

# The Study of the Currently Practiced Assessment Techniques in Conversation Courses for English Translation Students in Iranian Universities

Mahdieh Shafipoor

Department of English Language Translation, Shahr-e-Qods Branch, Islamic Azad University, Tehran, Iran

Farnaz Latif

Department of English Language, Shahr-e-Rey Branch, Islamic Azad University, Tehran, Iran

**Abstract**—Considering “Conversion courses” as fundamental prerequisite in English Language Translation major in Iran, it is crystal clear that the quality of relevantly employed teaching methodology and subsequent adopted assessment procedures will directly influence the prospect achievement of the target learners. As a result, the researchers in the realm of pedagogical assessment have constantly attempted to explore the most applicable methods. Thus, in the current research project, the researchers aimed at delving into the current statuesque and problems in order to present an innovative model of oral assessment as a practical solution. To this end, the present study was conducted in 4 Conversation classes with 4 instructors at Islamic Azad University, Shahr-e-Qods branch through 4 main phases which entailed a complete Mixed Method research design. Consequently, the results reflected on significant inconsistencies among the different raters’ Conversation Courses assessment regarding their conceptualization, pedagogical performance and the standard criteria globally considered to validate assessments.

**Index Terms**—conversation course, formative assessment, summative assessment, pedagogical performance

## I. INTRODUCTION

Teaching involves assessment. In making decisions about lesson content and sequencing, about materials, learning tasks and so forth, teachers have to determine the strengths and weaknesses of the alternatives available to them. They make selections based on their experience, on their understandings of learning, language development and of language proficiency itself, together with what they consider to be most appropriate and in the best interests of those they teach. Equally, as part of their professional practice, they are always involved in the observation of their learners, which leads to the development of insights about learner progress and judgments about specific learning outcomes and overall performance (Rea-Dickins, 2004).

To be more specific, during 1990s, language testers were by large interested toward test developments in two related fields: educational measurement and language teaching (McNamara, 1997).

In educational measurement, there has been a movement towards what has been referred to variously as ‘alternative’ (e.g., Herman et al., 1992), ‘authentic’ (e.g., Wiggins, 1989; 1993; Newman et al., 1998; Terwilliger, 1998, O’Malley & Valdez Pierce, 1996; TESOL, 1998; 2000; Alderson & Banerjee, 2001) or ‘performance’ assessment (e.g., Aschbacher, 1991; Shavelson et al., 1992; Swanson et al., 1995; Solano-Flores & Shavelson, 1997), which has been moved quickly by widespread dissatisfaction with standardized multiple-choice tests, on the one hand, and by a movement towards the development of standards-based assessment, on the other.

However, when teachers are asked about assessment in general and oral assessment in particular, they do not have much to say especially at the tertiary level and they will tell you first and foremost about the formal mechanisms that are in place to monitor language achievement, or about the specific assessment procedures that they use. There is a tendency to prioritize the ‘formal’ and the ‘procedural’ and to underplay the observation-driven approaches to assessment which is strongly in evidence in their everyday classroom practice, such as language sampling (see Gardner & Rea-Dickins, 2002; Rea-Dickins, 2002).

Consequently, considering the oral assessment, some researches have been done so done investigating the oral proficiency assessment (e.g., Fulcher, 1996; Kormos, 1999; Chulhoub-Deville & Fulcher, 2003), raters’ effects on oral assessment (e.g., Brown, 2003, Brown & Hill, 1998; Shafipoor & Latif, 2012), the role of task difficulty in oral assessment (e.g., May, 2010). In order to shed more light in this regard, the present paper aims at exploring more about the oral assessment by studying the current assessment techniques of Conversations I and II in English translation course at tertiary level with the purpose of designing and presenting alternative assessment techniques.

## II. METHOD

### A. Participants

In this study, four university instructors teaching conversation I and II were randomly selected by the researchers. Each instructor had one class with almost 20 students. Then, there were 4 classes with total of 80 students who attended conversation I and II classes. The classes were held in the department of English Language Translation at Islamic Azad University Shahr-e- Qods Branch.

### B. Instruments

In the present study, initially in the first phase, 10 TEFL experts were interviewed to explore their conceptualizations toward the issue of assessment for Conversation courses at universities in Iran.

Then, to run the second phase, a semi-structured interview was devised based on the coded ideas and responses obtained from the first phase's interview. Consequently, the semi-structured interview questions were used to interview total number of 4 instructors from 4 classes teaching conversation I and II courses at Islamic Azad University Shahr-e-Qods Branch to find out their ideas toward assessment and the procedures employed by them.

### C. Procedure

The design adopted to carry out the present investigation is called "Mixed Method" which is a combination of qualitative and quantitative methods.

A mixed methods study involves the collection or analysis of both quantitative and qualitative data in a single study with some attempts to integrate two approaches at one or more stages of the research process. In other words, mixed methods research involves the mixing of quantitative and qualitative research methods or paradigm characteristics (Johnson & Christensen 2004).

According to Sandelowski (2003), there are two main and somewhat conflicting purposes for combining methods: (a) to achieve a fuller understanding of a target phenomenon and (b) to verify one set of findings against the other. To this end, in the 1970s, scholars who conduct mixed methods research desired to combine qualitative and quantitative research methods through introduction of a new concept called "Triangulation". The term "Triangulation" became synonymous with combining data sources through multiple methods. This method maximizes both the internal and external validity of research and helps reduce the inherent weaknesses of individual methods (Dornyei, 2007).

In the present study, at the very first qualitative stage, a field study was conducted. To do so, the researchers interviewed 10 TEFL experts to explore their conceptualizations toward the issue of assessment for Conversation courses at universities in Iran. As a result of this phase of the study, the researchers coded the ideas and responses to devise a semi-structured interview to be employed at the second phase of the study.

Then, the questions were used to interview total number of 4 instructors from 4 classes teaching conversation I and II courses at Islamic Azad University Shahr-e-Qods Branch to find out their ideas toward the concept of assessment theoretically and at the same time, practically.

To carry out the third phase of this study quantitatively, an external assessor joined each of the four classes for eight sessions and evaluated the students personally. Then, the assessor was present at final assessment sessions of all the instructors as well. Finally, the correlation coefficient between the final scores assigned by each of the four participant instructors and the final scores assigned by the external assessor was estimated.

At the final phase which was again, a qualitative one, a comprehensive comparison was performed to compare the standard assessment techniques used globally and the ones used in Iran.

### D. Assessment Setting

In the present study, both the formative and summative assessments were considered. Four instructors were selected for this study each having one conversation I or II class. The classes were held in the Islamic Azad University Shahr-e-Qods branch. The dedicated time for each class was three hours per week and there were 14 class sessions and 2 final exam sessions. An external rater took part in all the four classes every session as well as the final exam sessions. She observed the assessment procedure and assessed the students, too.

The followings are each instructor's assessment procedure:

#### **Instructor 1**

He only believed in the formative assessment not summative one. As a result, each session the students were assessed through their class participation, their projects, as well as the lectures they delivered. The scoring was out of 20 for each session.

#### **Instructor 2**

She employed both formative and summative assessment. Formative assessment was conducted each session through observing the students' class participation and checking their assignment.

Then, the students' assessment results were designated by (+, —, ×) for good performance, average performance, and weak performance, respectively. The total score for this section was 3.

The summative assessment which was run at the last two sessions was both oral and written. The written part consisted of listening, and use of English sections with 5 scores. The oral section had total score of 12 and the students were evaluated based on their accuracy, fluency, pronunciation, and vocabulary. They performed in pairs and were provided with a prompt prepared by their instructors in advance.

### ***Instructor 3***

The 3rd instructor also used both formative and summative assessment. In her formative assessment, she considered the students' class participation and their assignment for assessment. The total allocated score for this section was 10.

She also considered both written and oral exam for summative assessment. In the written part, the exam consisted of listening, vocabulary, grammar, and use of English sections with total score of 14. In the oral part, the students in pairs were provided with some prompts which were prepared in advance by their instructor. The topics of the prompt were similar to the topics in their course books. In this section, the total score was considered 16 and the instructor evaluated the students based on a four-scaled table of scoring entitled as Poor, Fair, Good, and Very good and the following scores of 1, 2, 3, and 4 were assigned to the students, respectively based on the above mentioned scales. The items being assessed were vocabulary, grammar, conversation strategy, and fluency.

### ***Instructor 4***

Similar to the 2nd and 3rd instructors, the 4th instructor also employed both formative and summative assessment. In addition, self-assessment and peer assessment were observed in her classes in some sessions as well. In the formative assessment, she considered students' class participation, lectures, individual oral presentations, and listening transcription tasks. The total score for this section was 80.

The summative assessment included both oral and written. The written part only consisted of listening and the allotted score was 10. The oral section with the total score of 10, involved both individual and pair work. The topics which were similar to their course book topics were shown on screen in power point format through a video projector. Coherence, pronunciation, using learned materials, and accuracy were the criteria for assessment in this section.

### ***External rater***

An external rater also took part in the study and accompanied the instructors both during the class assessment and in final exam assessment. She scored the students on her own every session as well as on the final exam. The total scores for class participation and final evaluation were considered 10 each. The items which were considered both for formative and summative assessment were accuracy, fluency, pronunciation, vocabulary, and using the learnt materials.

## **III. RESULT**

In the proceeding sections to answer the two research questions mentioned in the study, the collected data through both qualitative and quantitative phases will be analyzed and the results will be presented and evaluated separately.

### ***A. Phase 1(Qualitative): Filed study***

At this stage the researchers interviewed 10 TEFL experts to explore their conceptualizations toward the issue of assessment for Conversation courses at universities in Iran (table 4.1). As a result of this phase of the study, the researchers coded the ideas and responses to device a semi-structured interview to be employed at the second phase of the study.

TABLE I.  
INTERVIEW WITH 10 TEFL EXPERTS

<b>Interview questions with 10 TEFL experts</b>
1- How do you evaluate your students in Conversation Courses at university?
2- How successful has this procedure been?
3- What are your personal criteria for this evaluation?
4- Describe your evaluation procedure in details?
5- What different types of evaluation do you employ?
6- Do you think there is a need for revision to your evaluation procedure?
7- How often do you update and revise your evaluation procedure, criteria and techniques?
8- Do you ever collaborate and swap your evaluation experience with your colleagues?
9- Do you believe in the role of external raters in evaluating your students' performance in order to reduce the possibility of test bias?
10- How you apply self-evaluation techniques in your classes?

### ***B. Phase 2(Qualitative): Interview with 4 Conversation Instructors***

At this stage, based on the coded data extracted from 10 expert interviews in the first phase, the following semi-structured interview questions were devised. Then, the questions were used to interview total number of 4 instructors from 4 classes teaching conversation I and II courses at Islamic Azad University Shahr-e- Qods Branch.

TABLE II.  
INTERVIEW WITH 4 CONVERSATION INSTRUCTORS

<b>Interview with 4 Conversation Instructors</b>
1- Describe your assessment procedure in details?
2- How much time do you usually allocate to your assessment procedure?
3- What are your personal criteria for this assessment? Do you have a table of specification?
4- Are you satisfied with the results of your assessment procedure?
5- What are the problems you usually encounter during assessing Conversation Courses?

### C. Phase 3(Qualitative): Semi-structured Interview Coding Results

The results of the interview with 4 conversation instructors are presented in the following tables:

TABLE III.  
1<sup>ST</sup> INSTRUCTOR'S INTERVIEW CODING RESULTS

1 <sup>st</sup> instructor's interview coding results
<ul style="list-style-type: none"> <li>• Focus on Practice Conversation as a criterion for assessment</li> <li>• Focus on Real Situations</li> <li>• Focus on learners' vocabulary range</li> <li>• Personal scoring criteria ( +, -, × )</li> <li>• Scoring is carried out per session</li> <li>• Have no idea of " table of specification"</li> <li>• Not satisfied with the results. Believed in the objective nature of assessing reliability.</li> <li>• Called physical environment as a problem</li> <li>• Large number of students as an obstacle</li> <li>• No specific format for scoring was the main problem</li> <li>• Couldn't recall the learners' names</li> </ul>

TABLE IV.  
2<sup>ND</sup> T INSTRUCTOR'S INTERVIEW CODING RESULTS

2 <sup>nd</sup> instructor's interview coding results
<ul style="list-style-type: none"> <li>• General evaluation by asking students personal questions at the very first session of the semester and assigning scores out of 20 without any specific criteria.</li> <li>• Emphasizing on making a positive atmosphere and taking care of the students' motivation more than anything else.</li> <li>• Disagreed with one shot evaluation and believed in continuous assessment.</li> <li>• Confused the concept of instruments for assessment (e.g. presentation, seminars, films, listening tracks, etc.) by criteria for assessment.</li> <li>• He believed that beginner students cannot be evaluated appropriately.</li> <li>• Eclectic approach toward evaluation ( subjective and objective)</li> <li>• Underestimating and anticipating low performance for Azad university students.</li> </ul>

TABLE V.  
3<sup>RD</sup> INSTRUCTOR'S INTERVIEW CODING RESULTS

3 <sup>rd</sup> instructor's interview coding results
<ul style="list-style-type: none"> <li>• Continuous assessment ( formative as a base but along with summative one )</li> <li>• Decision making based on every session feedback( self- evaluation)</li> <li>• Session by session evaluation and scoring on specific basis.</li> <li>• A rather detailed and practical table of specification</li> <li>• Adjusted the criteria for evaluation with the course book and the learners'</li> <li>• Heterogeneous classes were a concern to her</li> <li>• Lack of Level adaptation of the course book and the learners</li> </ul>

TABLE VI.  
4<sup>TH</sup> INSTRUCTOR'S INTERVIEW CODING RESULTS

4 <sup>th</sup> instructor's interview coding results
<ul style="list-style-type: none"> <li>• Integrated techniques are employed ( self-evaluation, teacher evaluation, peer –evaluation)</li> <li>• Focus on the target content and expected performance</li> <li>• Focus on the learners' competence and adapt the tasks to accomplish the goals</li> <li>• Ongoing assessment for class participation</li> <li>• Three –phase assessment for end of the semester evaluation which entails various tasks and performances</li> <li>• Clarified table of specification for each class which is presented to the learners at the very first session of the course.</li> <li>• Providing learners with report forms filled up with their cooperation.</li> <li>• Considering heterogeneity as a tool to encourage challenge, cooperation and collaboration in the class.</li> <li>• Believe in the fairness and accuracy of the results since the learners were satisfied as well.</li> <li>• Lack of facilities on behalf of administrators</li> </ul>

### D. Phase 4(Quantitative): Correlation Coefficients

To carry out this phase, an external assessor joined each of the four classes for eight sessions and evaluated the students personally. Then the assessor was present at final assessment sessions of all the instructors as well. Finally, the correlation coefficient between the final scores assigned by each of the four participant instructors and the final scores assigned by the external assessor was estimated. The following tables show the results.

TABLE VII.  
CORRELATION COEFFICIENTS BETWEEN THE INSTRUCTOR 1 AND THE EXTERNAL RATER

Variable	Instructor 1	External Rater
Instructor 1	1	*
External Rater	.32	1

TABLE VIII.  
CORRELATION COEFFICIENTS BETWEEN THE INSTRUCTOR 2 AND THE EXTERNAL RATER

Variable	Instructor 2	External Rater
Instructor 2	1	*
External Rater	.67	1

TABLE IX.  
CORRELATION COEFFICIENTS BETWEEN THE INSTRUCTOR 3 AND THE EXTERNAL RATER

Variable	Instructor 3	External Rater
Instructor 3	1	*
External Rater	.83	1

TABLE X.  
CORRELATION COEFFICIENTS BETWEEN THE INSTRUCTOR 4 AND THE EXTERNAL RATER

Variable	Instructor 4	External Rater
Instructor 4	1	*
External Rater	.93	1

As tables 4.7 shows, the correlation between the 1st instructor and the external rater is .28 which as compared to the others ranks as the lowest correlation in the present study. Whereas, the correlation between the 4th instructor's scores and that of the external rater's (table 4.10) reveals an extremely high correlation of .93 which is the highest among all the four correlation studies. Eventually, it should be pointed out that the two other remaining correlation studies (tables 4.8 & 4.9) indicate closer correlation as compared to the external rater.

#### IV. CONCLUSIONS AND DISCUSSIONS

In the present research project to initially delve into the instructors' mental conceptualization toward assessment and explore their relevant associated pedagogical performance, the researchers attempted to run a filed study as a result of which the following concepts were gathered that can well reflect on the TEFL society's perception in this regard.

TABLE XI.  
INSTRUCTORS' FOCUS OF CONCERN TOWARD ASSESSMENT

Instructors' focus of concern toward assessment
The concern for assessment procedure
The concern for time management relevant to assessment procedure
The concern for having personal criteria and table of specification for assessment
The concern for instructor's satisfaction toward the assessment procedure
The concern for encountering the problems while carrying out the conversation courses assessment

The above table illustrates the main concentration of the Iranian TEFL instructors concerning the issue of assessment. As it is obviously illustrated the issues are rather general and confined to some while the assessment procedure in its standard meaning entails wide range of criteria which follow more detailed concepts and will embrace a longer list than achieved.

Secondly, to elaborate more on the practically performed assessment techniques four TEFL instructors were selected to carry out a semi-structured interview, whose results presented in chapter 4, reveals the following facts practiced in Iranian Conversation courses at Islamic Azad University Shahr-e- Qods branch.

Due to the confined and pedagogical knowledge of the participant instructors the rationale behind their practice assessment techniques was not vividly clarified. This had further led to tangible and observable inconsistencies which directly affect the validity and reliability of the employed assessment procedure and assessment results, as well.

As it was observed and studied, the instructors did not have a homogenized assessment policy to follow. In addition, the course syllabi as well as the assessment types were totally varied from one instructor to another. Therefore, this could under question the efficiency of the assessment procedure, its reliability and the expected validity.

Since the assessment procedure is by itself time consuming an energy taking, more innovation and time management strategies consistent with the standard criteria are required to objectively operationalize the procedure.

As the conversation courses are prerequisite to the proceeding courses to accomplish the BA degree then more coordination and more logical linkage between the decisions made inside the classes and the policies run in the faculty required to be observed and carefully attended. However, in the present study all the aforementioned concerns were ignored or underestimated.

#### *Proposals for alternative assessment*

Due to the fact that the assessment procedure in general and assessment for conversation courses in particular demand a rather flexible and objective while clearly defined procedure with localized and situation-specific criteria, it is alternatively proposed to the Iranian university instructors in this area to reconsider their perspectives, carry out a self-observation and assessment to scrutinize their personal outlook and practical performance in this regard. In addition, the students are also expected to have a clear road map of how they will be assessed and thus how they are expected to perform tasks during the course.

To this end and according to the similar studies carried out on assessment in this regard, the researchers propose two checklists (Appendices B & C) as compared to the standard ALTE scale (appendix A) and therefore, devised as a localized checklist for Iranian universities. The first checklist is distributed and presented to the students of Conversation courses at the beginning of every semester to give them the opportunity to reflect on their progress and go through self- assessment procedure. Whereas the second checklist which is designed as the class list given to instructors to employ both formative and summative assessment procedures in order to have a united model and compensate the inconsistencies observed in the study.

Since they need to perceive their deficiencies and confined pedagogical knowledge level and up-to date themselves to promote toward an assessor- instructor role which is more comprehensive and more operationalized at the current situation.

In conclusion, the aforementioned concerns need to be responded as promptly as possible to avoid such devastating threats ahead of the TEFL educational society. Thus, the practitioners and policy makers as well as the sole authoritarian decision makers are highly called for adopting the proposed checklist as the base and then after deployment of localized needs analysis and situation studies develop their context specific criteria which works best too the benefit of the system, learners, instructors and the whole educational world.

#### APPENDIX A THE STANDARD ALTE SCALE

1. Overall general ability			
LEVELS	Listening/Speaking	Reading	Writing
<b>C2</b> (ALTE 5)	CAN advise on or talk about complex or sensitive issues, understanding colloquial references and dealing confidently with hostile questions.	CAN understand documents, correspondence and reports, including the finer points of complex texts.	CAN write letters on any subject and full notes of meetings or seminars with good expression and accuracy.
<b>C1</b> (ALTE 4)	CAN contribute effectively to meetings and seminars within own area of work or keep up a casual conversation with a good degree of fluency, coping with abstract expressions.	CAN read quickly enough to cope with an academic course, to read the media for information or to understand non-standard correspondence.	CAN prepare/draft professional correspondence, take reasonably accurate notes in meetings or write an essay which shows an ability to communicate.
<b>B2</b> (ALTE 3)	CAN follow or give a talk on a familiar topic or keep up a conversation on a fairly wide range of topics.	CAN scan texts for relevant information, and understand detailed instructions or advice.	CAN make notes while someone is talking or write a letter including non-standard requests.
<b>B1</b> (ALTE 2)	CAN express opinions on abstract/cultural matters in a limited way or offer advice within a known area, and understand instructions or public announcements.	CAN understand routine information and articles, and the general meaning of non-routine information within a familiar area.	CAN write letters or make notes on familiar or predictable matters.
<b>A2</b> (ALTE 1)	CAN express simple opinions or requirements in a familiar context.	CAN understand straightforward information within a known area, such as on products and signs and simple textbooks or reports on familiar matters.	CAN complete forms and write short simple letters or postcards related to personal information.
<b>A1</b> (ALTE Breakthrough)	CAN understand basic instructions or take part in a basic factual conversation on a predictable topic.	CAN understand basic notices, instructions or information.	CAN complete basic forms, and write notes including times, dates and places.

#### APPENDIX B FORMATIVE ASSESSMENT CHECKLIST FOR STUDENTS' SELF-EVALUATION

FORMATIVE ASSESSMENT CHECKLIST FOR STUDENTS' SELF-EVALUATION						
TOTAL SCORE: 50						
STUDENT :	INSTRUCTOR :	COURSE :		SEMESTER :		
CRITERIA		POOR =1	FAIR =2	GOOD =3	VERY GOOD=4	EXCELLENT =5
PRONUCIATION						
USING THE LEARNED VOCABULARY						
ACCURACY						
FLUENCY						
COMMUNICATING THE MESSAGE						
CONVERSATION STRATEGIES						
ACCOMPLISHING THE ASSIGNED TASK						
INTERACTION WITH PEERS						
VOLUNTEERISM						
ATTENDANCE						

## APPENDIX C FORMATIVE AND SUMMATIVE ASSESSMENT'S CHECKLIST

			Summative assessment (50 scores)									Formative assessment (50 scores)											
			Written			Oral																	
Total Mark	<div></div>		writing	use of English	listening	communicating the message	Using the learnt vocabulary	communication strategies	Fluency	Accuracy	Pronunciation	Attendance	volunteerism	interaction with peers	accomplishing the assigned tasks	conversation strategies	communicating the message	Fluency	accuracy	using the learnt vocabulary	pronunciation	Name	Row
			summative	formative	5	5	10	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	<div></div>
																							1
																							2
																							3
																							4
																							5
																							6
																							7
																							8

## ACKNOWLEDGMENT

The authors wish to thank Islamic Azad University, Shahr-e-Ghods branch to support this work by a grant.

## REFERENCES

- [1] Alderson, J.C., & Banerjee, J. (2001). Language teaching. Cambridge: Cambridge University Press.
- [2] Aschbacher, P.R. (1991). Performance assessment: state of activity, interest, and concerns. *Applied Measurement in Education* 4, 275–88.
- [3] Brown, A. (2003). Interviewer variation and the co-construction of speaking proficiency. *Language Testing* 20.1, 1-25.
- [4] Brown, A., & Hill, K. (1998). Interviewer style and candidate performance in the IELTS oral Interview. In S. Wood (ed.), *IELTS Research Reports* 1, 1-19.
- [5] Chulhoub-Deville, M. & Fulcher, G. (2003). The oral proficiency interview: A research agenda. *Foreign Language Annals* 36.4.
- [6] Dornyei, Z. (2007). Research methods in applied linguistics. Oxford: Oxford University Press.
- [7] Fulcher, G. (1996). Testing tasks: Issues in task design and the group oral. *Language Testing* 13, 23-49.
- [8] Gardner, S. & Rea-Dickins, P. (2002). Focus on language sampling: a key issue in EAL assessment. Occasional Paper 15. National Association for Language Development in the Curriculum: NALDIC Publications Group.
- [9] Herman, J.L., Aschbacher, P.R. & Winters, L. (1992). A practical guide to alternative assessment. Alexandria, VA: Association for Supervision and Curriculum Development.
- [10] Kormos, J. (1999). Simulating conversations in oral-proficiency assessment: a conversation analysis of role plays and non-scripted interviews in language exams. *Language Testing* 16, 163–88.
- [11] McNamara, T.F. (1997). 'Interaction' in second language performance assessment: Whose performance? *Applied Linguistics* 18, 444-446.
- [12] May, L. (2010). Developing speaking assessment tasks to reflect the 'social turn' in language Testing. *University of Sydney Papers in TESOL* 5, 1-30.
- [13] Newman, F., Brandt, R. & Wiggins, G. (1998). An exchange of views on 'semantics, psychometrics, and assessment reform: a close look at "authentic" assessments'. *Educational Researcher* 27.6, 19-22.
- [14] O'Malley, J.M. & Valdez Pierce, L. (1996). Authentic assessment for English language learners. Reading, MA: Addison-Wesley.
- [15] Rea Dickins, P.M. (2002). Exploring the educational potential of assessment with reference to learners with English as an additional language. In Leung, C., editor, *Language and additional=second language issues for school education: a reader for teachers*. York: York Publishing Services, 81-93.
- [16] Rea-Dickins, P. (2004). Understanding teachers as agents of Assessment. *Language Testing* 21.3, 249–258.

- [17] Sandelowski, M. (2003). 'Tables of tableaux? The challenges of writing and reading mixed methods studies' in A. Tashakkori and C. Teddlie(eds.). *Handbook of mixed methods in Social and Behavioral Research*. Thousands Oaks, Califs: Sage.
- [18] Shafipoor, M. & Latif, F. (2012). The impact of the assessors' attitude on the learners' oral assessment score pollution. Akdeniz Language Studies Conference. Antalya, Turkey.
- [19] Shavelson, R.J., Baxter, G.P. & Pine, J. (1992). Performance assessments: political rhetoric and measurement reality. *Educational Researcher* 21.4, 22–27.
- [20] Solano-Flores, G. & Shavelson, R.J. (1997). Development of performance assessment in science: conceptual, practical, and logistical concerns. *Educational Measurement: Issues and Practice* 16.3, 16–25.
- [21] Swanson, D.B., Norman, G.R. & Linn, R.L. (1995). Performance-based assessment: lessons from the health professions. *Educational Researcher* 24.5, 5–11.
- [22] Terwilliger, J. (1998). Rejoinder: response to Wiggins and Newman. *Educational Researcher* 26.6, 22–23.
- [23] TESOL. (1998). Managing the assessment process. Alexandria, VA: TESOL Inc.
- [24] TESOL. (2000). Integrating the ESL standards into classroom practice: grades pre-K\_12. Alexandria, VA: TESOL Inc.
- [25] Wiggins, G. (1989). A true test: toward more authentic and equitable assessment. *Phi Delta Kappan* May, 703–713.
- [26] Wiggins, G. (1993). Assessment: authenticity, context and validity. *Phi Delta Kappan* November, 200–14.

**Mahdieh Shafipoor** is a TEFL M.A. holder and a faculty member at Islamic Azad University Shahr-e-Qods branch. She is a lecturer at research, testing, teaching methodology, and contrastive analysis. She also has totally 14 years' experience of teaching English in different fields. She has carried out research studies on applied linguistics. Her fields of interest are testing and assessment, syllabus design, and Materials Development.

**Farnaz Latif** is a PhD holder in TEFL who graduated from Department of Foreign Languages, Khorasgan Branch, Islamic Azad University, Isfahan, Iran. At the moment, she is the head of language Department in Islamic Azad University Shahr-e-Rey branch, and a General English Course instructor at university. In addition she is supervising and instructing Business English Courses in SAPCO (an automotive part supplier company). She has totally 14 years experience of teaching English in different fields. Her research interests are Teacher Education, ESP, CALL, and Materials Development.