The Influence of TESOL Teacher Assessment Literacy on the Ethicality of English Language Assessment and the Position of Kurdish TESOL teachers

Dler Abdullah Ismael

University of Sulaimani, Kurdistan Region - Iraq
Email: dler276@gmail.com /dlerkanibard@gmail.com

_____

Abstract

In this paper, I start by demonstrating the importance of language assessmentespecially of English language. Then, in the second section,I attempt toconceptualize language testing and assessment by providing various definitions. Also in this section, there isa literature review of the major mainstream tenets of language assessment addedwith my critical views about their working, ethicality, validity and reliability. While the thirdsection is on CLT and allocated, firstly, for critical tenets of language testing and assessment particularlyofEnglish language, secondly, language assessment literacy, thirdly, ethicality of language assessment and fourthly the role of language assessment literacy on the language assessment ethicality; they are elucidated in different subsections. Also throughout this section, I explain CLTand its workingin the field, in what aspects and to what extent itprovidesethicality, reliability and validity with giving my critical views within CALxand critical pedagogy framework. Section four is for explaininghow appropriate and practicable to introduce criticalEnglish language testing and assessment to my professional context as a TESOL teacher in some English departments at the universities of Kurdistan Region/Iraq,making clear to what extent it is feasible and what are the educational challenges and obstacles to it.For explicating those issues, I refer to my recent experience as a TESOL teacherin my professional context.

Keywords: language assessment, critical testing, teacher assessment & evaluation, Teachers' Assessment Literacy, Language Assessment ethicality
_____

1. Introduction

Generally speaking, nowadays, teaching in many parts of the world is in the midst of a great transformation becauseteachers'expectations to get their students to high standards of performance and to ensure their learning are continually escalating (Hargreaves, 2000).Probably one of the most important aspects of the teaching process is assessment procedures since 250 studies discovered that the use of assessment to promote learning in the classroom improved student achievements (Earl and Katz, 2006). Therefore, evaluation of students' progress is considered as a major side of teachers' job (Brumfit and Johnson, 1979). However, until recently, assessment processes was based on traditional standard pen and paper tests only (Ataç, 2012) which I and probably many TESOL practitioners believe that it is inadequate butfortunately, the inquiry of language assessment isnow widened to recognize various factors other than language ability that impact test performance and to account for ethics and professionalism issues.Currently, language testers investigatethe use of language assessment, test developers' and users'ethical responsibilities, language assessment fairness and the impact and consequences of assessment useon instructional practice and societal values (Bachman, ND). Thus, establishing a valid assessment system is urgent for ELLs to requirecontinuous validation because assessment results are not only used to make academic decisions (Wolf, 2008), as one of the goals of assessing ELLs English proficiency is to measure their acquired skills to access content learning (Bailey and Wolf, ND) but also they have effective implications for improving the overall

teaching and learning qualities (Wolf, 2008).As engines of reform,tests can influence students, curricula, and educational systems (Winke, 2011). Also, because test consequences is associated with the usefulness of test results, there should be actions to reform learning byimproving feedback to students, theircomplaints or seeking re-assessments and assessment result consequences forthemwith considering test washback effects thatpossibly initiatesome changes in the current practices (EALTA Executive Committee, ND). To analyse and understand language assessment and its role in reforming teaching practice and critically exploring the effect of language assessment literacy on the ethicality of testing, we should first discuss the conceptualization of assessment and testing within their mainstream aspects.

2. Mainstream Language Assessment

2.1 Conceptualization of Language Testing and Assessment

Testing is a universal facet of social life. Throughout history people were tested toprove their abilities. However, currently tests have proliferated rapidly (McNamara 2000: 3) resulting in the appearance of many test types and procedures because of several educational issues (see section 1). To understand those test procedures, we should first conceptualize assessment and testing especially of language. Let us begin with the timing of the test that technicallydistinguishes three kinds of testing such as quizzes that are short, 5-10 minutes and includes just the current and tests that last 30-60 minutesand covers one or more units; whereas, exams are two hours or longer and contains at least half of the course content (Brumfit and Johnson, 1979). Whatsoever the length of time, assessment is basically a measurement process and measurement itself is used to determine the degree of something (Earl and Katz, 2006) by obtaining numerical descriptions (Brumfit and Johnson, 1979). So, language testing is a tool of measurement for determining test-takers' language proficiency (Piggin, 2012). Simply, the next step will be evaluation which uses the information collected by measurement. So, tests measure the achievement of students but the grades assigned depending on test results are evaluations of students' achievements (Brumfit and Johnson, 1979).At the end, there are three reference points for considering students'achievements and performance; firstly, students' performance in relation to pre-determined criteria (criteria-referenced); secondly, students' performance compared to other students'performance (norm-referenced); and thirdly, students' performance in relation to his/her previous performance (self-referenced) (Earl and Katz, 2006). In other words, testing demonstrates students' proficiency in a set of skills by comparing students' performance to a criterion and it determines students' ranks by comparing the students' performance to each other (Brumfit and Johnson, 1979) and then each student's progress over time. In the course of assessment, there are two assessment types: formative assessment which happens during teaching to make adjustments to the teaching process while summative assessment is at the end of a term to show students' progress by assessing a wider range of students' work, practical tasks, coursework and presentations (Earl and Katz, 2006). Another classification of testing is that of classroom teststhat are prepared by one teacher whose objectives can be based on course contents; whereas, standardized tests are designed for hundreds of thousands of subjects globally which are prepared by testing specialists(Brumfit and Johnson, 1979).

I believe that the recently developed alternative assessment procedures are the most valuable in students' learning process. Alternative assessments are the multiple techniques of assessment that show students' learning, achievement and motivation in instructional classroom activities in real-life situations including their attitudes on them. Many terms emerge for that sort of assessment likealternative, authentic, performanceor portfolio (Ataç, 2012).In such assessments, instruction and assessment are intrinsically integrated (Ripley, 2012).For example, portfoliosarethe collections of students' work accumulated over time to be organized to assess their competencies in a given standard or objective (Ripley, 2012) throughout their learning process. Another influential kind of alternative assessment is self-assessmentwhich is regarded as an integral component of language learning as well thatbasically indicates that no self-assessment means no self-awareness. Self-awareness is the knowledgeof one's level, strengths, weaknesses and preferred way of learning

(Cummins and Davesne, 2009). Additionally, self-assessment is to encourage learner awareness to obtain confidence and acquire a view of evaluation and see errors as helpful (Alderson and Banerjee, 2001). Additionally, there are many alternative assessment methods such as conferences, debates, demonstrations, diaries/journals, dramatizations, exhibitions, games, observations, peer-assessment, portfolios, projects, self-assessment, story retelling and think-alouds (Tsagari, 2004) alongside with many ways of recordinglike anecdotal records, checklists, learner profiles, progress cards, questionnaires and rating scales (Tsagari, 2004). Regarding the aims of language assessment,there are three basic purposes: assessment for learning, assessment as learning and assessment of learning. Assessment for learning isto give information to modify learning activities, to target instruction and resources and to give feedback to students. Assessment as learning is to develop and support metacognition for students by focusing on their roles as critical connectors between assessment and learning by reflection and critical analysis of their learning. While assessment of learning is the summative outcome used to confirm students' knowledge and abilities (Earl and Katz, 2006). After the definitions and conceptualization of testing and assessment, let's elucidate some of their mainstream sides.

2.2 Major Mainstream Tenets of Language Assessment

First of all, there are two contrasting assessment approaches, traditional and alternative, that are based on two different underpinning paradigmatic assumptions. It is said that traditional language assessment follows positivismwhich sees language ability as parallel to physical world objects;while, alternative assessment which is informed by interpretivismconsiders language proficiency as part of the social world. Acknowledging the latter, language knowledge can be pursued in ways other than the scientific method. Another distinguishing characteristic is related to the assessor'srelationship with the assessed; positivist testers use objectivity with the assessor that he/she should stay neutral and disinterested in the object of inquiry. But interpretivist testers find it impossible to separate facts from the subjectivity of values in relation to people and the social world. Thus, the abilities that an assessor tries to assess are seen as socially constructed not as external and independent of the assessor (Lynch and Shaw, 2005). Concerning the assessed, Shohamy(1998 cited in Piggin, 2012) urges that the definition of language proficiency should be critically examined because simply the test decisions can change test-takers' lives. Therefore, assessing both academic and social English language skills that provide a clearer picture of students' English proficiency (Stephenson et al., 2004) might follow Shohamy'surge. I think that this sort of assessment for more comprehensive language proficiency can be fulfilled by alternative assessment which happens at various points in time and in various ways both inside and outside the classroomthatmatches TESOL Inc.(2010) recommendationtouse various performance-based assessment tools and techniques. Moreover, alternative assessment results could be used to improve instruction (Tsagari, 2004)since major consequences of test use is itsimpact on instructionor washback (Bachman, 2005) by using tests as techniques for getting systematic evidence to base instructional decisions on them that can effectively enhance educational processes (Brumfit and Johnson, 1979).Being familiar with the two contrasting aspects of language assessment, it is time to discuss both in relation to assessment literacy and ethicality within the framework of CLT, critical pedagogy and CALx.

3. Critical Language Testing

CLT begins with the assumption that language testing is not neutral but it is a product of cultural, social, political and educational agendas that affect teachers' and learners' lives. Hence, test-takers are probably political subjects in political contexts. CLT explores whose agendas are behind tests, what visions of society tests presuppose, whose knowledge tests are based on and whether it is negotiable. CLT also inquires test meanings and scores and their openness to interpretations(Quang, 2007). Personally, I advocate CLT which attempts to challenge psychometrics of traditional language testing and supports interpretive approaches to language assessment(Quang, 2007). This signifies a

paradigmatic shift in which many new criteria for understanding validity are considered like consequential, systemic and interpretive(Quang, 2007) and more importantly, I believe, is the consideration of ethical issues that are associated with the effects (consequential validity) of tests. Here, CLT possibly supports alternative assessments to language testing like portfolios for developing more democratic testing methods in which test-takers and local bodies are more active (Quang, 2007) because for the quality of assessment, collecting information by piloting tests to get students' opinions is considered very important(The EALTA Executive Committee, ND). Democratic model might be the best way for English language assessment; next, I will consider some other sides of CLT.

3.1 Main Critical Aspects of English Language Assessment

It is realized that testing has impact on test-takers, stakeholders and the society (Hamp-Lyons, ND). However, testing as a criterion for graduation affects students differentlyfrom testing designed for informing instructional decisions (Educational Testing Service, 2009). To eliminate negative effects, testers should regard some basic principlesof testing that are significant in classroom assessment like reliability, validity,reference-points, record-keeping (Earl and Katz, 2006) fairness and washback (The EALTA Executive Committee, ND). This is to maketeachers'inferences about students' learning be credible, fair and free from bias (Earl and Katz, 2006). Regarding bias which I think it is more important than the other principles of language testing because it directly influences fairness and ethicality of language assessment. Bias isthe systematic unequal effect on specific subgroups in assessments concerning whether they are fully included, how they are treated, how they perform, how their performance is scored and how their test scores are used (Bachman, 2005).

My understanding is that ethical assessment cannot be achieved by traditional testingthat emphasizes the quality of testing by adopting accepted models and procedures for testing accuracy but little attention is paid to the test uses and their importance in test-takers' lives and their status in societysincethe power of tests is evidenced by listening to the voices of test-takersyet traditional testing does not focus much on the testing experience and the meanings and feelings that tests create in test-takers minds (Shohamy, 2001). Therefore, I believe that ethicality of traditional language testing is possibly violated by ignoring the impact testing brings about to students' lives because its major purpose and focus is creating qualitytests that can accurately measure the knowledge and it is a professional area that has strict rules and applications for appropriate practices (Shohamy, 2001). Furthermore, traditional testing techniques such as multiple-choice, fill-in-the-gaps, matching, etc. are often incompatible with the current ESL/EFL classroom practices. For example,high-stakes standardised tests are possibly the same as traditional testing in that it is done at one point of time and tests huge material which is the opposite of alternative assessments. Moreover, there is also negative criticism to the washback effects and consequences of high-stakes standardised tests that resulte in narrowing the curriculum by focusing only on those subjects and skills included in the examinations; thus, such tests might dominate and distort the whole curriculum. Additionally, this leads to that teachers restrict the methods to employ various exam preparation practices; this is at the expense of other learning activities that do not contribute directly to passing the exams. High-stakes standardised tests have effects on students' psychology as well that give them the role of passive recipients of knowledge with ignoring their needs and intentions; they have also detrimental consequences on students' intrinsic motivation, self-confidence, effort, interest and involvement in the language learning process and induce negative feelings in students' minds like anxiety, boredom, worry and fear. Moreover, teachers' psychology is also affected by the dictates of high-stakes tests that reduce their professional knowledge and status by pressuring on them to improve test scores which consequently make their teacher-made tests generate faulty results (Tsagari, 2004).

As a TESOL teacher,I thinkthat what is recommended by Assessment and Reporting Unit(2005) that assessment must be an integral part of course designs not just something to add afterwards can be certainly achieved by alternativeassessment procedures which are integral components of teaching and learning and that each learner is treated as a unique person with more informative views, with

more focus on strengths, progress, on-going assessment, culture-fair, possibility of several perspectives, improving and guiding learning and collaborative learning,with comparing learners to both theirpast performances and the aims with intrinsic learning for its own sake (Ataç, 2012). Alternative assessment means assessment procedures which are less formal, gathered over a period of time, formative in function, often low-stakes in terms of consequences and have beneficial washback effects. Also its advantages possibly are that they provide easily understood data, more integrative than traditional tests and more easily integrated into the classroom (Alderson and Banerjee, 2001) to promote learning and to enhance educational access and equity (McNamara, 1998 as cited in Alderson and Banerjee, 2001).That is why I do not think that classroom achievement tests that follow traditional approaches to testing (they in turn follow behaviouristic language-learning theories)(Brumfit and Johnson, 1979) that are possibly informed by positivism (see section 2.2) can fulfil the above mentioned recommendation about the integration of instruction and assessment.

Regarding the relation of interpretivismand alternative assessments, Moss (as cited in Lynch and Shaw, 2005)argues for an interpretivist hermeneutic approach that probably responds to many alternative assessment essential qualities as it acknowledges the effectiveness of the context of assessment and formulates validity as a consensus by dialogue between stakeholders: teachers, students and parents, not between disinterested external experts. Moss also challenges the generalizability principle of traditional testing in which educators should generalize from one performance to all similar performances and contexts (Lynch and Shaw, 2005). This is also a clear disapprovement of high-stakes tests and suggests that alternative assessment informed by interpretivism surely provides what Shohamy called democratic testing (see below) in which I believe there would be high degree of testing ethicality.

Concerning the consequences of testuse, in Shohamy's critical language testing, language testers are warned of the political uses and potential test abuses (Bachman, 2005) which is considered to be one of the main aspects of CALxthat draws connections between tests and issues of gender, class, ethnicity, culture, identity and politics (Pennycook 2008, cited in Talmy, ND). Therefore, she elucidates that test developers and test users need to be self-critical of the ways tests are used (Bachman, 2005) because it is quite clear that the use of tests to make decisions is fundamental and underlies virtually all uses of language tests (Bachman, 2005).For example, many Japanese students are worried about testing because Japan is an edumetric society in which test results influence life outcomes significantly (Newfields, 2006).This necessitates thatscore-based interpretations must be considered from the beginning of test design (Bachman, 2005). Then, it is urgent to develop a discourse of ethics as a foundation for decisions about classroom knowledge and pedagogy (Wardekker and Miedema, 1997) that certainly includeslanguage assessment.Here, I come back to Shohamy's democratic model of testing,she provides some guidelines for making testing more democratic as they might limit and control the powerful uses of tests (Broad, ND) and might achieve such ethical discourse by adopting critical pedagogy whichhelps students to question and challenge the dominating beliefs and practices (Riasati and Mollaei, 2012) including language testing in a democratic context. This is what Freire advocates asan anti-authoritarian, dialogical and interactive approach to examine relational power issues for students (Chandella and Troudi, 2013). To foster such a discourse of ethicality,Shohamy'sprinciples of critical language testing about questioning the informing values, agendas, goals, needs, purposes and actual uses of tests should be considered in the language assessments (Bachman, 2005). This is surely informed by CALxwhich tries to relate micro-relations of applied linguistics to macro-relations of society with constant questioning of all assumptions (Pennycook, 2001). Concerning democratization of assessment, Foucault regards examination as making individuals like objects of power and knowledge but substituting an alternative assessment for a traditional test does not necessarily change the power relations, if students do not control the aspects of the portfolios, for instance. In power relations, domination, exploitation and subjection are forms resulted from immobilizing and preventing any reversibility of movement. Foucault considers ethics as the practice of freedom; this can be found in portfolio assessment in which the freedom lies in the students' abilities to shape the portfolio process and form (Lynch and Shaw, 2005). Portfolio assessment focuses on validity and ethics affected by power relations (Lynch and Shaw, 2005).I

personally believe in the democratic model of language assessment that can be embodied by alternative assessments such as portfolios(Foucault as cited inLynch and Shaw, 2005)and I think it must be further emphasized because Howe (1994 as cited in Lynch, 1997) identifies that the democratic model is the most moral approach that provides a viable alternative by including the voices that have historically been banned from negotiating educational issues. This democratic framework also replaces the economic efficiency goals and the mastery of traditional domains of knowledge while enhancing the habits of mind that render individuals capable of to secure each individual's effective participation in the political processes of his/her society and to achieve a sense of self-esteem. (Lynch, 1997). This matches with what critical theories, pedagogy and research attempt to enlighten, empower and emancipate people from oppression (Brown and Jones 2001: 101-102) that sometimes constrain social and educational practice and produce results contrary to those desired by participants (Reason and Bradbury 2001: 95). To reduce negative power relations and unfairness in language assessment, we should pay a special attention to assessment literacy.

3.2 Language Assessment Literacy

At the beginning, TESOL teachers should know why language assessment literacy is important because, first, assessment is a widespread characteristic of educational systems, second, it is estimated that teachers spend 10%-50% on assessment activities, third, it enables teachers to share their classroom resultswith other teachers to develop a community of teachers that fosters learning and more importantly, fourth, it is suggested that assessment literacy is an influential aspect of teachers'professional development (Newfields, 2006).Additionally, one of the eleven TESOL standards of TESOL Inc. is about ESL/EFL assessment which is regarded asapplications that need pedagogical knowledge (Thibeault et al., 2011); hence, pedagogical knowledge for obtaining assessment literacy. What I discuss in this paper is concerned with TESOL teachers' assessment literacyrather than others because assessment literacy is conceptualized from three contrasting views. For students, the concept of assessment literacy means knowing how to perform well on exams; while for teachers, it is the ability to ethically and accurately grade students; whereas for professional test developers, every side of their work rests on assessment literacy (Newfields, 2006).

Assessment literacy for language teachers includes having: 1. The ability to employ a variety of assessment measures with minimal bias, 2. The ability to construct, administer and score tests, 3. The ability to evaluate the reliability, item difficulty, item facility and content validity of tests, 4. The ability to statistically determine the cut-off point of examination, 5. The ability to appropriately intervene when students engage in unethical behaviour during tests and 6. The skill in communicating assessment results to parents, peers and students. These assessment literacy items are based on the 1990 Standards for Teacher Competence in the Educational Assessment of Students that are published jointly by the American Federation of Teachers, the National Council on Measurement in Education and the National Education Association (Newfields, 2006).TESOL teachers should become aware of those assessment literacy points to assess the ESL/EFLstandards that are related to the goals of, firstly, to use English to communicate in social settings, secondly, to use English to achieve academically in all content areas, and thirdly, to use English in socially and culturally appropriate ways (Short, 2000). Those standards should be adhered to for ensuring the qualityof teaching and assessment (Aldersonand Banerjee, 2001).I believe that those assessment standards should be followed by all TESOL teachers across different cultures and education systems because it is directly related to assessing ESL/EFL proficiency in all its aspects and I agree with what explained in Newfields (2006) concerning the appropriacy of the application of those standards in Japanese contexts in spite of the different testing cultures. However, we have to know that in some contexts, teachers are not involved in assessment decision-making processes due to top-down managerial approaches (Troudi et al., ND).That is why in such cases, even if they have agreeable assessment literacy, it does not work. Therefore, local expertise should be supported by considering teachers' knowledge and views and by providing professional development opportunities for them (Troudi et al., ND).

Because of its significancefor TESOL teachers, there are some recommendations that might enhance assessment literacy as follows: 1. Beta-testing the test and revising it after examining thetestee responses, 2. Paying special attention to the cut-off points and questions that over 90% are answered correctly or incorrectly when grading a test, 3. Explaining the descriptive statistics clearly when mentioning test scores to students, 4. Making assessments educationally valid and sufficiently clear to stakeholders when deciding how to grade a course, 5. Improving micro-assessment and daily feedback skills and 6. Consulting with peers and thinking to rectify the problem of unethical assessment practices (Newfields, 2006). I think that ethicality involves all the above points in some way or other because basically these are followed to maximize ethicality of language assessment or other assessments; hence, the main purpose of education which is serving humans.

3.3 Ethicality of Language Assessment

Semantically speaking, ethics,morality and fairness are members of the same semantic set (Hamp-Lyons, ND). Ethicality includes the issues of harm, consent, fairness, deception, privacy and confidentiality(Lynch, 1997) andalso covers validity, absence of bias, access, administration and social consequences (Kunnan, 2003 cited in Bachman, 2005).Moreover, ethics, for language testers, involves as well whose voices are heard, whose needs be met and how society determines what is best when fairness is in conflict (Hamp-Lyons, ND). Concerning consent, tested students are rarely consulted aboutfor what they are being tested and what will bethe consequences. Deception is that test-takers' abilities are measured and tested indirectly; here, face validity is considered a moral issue. Privacy and confidentiality is related to the concern for public humiliation as a result of being tested. Therefore, it is identified that the core principle of ethical language testing is no test-taker will be harmed by the tests(Lynch, 1997). This is also related to test validity since ethical issues and practices to protect participants' rights not to be harmed, coerced or manipulated socially, psychologically, emotionally and physically is certainly an essential part of test validity(Lynch and Shaw, 2005).To make more account of ethicality of language testing, we should come back to CLT which addresses questionsof how and why language assessments are used, societal values, consequences, and test developers' and users' ethical responsibilities (Bachman, ND) because test-takers' rights not to be harmedremain constantwhether in traditional or alternative assessmentdespite the fact that educators' definition ofcoercion and ethical responsibility might differ (Lynch and Shaw, 2005). However, I feel in agreement with what is discussed in Lynch(1997)that if denying some test-takers entrance to university or preventing some from accessing specific social and economic resources is harmful, so no test would be moral at all. Whatsoever, tests are regarded as means for finding differences in abilities but they do not create those differences (Lynch, 1997). Furthermore, we should also base our understanding and analysis onthecontradictory strategies for considering fairness from different stakeholders' perspectives such as taxpayers, education department officials, businessmen, political parties and governments. That is why making tests fit these fairnesses is quite difficult and what makes it more complex is that fairness is a concept that there is no one standpoint from which to consider a test as being fair or not. Moreover, the language testers have no more right than anyone else to decide what is fair or not but they have the responsibility to make tests as fair as possible. Some examples of possible misuses of language testing might be the use of IELTS with applicants for immigration to New Zealand and probably using TOEFL and other proficiency tests to measure test-takers' achievements and growth in instructional programmes (Alderson and Banerjee, 2001). This heightened interest in ethics and tests role in society as McNamara (1998 as cited in Alderson and Banerjee, 2001) anticipates a renewed awareness of the socially constructed nature of test performance and test score interpretations and an awareness of the issues of testing in the context of English as an international language. Shohamy (1997a as cited in Alderson and Banerjee, 2001) explicates directly that language tests that contain unfair content or methods are unethical. She also argues that test uses to control and manipulate stakeholders rather than showing proficiency levels are unethical as well; hence, she advocates CLT. Thus, ethically speaking, the strength or credibility of the evidence on which we make test decisions (Bachman, 2005) is, I believe, extremely important for

language testers who are considered as independent moral agents that can refuse participation in procedures that violatetheir personal moral beliefs; this goes with the basic intent of the Code of Ethics which is that simply testers must adopt ethical practices (Alderson and Banerjee, 2001). What also maximize fairness and accessibilityof tests is that the influence of construct-irrelevant knowledge of test materials shouldbe minimized and unnecessarily controversial, inflammatory, offensive or upsetting test material must be avoided (Educational Testing Service, 2009).Concerning equal chances to students, probably in moral tests everyone has a fair chance to demonstrate his/her ability. However, if broader sense of test validity and ethicality is considered, one cannot limit his/her attention to providing onlyfair chances to test-takers because this does not take into account the fairness in the consequences of language testing decisions. Therefore, there should be ways to justify those decisionsdepending on differential test performances in accordance with the requirements of equality(Lynch, 1997) that consequently maximize ethicality.Regarding legalization of ethicality, the impetus for appropriateness and responsiveness of assessing young ELLs is supported by some legal requirements and ethical guidelines like case law, public law, and ethical codes from professional organizations that support using sound assessment tools, practices, and interpretations (Garcia et al., ND),and last century, because it became obviousthat schooling was a key to social mobility and that achievements in school was used for entry into the higher education or workplace, many jurisdictions instituted standardized testing with classroom assessment to ensure fair, accurate, and consistent opportunities for all students (Earl and Katz, 2006). All these actions indicate the fundamentality and significance of ethicality and fairness in English language assessment. Additionally, almost all codes (which represent the widely accepted beliefs about validity, reliability, washbackand fairness of language tests)consist of a set of expectations to potentially judge language testers and to improve the quality of language assessment and test fairness (Jia, 2009).

Another effective aspect of teachers' assessment literacy is that teachers should be aware of using multiple methodsin assessments (alternative Assessment) to ensure fairness for students (Troudi et al., ND).Allowing students to show their knowledge in many ways by a range of assessment approaches and multiple measures to establish a composite picture of student learning (Earl and Katz, 2006)possibly is the best way to achieve that language testing fairness.For the fairness of testing, validation which is the accumulation of evidences to justifytest interpretations or uses (Lynch and Shaw, 2005) should be considered as well.Lynch (2001 as cited in Bachman, 2005) describes validity framework as includingfairness, ontological authenticity, cross-referential authenticity, impact-consequential validity, and evolved power relations (Bachman, 2005) among them consequential validity encompasses ethical, social, and practical considerations (Winke, 2011). Wolf's(2008) discussion of validity covers some major validation components of test validity such as validity, reliability and testing-system adequacy. Test validity also includes some certain criteria such as field testing, design, freedom from bias, content, construct, criterion and consequential validities. Freedom from bias covers content, ethnicity, gender, disabilities, universal design, and linguistic, socioeconomic and geographic factors. Additionally, the validation process also considers, first, the construct to be measured like ELP, then, the interpretations of the test as what level of language proficiency students reach, after that, the purposes of tests, e.g. placement or determining progress (Wolf, 2008).Regarding the relation of validity and reliability, they could be viewed as complementary sides of identifying and interpreting variance sources in test scores; yet, a reliable test can be invalid, e.g. a speaking test comprised of multiple-choicesmight have high reliability but not be a valid test of speaking since its scores may not exactly represent the test-takers' real-world speaking abilities (Winke, 2011). After discussing assessment literacy and ethicality of language testing, let's investigate the possibility of the influence of language assessment literacy on language testing ethicality.

3.4 The Influence of Language Assessment Literacy on Language TestingEthicality

Concerning the assessment literacy of TESOL teachers, I think that it has a major role in assessment ethicality and fairness because it is urgent for TESOL teachers to be benchmarked to lead students to have better self-knowledge and not to be misled by their teachers' encouragement or criticism to see

themselves as more successful than they are or less successful than they are. Also, teachers entering new teaching situations, school years and kinds of learners possibly need re-benchmarking courses (Hamp-Lyons, ND). This goes with what TESOL Inc.(2012) recommends that first TESOL teachers should beequipped with knowledge of assessment, i.e. raising their language assessment literacy to empower students and measuretheir developments properly (TESOL Inc., 2012). This proper measurement with empowered students guarantees, to a good extent, the ethicality and fairness of ESL/EFL assessment in a democratic testing context, and I think that this reminds all stakeholders and practitioners to be aware of the power of testing and comprehend its ethical issues (Shohamy, 2001).Thus, when they know about that language testing has influence on students and society in general, language testing policymakers, specialists, and test users are forced to attempt to minimize the negative consequences and to maximize the positive consequences of using high-stakes tests of L2 ability (Stoynoff, 2008), of course, to foster the ethicality and fairness of testing and assessment.Another contributing factor to teachers' awareness is that of the development of language testing standards and professional morality among language testers which are to protect all from the misuse and abuse of tests since tests are used as instrumentsof social policy and control (Alderson and Banerjee, 2001).

Moreover, TESOL teachers should be aware of assessment procedures that include the issues of who designs assessment procedures, appropriateness of assessment procedures to the learners, how information on students' learning is collected, how it is assessed and stored, accuracy and fairness of assessment results and students' views on assessment procedures (The EALTA Executive Committee, ND). Additionally, TESOL teachers should understand the issues and concepts of assessment like accountability bias, language proficiency, testing accommodations anddifferent purposes of assessment for measuring language knowledge or ability to be equitable, accurate, consistent, and practical to administer, i.e. to be fair, valid, reliable and easy respectively. In this case, performance-based assessments can best measure these criteria that possibly cover formative and summative assessments (TESOL Inc., 2010).Furthermore, TESOL teachers must understand that assessments for English native-speakers and ELLs differ. Assessments for ELLs might contain cultural bias such as unfamiliar-to-ELLs images or references or contain linguistic bias likesome language items are more difficult for ELLs due to their complex language (TESOL Inc., 2010). This knowledge enhances the ethicality of their language testing and assessment in favour of both sides. Additionally, they should know that these assessments can be used to show language growth over time and to find areas that need more focus.To this end, they must know about portfolio assessment which is a collection of students' work that reflects progress over time and its samples are based on class activities or home assignments (TESOL Inc., 2010). Also to maximize ethicality,TESOL teachers have to be sure about their students' prior experience with the test question and answer formats as well (TESOL Inc., 2010).They should also understand that self-assessment and peer-assessment methods must be used regularly to push students monitor and control their learning (TESOL Inc., 2010).Then, these sorts of awareness can certainly be contributing factors to ensure fairness and ethicality of language assessment in general and English language assessment in particular. So far in this paper, literacy and ethicality of language assessment especially of English is talked about in relation to mainstream and critical tenets of language assessment. Next, the critical sides of the literacy and ethicality of English language assessment will be considered with reference to Kurdish TESOL teachers' educational context.

4. Feasibility of Critical English Language Assessment in the Context of English Departments at KurdistanRegion Universities and the Educational Challenges

4.1 Introduction

In this section, I will attempt to elucidate how appropriate and practicable to introduce critical English language testing and assessment to my professional context as a TESOL teacher in some English departments at the universities of Kurdistan Region/Iraq,regarding in what aspects and to what extent

this is feasible and whateducationalchallenges are predicted to obstacle it.I start by describing the current situation of English language assessment in my professional context in which all English language testing procedures are still based on traditional standardisedpen and paper tests only (Ataç, 2012).This shows that it has not reached the current widened scope of language assessment to pay attention to various factors that impact test performance, ethics and professionalism (see section 1) and how assessment results have effective implications for improving overall teaching and learning qualities (Wolf, 2008) andthat this can be fulfilled by alternative assessments but possibly not by traditional testing which is currently adopted by Kurdish TESOL teachers. I believe that Kurdish test specialists, TESOL teachers and EFL undergraduate students do not have the knowledge of the current advances in English language assessment initiated by TESOL Inc. assessment standards and CLT that shows abig gap and inadequacy of their English language assessment literacy.Kurdish TESOL teachers'lack ofassessment literacy is obvious from their assessment practices based on traditional methods of language testing like multiple-choice, fill-in-gaps and matching items through pen and paper exams that I believe in many cases;they are not comprehensive, valid and fair.

4.2 The Importance of Alternative Assessments for a More Comprehensive and Fairer English Language Testing in the Context of Kurdish TESOL Teachers

The test decisions that can change test-takers' lives (Shohamy, 1998 cited in Piggin, 2012) necessitates assessing both academic and social English language skills to provide a clearer picture of students' English proficiency (Stephenson et al., 2004). So, I think that alternative assessments which happen at various points in time and in various ways both inside and outside the classroom (Tsagari, 2004) are the best procedures to assess academic and social English as TESOL Inc.(2010)also recommends the use of various performance-based assessment tools and techniques. This is to usethe results to improve instruction based on the learner progress overtime (Tsagari, 2004). However, concerning the current assessment procedures of Kurdish TESOL teachers, I believe that they do not have the knowledge of alternative assessmentmethods and result recordingtechniques (Tsagari, 2004) that are widespread nowadays.Therefore, I think that the ways of assessing English language proficiency in the English departments in Kurdistanare not adequate and possibly not fair since it is chiefly done by traditional testing whether monthly or a sort of high-stakes final year exams which are not multi-dimensional and does not assess all sides of Kurdish students' proficiency of English. Thus, their classroom performance-assessments of the students are very restricted and there are not equal opportunities for all students. I believe that even in those restricted practices, the TESOL teachers cannot be fair enough as some students are more engaged by the teachers than others possiblydue to the large number of students nearly 40-50 students in a class or related to gender discrimination when females are engaged more by male single teachers or vice-versa which demonstrates bias to a certain gender during assessment. In such cases, sometimes, assessment is built on teacher-student personal relationships that affect even giving marks as well which is surely unethical.

Interpretive alternative assessments to language testing are possibly supported CLT to develop more democratic tests in which test-takers and local bodies are more active (Quang, 2007) through piloting testsby students' opinions (The EALTA Executive Committee, ND).To this end, Shohamy's guidelines for making testing more democratic(Broad, ND), Freire'santi-authoritarian, dialogical and interactive approach (Chandellaand Troudi, 2013)and Foucault's consideration ofethics as the practice of freedom can be achieved in portfolio assessment that in which freedom lies in the students' abilities to shape the portfolio process and form (Lynch and Shaw, 2005). I believe that this democratic model by portfolio assessment is quite suitable for my professional context. However, in my context,TESOL teachers have no knowledge or literacy of that model as they mostly use multiple-choice, fill-in-the-gaps, matching, etc.that are often incompatible with the current ESL/EFL classroom practices and furthermore cannot integrate instruction and assessment at all.Concerning this democratic approach, my understanding is that some Kurdish EFL students might not have creative views of the process of their learning based on alternative assessments because they just want to graduate and be employed. This can be two-side educational obstacle, the students' indifference before graduation and school

managers' indifference about their level of proficiency since they are not expected to have much knowledge and skills of English language when they become TESOL teachers at the schools.I think thatthis indifference is resulted from the lack of teachers' assessment literacy and unawareness of the importance of language proficiency that in turn affects testing ethicality negatively.

4.3 The Importance of Assessment Literacy for the Ethicality of Language Assessment in the Context of Kurdish TESOL Teachers

TESOL Inc.(2012) recommends that first it is urgent to equip teacherswith knowledge of assessment to be able to empower students and measure their developments properly (TESOL Inc., 2012)that possibly guarantees the ethicality and fairness of ESL/EFL assessment in a democratic testing context (Shohamy, 2001). Concerning programs to equip teachers with assessment literacy, I am sure that there are no such programs in KurdistanRegion/Iraq and almost all teachers have not had any opportunity to attend those programs. As a result, nearly all Kurdish TESOL teachers have the least of assessment literacy that probably affect the fairness of their assessments. For example, I was not aware of thelanguage testing standards, professional morality(Alderson and Banerjee, 2001), significance of language assessment literacy,the educational factors,various abilitiesfrom teachers'assessment literacy (see section 3.2) and most importantly the use of multiple methodsin language assessments. This is to ensure fairness for students (Troudi et al., ND) and to protect all from the misuse and abuse of tests (Alderson and Banerjee, 2001).Actually, I believe that this is right for almost all Kurdish TESOL teachers as I remember we were nearly equal in having knowledge of language education issues. We are also unaware of the assessment procedures,designs, appropriateness, accuracy, fairness and students' views on them and that assessments for English native-speakers and ELLs are different (TESOL Inc., 2010).As a result, sometimes our assessments of students might be incomplete and unethical, for example, difficult exam questions make many students suffer from not understanding them fully because of language complexity.As I elaborated on it above (see section 3.3), logically and educationally, having language assessment literacy influences the ethicality of language testing as it leads teachers to assess ESL/EFL students' proficiency more comprehensively and fairly by following various techniques anddoing more frequently in successive times throughout the academic years, inside or outside classrooms. Assessment literacy also makesthem aware of ethical issues such as harm, consent, fairness, deception, privacy, confidentiality(Lynch, 1997) validity, absence of bias, access, administration, and social consequences (Kunnan, 2003 cited in Bachman, 2005).Therefore, any teacher who has knowledge of those issues assesses language proficiency so differently (possibly more ethically) than others. So, as Kurdish TESOL teachers have the least knowledge of these assessment issues, they are not expected to assess Kurdish EFL learners properly, validly, fairly and ethically.

5. Conclusions

5.1 The Possible Impacts of TESOL Teacher Assessment Literacy on TestingEthicality

In the conclusion, I realize that language assessment literacy of TESOL teachers playsa major role in language assessment ethicality and fairness because first teachers need to be benchmarked to be able to makestudents have better self-knowledge of their proficiency;even when teachers enter new teaching situations, they need benchmarking courses again (Hamp-Lyons, ND). This goeswith what TESOL Inc.(2012) recommends that it is urgent to raise teachers' language assessment literacy to first empower students and then measure their developments properly (TESOL Inc., 2012).This is all to maximize assessment ethicality. I also conclude that the most appropriate measurement that guarantees the ethicality of ESL/EFL assessment in a democratic testing context can be achieved by alternative assessments of language proficiency. Another effective aspect that I conclude is that the development of language testing standards and professional morality among language testers are

influential in protecting all from the misuses of tests (Alderson and Banerjee, 2001). This clearly shows the impact of assessment literacy on assessment ethicality.

I also realize thatTESOL teachers should be aware of many assessment procedures,ethicality and fairness issues thatplay an important role in languageassessmentto be fair, valid, reliable and easy.Another significant side of TESOL teachers'assessment literacy that affects the testing ethicality is their understandingof thatassessments for English native-speakers and ELLs are different. Knowingthese assessmentissues in ESL/EFL testingsurely enhancesassessment ethicality.Another conclusion is about adopting portfolios as possibly the most democratic method of language assessment which is quite influential in fostering the ethicality of language testingand which is based on a collection of students' work (samples are collected from class activities or home assignments)that reflects progress over time (TESOL Inc., 2010).TESOL teachers'awareness of their students'rights to know about test questions, answer formats and to do self-assessment and peer-assessment methods(TESOL Inc., 2010) also increases the testing ethicality. Finally, it can be concluded that all those sorts of awareness and knowledge of language assessment can certainly be contributing factors to ensure fairness and ethicality of English language assessment for ELLs.

5.2The Possible Impacts of TESOL Teacher Assessment Literacy on AssessmentEthicality in the Educational Contextof Kurdish TESOL teachers

Firstly, I conclude that the current assessment practices of Kurdish TESOL teachersare still based on the traditional standardisedpen and paper tests only (Ataç, 2012) which is certainly inadequate and indicates thatthey have not yet reached the current advances of language assessment which coversvarious factors that impact testing, ethicality and professionalism and CLT concerns, TESOL Inc. standards on assessment and how assessment possibly improves overall teaching and learning qualities. This shows the small amount of Kurdish TESOL teachers' English language assessment literacy. I also realize that they do not have the knowledge of alternative assessment methods and their result recordingtechniques (Tsagari, 2004) as their current assessment practices are traditional methods of testing which cannotassess all sides of KurdishEFL students' proficiency and this is possibly inadequate and unethical. Also,in their restricted classroom performance-assessments, there are not equal opportunities for all students as some are engaged more by the teachers than others. This might be due to the large number of students in a class or because of gender discrimination.

I can suggest that the democratic model of assessment is quite suitable for most TESOL teachers and EFL students in my professional context; however, this model is not known and followed currentlyinstead always traditional testing techniques are adopted for classroom achievement tests that are often incompatible with current ESL/EFL classroom practices and that cannot integrate instruction and assessment together.However, if that democratic model of assessment is practiced in my professional context, my understanding is that some Kurdish EFL students might not have creative viewson the process of their learning based on alternative assessments as they just want to graduate and be employed, because school managers are indifferent about their level of proficiency that leads to teacher students' indifference about their learning progress. Overall, I can say that English language assessment in my country is, to some extent, inadequate, unfair, and unethical and lacks many influential developments that have been achieved through alternative assessments, especially portfolios, peer assessment and self-assessment. Therefore, I recommend some courses on the developments of TESOL assessment standards, alternative assessments and ethicality to train Kurdish TESOL teachers to make them obtain English language assessment literacy,then, to practice those assessments with the Kurdish EFL students for their assessments to be more valid, fair and ethical to the students.

References

AksuAtaç, B. (2012). Foreign language teachers' attitude toward authentic assessment in language teaching. *The Journal of Language and Linguistic Studies*, 8 (2), 7-19.

Alderson, J. C., & Banerjee, J. (2002).Language testing and assessment(Part 1). *Language Teaching*, *35*(2), 79-113.

Bachman, L. F. (2005). Building and supporting a case for test use. *Language Assessment Quarterly: An International Journal*, *2*(1), 1-34.

Bachman, L. F. (2007, April). Language assessment: Opportunities and challenges.
   In *meeting of the American Association of Applied Linguistics (AAAL), Costa Mesa,CA*.

Bailey, A. L., & Wolf, M. K. (2012). The challenge of assessing language proficiency aligned to the common core state standards and some possible solutions. In *Understanding Language Conference, Stanford University, Stanford, CA*.

Broad, B. Review of Shohamy, E. G. (2001). *The power of tests: A critical perspective on the uses of language tests.* Pearson Education.

Brown, T. & Jones, L. (2001).*Action Research and Postmodernism, Congruence and Critique.*Buckingham: Open University Press.

Brumfit, C. J., & Johnson, K. (Eds.). (1979). *The communicative approach to language teaching* (p. 2). Oxford: Oxford University Press.

Chandella, N. &Troudi, S. (2013). *Critical Pedagogy in Language Education: Challenges and Potentials* in Akbari, R. and Cumbe, C. (2013). Middle East Handbook of Applied Linguistics. Dubai, United Arab Emirates: TESOL Arabia.

Cheng, L., & DeLuca, C. (2011). Voices From Test-Takers: Further Evidence for LanguageAssessment Validation and Use. *Educational Assessment*, *16*(2), 104-122.

Cox, M. &Assis-Peterson, A. (1999). Critical Pedagogy in ELT: Images of Brazilian Teachers of English,*TESOL Quarterly*, 33(3), Autumn 1999.

Crookes, G. (2010). The practicality and relevance of second language critical pedagogy. *Language Teaching*, *43*(03), 333-348.

Cummins, P. W., &Davesne, C. (2009). Using electronic portfolios for second language assessment. *The Modern Language Journal*, *93*(s1), 848-867.

Dai Quang, V. (2007). *Critical applied linguistics: concerns and domains.* Scientific Research Management Office, College of Foreign Languages - VNU.

Pennycook, A. (2004). Critical Applied Linguistics in Davies, A. and Elder, C. (2004).*The Handbook of Applied Linguistics*. Oxford: Blackwell Publishing Ltd.

EALTA Executive Committee (ND). *EALTA Guidelines for Good Practice in Language Testing and Assessment.*

Earl, L. M., & Katz, M. S. (2006). *Rethinking classroom assessment with purpose in mind: Assessment for learning, assessment as learning, assessment of learning*. Manitoba Education, Citizenship & Youth.

Educational Testing Service (ETS) (2009). *Guidelines for the Assessment of English Language Learners.*

Farnsworth, T., & Chang, S. M. (2008). Issues in assessing English language learners: English language proficiency measures and accommodation uses.

Garcia, E., Lawton, K., &Diniz de Figueiredo, E. H. (2010). Assessment of young English language learners in Arizona: Questioning the validity of the state measure of English proficiency.

Hamp-Lyons, L. (2001). Ethics, fairness (es) and developments in language testing. *Experimenting with uncertainty: Essays in honour of Alan Davies*, 222-227.

Hargreaves, A. (2000). Four Ages of Professionalism and Professional Learning. *Teachers and Teaching: History and Practice.* 6(2), 2000.

Jia, Y. (2009). Ethical Standards for Language Testing Professionals: An Introduction to Five Major Codes.*Shiken: JALT Testing & Evaluation SIG Newsletter.* 13 (2) May 2009 (p. 2 - 8).

KhoaHäc, T. (2007).Critical Applied Linguistics: Concerns and Domains.*Scientific Research Management Office.*

Kozulin, A., Gindis, B., Ageyev, V., &Miller, S. (2003).*Vygotsky's educational theory and practice in cultural context.* Chaiklin, S. The ZPD in Vygotsky's analysis of learning and instruction. Cambridge: Cambridge University Press.

Lynch, B. K. (1997). In search of the ethical test. *Language Testing*, *14*(3), 315-327.

Lynch, B., & Shaw, P. (2005). Portfolios, power, and ethics. *Tesol Quarterly*,*39*(2), 263- 297.

McNamara, Tim (2000) Language Testing. Oxford, Oxford University Press.

Newfields, T. (2006).Teacher development and assessment literacy. *Authentic Communication: Proceedings of the 5th Annual JALT Pan-SIG Conference, (p. 48 - 73),*

May 13-14, 2006. Shizuoka, Japan: Tokai University College of Marine Science.

Pennycook, A. (2001). *Critical Applied Linguistics, A Critical Introduction.* London: LEA.

Pennycook, A. (2010). Criticaland Alternative Directions in Applied Linguistics, *Australian Review of Applied Linguistics,* Volume 33, Number 2, 2010.

Pessoa, R. &Freitas,M. (2012).Teaching.*TESOLQuarterly,*Vol. 0, No. 0, xxxx 2012.

Piggin, G. (2012). What are our tools really made out of? A critical assessment of recent models of language proficiency. *Polyglossia: the Asia-Pacific's voice in language and language teaching*, *22*, 79-87.

Reason, P. & Bradbury, H. (ed) (2001).*Handbook of Action Research.*London: SAGE Publications.

Riasati, M. J., &Mollaei, F. (2012).Critical Pedagogy and Language Learning. *International Journal of Humanities and Social Science ,2(21)*

Richards, J. & Schmidt, R. (2010).*Longman Dictionary of Language Teaching and Applied Linguistics.*Edinburgh: Pearson Education Limited.

Ripley, D. (2013). Implementing Portfolio-Based Language Assessment in LINC Programs: Benefits and Challenges. *TESL Canada Journal*, *30*(1), 69.

Shohamy, E. G. (2001). *The power of tests: A critical perspective on the uses of language tests*. Pearson Education.

Short, D. J. (2000). *The ESL standards: Bridging the academic gap for English language learners*. ERIC Clearinghouse on Languages and Linguistics.

Stephenson, A., & Johnson, P. D. F. (2003). Assessing English language proficiency: Using valid results to optimize instruction.

Stoynoff, S. J. (2008). *Developments in English Language Assessment APEC Strategic Plan for English and Other Languages APEC Human Resources, Development Working Group.*USA: Minnesota State University.

Talmy, S (ND). *Research Methods in Applied Linguistics,* Critical Research in Applied Linguistics (chapter 10).

TESOL Inc. (2012).*A Principles-Based Approach for English Language Teaching Policies and Practices.* A TESOL White Paper, March 2012.

TESOL International Association. (2010). TESOL/NCATE standards for the recognition of initial TESOL programs in P-12 ESL teacher education.

Thibeault, C., Kuhlman, &N., Day, C. (2011). ESL Teacher-Education Programs: MeasuringUp to the TESOL/NCATE Yardstick.*The CATESOL Journal.*22.1, 2010/2011.

Troudi, S., Coombe, C., &AlHamly, M. (2009). EFL Teachers' Views of English Language Assessment in Higher Education in the United Arab Emirates and Kuwait. *TESOL Quarterly*, *43*(3), 546-555.

Tsagari, D. (2004). Is there life beyond language testing? An introduction to alternative language assessment. *Center for Research in Language Education, CRILE Working Papers*, (58).

Wardekker, W. L., &Miedema, S. (1997). Critical pedagogy: An evaluation and a direction for reformulation. *Curriculum Inquiry*, *27*(1), 45-61.

Weideman, A. (ND).*Applied linguistics beyond postmodernism.*Department of English University of the Free State.

Winke, P. (2011). Evaluating the Validity of a High-Stakes ESL Test: Why Teachers' Perceptions Matter. *TESOL Quarterly*, *45*(4), 628-660.

Wolf, M. K., Herman, J., Bachman, L., Bailey, A., & Griffin, N. (2008). Recommendations for assessing english language learners: English language proficiency measures and accommodation uses. *Recommendations Report. Los Angeles: National Center for Research on Evaluation, Standards, and Student Testing, University of California*.