (Bartlett, 2015). Different assessment types are used in different situations for different intentions. The overall aim of assessment is for collecting data about students' performance.

2.3.1 Summative assessment

Summative assessment is that conducted in the end of accomplishment; therefore, it is hard to change or modify what has previously happened. (Airsian,

1994). The purpose of summative assessment is to grade, or to sum up what students have achieved and usually happen after the instruction (Brown, 2004). The purpose of summative tests is to know students level of achievement or proficiency at a definite time like a snapshot of what have gained. In summative assessment, there is no or little concern about students' progress or feedback for instructional purpose. Summative assessment results may be used to decide about program level like proficiency and placement tests or can be used to make decisions about classroom level such as diagnostic and achievement tests (Brown, 2005).

The primary purpose of summative assessment is to insure the intended or the desired level of performance has obtained (21st century skills assessment, 2007). One of the distinguishing features of summative assessment is the measure of status of achievement at the end of a course for the purpose of certification (Sadler, 1989). Other characteristic of summative assessment is that it is administered in certain times to report achievement and connects learning progress to public criteria (Harlen & James, 1997).

The problem of the use of summative assessment is addressing all students' ability, or performance uniformly, i.e. all students judged by a standardized scale equally for everyone. Even if the scales are objective and valid, it neglects individual differences of students (Harpen, 1994). Individual uniqueness cannot possibly be taken into account in summative tests, which are dominantly paper-pencil tests at the end of a course. Another problem may appear because of the snapshot nature of the summative tests. It has been reported that it sometimes provide inaccurate information about test takers' ability (Hugues, 1989)

Summative tests use in classroom needs less time to be administered, easily objective and applicable for most of the context and provide grading rationality; therefore, teachers apply in classroom as a replacement of assessment for learning. Another factor may be fairness issue (Douglas, 2010). Fairness is not gained if students, who have different learning styles and performance ability, were assessed with a unified test. The best use of summative tests may be for proficiency and placement purpose. The value of a test may depend on the test takers' performance in that given time (Bachman, 2004).

2.3.2 Formative assessment

Formative assessment was defined by Black & Wiliam (1998) and then Assessment Reform Group (2002). Both definitions were combined and redefined in a slightly different way by Black & Wiliam (2009) as:

Practice in a classroom is formative to the extent that evidence about student achievement is elicited, interpreted, and used by teachers, learners, or their peers, to make decisions about the next steps in instruction that are likely to be better, or better founded, than the decisions they would have taken in the absence of the evidence that was elicited (p. 9).

Formative assessment has gone through historical improvement by researchers and specialists in the aspect of terminology, purpose and principles. Stiggins (2005) explains the historical development of formative assessment under three approaches. First approach is using diagnostic tests repeatedly by teachers to improve the quality of teaching. Second approach arises from the perception that successful formative assessment is not how information gathered, but how the results are interpreted and used for instructional decision-making. The third approach is the assessment for learning and more improved recently for using multi-methods of data collection about students' progression, as a constant feedback for teachers, students and parents for directing students towards success.

Assessment for learning, or formative assessment, is an ongoing day-to-day process multi-method of gathering information about students' progression during learning process (Black & Wiliam, 1998, 2009; Mcmillan, 2005; Popham, 2008; Scarino, 2013; Stiggins, 2005, 2008). Formative assessment decreases the gap between teaching and learning and builds strong prevailing connections between them (Anderson, 1998). Formative assessment seems a teaching application rather than testing (Davies, 2000).

The evidences gathered in assessment for learning are used for feedback for teachers to adapt their teaching; for students to be aware of their learning progression.

The aim of feedback is to enhance further learning (Black & Wiliam, 1998a; Brookhart, 2006; Heritage, 2008; Popham, 2008). The feedback is sometimes shared with parents (Stiggins, 1988). Students, beside teachers, involve in assessment and both of them use assessment results for instructional decision-making (Stiggins, 2005, 2008). Students' active participation in assessment for learning comes after having necessary feedback on their own learning is not coincidence; it is an evitable part of formative process (Black & Wiliam, 1998b).

Many researchers have addressed the issue of formative assessment regarding, classroom use, possible benefits and methods of applying. Engaging learners as active participants of assessment in formative assessment affects student motivation and achievement in several aspects including: providing students with the opportunity to monitor their own learning progression. Students evaluate their own performance and they decide about their own weakness and strength. Students become more reliable on themselves and more confident in their learning. This also improves their self-efficacy and self-efficacy make students regulate their own learning, which leads to motivation and success (Arter, 2001; Cauley & Mcmillan, 2010; Cahpuis & Stiggins, 2006; Hatie & Timperley, 2007; Stiggins, 2005). Assessment for learning has a positive motivational impact on learning and academic achievement (Cauley & Mcmillan, 2010). If formative assessment is used properly in classroom, it will have practical consequences on learning and raises students' standard level (Black & wiliam, 1998a, 1998b; Wiliam, 2004; Mcmillan, 2005; Stiggins, 2007). Formative assessment as an instructional mechanism holds students responsible for their own improvement and achievement, by providing them necessary feedback on their progression (Black & Wiliam, 1998b; Harlen & James, 1997). Formative assessment enhances higher level of thinking and deep understanding (Brooks, 1990). Assessment for learning changes students' role from passive learners to active ones (Valencia, 2002). Teachers are only responsible for providing formative assessment and engage students in the learning assessment (Black & Wiliam, 2009).

The methods of collecting information in formative assessment are various. Most used methods in literature are portfolio, self-assessment, peer assessment, projects, seminars, daily assignments, essays, interviews, discussions and sharing opinions etc. All these methods of gathering evidence are presented in literature;

however, it is generally claimed that objectivity and fairness of these methods may still need investigation and improvement and there are no clear suggestions or procedures for grading purpose. The multi-methods of gathering information do not affect learning if the results misinterpreted (Stiggins, 2012).

Assessment for learning has several distinguishing features presented in the literature by scholars (Black & William, 2006, 2009; Black & Thompson, 2007; Lee & Wiliam, 2005; Shepard, 2000; Stiggins, 1988) with a slightly different way, which can be summarized as:

1. Explaining and sharing learning objectives and success criteria with students; for example sharing previous works and discussing strong and weak points of the given work.
2. Influential classroom interaction, discussion and observation like classroom questioning for all the individuals not elites to provoke evidence that learning takes place.
3. Providing feedback for teachers, students and parents not on grade scales but on the bases of comment-only-marking.
4. Students' active involvement in assessment to be responsible for their own learning as in self-assessment and to be a source of instruction for other students as in peer-assessment.

Assessment Reform Group (2002) published ten principles of assessment for learning. These principles are as follows:

1. Assessment for learning should be part of effective planning of teaching and learning.
2. Assessment for learning should focus on how students learning.
3. Assessment for learning should be recognized as central to classroom practice.
4. Assessment for learning should be regarded as a key professional skill for teachers.
5. Assessment for learning should be sensitive and constructive because any assessment has an emotional impact.
6. Assessment should take account of the importance of learner motivation.

7. Assessment for learning should promote commitment to learning goals and a shared understanding of the criteria by which they are assessed.
8. Learners should receive constructive guidance about how to improve.
9. Assessment for learning develops learners' capacity for self-assessment so that they can become reflective and self-managing.
10. Assessment for learning should recognize the full range of achievements of all learners. (ARG, 2002, p 2)

It has been generally accepted that the best approach for classroom assessment is assessment for learning. Nevertheless summative and formative, or assessment of learning and assessment for learning are not the opposite or alternative (Hargreaves, 2005). Summative assessment should be minimized and the focus should be on assessment for learning to dominate education process, which makes the learning standards raise due to student motivation and anxiety reduction.

2.4 The role of classroom assessment

Classroom assessment has a powerful influence on teaching and learning process (Cheng, Rodgers & Wang, 2008; Stiggins, 2012). The importance of classroom assessment is giving constant feedback on students' progression for teachers, students and parents (Stiggins, 2012). Classroom assessment is one of the most challenging aspects of teachers' profession (Mertler, 2003, 2009). Teachers spend more than one third of their professional times engaging with assessment (Stiggins, Conkling & Bridgeford, 1986; Stiggins, 1988).

Literature supports the claim that despite the complexity and importance of classroom assessment the current level of practice is not adequate. Several factors influence the quality of classroom assessment including external tests, teachers' assessment literacy and classroom realities. The prominent power of external tests as: national standardized affects teaching plans, learning objectives and assessment performance accordingly. Teachers' planning is affected by high-stake tests in a way they almost ignore classroom assessment (Mcmillan, 2005; Stiggins, 2007).

Other obstacles can be the miss-conception of assessment among teachers, students and education community as a general, objective and accurate scientific measurement administered in certain times non-related to teaching. This classical view

of assessment does not support students' motivation and confidence in learning (Cauley & Mcmillan, 2010; Shepard, 2000; Stiggins, 2001b). Teachers' lack of knowledge is another obstacle in front of adequate practice of quality assessment in classroom education because teacher preparation courses and professional trainings are not preparing teachers well to challenge classroom assessment (Mcmillan, 2005; Rogier, 2014; Shepard, 2000; Stiggins, 2001b, 2004; Zhang & Burrey-Stock, 2003)

Additionally there are some classroom realities and learning environment that affect classroom assessment practice negatively such as inappropriate behaviors of learners, heterogeneity of learners' abilities (Mcmillan, 2005), crowded classrooms and very limited classroom time especially in public schools.

In order to accomplish the improvement in learning achievement and raise students' academic standards, necessary actions should be taken to promote the quality of classroom assessment. Training programs need to have courses that are more intensive on assessment literacy to prepare teachers well for the challenge of classroom assessment. The practice of assessment in classrooms should change to the assessment for learning not of learning for obtaining learning goals (Mcmillan, 2005; Stiggins, 2002, 2008, 2012; Valencia, 2002). Any improvement in the quality of assessment for learning leads to students' greater achievement (Stiggins, 2012). The gap between external factors on classroom assessment and principles of assessment literacy need to be minimized.

2.5 Assessment reform

The traditional view of assessment and the purpose of assessment has changed a lot recently. Educational experts have realized the importance of assessment in education; therefore, they have shifted their focus towards modern visions of assessment. It has been universally accepted that the concept of assessment, as only summative and judgment of learning in a certain time, is not adequate to the assessment process and need to be changed. Several terms have been used by scholars and experts in this field to imply the need of conceptual change of assessment.

Some researchers believe in radical change in assessment system and impose the term 'alternative assessment' at least in classroom education. Alternative assessment has various understandings (Fox, 2008). Some believe formative

assessment methods should be alternative of classical summative testing, while others believe that formative assessment and summative assessment should be balanced in education system. Alternative assessment needs more effort by teachers. A possible benefit of alternative assessment is minimizing students' anxiety, which result in higher levels of thinking and metacognitive ability (O'neil & Abedi, 1996). Constructive theory of learning supports alternative assessment but is at odds with traditional assessment (Anderson, 1998).

On the other hand, most of the scholars believe that the modern methods and purpose of assessment are necessary alongside the old methods and purpose of assessment. Both can have its importance and role in education. In a balanced assessment system, it can be concluded that no single type of assessment can achieve the duty of evaluation alone (Stiggins, 2012). The concept of balanced assessment system covers classroom assessment, benchmark assessment and annual testing and each has its own importance and help different assessment purpose (Stiggins, 2012). Balanced assessment system both supports and certifies learning. Quality assessment needed in all three levels of assessment: classroom, benchmark and annual assessment and fruitful assessment dynamics to assist with the emotional experience of assessment (Stiggins, 2012). Assessment for supporting and developing learners' ability described as dynamic assessment that covers summative and formative assessment (Poehner, 2010), which have the same implications of balanced assessment.

In both alternative assessment and balanced assessment, system experts have directed their attention towards multi-methods of gathering information as a feedback for different purpose and use. All the scholars in the field of assessment focus on assessment reform to cover formative classroom assessment, or assessment for learning.

Assessment reform, which has been universally claimed by experts and researchers, includes the change of traditional concept, principle and practice of assessment. Assessment seen as formally administered in definite time is not valid any more in the context of teaching and learning process (Brown, 2004). Some testing systems have negative consequences on the learning process: like high-stakes testing or standardized national testing, any pressure may rise test scores but no progression in academic achievement realized (Shepard, 2000). Especially when standardized

national annual testing is not quality assessment, which has a bad consequences on teachers' skills and professional abilities (Shepard, 2000).

Traditional assessment requires amendment from memorizing and reciting facts to assessing critical thinking and ability. In the Act of no child left behind in the United States of America (2001) it has been made clear that the twenty first century assessment has to focus on the learners' functioning talents in dealing with different sources properly, rather than submitting correct response (21st century skills, 2007).

The common miss-conception about teachers' role in assessment as the only assessors of students also must be corrected. Teachers should be aware of the elements of modern assessment concepts. In modern assessment, students have an important role in assessing themselves, they all have ability to succeed, assessment help students to enhance self-confidence in their abilities. These features make students feel more responsible and motivated and the sole goal of assessment is not only grading but also supporting learning by providing feedback for both teachers and students to make adjustments about teaching/learning progress (Black & Wiliam, 1998). These elements can be claimed as the bases of assessment reform (Boud, 2000). To accomplish the favorable transformation of assessment system to serve learning process, the social meaning of evaluation needs correction (Shepard, 2000). The naive view of assessment, as the source of fear of public disapproval, individual anxiety and stress, need to be changed to the source of motivation and self-confidence for further learning (Chapuis & Stiggins, 2002; Stiggins, 2005).

Assessment plays its important role in education only when it is modified to be an ongoing process to collect information about individual students before, during and after the learning process for educational decision-making (Mcmillan, 2005; On'eil & Abedi, 1996; Scarino, 2013). When quality assessment is not practiced day-to-day, learning achievement will remain below the desired level (Stiggins, 2005, 2012). Assessment should be more like a teaching tool to promote learning than a grading mechanism (Davies, 2000). Traditionally assessment is seen as a grading scale to judge students' learning achievement and comes after teaching (Assessment Reform Group, 1999; Chapuis & Stiggins, 2002; Crooks, 2001; Shepard, 2000). The traditional assessment concept requires modification.

Assessment reform can be achieved by expanding teachers' quality requirements to cover adequate level of assessment literacy, increasing investment in teachers' professional development, decreasing the restrictive factors of classroom assessment that limit teachers and students concept and practice of assessment (Assessment Reform Group, 1999). The restrictive factors of classroom assessment reform may include first, resources related to assessment literacy are very limited in both teacher preparation programs and professional trainings for teacher development. Second, high stake assessment or annual national testing has negative impact on teachers' and students' view and practice of assessment. In the presence of these restrictions, assessment for learning will not be easily achieved; neither will the desired level of learning progression (Mcmillan, 2005).

In conclusion, for controlling students' learning and obtaining an effective schooling system assessment reform must occur in conjunction with education reform. Diverse reliable mechanisms of gathering information about students' progress should replace end- course pen and paper tests. Multiple methods of gathering information can address different learning styles more accurately (Harpen, 1994).

2.6 The impact of assessment on learning

Assessment has a controlling power on teaching learning process. It is a decisive part of education. The impact of assessment on learning could be various depending on the types, or purpose of assessment and the influence is far more than study patterns, teaching methods and course books (White, 2004). Assessment is closely tied to student achievement and could have either positive or negative wash back on learning (Marzano, 1993; Crooks, 1988). Psychological or emotional response of the students, who are in charge of learning, towards assessment results make them eagerly or hopelessly decide about their success opportunity (Stiggins, 2005, 2008).

Assessment of learning, as a success/failure measurement for judging students, increase learners' anxiety and influence negatively on students' learning. Assessment for grading purpose leads students to focus on marks rather than learning (Boud & Falchikov, 2006). Thus, students who get low grades will be demotivated and lose confidence in their ability (Stiggins, 2005). Education and schooling system does

not have the right direction of serving community; because of failure to provide assessment balance in classroom.

On the other hand, there is plenty of evidence that assessment for learning has a positive impact on learning and raise standards of student achievement (Black & Wiliam, 1998a; Brown, 2004; Campbell & Collins, 2007; Wiliam, 2004, 2007, 2011). Assessment affects both learning and teaching methods, procedures that taken into account for educational objectives and goals (Black & Wiliam, 2003; Marzano, 1993; Popham, 2007, 2011; Scarino, 2013; Stiggins, 2002, 2008; Valencia, 2002; Zhang & Burrey-Stock, 2003). It has been reported that teachers use more than one third of their professional time on assessment (Stiggins, 2005) and students use a lot of time reviewing what has been assessed and will be assessed (Crooks, 1988; Shepard, 1990). Formative assessment raises achievement if it is applied properly (Shepard, 2000, 2001) and can double the learning achievement (Wiliam, 2007).

Literature review about the impact of assessment on learning gives evidence to state that assessment shapes learning and teaching; the influence could be constructive and destructive to instructional objectives and educational goals. The impact of assessment on students and teachers make alteration in view and practice of learning/ teaching process. Any attempt for education reform need to cover assessment reform due to the assessment power on learning achievement (Arter, 2001).

2.7 Assessment literacy

The knowledge of assessment was previously known as assessment competence; later the term assessment literacy was used by Stiggins (1991). Assessment literacy is defined as “individuals understanding of the fundamental assessment concepts and procedures” (Popham, 2011). Similarly, assessment literacy is identified as “understanding principles of sound assessment” (Stiggins, 2002). Precisely assessment literacy defined as the knowledge, skill and ability to form and improve all types of assessment for different purposes, knowing the procedures, principles and concepts to serve further performance and learning, familiarity with the role of assessment and its effect on individuals, institutions and society (Fulcher, 2012).

Assessment literacy includes the knowledge of the concept, principle and procedures of all types of assessment for different purposes and in different context. Assessment literacy needs skills, knowledge and principles (Davies, 2008) to be applied properly. It is also known as teachers' ability to use assessment as a feedback to improve teaching and learning (Rogier, 2014).

The characteristics of assessment literate instructors are the awareness of objectives they want to assess, their purpose, and know how to apply the assessment. The other assessment literacy distinguishing features can be regarded as the familiarity of the possible obstacles that may happen in assessment process and the skill of preventing those problems. Another important feature of assessment literacy is being conscious about the negative consequences of unsound assessment (Stiggins, 1995).

Professional organizations, education and assessment reformists and scholars have tried to set standards for assessment literacy. These attempts to make clear criteria for assessment literacy are needed to be identified and applied by teacher preparation programs, professional training courses, institutions and teachers. The result of these attempts is the joined work of American Federation of Teachers (AFT), National Council of Measurement in Education (NCME) and National Education Association (NEA). The professional organizations have set seven Standards for Teacher Competence in Educational Assessment of Students (AFT, NCME, & NEA, 1990).

These seven standards are as follows:

1. Teachers should be skilled in choosing assessment methods appropriate for instructional decisions.
2. Teachers should be skilled in developing assessment methods appropriate for instructional decisions.
3. The teacher should be skilled in administering, scoring and interpreting the results of both commercially produced and teacher-produced assessment methods.
4. Teachers should be skilled in using assessment results when making decisions about individual students, planning teaching, developing curriculum, and school improvement.

5. Teachers should be skilled in developing valid pupil grading procedures that use pupil assessments.
6. Teachers should be skilled in communicating assessment results to students, parents, other lay audiences, and other educators.
7. Teachers should be skilled in recognizing unethical, illegal, and otherwise inappropriate assessment methods and uses of assessment information.

These principles are considered the requirements for teachers' assessment literacy scales.

Stiggins (1999) has similarly asserted seven standards for classroom assessment literacy. The assessment literacy competences for teachers by Stiggins are:

1. Connecting assessments to clear purposes
2. Clarifying achievement expectations
3. Applying proper assessment methods
4. Developing quality assessment exercises and scoring criteria and sampling appropriately
5. Avoiding bias in assessment
6. Communicating effectively about student achievement
7. Using assessment as an instructional intervention (cited in Mertler & Campbell, 2005, p.7).

Mertler & Campbell (2005) compared The Standards for Teacher Competence in the Educational Assessment of Students (1990) with Stiggins Classroom Assessment Competence (1999). Table 1 shows the relationship between them:

Table 1.

The alignments of Stiggins' (1999) competence with the seven standards set by AFT, NCME, & NEA (1990)

Classroom assessment Competence (1999)	The Standards for Teacher Competence in the Educational Assessment of Students (1990)

1- Connecting assessments to clear purposes	Standard (1, 2 & 4) choosing, developing assessment method properly. Using assessment results in decision-making.
2- Clarifying achievement expectations	Standard (4) using assessment results in decision making
3- Applying proper assessment methods	Standard (1 & 2) choosing and developing assessment methods properly
4- Developing quality assessment exercises and scoring criteria and sampling appropriately	Standard (2 &5) developing assessment method appropriately. Developing valid grading procedure.
5- Avoiding bias in assessment	Standard (5 & 7) developing valid grading procedure. Recognizing unethical, illegal and in appropriate assessment method.
6- Communicating effectively about student achievement	Standard (6) communicating assessment results.
7- Using assessment as an instructional intervention	Standard (3 & 7) administering, scoring and interpreting. Recognizing unethical, illegal and in appropriate assessment method.

(Mertler & Campbell, 2005, p.8)

The standards of Assessment literacy and principles can be generally applied to language assessment literacy. Language assessment literacy consists of skills, knowledge and principles. Assessment literacy in language and the related textbooks has improved from covering only skills and knowledge to skills, knowledge and principles (Davies, 2008). Through perceiving these three components, language assessment literacy can be realized (Taylor, 2009). One of the problems of assessment

literacy is the lack of ability to translate knowledge into practice, therefore clear, relevant and practical guide with detailed activities and exercises as a model for language assessment literacy in the textbooks are definitely needed (Fulcher, 2012).

Assessment knowledge is of great importance for language teachers. Teachers should constantly update the assessment awareness through in-service development training courses. It is not only enough for assessment to be a part of pre-service teacher preparation programs (Malone, 2008).

2.8 Teachers' assessment literacy

Teachers' assessment literacy is a crucial part of teachers' quality and professional requirements. Teachers' assessment literacy can serve or limit education process and student achievement (Mertler, 2003). Teachers' lacks of belief, knowledge and practice of quality assessment have negative consequence on student achievement and education. Research showed that teachers' assessment literacy level is not satisfactory for assessing students (Brookhart, 2001; Chapuis & Stiggins, 2002; Mertler, 1998; Plake, 1993; Rogier, 2014; Stiggins, 2004). It has been made clear that teachers' preparation programs are not well organized to address all teachers need for classroom assessment (Mcmillan, 2005; Mertler, 2003, 2004, 2009; Popham, 2009; Shwartz & Lissitz, 1987; Stiggins, 1988, 2002, 2007, 2008; Taylor, 2009). Teacher preparation programs need modification to include assessment literacy courses because of teachers' necessity to improve classroom assessment skill (Brookhart, 2001; Chapuis & Stiggins, 2002; Mcmillan, 2005; Popham, 2011; Siggins, 1997, 2001; Zhang & Burrey-Stock, 2003).

Teachers' classroom practice in assessment is limited by their belief and knowledge about assessment literacy; however, teachers' assessment literacy level is different across classroom levels. With the increase of classroom level they teach, teachers tend to use tests that are more objective and care about the quality of assessment; teachers in lower levels rely on performance assessment more than tests (Stiggins, Conklin & Bridgeford, 1986).

Research in the assessment literacy field has not well developed yet (Fulcher, 2012); however, literature presents assessment literacy in three phases as:

1. Teachers' concept of assessment
2. Teachers' knowledge and skills
3. Teachers' classroom practice of assessment.

Teachers' concepts and value of assessment is the most significant factor in assessment interpretation and decision-making (Mcmillan, 2005). The misperceptions of assessment that restrict teachers' assessment literacy include the wrong view of assessment as non-related to teaching process (Brown, 2004; Shepard, 2000), the classical concept of scientific measurement (Shepard, 2000) and teachers underestimation of the role of assessment. Another element in teachers' assessment literacy is their knowledge. Most of teachers and principals engage in assessment performance without trusting their knowledge and ability (Stiggins, 2008).

Teachers' inappropriate belief and lack of knowledge leads to poor assessment practice. Teachers, who have no explicit view on students' performance, base their teaching methods and objectives on vague standards and lead to inadequate assessment practice (Stiggins, 1988). Research have shown that assessment practice lacks several important principles and procedures. Teachers collect a lot of evidence through classroom assessment without using them for instructional decisions (Valencia, 2002). Another problem is that the evidences used only for reporting learning achievement. Most of the teachers do not use assessment to boost learning; they are likely misinterpreting assessment results (Rogier, 2014). Teachers only focus on constructed response rather than critical thinking (Mcmillan, 2005). In most of the cases teachers do not refer to authorities, there is no significant application of experts' recommendations, and research findings in classroom practice (Mcmillan, 2005). There also may be a problem of translating knowledge and concept into practice by teachers for classroom assessment performance.

Suggestions have been cited in literature for the solution of teachers' assessment literacy and according to these suggestions; applying some principles can improve teachers' knowledge and skills (Davies, 2008). The recommendations are similar to the principles of the seven Standards to some extent and can be summarized as in the following:

1. Teachers should know how to apply assessment and solve assessment problems; how to interpret and decide about assessment results, they similarly need to be aware how these activities affect assessment performance (Airsian, 1994; Mcmillan, 2005; Stiggins, 2012).

2. Teachers should know that students' moderate efforts promote better learning and try to engage students through assessment to reach this level of effort (Mcmillan, 2005).

3. Self-assessment and portfolio assessment boosts students confidence, self-regulations, self-responsibility of learning and as a result increase motivation and learning achievement (Mcmillan, 2005; Stiggins, 2012)

4. Teachers are required to be knowledgeable of various aims of assessment and use them proficiently in appropriate context (Volante & Fazio, 2007).

5. Teachers have to have explicit visions of learning goals and teaching objectives to design quality assessment for learning accordingly (Stiggins, 2012).

6. Teachers should recognize possible obstacles of assessment, have ability to prevent them and be aware of the negative influences of unsound assessment (Stiggins, 1995).

7. Teachers have a duty to be capable of linking knowledge to practice properly (Scarino, 2013).

8. Teachers ought to be familiar with using ongoing assessment, making descriptive and formative feedback to provide opportunity of students' involvement in assessment process and learning decisions (Chapuis & stiggins, 2002; Stiggins, 2012).

It can be asserted that literature supports the assumption of this research. Teachers' assessment literacy level is not adequate. Teachers need more preparation for assessment knowledge, skill and principles to be able to conduct classroom assessment properly (Popham, 2006, 2011). Teachers' satisfaction, with the information they got from their own experience, may lead to misrepresenting students' learning (Stiggins, 2012).

The problem in the field of assessment literacy may well be that some claims and recommendations asserted by scholars are not supported by empirical evidences.

There is a real need for expanding empirical research in teachers' assessment knowledge and its impact on learning achievement; to support the principles and recommendations of the scholars.



CHAPTER III

Methodology

3.1 Presentation

In this chapter, the method of the research is presented. Methodology of the data collection using a survey for descriptive purpose and rationales of using this type of method were discussed. The population and sampling procedure of the study is shown. The procedures of data collection and justifications of the techniques were discussed. The survey instrument was explained and defined and final section describes the process of data analysis.

3.2 Research design

In this research, quantitative method was used. The present study is a descriptive study of English as a foreign language teachers' assessment literacy level regarding the seven standards of teachers' assessment competence for educational assessment of students (AFT, NCME, & NEA, 1990). The research tries to find out the strong and weak points of teachers' assessment literacy. The answer for the research questions could be found by applying a survey for collecting evidence for the proposed assumptions. The researcher used questionnaire to investigate the research questions. The result of survey using questionnaire in most cases are quantitative, the most frequent method of collecting data is questionnaire (Dörnyei, 2007). Questionnaire can provide researchers with immediate actual information to process (Lester & Lester, 2010).

The study is snapshot like description of teachers' assessment literacy levels. However, teachers' responses are generally based on all the information and experience they have gained in assessment literacy.

The purpose of using survey questionnaire is explained by Creswell (2009) as “provides a qualitative data or numeric description of trends, attitudes or opinions of population” (p.145). He also mentions that through studying samples of the population of the study, the findings could be generalized to the whole population.

Teachers of English as a foreign language reported about their own assessment level, by responding a close-ended multi-choice questionnaire; researchers’ bias issue could be easily addressed in the questionnaire survey besides the objective data collection, which makes the responses credible for the investigation.

Participants are more familiar with the questionnaire survey. They can easily be persuaded to participate, they do not need a lot of explanation, there is no need of specific appointment for instance they can respond in a time best suits them and they do not need to answer all the questions in one sitting. All these reasons may facilitate teachers’ cooperation with the researchers in survey questionnaire.

3.3 Context of the study

The population of this research is the teachers of English as a foreign language in schools in Northern Iraq. Most of the teachers are similar in terms of general background information, the preparation courses they have received and the working context. However, they are different in academic attainment, the levels they teach and experience, beside gender differences and age. All the teachers are speakers of Kurdish as a first language and most of them, if not all, are teaching English in Kurdish in public schools belong to the ministry of education in Northern Iraq.

Some of the teachers of English language have not graduated from faculties of education and teaching departments, therefore they have not prepared for teaching and assessing students, and this is true for teachers of other subjects. The teachers, who have not received teaching courses at the university and institutions, participate in one-month training course about psychology, pedagogy, education system and sometimes classroom management after a year of recruitment. This training course is called teachers’ suitability training courses; the aim is to make them suitable for teaching.

The context of this study includes teachers of English as a foreign language, teaching in different levels of education classes. The public education consist kindergarten two years of preschool education, basic education (1-9) and preparatory education (10-12). Basic education is subdivided into primary education (1-6) and secondary education (7-9). Teachers who graduated from institutions, holding diploma after two years of education, teach in basic schools specifically at primary levels. Teachers holding Baccalaureate certificate teach in basic schools at secondary level and at preparatory level. Teachers who hold Master's Degree are preferred to teach at preparatory levels. Therefore, academic attainment affects the working context and the class levels they teach.

Teachers of English as a foreign language have studied English for (7- 12) years in public schools basic and high school; and they have studied teaching English for (2-5) years according to academic attainment.

3.4 Sampling

Teachers of English as a foreign language participated in the survey in multiple schools from different geographical places of Sulaimaniyah and Erbil governorates. The schools and teachers were chosen from different parts of the region i.e. from the city centers and rural areas. Samples cover teachers holding different academic certificate, teaching at different classroom levels in order to represent the whole population. All participants are teachers of English as a foreign language and speakers of Kurdish as a first language.

The samples were taken from teachers of different gender, age and working experience. Teachers of both basic schools and preparatory schools in private and public education sectors are covered in data collection. Some participant teachers have not graduated from department of education and they have not studied teaching. The samples include teachers of English who have studied English language and literature and still they practice teaching at schools as a profession.

The sample design was single stage sampling. Samples were chosen randomly, hard copies of the survey form have been distributed to 150 teachers from different places of the region; 101 teachers have completed and returned the questionnaire. Therefore, the participants of this study are 101 teachers of English as

a foreign language. Choosing the samples is important for a study because of representing and forming the subsets of the whole population (Dörnyei, 2007). In this study, Classroom Assessment Literacy Inventory developed by Mertler (2002) was used to investigate teachers' assessment literacy levels.

3.5 Procedure

Before starting the survey, permission was taken from the author of the questionnaire form, Classroom Assessment Literacy Inventory (CALI) through email exchange. Teachers of English as a foreign language from different places were contacted or met by the researcher. Literature shows that questionnaire administration process has a significant effect in the excellence of elicited answers (Dörnyei, 2007); so, an introduction about the research and the significance of the research explained to the teachers. Participants should have impression that their ideas matter and about an important case (Dörnyei, 2007). Teachers were given necessary information about Classroom Assessment Literacy Inventory (CALI) to become familiar with the purpose of the survey and they were asked to choose one possible response to each question.

150 survey forms were distributed among EFL teachers and they were asked to respond in the best possible time that suits them. Contact information of the researcher was given to the participant teachers so as to give any further explanation in case of necessity. After collecting the forms, the answers were checked to ensure the accuracy of the responses; if any item left out, teachers were asked to review the item and choose one possible answer.

The participants were made sure that their responses were protected and only used for the educational research without reference to teachers' identity. Teachers' identities were represented in numbers for ethical consideration. At the end, 101 teachers from different places and levels completed and returned the questionnaire.

3.6 Instruments

Classroom Assessment Literacy Inventory (CALI) designed by Mertler (2003) was used in the process of data collection. The survey form was originally adapted from "Teacher Assessment Literacy Questionnaire" (1993), by Plake &

Impara. Necessary permission was taken from the author for using the questionnaire in this study.

The inventory consists of two parts: the first part consists of 35 items to investigate teachers' assessment literacy based on the seven Standards for Teacher Competence in the Educational Assessment of Students (AFT, NCME, & NEA, 1990). The second part consists of seven questions about personal and background information of teachers. Table 2 demonstrates how the 35 items in Classroom Assessment literacy Inventory (CALI) represents the seven standards.

Table 2

Alignment of teacher competence (1990) with CALI (2003)

<i>Standards for Teacher Competence</i>	<i>Number of items in</i>
Standard 1: choosing appropriate methods of assessment	1, 8, 15, 22, 29
Standard 2: developing appropriate methods of assessment	2, 9, 16, 23, 30
Standard 3: administering, scoring and interpreting assessment results	3, 10, 17, 24, 31
Standard 4: using assessment results for decision making	4, 11, 18, 25, 32
Standard 5: developing valid grading practice	5, 12, 19, 26, 33
Standard 6: communicating assessment results	6, 13, 20, 27, 34
Standard 7: recognizing unethical or illegal assessment practice.	7, 14, 21, 28, 35

Classroom Assessment Literacy Inventory consists of 35 multi choice questions with 4 options each. The questions measure teachers' assessment literacy basing on the seven standards, which means each five questions stand for one standard. The questions 1, 8, 15, 22 and 29 measure standard 1 *teachers' ability to choose proper methods of assessment as suggests*. The questions 2, 9, 16, 23 and 30 represents Standard 2 *developing appropriate methods of assessment*. The questions 3, 10, 17, 24

and 31 investigates Standard 3 *teachers' ability in administering, Scoring, and Interpreting the Results of Assessments*. The questions 4, 11, 18, 25 and 32 measure standard 4 *teachers' skill in using assessment results for educational decision making*. The questions 5, 12, 19, 26 and 33 explore standard 5 *teachers knowledge in developing valid grading procedure*. The questions 6, 13, 20, 27 and 34 examines Standard 6 *teachers' knowledge about communicating assessment results*. The questions 7, 14, 21, 28 and 35 measures standard 7 *teachers' knowledge in recognizing unethical or illegal practice and preventing them*.

The instrument, according to Plake, Impara, & Fager (1993), has shown the reliability of .54 for both pre-service and in-service teachers in the original survey. The mean score was 24.50 with a standard deviation of 4.92 in study conducted by Mertler (2003), CALI was given to 152 in-service teachers. In (2004) CALI was given to 249 pre- service teachers the mean score for this group was 22.98 with a standard deviation of 4.05. The overall reliability of the instrument was found to be .54 for in-service teachers and .74 for pre-service teachers. The original instrument was subjected to a thorough content validation in 1991, including reviews by members of the National Council on Measurement in Education (Mertler, 2003). The data collected in the current study from in-service English teachers (N=101) showed internal consistency reliability as .60. It can be claimed that the reliability in the current study is adequate and generalizable and it is a well-designed questionnaire, which is appropriate for the purpose of this study.

3.7 Data analysis

Descriptive analysis was conducted to explore teachers' assessment level according to the first part of the questionnaire. Each subscale based on the seven standards was investigated individually for determining the weak and strong aspects of teachers' assessment literacy. Means and standard deviation for each standards and the whole instrument were examined. Analysis of variance (ANOVA) was done to explore if there were any statistically significant differences among demographic variables, such as level of academic attainment, years of experience.

Analyses were calculated for the personal information of the participants. The number of participants and the percentages of the respondents for each variable was

presented. The variables include gender, years of experience, classroom level they teach, academic achievement and teachers thought about their preparation program.

All the statistical computations were calculated by using SPSS (v.20) to test the research questions of this study in order to investigate teachers' assessment literacy level and to discover the weak and strong areas of assessment knowledge.



CHAPTER IV

FINDINGS

4.1 Presentation

In this chapter personal and background, information of participant teachers is presented. Then descriptive statistical analyses of the independent variables are given as frequencies. After that, the results of the descriptive analysis of teachers' assessment literacy level were given. Finally, inferential analysis was presented to show whether there were differences among groups of the independent variables.

4.2 Analysis of background information

The survey, Classroom Assessment literacy Inventory (CALI) by Mertler (2003) was distributed among 150 teachers of English; only 101 completed and returned the form, which is equal to 67.3% of the intended participants. Participants were required to provide information about their gender, the classroom level they teach, educational level they have attained, classroom experience. Participants were also required to provide their own perception about their undergraduate preparation program to teach and assess students' performance. They were asked to give response whether they have received a standalone course in assessment or not.

Statistical analysis for the participants' gender frequency revealed 54.5% of participant teachers were male (n= 55), 42.6% female (n= 43), and three of the participants did not provide the information.

Table 5 describes the frequency of classroom levels that EFL teachers practice their teaching.

Table 3

Participants' teaching level frequency

		<i>Frequency</i>	<i>Percent</i>	<i>Valid Percent</i>
	primary level	19	18.8	20.4
	secondary level	17	16.8	18.3
Valid	preparatory level	50	49.5	53.8
	other	7	6.9	7.5
	Total	93	92.1	100.0
Missing	System	8	7.9	
Total		101	100.0	

As can be clearly seen in table 3, public schools are divided into basic (primary 1-6 and secondary 7-9) and preparatory schools (10-12). 18.8% of the participants teach at primary levels (n= 19), 16.8% teach at secondary schools (n= 17) and 49.5% of them teach at high schools (n=50). Seven teachers did not respond to this question. They may possibly teach in language summer schools or private courses but they are still teachers of English as a foreign language.

Table 4 shows a descriptive frequency of the academic attainment of the participants

Table 4

Frequency of academic attainment

		<i>Frequency</i>	<i>Percent</i>	<i>Valid Percent</i>
	M.A.	29	28.7	29.3
Valid	B.A.	62	61.4	62.6
	Diploma	8	7.9	8.1
	Total	99	98.0	100.0
Missing	System	2	2.0	
Total		101	100.0	

Table 4 clarifies that 28.7% English teachers hold Master degree (n=29). 61.4% of the respondents accomplished Bachelor degree (n= 62) and 7.9% of participants have Diploma degree after two years of education (n= 8).

Table 5 illustrates descriptive calculation of the frequency of the participants' classroom experience.

Table 5

Teachers' classroom experience

	<i>Frequency</i>	<i>Percent</i>	<i>Valid Percent</i>
	1-5 years	51	50.5
	6-10 years	33	32.7
	11-15 years	7	6.9
Valid	16-20 years	3	3.0
	21-25	2	2.0
	25-30	1	1.0
	Total	97	96.0
Missing	System	4	4.0
Total		101	100.0

Table 5 explains that 50.5% of participants have 1-5 years of teaching experience (n= 51), 32.7% of teachers have 6-10 experience (n= 33). 6.9% of teachers have 11-15 years of teaching experience (n= 3), 2% of participants have teaching experience 21-25 years (n= 2) and 1% of participant have above 25 years of classroom experience (n= 1). This calculation shows that most of the teachers in this study have less than 11 years of teaching experience.

Table 6 gives detailed information about frequency analysis of the question whether teachers have received a standalone course in classroom assessment as a part of their undergraduate program.

Table 6

Standalone course in classroom assessment

		<i>Frequency</i>	<i>Percent</i>	<i>Valid Percent</i>
	yes	43	42.6	42.6
Valid	no	58	57.4	57.4
	Total	101	100.0	100.0

Table 6 shows that 42.6% of the teachers (n= 43) took a standalone course in classroom assessment and 57.4% of participants (n= 58) did not take a standalone course in classroom assessment.

Table 7 displays description of the frequency level of teachers' perception about undergraduate preparation program for being classroom teacher.

Table 7

Perception of undergraduate preparation for classroom teaching

		<i>Frequency</i>	<i>Percent</i>	<i>Valid Percent</i>
	very unprepared to be classroom teacher	6	5.9	6.4
	somewhat unprepared to be classroom teacher	11	10.9	11.7
Valid	somewhat prepared to be classroom teacher	47	46.5	50.0
	very prepared to be classroom teacher	30	29.7	31.9
	Total	94	93.1	100.0
Missing	System	7	6.9	
Total		101	100.0	

In table 7 it is seen that only 16.8% of teachers (n= 17) consider themselves unprepared by the undergraduate preparation program, while 76.2% of teachers (n= 77) think that undergraduate preparation program prepared them for the job of being classroom teacher.

Table 8 illustrates teachers' perceptions of undergraduate preparation program for students' assessment in the classroom.

Table 8

Perception of undergraduate preparation for classroom assessment

		<i>Frequency</i>	<i>Percent</i>	<i>Valid Percent</i>
Valid	very unprepared for assessing student performance	8	7.9	8.2
	somewhat unprepared for assessing student performance	15	14.9	15.5
	somewhat prepared for assessing student performance	42	41.6	43.3
	very prepared for assessing student performance	32	31.7	33.0
	Total	97	96.0	100.0
Missing	System	4	4.0	
Total		101	100.0	

Statistical analysis presented in table 8 demonstrates that 7.9 % of teachers (n= 8) think very unprepared, 14.9 % of teachers (n= 15) think somewhat unprepared; while 41.6% of teachers (n= 42) reported somewhat prepared and 31.7% of teachers (n= 33) believe that they are very prepared for assessing students' performance. The descriptive frequency statistics show that 73.3% of English teachers (n= 74) feel prepared for assessing student performance.

4.3 Analysis of teachers' assessment literacy

Teachers' assessment literacy, in this study, was measured by Classroom Assessment literacy Inventory (CALI) was designed by Mertler (2003). The inventory is a modified version of Teachers Assessment Literacy Questionnaire (TALQ) originally invented by Plake & Impara (1993).

Statistical analysis was calculated to combine item variables to make new variables. First in order to explore each of seven standards individually then make one variable which include all standards. The aim of combining items was to identify overall assessment literacy performance and to decide about the weak and strong points of assessment knowledge among teachers of English as a foreign language. Each time five items in CALI are combined to make a new variable according to the standards.

Research question 1: what is assessment literacy level of Iraqi teachers of English as a foreign language measured by Classroom Assessment Literacy Inventory?

Descriptive analysis was done to investigate Iraqi EFL teachers' assessment literacy level according to the standards for Teacher Competence in the Educational Assessment of Students (AFT, NCME, & NEA, 1990). Table 9 below shows the descriptive statistics of their assessment literacy levels.

Table 9

Descriptive Statistics of teachers assessment literacy level

	<i>N</i>	<i>Minimum</i>	<i>Maximum</i>	<i>Mean</i>	<i>Std. Deviation</i>
Total correct answer	101	5.00	28.00	14.3465	4.45968
Valid N	101				

The average performance of teachers ($N=101$) was less than satisfactory. Teachers answered less than 15 questions correctly ($M=14.34$, $SD= 4.45$) out of 35 questions. The minimum correct response was 5 and the maximum correct response was 28. The statistics showed that Iraqi EFL teachers' assessment literacy level in the current study was found to be less than all the previous studies around the world.

Research question 2 what are the weak and strong aspects of Iraqi English language teachers' assessment literacy according to the standards of teacher competence?

Descriptive analysis was calculated for Classroom Assessment Literacy Inventory, according Standards for Teacher Competence in the Educational

Assessment of Students in order to explore the strong and weak points of teachers' assessment literacy. Details about strong and weak points of teachers' assessment literacy are explained in table 10.

Table 10

Statistical analysis of teachers' performance for each standard

	<i>Mean</i>	<i>Std. Deviation</i>
Standard 1- Choosing appropriate methods of assessment	2.2475	1.12611
Standard 2- Developing appropriate methods of assessment	2.3465	1.12637
Standard 3- Administering, scoring and interpreting assessment results	2.3366	1.07030
Standard 4- Using assessment results for decision making	2.4059	1.12408
Standard 5- Developing valid grading practice	1.8317	1.15818
Standard 6-Communicating assessment results	1.6931	.99742
Standard 7- Recognizing unethical or illegal assessment practice	1.4851	1.01601
Total correct answer	14.3465	4.45968

Note $N= 101$

The result showed that teachers' highest performance was on standard 4: *Using assessment results for decision-making* ($M= 2.4$, Maximum possible answer 5, $SD= 1.12$), and then for standard 2: *Developing appropriate methods of assessment* ($M= 2.34$, Maximum possible answer 5, $SD= 1.13$). The lowest performance was found in standard 7: *Recognizing unethical or illegal assessment practice* ($M= 1.48$, Maximum possible answer 5, $SD= 1.01$), and then standard 6: *Communicating assessment results* ($M= 1.69$, Maximum possible answer 5, $SD= .99$).

Research question 3: Does years of classroom experience make difference in teachers' assessment literacy level?

Analysis of variance ANOVA was done for all the standards as dependent variable, and years of classroom experience as a factor. Table 11 shows the detail of total correct answers and years of classroom experience.

Table 11

Years of classroom experience with total correct answers

	<i>Sum of Squares</i>	<i>df</i>	<i>Mean Square</i>	<i>F</i>	<i>Sig.</i>
Between Groups	88.130	5	17.626	.846	.521
Within Groups	1896.241	91	20.838		
Total	1984.371	96			

The result presented in table 11 showed no statistically significant difference for the variables ($p > .05$). Thus, we can say that years of classroom experience in the current study make no significant difference in teachers' assessment literacy.

Research question 4: Does academic attainment make difference in teachers' assessment literacy level?

Analysis of variance ANOVA was also calculated for all the standards as dependent variable, and academic attainment as a factor. The detail presented in table 12.

Table 12

Academic attainment with total correct answers

	<i>Sum of Squares</i>	<i>df</i>	<i>Mean Square</i>	<i>F</i>	<i>Sig.</i>
Between Groups	103.257	2	51.628	2.633	.077
Within Groups	1882.400	96	19.608		
Total	1985.657	98			

Inferential analysis showed that no statistically significant difference was found in terms of teachers academic attainments ($p>.05$). The findings of the current study showed that academic attainment makes no statistically significant difference in teachers' assessment literacy.



CHAPTER V

DISCUSSION

5. 1 Presentation

In this chapter, comparison of the findings of this study was presented and compared to the other studies in the field around the world. The strong and weak points of Iraqi EFL teachers' assessment literacy were discussed and demonstrated. Possible reasons have been explained depending on the background information that teachers provided.

5. 2 Iraqi EFL teachers' assessment literacy level

The findings of the current research showed that Iraqi EFL teachers' assessment literacy level was much less than satisfactory. Teachers responded less than 15 items correctly out of 35 items. However, teachers' correct response ranged from slightly more than 17 to 24 in the previous studies (Campbell, Murphy, & Holt, 2002; Davidheiser, 2013; Mertler, 2003; Perry, 2013; Plake, 1993; Plake, Impara, & Fager, 1993; Yantim & Wongwanich, 2013). A study was conducted by (Campbell, Murphy, & Holt, 2002) in the united stated showed that pre service teachers ($M = 21$) and In-service teachers' performance ($M=23$). Another study was conducted by Plake (1993) found ($M=23$) in assessment literacy performance. Plake, Impara, & Fager (1993) originally found assessment literacy level of teachers ($M= 22$).

Another study by Mertler (2003) showed assessment literacy level ($N= 197$) for in-service teachers ($M = 21, 96, SD = 3.44$), while pre-service teachers ($N= 67$) performance was less than in-service teachers ($M= 18.96, SD=4.67$). In a study was conducted in Pennsylvania by (Davidheiser, 2013), presented to Drexel University participants teach different core subjects ($N= 102$); assessment literacy level was found higher than all the previous studies and there was a huge gap with the current study findings ($M= 24.51, SD= 4.99$). Further a study presented to the university of

Montana by Perry (2013) teachers assessment literacy ($N= 14$), ($M=21.93$, $SD= 3.27$) was found. In all the above-mentioned research in the United States of America, teachers' assessment literacy performance was higher than the findings of the current study.

On the other hand in a study conducted by Yamtim & Wongwanich (2013) in Taiwan, Primary teachers assessment literacy level ($N= 19$) was found ($M= 17.11$, $SD= 3.62$). Teachers' assessment literacy performance. In the comparison of the findings of this study with the other studies about teachers' assessment literacy, conclusion can be drawn that Iraqi EFL teachers have the poorest level of assessment literacy so far.

Iraqi EFL teachers' low performance in assessment literacy may result from poor preparation program and lack of in-service training courses for teacher development. Despite of the limited resources in assessment and the dominant role of national testing on the teaching/ learning process.

5. 3 Strong and weak aspects of Iraqi EFL teachers' assessment literacy

The weakest aspect of teachers' assessment literacy in this research was found in standard7 of Teacher Competence in the Educational Assessment of Students (AFT, NCME, & NEA, 1990). According to standard 7 *teachers should be skilled in recognizing unethical and illegal and otherwise inappropriate assessment methods and uses of assessment information*. The result showed that teachers' highest performance was on standard 4. *Teachers should be skilled in using assessment results when making decisions about individual students, planning teaching, developing curriculum and school improvement*.

The findings revealed that teachers' assessment strong aspect in the current research is similar to the study by Perry (2013) in which teachers highest performance was also in standard 4 ($M= 4.07$, $SD= 1.00$). The weak points of teachers' assessment literacy were also similar to Perry's study in standard 7 ($M= 1.29$, $SD= .91$). The second weakest aspects of teachers assessment literacy, found in this study is slightly different from the study by Plake, Impara, & Fager (1993) in which teachers' weakest point was in standard-6 ($M=2.70$). The strong aspect of this research was found to be totally different from the study that was done by Plake, Impara, & Fager (1993) in

which the highest point of performance was in standard 3: *Administering, scoring and interpreting assessment results*, found ($M= 3.96$). Mertler (2003) found strong points of performance in standard three ($M= 3.95$) and lowest performance was in standard five ($M= 2.06$), which was different from the current. Yamtim & Wongwanich (2013) found different strong and weak points of assessment literacy as well; the weakest point was in standard five ($M= 1.79$, $SD= .79$) and the strongest aspect was on standard one ($M= 3.11$, $SD= .99$).

The strongest and the weakest aspects of teachers assessment literacy is moderately various, across study results. The various results may due to the difference in courses related to assessment in preparation programs, curriculum differences and education policy. Another reason may come from the fact that teachers learn from colleagues (Stiggins, 1988) and revise their own samples as a result the strong and weak points may be shared among them in similar contexts.

5. 4 Iraqi EFL teachers' conception on assessment

The findings of the present study showed that more than 57% of participants reported that they had not received a standalone course in assessment as undergraduate preparation. Scholars constantly claim the importance of assessment. Improvement in learning is inevitably connected to teachers' assessment literacy (Arter, 2001). In the context of this study, teachers have not received any in-service training course on assessment, at least for the last 10 years. Even there are rarely any courses for in-service teachers' quality improvement while teachers' quality is one of the most crucial factor of learning achievement (Lee & Wiliam, 2005).

This poor level of assessment literacy may result from the misconception of assessment practice and role. Teachers have performed the least in comparison to all the studies have been done in assessment literacy so far; yet more than 73% of teachers reportedly believed that the undergraduate program has prepared them for classroom assessment. Teaches' claim of being prepared can be regarded as false statement, the illogical contradiction between the claim of being prepared and the real performance may due to teachers' misunderstanding of classroom assessment; Particularly most teaches who have performed much better, reported that they have not prepared well for students' assessment (Mertler, 2009; Stiggins, 2002, 2007). The results showed

that teachers in the context of this study are unaware of the principles of assessment literacy. Therefore, preparation program is to blame for not making sure that teachers perceived the necessary knowledge before starting the job (Mertler, 2004).

Teachers may have not based their response on scientific background about assessment, since most teachers learned the strategies they use in assessment from colleagues and their own experience (Stiggins, 1988). If teachers have not based their knowledge on scientific procedures in assessing students, then they mislead themselves and the students (Stiggins, 2012). Teachers are not aware of the principles of quality assessment therefore they cannot represent true learning achievement and performance (Stiggins, 2012).

Despite of the limitations of this study, in which the data collection was a single stage survey and only teachers of English participated in the data collection process, there is enough evidence in hand to claim that teachers assessment literacy is much less than satisfactory. Teachers need in-service training courses for professional development. Teachers are not aware of the principles of assessment literacy and even they are not aware of their need to improve. Teachers need to be taught about unethical and illegal practice of classroom assessment more than other seven aspects.

In conclusion, the assessment literacy level of Iraqi English language teachers is not satisfactory due to several reasons. First the misconception of assessment practice and its impact on students achievement. Second teachers' unawareness of principles of classroom assessment because of insufficient or lack of information they have given in preparation program about assessment. Third teachers are not conscious of their own level of assessment literacy and their need to improve in assessment literacy.

CHAPTER VI

CONCLUSION AND RECOMMENDATIONS

6.1 Conclusion

It is clear that assessment is one of the most important factors of promoting learning and it has received considerable attention by professional organizations and scholars. Teachers' assessment literacy is claimed to be a necessary characteristic teacher and it has proven to be one of the factors that affect learning achievement; therefore, assessment literacy has become an important field of education research. The current research was a part of the global effort to raise the awareness of teachers' assessment level of knowledge and to shed some lights on what needs to be done.

The study was an attempt to determine Iraqi EFL teachers' assessment literacy level. Classroom Assessment Literacy Inventory (Mertler, 2003) was used to test the research questions of this study. The research was a descriptive qualitative description of teachers' assessment literacy; population and sampling were restricted to northern part of Iraq, specifically in Sylaimaniyah and Erbil governorate.

The findings provide empirical evidence in hand supported by 101 survey samples to claim that Iraqi EFL teachers' level of assessment knowledge is very low; according to Standards for Teacher Competence in the Educational Assessment of Students (AFT, NCME, & NEA, 1990). The results indicated that Iraqi EFL teachers perform the weakest level of knowledge as compared to all the previous research in this field (Campbell, Murphy, & Holt, 2002; Davidheiser, 2013; Mertler, 2003; Perry, 2013; Plake, 1993; Plake, Impara, & Fager, 1993; Yamtim & Wongwanich, 2013). Teachers' weakest performance was found in *recognizing unethical or illegal assessment practice* but they performed somehow better *in using assessment results for decision-making*. Iraqi EFL teachers lack knowledge about their assessment level since most of them reported being prepared and the findings proved the opposite.

They underestimate or misunderstand the role of assessment and their own potential need for improvement. Inferential analysis revealed no statistically significant difference for teachers' academic attainment and years of classroom experience. The findings illustrated that years of classroom experience and academic attainment have not affected teachers' level of assessment literacy as expected.

It is clearly proved that Iraqi EFL teachers are not prepared well to assess students' performance adequately and determine the true level of achievement. Teachers need improvement in assessment through reform in preparation programs and training courses. This research had implications for increasing interest in teacher quality and its impact on learning process.

6.2 Suggestions for authorities

The findings of the study gives empirical evidence in hand to claim that the preparation programs are not sufficient to prepare teachers for classroom assessment, especially more than 57% participants reported that they had not received a standalone course in assessment as undergraduate preparation. Revision and improvement of teachers preparation program in Northern Iraq is definitely needed in order to achieve the ultimate goal of education and schooling system. Most teachers have not received a standalone course in assessment, while assessment has a determinant role in learning achievement and has a controlling impact on classroom teaching. Therefore undergraduate preparation program should be extended to cover intensive courses to familiarize teacher candidates with the principles of quality assessment for learning. There is also an urgent need of appropriate in-service training courses for teachers' professional development. The training courses need to address all the principles and procedures of sound assessment. There should be a standard for teachers' quality and a clear policy for recruitment to raise the standards of teaching and learning.

The concept of assessment as a judgmental scale and the social impact of assessment have to be reviewed on and modified, this alteration need the collaboration work of education reformists, researchers and policy makers. Updated information about principles and procedures of assessment ought to be shared with teachers and school principals instantly.

Education reform should not only focus on changing course books and pedagogy. Curriculum reforms need to cover assessment concept, principle and procedures alongside with the other aspects. The important role of assessment in education should receive enough consideration by the schooling system.

Publications, workshops and seminars should be held in schools and directorates of education to provide teachers with the necessary information. Publications about obstacles and improvement of education can reach almost every school and teachers and may be a good manner to make teachers and principals aware of what they need.

6.3 Suggestions for teachers

Teachers should consider a balance between formative and summative assessment. Teachers need to involve students in assessment for learning to reduce the anxiety of testing process and make assessments a part of teaching process to support learning. It is better for teachers to use multi methods of assessment in collecting evidence about students learning progression constantly to provide a fair opportunity for all students and address individual differences. Results had better interpreted to design the need of the students for further learning rather than judging their ability. The aim should not be only on high marks except the real learning has occurred

Language assessment in public schools should be expanded to cover all the language skills in a way that service the curriculum. Language communicative role should be considered especially when the curriculum is based on communicative approach, the current testing practice that only cover grammar and vocabulary need revision and amendment.

6.4 Suggestions for further studies

The implication of the findings of this research may promote further research about the obstacles in front of education and raising standards of learning. Research need to focus on what makes education stuck in theoretical framework. There is a necessity of research addressing the quality of teachers since teachers' quality affects learning and education development directly. Obstacles in front of education and schooling system can possibly be curriculum, classroom reality, social appreciation

and teachers' quality. All these aspects need intention of the researchers to vanish the difficulties facing education improvement.

One of the limitations of this research is survey data collection, while it would have been more effective if the evidence collection had prepared as an ongoing process. It can be suggested for further studies to include classroom observation and assessment interpretation, since multi-method data collection could be more reliable for gathering evidence.



REFERENCES

- 21st century skills assessment. (2007). Tucson. Retrieved 11 October 2016, from http://www.p21.org/storage/documents/21st_Century_Skills_Assessment_e-paper.pdf
- Airasian, P. W. (1994). *Classroom assessment (2nd ed.)*. New York: McGraw Hill McGraw-Hill.
- American Federation of Teachers, National Council on Measurement in Education, & National Education Association. (1990). *Standards for Teacher Competence in Educational Assessment of Students*. Retrieved 10 January 2016, from <http://buros.org/standards-teacher-competence-educational-assessment-students>
- Anderson, R. (1998). Why Talk About Different Ways to Grade? The Shift from Traditional Assessment to Alternative Assessment. *New Directions for Teaching and Learning*, 1998(74), 5-16. <http://dx.doi.org/10.1002/tl.7401>
- Arter, J. (2001). Learning Teams for Classroom Assessment Literacy. *NASSP Bulletin*, 85(621), 53-65. <http://dx.doi.org/10.1177/019263650108562107>
- Assessment Reform Group (2002). *Assessment for learning: 10 principles*. 10 October 2016, <http://www.assessment-reform-group.org/CIE3.PDF>
- Bachman, L. (2004). *Statistical analyses for language assessment (1st ed.)*. Cambridge: Cambridge University Press.
- Bartlett, J. (2015). *Outstanding Assessment for Learning in the Classroom (1st ed.)*. New York: Taylor and Francis.
- Black, P. & D. Wiliam (2003). In Praise of Educational Research: Formative Assessment. *British Educational Research Journal*, 29 (5), 623-637.
- Black, P., & Wiliam, D. (1998 a). Assessment and Classroom Learning. *Assessment in Education: Principles, Policy & Practice*, 5(1), 7-74. <http://dx.doi.org/10.1080/0969595980050102>
- Black, P., & Wiliam, D. (1998 b). Inside the Black Box: Raising Standards through Classroom Assessment. *Phi Delta Kappan* 80, (2), 139-144, 146-148.
- Black, P., & Wiliam, D. (2005). Classroom Assessment Is Not (Necessarily) Formative Assessment (and Vice-versa). *Yearbook of the National Society for the Study of Education*, 103(2), 183-188. <http://dx.doi.org/10.1111/j.1744-7984.2004.tb00054.x>
- Black, P., & Wiliam, D. (2009). Developing the theory of formative assessment. *Educational Assessment, Evaluation and Accountability*, 21(1), 5-31. <http://dx.doi.org/10.1007/s11092-008-9068-5>

- Boud, D., & Falchikov, N. (2006). Aligning assessment with long-term learning. *Assessment & Evaluation in Higher Education*, 31(4), 399-413. <http://dx.doi.org/10.1080/02602930600679050>
- Brookhart, S. (2011). Educational Assessment Knowledge and Skills for Teachers. *Educational Measurement: Issues and Practice*, 30(1), 3-12. <http://dx.doi.org/10.1111/j.1745-3992.2010.00195.x>
- Brooks, J. G., and Brooks, M. G. *In Search of Understanding: The Case for Constructivist Classrooms*. Alexandria, Va.: ASCD, 1993.
- Brown, G. (2004). Teachers' conceptions of assessment: implications for policy and professional development. *Assessment in Education: Principles, Policy & Practice*, 11(3), 301-318. <http://dx.doi.org/10.1080/0969594042000304609>
- Brown, H. (2010). *Language assessment* (1st ed.). White Plains, NY: Pearson Education.
- Campbell, C., & Collins, V. (2007). Identifying Essential Topics in General and Special Education Introductory Assessment Textbooks. *Educational Measurement: Issues and Practice*, 26(1), 9-18. <http://dx.doi.org/10.1111/j.1745-3992.2007.00084.x>
- Campbell, C., Murphy, J. A., & Holt, J. K. (2002). *Psychometric analysis of an assessment literacy instrument: Applicability to pre-service teachers*. Paper presented at the annual meeting of the Mid-Western Educational Research Association, Columbus, OH.
- Cauley, K., & McMillan, J. (2010). Formative Assessment Techniques to Support Student Motivation and Achievement. *The Clearing House: A Journal of Educational Strategies, Issues and Ideas*, 83(1), 1-6. <http://dx.doi.org/10.1080/00098650903267784>
- Chappuis, S., & Stiggins, R. (2002). Classroom Assessment for Learning. *Educational Leadership*, 60(1), 40-43. Retrieved 10 October 2016, from <http://www.ascd.org/publications/educationalleadership/sept02/vol60/num01/toc.aspx>
- Cheng, L., Rogers, W., & Wang, X. (2008). Assessment purposes and procedures in ESL/EFL classrooms. *Assessment & Evaluation in Higher Education*, 33(1), 9-32. <http://dx.doi.org/10.1080/02602930601122555>
- Creswell, J. (2009). *Research design* (3rd ed.). Los Angeles, Calif.: SAGE.
- Crooks, T. (1988). The Impact of Classroom Evaluation Practices on Students. *Review of Educational Research*, 58(4), 438. <http://dx.doi.org/10.2307/1170281>
- Davidheiser, S. (2013). *Identifying Areas for High School Teacher Development: A Study of Assessment Literacy in the Central Bucks School District*. Drexel University.

- Davies, A. (2008). Textbook trends in teaching language testing. *Language Testing*, 25(3), 327-347. <http://dx.doi.org/10.1177/0265532208090156>
- Dörnyei, Z. (2007). *Research methods in applied linguistics* (1st ed.). Oxford: Oxford University Press.
- Fulcher, G. (2012). Assessment Literacy for the Language Classroom. *Language Assessment Quarterly*, 9(2), 113-132. <http://dx.doi.org/10.1080/15434303.2011.642041>
- Fulcher, G., & Davidson, F. (2010). *Language testing and assessment* (1st ed.). Abingdon: Routledge.
- Hargreaves, E. (2005). Assessment for learning? Thinking outside the (black) box. *Cambridge Journal of Education*, 35(2), 213-224. <http://dx.doi.org/10.1080/03057640500146880>
- Harlen, W., & James, M. (1997). Assessment and Learning: differences and relationships between formative and summative assessment. *Assessment in Education: Principles, Policy & Practice*, 4(3), 365-379. <http://dx.doi.org/10.1080/0969594970040304>
- Hattie, J., & Timperley, H. (2007). The Power of Feedback. *Review of Educational Research*, 77(1), 81-112. <http://dx.doi.org/10.3102/003465430298487>
- Heritage, M. (2008). *Learning progressions: Supporting Instruction and Formative Assessment*. In State Collaborative on Assessment and Student Standards (SCASS). Washington, DC: CCSSO. 8 June 2016, from http://www.ccsso.org/Resources/Publications/Learning_Progressions_Supporting_Instruction_and_Formative_Assessment.html
- Lee, C., & Wiliam, D. (2005). Studying changes in the practice of two teachers developing assessment for learning. *Teacher Development*, 9(2), 265-283. <http://dx.doi.org/10.1080/13664530500200244>
- Lester, J., & Lester, J. *Writing research papers* (15th ed.). Pearson. London: Sage Publications.
- Malone, M. (2013). The essentials of assessment literacy: Contrasts between testers and users. *Language Testing*, 30(3), 329-344. <http://dx.doi.org/10.1177/0265532213480129>
- Marzano, R., Pickering, D., & McTighe, J. (1993). *Assessing Student Outcomes* (1st ed.). Alexandria: ASCD.
- McMillan, J. (2005). Understanding and Improving Teachers' Classroom Assessment Decision Making: Implications for Theory and Practice. *Educational Measurement: Issues and Practice*, 22(4), 34-43. <http://dx.doi.org/10.1111/j.1745-3992.2003.tb00142.x>
- Mertler, C. (1998). Classroom assessment practices of Ohio teachers. *In the Annual Meeting of the Mid-Western Educational Research Association*. Chicago: Mid-

- Western Educational Research Association. 10 October 2016, from <http://files.eric.ed.gov/fulltext/ED428085.pdf>
- Mertler, C. (2002). Demonstrating the Potential for Web-based Survey Methodology with A Case Study. *American Secondary Education*, 30(2), 49-61. 9 June 2016, from <http://www.jstor.org/stable/41064450>
- Mertler, C. (2002). Job Satisfaction and Perception of Motivation among Middle and High School Teachers. *American Secondary Education* Vol. 31, No. 1 (Fall 2002), Pp. 43-53, 31(1), 43-53. 8 June 2016, from <http://www.jstor.org/stable/41064589>
- Mertler, C. (2003). Patterns of Response and Nonresponse from Teachers to Traditional and Web Surveys. *Practical Assessment, Research and Evaluation*, 8(22). 8 June 2016, from <http://PAREonline.net/getvn.asp?v=8&n=22>.
- Mertler, C. (2004). Secondary Teachers' Assessment Literacy: Does Classroom Experience Make A Difference? *American Secondary Education*, 33(1), 49-64. 9 June 2016, from <http://www.jstor.org/stable/41064623>
- Mertler, C. (2009). Teachers' assessment knowledge and their perceptions of the impact of classroom assessment professional development. *Improving Schools*, 12(2), 101-113. <http://dx.doi.org/10.1177/1365480209105575>
- Mertler, C. A. (2003a). Classroom assessment: A practical guide for educators. Los Angeles, CA: Pyrczak.
- Mertler, C. A., & Campbell, C. (2005, April). *Measuring teachers' knowledge and application of classroom assessment concepts: Development of the assessment literacy inventory*. Paper presented at the annual meeting of the American Educational Research Association, Montreal, Canada
- Mertler, C.A. (2003b, October). Preservice versus inservice teachers' assessment literacy: Does classroom experience make a difference? *Paper presented at the annual meeting of the Mid-Western Educational Research Association*, Columbus, OH. Retrieved 10 October 2016, from <http://wps.ablongman.com/wps/media/objects/1530/1567095/tchassesslit.pdf>.
- Moss, C., & Brookhart, S. (2009). *Advancing formative assessment in every classroom* (1st ed.). Alexandria, Va.: Association for Supervision and Curriculum Development.
- O'Neil, H., & Abedi, J. (1996). Reliability and Validity of a State Metacognitive Inventory: Potential for Alternative Assessment. *The Journal of Educational Research*, 89(4), 234-245. <http://dx.doi.org/10.1080/00220671.1996.9941208>
- Perry, M. (2013). *Teacher and Principal Assessment Literacy*. The University of Montana.
- Plake, B. S. (1993). Teacher assessment literacy: Teachers' competencies in the educational assessment of students. *Mid-Western Educational Researcher*, 6(2), 21-27.

- Plake, B., Impara, J., & Fager, J. (1993). Assessment Competencies of Teachers: A National Survey. *Educational Measurement: Issues and Practice*, 12(4), 10-12. <http://dx.doi.org/10.1111/j.1745-3992.1993.tb00548.x>
- Popham, W. (2009). Assessment Literacy for Teachers: Faddish or Fundamental? *Theory into Practice*, 48(1), 4-11. <http://dx.doi.org/10.1080/00405840802577536>
- Popham, W. (2011). Assessment Literacy Overlooked: A Teacher Educator's Confession. *The Teacher Educator*, 46(4), 265-273. <http://dx.doi.org/10.1080/08878730.2011.605048>
- Rogier, D. (2014). Assessment Literacy: Building a Base for Better Teaching and Learning. *English Teaching Forum*, 52(3), 2-13. 8 June 2016, from http://americanenglish.state.gov/files/ae/resource_files/etf_52_3_02-13.pdf
- Sadler, D. (1989). Formative assessment and the design of instructional systems. *Instructional Science*, 18(2), 119-144. <http://dx.doi.org/10.1007/bf00117714>
- Scarino, A. (2013). Language assessment literacy as self-awareness: Understanding the role of interpretation in assessment and in teacher learning. *Language Testing*, 30(3), 309-327. <http://dx.doi.org/10.1177/0265532213480128>
- Shepard, L. (2000). The Role of Assessment in a Learning Culture. *Educational Researcher*, 29(7), 4. <http://dx.doi.org/10.2307/1176145>
- Stiggins, R. (1988). Revitalizing Classroom Assessment: The Highest Instructional Priority. *Phi Delta Kappan*, 69 (5), 363-368. Retrieved from <http://www.jstor.org/stable/20403636>
- Stiggins, R. (1991). Relevant Classroom Assessment Training for Teachers. *Educational Measurement: Issues and Practice*, 10(1), 7-12. <http://dx.doi.org/10.1111/j.1745-3992.1991.tb00171.x>
- Stiggins, R. (1999). Assessment, Student Confidence, and School Success. *Phi Delta Kappan*, 81(3), 191-198. Retrieved 9 June 2016, from <http://www.jstor.org/stable/20439619>
- Stiggins, R. (2002). *Assessment Crisis: The Absence of Assessment for Learning*. *Phi Delta Kappan*, 83(10), 758-765. <http://dx.doi.org/10.1177/003172170208301010>
- Stiggins, R. (2005). Evaluating Classroom Assessment Training in Teacher Education Programs. *Educational Measurement: Issues and Practice*, 18(1), 23-27. <http://dx.doi.org/10.1111/j.1745-3992.1999.tb00004.x>
- Stiggins, R. (2005). From Formative Assessment to Assessment for Learning: A Path to Success in Standards-Based Schools. *Phi Delta Kappan*, 87(4), 324-328. <http://dx.doi.org/10.1177/003172170508700414>

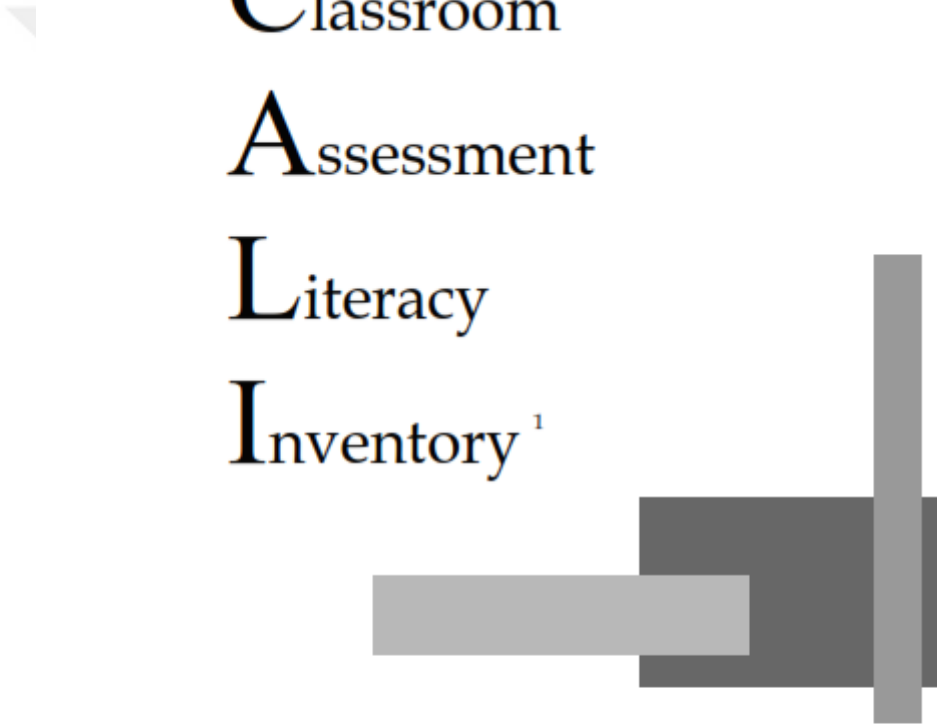
- Stiggins, R. (2007). Five Assessment Myths and Their Consequences. *Education Week*, 27(8), 3. Retrieved 11 October 2016, from http://www.ets.org/research/policy_research_reports/publications/article/2007/ddtg
- Stiggins, R. (2008). Assessment FOR Learning, the Achievement Gap, and Truly Effective Schools. In *Educational Testing Service and College Board conference, Educational Testing in America: State Assessments, Achievement Gaps, National Policy and Innovations*. Portland: ETS Assessment Training Institute. Retrieved 11 October 2016, from https://www.ets.org/Media/Conferences_and_Events/pdf/stiggins.pdf
- Stiggins, R. (2008). *Assessment Manifesto A Call for the Development of Balanced Assessment Systems* (1st ed.). Portland: ETS Assessment Training Institute. Retrieved 10 October 2016, from <http://www.ets.org/ati>
- Stiggins, R. (2009). Assessment for Learning in Upper Elementary Grades. *Phi Delta Kappan*, 90(6), 419-421. <http://dx.doi.org/10.1177/003172170909000608>
- Stiggins, R. (2012). *Classroom Assessment Competence: The Foundation of Good Teaching*. Retrieved 11 October 2016, from http://images.pearsonassessments.com/images/NES_Publications/2012_04Stiggins.pdf
- Stiggins, R., & Duke, D. (2008). Effective Instructional Leadership Requires Assessment Leadership. *Phi Delta Kappan*, 90(4), 285-291. <http://dx.doi.org/10.1177/003172170809000410>
- Stiggins, R., Conklin, N., & Bridgeford, N. (1986). Classroom Assessment: A Key to Effective Education. *Educational Measurement: Issues and Practice*, 5(2), 5-17. <http://dx.doi.org/10.1111/j.1745-3992.1986.tb00473.x>
- Valencia, S. (2002). *Understanding Assessment: Putting Together the Puzzle* (1st ed.). Boston: Houghton Mifflin. Retrieved 10 October 2016, from <https://www.eduplace.com/state/author/valencia.pdf>
- Volante, L., & Fazio, X. (2007). Exploring Teacher Candidates' Assessment Literacy: Implications for Teacher Education Reform and Professional Development. *Canadian Journal of Education / Revue Canadienne De L'éducation*, 30(3), 749. <http://dx.doi.org/10.2307/20466661>
- White, E. (2004). Are You Assessment Literate? Some Fundamental Questions Regarding Effective Classroom-based Assessment. *Oncue Journal*, 3(1), 3-25.
- Wiliam, D. (2007). Changing Classroom Practice. *Educational Leadership*, 65(4), 36-42. Retrieved 10 October 2016, from http://www.ascd.org/publications/educational_leadership/dec07/vol65/num04/Changing-Classroom-Practice.aspx
- Yantim, V., & Wongwanich, S. (2014). A Study of Classroom Assessment Literacy of Primary School Teachers. *Procedia - Social and Behavioral Sciences*, 116, 2998-3004. <http://dx.doi.org/10.1016/j.sbspro.2014.01.696>

Zhang, Z., & Burry-Stock, J. (2003). Classroom Assessment Practices and Teachers' Self-Perceived Assessment Skills. *Applied Measurement In Education*, 16(4), 323-342. http://dx.doi.org/10.1207/s15324818ame1604_4



Appendices

Appendix I: Classroom Assessment Literacy Inventory (CALI)



Classroom Assessment Literacy Inventory¹

Craig A. Mertler, Ph.D.
Bowling Green State University

¹ Adapted from the "Teacher Assessment Literacy Questionnaire" (1993), by Barbara S. Plake & James C. Impara, University of Nebraska-Lincoln, in cooperation with The National Council on Measurement in Education & the W.K. Kellogg Foundation

Description of the inventory:

This inventory consists of two parts. Part I consists of 35 items related to the seven "Standards for Teacher Competence in the Educational Assessment of Students." Some of the items are intended to measure general concepts related to testing and assessment, including the use of assessment activities for assigning student grades and communicating the results of assessment to students and parents; other items are related to knowledge of standardized testing and the remaining items are related to classroom assessment. Part II consists of items related to your background as a classroom teacher.

Directions:

Please read each item carefully and select the response you think is the best one by shading the corresponding circle. Even if you are not sure of your choice, but you *think* you know which is best, mark that response.

PART I

1. What is the most important consideration in choosing a method for assessing student achievement?
 - The ease of scoring the assessment.
 - The ease of preparing the assessment.
 - The accuracy of assessing whether or not instructional objectives were attained.
 - The acceptance by the school administration.
2. When scores from a standardized test are said to be "reliable," what does it imply?
 - Student scores from the test can be used for a large number of educational decisions.
 - If a student retook the same test, he or she would get a similar score on each retake.
 - The test score is a more valid measure than teacher judgments.
 - The test score accurately reflects the content of what was taught.
3. Mrs. Bruce wished to assess her students' understanding of the method of problem solving she had been teaching. Which assessment strategy below would be most valid?
 - Select a textbook that has a "teacher's guide" with a test developed by the authors.
 - Develop an assessment consistent with an outline of what she has actually taught in the class.
 - Select a standardized test that provides a score on problem solving skills.
 - Select an instrument that measures students' attitudes about problem solving strategies.
4. What is the most effective use a teacher can make of an assessment that requires students to show their work (e.g., the way they arrived at a solution to a problem or the logic used to arrive at a conclusion)?
 - Assigning grades for a unit of instruction on problem solving.
 - Providing instructional feedback to individual students.
 - Motivating students to attempt innovative ways to solve problems.
 - None of the above.
5. Ms. Green, the principal, was evaluating the teaching performance of Mr. Williams, the fourth grade teacher. One of the things Ms. Green wanted to learn was if the students were being encouraged to use higher order thinking skills in the class. What documentation would be the most valid to help Ms. Green to make this decision?
 - Mr. Williams' lesson plans.
 - The state curriculum guides for fourth grade.
 - Copies of Mr. Williams' unit tests or assessment strategies used to assign grades.
 - Worksheets completed by Mr. Williams' students, but not used for grading.

6. A teacher wants to document the validity of the scores from a classroom assessment strategy she plans to use for assigning grades on a class unit. What kind of information would provide the **best** evidence for this purpose?
- Have other teachers judge whether the assessment strategy covers what was taught.
 - Match an outline of the instructional content to the content of the actual assessment.
 - Let students in the class indicate if they thought the assessment was valid.
 - Ask parents if the assessment reflects important learning outcomes.
7. Which of the following would most likely **increase** the reliability of Mrs. Lockwood's multiple choice end-of-unit examination in physical science?
- Use a blueprint to develop the test questions.
 - Change the test format to true-false questions.
 - Add more items like those already on the test.
 - Add an essay component.
8. Ms. Gregory wants to assess her students' skills in organizing ideas rather than just repeating facts. Which words should she use in formulating essay exercises to achieve this goal?
- compare, contrast, criticize
 - identify, specify, list
 - order, match, select
 - define, recall, restate
9. Mr. Woodruff wanted his students to appreciate the literary works of Edgar Allen Poe. Which of his test items shown below will **best** measure his instructional goal?
- "Spoke the raven, nevermore." comes from which of Poe's works?
 - True or False: Poe was an orphan and never knew his biological parents.
 - Edgar Allen Poe wrote:
 1. Novels
 2. Short stories
 3. Poems
 4. All of the above.
 - Discuss briefly your view of Poe's contribution to American literature.
10. Several students in Ms. Atwell's class received low scores on her end-of-unit test covering multi-step story problems in mathematics. She wanted to know which students were having similar problems so she could group them for instruction. Which assessment strategy would be **best** for her to use for grouping students?
- Use the test provided in the "teacher's guide."
 - Have the students take a test that has separate items for each step of the process.
 - Look at the student's records and standardized test scores to see which topics the students had not performed well on previously.
 - Give students story problems to complete and have them show their work.
11. Many teachers score classroom tests using a 100-point percent correct scale. In general, what does a student's score of 90 on such a scale mean?
- The student answered 90% of the items on this test correctly.
 - The student knows 90% of the instructional content of the unit covered by this test.
 - The student scored higher than 90% of all the students who took the test.
 - The student scored 90% higher than the average student in the class.

12. Students in Mr. Jakman's science class are required to develop a model of the solar system as part of their end-of-unit grade. Which scoring procedure below will maximize the objectivity of assessing these student projects?
- When the models are turned in, Mr. Jakman identifies the most attractive models and gives them the highest grades, the next most attractive get a lower grade and so on.
 - Mr. Jakman asks other teachers in the building to rate each project on a 5-point scale based on their quality.
 - Before the projects are turned in, Mr. Jakman constructs a scoring key based on the critical features of the projects as identified by the highest performing students in the class.
 - Before the projects are turned in, Mr. Jakman prepares a model or blueprint of the critical features of the product and assigns scoring weights to these features. The models with the highest scores receive the highest grade.
13. At the close of the first month of school, Mrs. Friend gives her fifth grade students a test she developed in social studies. Her test is modeled after a standardized social studies test. It presents passages and then asks questions related to understanding and problem definition. When the test was scored, she noticed that two of her students—who had been performing well in their class assignments—scored much lower than other students. Which of the following types of additional information which would be most helpful in interpreting the results of this test?
- The gender of the students.
 - The age of the students.
 - Reliability data for the standardized social studies test she used as the model.
 - Reading comprehension scores for the students.
14. Frank, a beginning fifth grader, received a G. E. (grade equivalent score) of 8.0 on the Reading Comprehension subtest of a standardized test. This score should be interpreted to mean that Frank
- can read and understand 8th grade reading level material.
 - scored as well as a typical beginning 8th grader scored on this test.
 - is performing in Reading Comprehension at the 8th grade level.
 - will probably reach maximum performance in Reading Comprehension at the beginning of the 8th grade.
15. When the directions indicate each section of a standardized test is timed separately, which of the following is acceptable test-taking behavior?
- John finishes the vocabulary section early; he then rechecks many of his answers in that section.
 - Mary finishes the vocabulary section early; she checks her answers on the previous test section.
 - Jane finishes the vocabulary section early; she looks ahead at the next test section but does not mark her answer sheet for any of those items.
 - Bob did not finish the vocabulary section; he continues to work on that section when the testing time is up.
16. Ms. Camp is starting a new semester with a factoring unit in her Algebra I class. Before beginning the unit, she gives her students a test on the commutative, associative, and distributive properties of addition and multiplication. Which of the following is the most likely reason she gives this test to her students?
- The principal needs to report the results of this assessment to the state testing director.
 - Ms. Camp wants to give the students practice in taking tests early in the semester.
 - Ms. Camp wants to check for prerequisite knowledge in her students before she begins the unit on factoring.
 - Ms. Camp wants to measure growth in student achievement of these concepts, and scores on this test will serve as the students' knowledge baseline.

17. To evaluate the effectiveness of the mathematics program for her gifted first graders, Ms. Allen gave them a standardized mathematics test normed for third graders. To decide how well her students performed, Ms. Allen compared her students' scores to those of the third-grade norm group. Why is this an **incorrect** application of standardized test norms?
- The norms are not reliable for first graders.
 - The norms are not valid for first graders.
 - Third grade mathematics items are too difficult for first graders.
 - The time limits are too short for first graders.
18. When planning classroom instruction for a unit on arithmetic operations with fractions, which of these types of information have **more** potential to be helpful?
- norm-referenced* information: describes each student's performance relative to a other students in a group (e.g., percentile ranks, stanines), or
- criterion-referenced* information: describes each student's performance in terms of status on specific learning outcomes (e.g., number of items correctly answered for each specific objective)
- Norm-referenced information.
 - Criterion-referenced information.
 - Both types of information are equally useful in helping to plan for instruction.
 - Neither, test information is not useful in helping to plan instruction.
19. Students' scores on standardized tests are sometimes inconsistent with their performances on classroom assessments (e.g., teacher tests or other in-class activities). Which of the following is **not** a reasonable explanation for such discrepancies?
- Some students freeze up on standardized tests, but they do fine on classroom assessments.
 - Students often take standardized tests less seriously than they take classroom assessments.
 - Standardized tests measure only recall of information while classroom assessments measure more complex thinking.
 - Standardized tests may have less curriculum validity than classroom assessment.
20. Elementary school teachers in the Baker School system collectively designed and developed new curricula in Reading, Mathematics, and Science that is based on locally developed objectives and objectives in state curriculum guides. The new curricula were not matched directly to the content of the fourth grade standardized test. A newspaper reports the fourth grade students in Baker Public Schools are among the lowest scoring districts in the State Assessment Program. Which of the following would **invalidate** the comparison between Baker Public Schools and other schools in the state?
- The curriculum objectives of the other districts may more closely match those of the State Assessment.
 - Other school systems did not design their curriculum to be consistent with the State Assessment test.
 - Instruction in Baker schools is poor.
 - Other school systems have different promotion policies than Baker.
21. Which of the following choices typically provides the **most** reliable student-performance information that a teacher might consider when assigning a unit grade?
- Scores from a teacher-made test containing two or three essay questions related directly to instructional objectives of the unit.
 - Scores from a teacher-made 20 item multiple-choice test designed to measure the specific instructional objectives of the unit.
 - Oral responses to questions asked in class of each student over the course of the unit.
 - Daily grades designed to indicate the quality of in-class participation during regular instruction.

22. A teacher gave three tests during a grading period and she wants to weight them all equally when assigning grades. The goal of the grading program is to rank order students on achievement. In order to achieve this goal, which of the following should be closest to equal?
- Number of items.
 - Number of students taking each test.
 - Average scores.
 - Variation (range) of scores.
23. When a parent asks a teacher to explain the basis for his or her child's grade, the teacher should
- explain that the grades are assigned fairly, based on the student's performance and other related factors.
 - ask the parents what they think should be the basis for the child's grade.
 - explain exactly how the grade was determined and show the parent samples of the student's work.
 - indicate that the grading scale is imposed by the school board and the teachers have no control over grades.
24. Which of the following grading practices results in a grade that least reflects students' achievement?
- Mr. Jones requires students to turn in homework; however, he only grades the odd numbered items.
 - Mrs. Brown uses weekly quizzes and three major examinations to assign final grades in her class.
 - Ms. Smith permits students to redo their assignments several times if they need more opportunities to meet her standards for grades.
 - Miss Engle deducts 5 points from a student's test grade for disruptive behavior.
25. During the most recent grading period, Ms. Johnson graded no homework and gave only one end-of-unit test. Grades were assigned only on the basis of the test. Which of the following is the major criticism regarding how she assigned the grades?
- The grades probably reflect a bias against minority students that exists in most tests.
 - Decisions like grade assignment should be based on more than one piece of information.
 - The test was too narrow in curriculum focus.
 - There is no significant criticism of this method providing the test covered the unit's content.
26. In a routine conference with Mary's parents, Mrs. Estes observed that Mary's scores on the state assessment program's quantitative reasoning tests indicate Mary is performing better in mathematics concepts than in mathematics computation. This probably means that
- Mary's score on the computation test was below average.
 - Mary is an excellent student in mathematics concepts.
 - the percentile bands for the mathematics concepts and computation tests do not overlap.
 - the mathematics concepts test is a more valid measure of Mary's quantitative reasoning ability.
27. Many states are revising their school accountability programs to help explain differences in test scores across school systems. Which of the following is not something that needs to be considered in such a program?
- The number of students in each school system.
 - The average socio-economic status of the school systems.
 - The race/ethnic distribution of students in each school system.
 - The drop-out rate in each school systems.

28. The following standardized test data are reported for John.

Subject	Stanine Score
Vocabulary	7
Mathematics Computation	7
Social Studies	7

Which of the following is a **valid** interpretation of this score report?

- John answered correctly the same number of items on each of the three tests.
 - John's test scores are equivalent to a typical seventh grader's test performance.
 - John had the same percentile rank on the three tests.
 - John scored above average on each of the three tests.
29. Mr. Klein bases his students' grades mostly on graded homework and tests. Mr. Kaplan bases his students' grades mostly on his observation of the students during class. A major difference in these two assessment strategies for assigning grades can **best** be summarized as a difference in
- formal and informal assessment.
 - performance and applied assessment.
 - customized and tailored assessment.
 - formative and summative assessment.
30. John scored at the 60th percentile on a mathematics concepts test and scored at the 57th percentile on a test of reading comprehension. If the percentile bands for each test are five percentile ranks wide, what should John's teacher do in light of these test results?
- Ignore this difference.
 - Provide John with individual help in reading.
 - Motivate John to read more extensively outside of school.
 - Provide enrichment experiences for John in mathematics, his better performance area.
31. In some states testing companies are required to release items from prior versions of a test to anyone who requests them. Such requirements are known as
- open-testing mandates.
 - gag rules.
 - freedom-of-information acts.
 - truth-in-testing laws.
32. Mrs. Brown wants to let her students know how they did on their test as quickly as possible. She tells her students that their scored tests will be on a chair outside of her room immediately after school. The students may come by and pick out their graded test from among the other tests for their class. What is wrong with Mrs. Brown's action?
- The students can see the other students' graded tests, making it a violation of the students' right of privacy.
 - The students have to wait until after school, so the action is unfair to students who have to leave immediately after school.
 - Mrs. Brown will have to rush to get the tests graded by the end of the school day, hence, the action prevents her from using the test to identify students who need special help.
 - The students who were absent will have an unfair advantage, because her action allows the possibility for these students to cheat.

33. A state uses its statewide testing program as a basis for distributing resources to school systems. To establish an equitable distribution plan, the criterion set by the State Board of Education provides additional resources to every school system with student achievement test scores above the state average. Which cliché best describes the likely outcome of this regulation?
- Every cloud has its silver lining.
 - Into each life some rain must fall.
 - The rich get richer and the poor get poorer.
 - A bird in the hand is worth two in the bush.
34. In a school where teacher evaluations are based in part on their students' scores on a standardized test, several teachers noted that one of their students did not reach some vocabulary items on a standardized test. Which teacher's actions is considered ethical?
- Mr. Jackson darkened circles on the answer sheet at random. He assumed Fred, who was not a good student, would just guess at the answers, so this would be a fair way to obtain Fred's score on the test.
 - Mr. Hoover filled in the answer sheet the way he thought Joan, who was not feeling well, would have answered based on Joan's typical in-class performance.
 - Mr. Stover turned in the answer sheet as it was, even though he thought George, an average student, might have gotten a higher score had he finished the test.
 - Mr. Lund read each question and darkened in the bubbles on the answer sheet that represented what he believed Felicia, a slightly below average student, would select as the correct answers.
35. Mrs. Overton was concerned that her students would not do well on the State Assessment Program to be administered in the Spring. She got a copy of the standardized test form that was going to be used. She did each of the following activities to help increase scores. Which activity was unethical?
- Instructed students in strategies on taking multiple choice tests, including how to use answer sheets.
 - Gave students the items from an alternate form of the test.
 - Planned instruction to focus on the concepts covered in the test.
 - None of these actions are unethical.

PART II

36. What is your gender?
- female
 - male
37. Which of the following is the most appropriate description of the level at which you teach?
- elementary - primary (K - grade 3)
 - elementary - intermediate (grades 4 - 6)
 - elementary (K - 6)
 - middle (grades 6 - 8)
 - high (grades 9 - 12)
 - secondary (grades 6 - 12)
 - K - 12
 - other
38. Which best describes the educational level you have attained?
- B.A. or B.S.
 - M.A. or M.S.
 - Specialist
 - Ed.D.
 - Ph.D.

Appendix II: answers of Classroom Assessment Literacy Inventory (CALI)

Answers to CALI

- | | | | |
|-----|---|-----|---|
| 1. | C | 28. | D |
| 2. | B | 29. | A |
| 3. | B | 30. | A |
| 4. | B | 31. | D |
| 5. | C | 32. | A |
| 6. | B | 33. | C |
| 7. | C | 34. | C |
| 8. | A | 35. | D |
| 9. | D | | |
| 10. | B | | |
| 11. | A | | |
| 12. | D | | |
| 13. | D | | |
| 14. | B | | |
| 15. | A | | |
| 16. | C | | |
| 17. | B | | |
| 18. | B | | |
| 19. | C | | |
| 20. | B | | |
| 21. | B | | |
| 22. | A | | |
| 23. | C | | |
| 24. | D | | |
| 25. | B | | |
| 26. | C | | |
| 27. | A | | |

Appendix III: Permission from the author of CALI



Dr. Craig Mertler <craig.mertler@gmail.com>
to me 

 25/01/2016

Yes, you may have my permission to use the inventory in your research. All I ask is that you cite me appropriately.

In case you need it, I have attached the instrument, along with its scoring guide.

Best of luck!

Thank you,

Dr. Craig A. Mertler

CRAIG A. MERTLER, Ph.D.

Associate Professor

Leadership & Innovation EdD Program Coordinator

Arizona State University | Mary Lou Fulton Teachers College

P.O. Box 37100 | Phoenix, Arizona | 85069-7100 | Mail Code 3151

Office: FAB N278 | PH: 602.543.2829 | E-mail: Craig.Mertler@asu.edu

<http://www.craigmertler.com>



CURRICULUM VITAE

Fryad H N. Muhammad was born in Sulaimanyeah in 1983. He has achieved BA in English Language and Literature (2008) in college of languages-University of Sulaimanyah. He has been Teaching English as a Foreign Language (TEFL) for eight years. He was also the manager of safeen high school (2013-2014) which belongs Darbandikhan directory of education in sulaimaniyah governorate, and he is director of IDEL private school- Darbandikhan branch.

ÖZGEÇMİŞ

Fryad H N. Muhammad 1983 yılında Süleymaniye’de doğdu. Lisans derecesini Süleymaniye Üniversitesi İngiliz Dili ve Edebiyatı bölümünden aldı (2008). Sekiz senedir yabancı dil olarak İngilizce (TEFL) öğretimi yapmaktadır. Süleymaniye ili Darbandikhan eğitim idaresine bağlı Safeen Lisesi’nde (2013-2014) müdürlük yaptı. Günümüzde IDEL Özel Okulu- Darbandikhan şubesinde müdürlük yapmaktadır.