The Relationship between EFL Teachers' Assessment Literacy, Their Teaching Experience, and Their Age: A Case of Iranian EFL Teachers

Samaneh Zolfaghari

Department of English, Torbat-e-Heydarieh Branch, Islamic Azad University, Torbat-e-Heydarieh, Iran

Hamid Ashraf

Department of English, Torbat-e-Heydarieh Branch, Islamic Azad University, Torbat-e-Heydarieh, Iran

Abstract—In recent decades, issues of assessment design and implementation has gained increased attention in education, (Saad, et al 2013). It is believed that ability in deciding correct, practical, and fair assessment methods are requirements to use information properly in order to support instructional decisions. Teachers are required to be aware of various kinds of information gathered from an extensive series of assessment alternatives and their strengths and weaknesses (Mertler, 2009). The present study, aimed to shed light on the relationship between Iranian EFL teachers' assessment literacy, their teaching experience, and age. To this end, 658 EFL teachers were selected to fill out a researcher made assessment literacy inventory which has gone through the validation and reliability processes. The obtained data were analyzed using SPSS program (version21). The results of data analysis indicated that there is a highly significant positive relationship between Iranian EFL teachers' assessment literacy and teaching experience. There also found a positive relationship between Iranian EFL teachers' assessment literacy and their age.

Index Terms—assessment literacy, teachers' teaching experience, age, EFL teacher

I. INTRODUCTION

Assessing students' performance is one of the most important duties of classroom teachers. Assessment is historically used as an important factor in determining students' educational success. It is also used to show differences in students' learning and rank students according to their level of achievement. Actually, assessment included relying on norm-referenced interpretations of students scores obtained from tests. But there seemed to be a limitation in this method, since these interpretations made determining specific areas of weakness and strength in students' learning difficult. Plake (1993) reported that "teachers spend up to 50 percent of their time on assessment-related activities." (Plake, 1993).

Recently, a great portion of professional development programs is connected with assessment literacy for teachers and/or administrators. Popham,W.J (2009), denotes that ''after dividing educators' measurement-related issues into both classroom assessments, and accountability assessments, it is found that instructors' insufficient consideration in each of these firms could influence the value of education. Assessment literacy is seen, therefore, as a sine qua non for today's competent educator ''.

Assessment could be described as any technique, tool or strategy that teachers use to elicit evidence of students' progress towards the stated goals (Chen, 2003; Wishon, Crabtree, & Jones, 1998). Here, the common side of the definition relates to the way of assessing students' performance that is "the procedure of gathering information related to educational issues of students to support decision making about the improvement and language development of the student" (p. 363). This useful definition demonstrates that assessment is an implementation of control that is caught up in a range of topics about testers' and test-takers' voices, roles, and beliefs. Shohamy (2000), asserted that instructors, students, and other educators "create the assessment awareness by trying to make sense of the knowledge in a dialogical and co-operative way" (Shohamy, 2000, p. 136).

It is considered that teacher' assessment competence is an essential aspect that affects their assessment practices (Cheng et al., 2004). According to Brookhart (2011) there are a set of knowledge and skills that teachers should be aware of. Particularly, teachers should be able to understand the "learning progression" in the content area. This as states by Brookhart (2011), "permits educators to recognize where a student is set regarding learning objectives and assist teachers understand students'' effort, and design steps in education and assessment, while assessing the students to reach the objective" (Brookhart, 2011, p. 7).

A. Research Questions

Q1: Is there any significant relationship between Iranian EFL teachers' assessment literacy their teaching experience?
Q2: Is there any significant relationship between Iranian EFL teachers' assessment literacy and their age?

B. Research Hypotheses

H02: There is no significant relationship between Iranian EFL teachers' assessment literacy and their teaching experience.

H03: There is no significant relationship between Iranian EFL teachers' assessment literacy and their age.

II. REVIEW OF THE RELATED LITERATURE

Research in different parts of the world has demonstrated that many teachers are unsatisfactorily trained to understand, manage and understand the results of various kinds of assessments (e.g., Bol, Stephenson, O'Connell, & Nunnery, 1998; Stiggins& Conklin, 1992; Wiggins, 1989). Instructors who were not sufficiently trained and insufficiently talented in managing valid assessments, in general, recognized these to be more complicated to build up than conventional paper-and-pencil examinations. Furthermore, instructors' assessment performance were frequently not well supported their educational objectives and tended to require a low level of cognitive processing. A large number of instructors were observed not to be fine critics of the quality of their own assessment assignments (Bol&Strage, 1996).

In a related study, Zhang and Burry-Stock (2003) examined "assessment history and the norms for teacher ability in the classroom assessment of learners (AFT, NCME, & NEA, 1990) in order to present a self-report measurement aimed at measuring instructor assessment observations and self-recognized assessment ability. It is found that, secondary school teachers use more paper-and-pencil tests compared to elementary school teachers who occasionally used performance assessments.

In another study, Alkharusi.H, (2011) examined instructors' self-recognized assessment abilities as a purpose of gender, subject area, mark, teaching practice, and the assessment training. Applicants were 213 instructors of Oman teaching in Muscat public schools. To this end, a 25-item Self-recognized assessment ability inventory was designed and applied in the observation. The findings of the study demonstrated important variations on the self-recognized assessment abilities related to instructors' gender, subject area, scores, teaching practice, and assessment preparation.

III. METHODOLOGY

To collect the required data 658 Iranian EFL teachers teaching in different cities of Iran were selected. Having considered the whole number of English Language institutes in Iran 5000, and each institute has an average of 15 teachers the society of this study was regarded as the whole number of 75,000 English language teachers. Sample size was calculated according to Krejcie and Morgan' sample size table; considering the 99% level of confidence and 0.05 degree of accuracy. They were considered experienced teachers having university education (Bachelor or Master or PhD degree). The participants were both males and females from different age groups and different years of experiences. Their fields of study were English language teaching, Translation, and English literature. In present study, 658 teachers who were all Iranian English as a Foreign Language Instructors, teaching in different English institutes of different cities of Iran were selected randomly. They were from both genders and ranging from different ages with different years of teaching experiences. They were selected according to Morgan's table of sampling. For collecting the data, questionnaires of assessment literacy and self-efficacy (see appendix A and B) in both forms: papers and online (using Google drive) were represented to different English language teachers to answer the questions. Collected data were analyzed in SPSS software (version21). Then, the correlation between these two variables was also analyzed.

A. Instrumentation

The needed data for this study were gathered through the application of one researcher's made questionnaire for assessing EFL teachers' assessment literacy. The questionnaire was made based on the assessment literacy standards of Michigan university (March 2013 V. 4.0). This 50 item researchers made questionnaire is a self report inventory which was meant to measure EFL teachers' knowledge and skill of assessment. The reliability was estimated using Cronbach's Alpha .86.

B. Study Design

The design of the study was a correlation one. The two major variables were Iranian EFL teachers' assessment literacy and their teaching experience, and their age.

IV. RESULTS AND DISCUSSION

The first question of the study was aimed to find out the relationship between Iranian EFL teachers' assessment literacy and their teaching experience. In order to answer this research question, one-way analysis of variance (ANOVA) was used for each of the subscales of assessment literacy. Teachers were grouped into three groups of low (1-5 years old), mid (6-10 years old), and high (above 11) teaching experience groups.

TABLE 1
DESCRIPTIVE STATISTICS FOR TEACHING EXPERIENCE

	N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean		Minimum	Maximum
					Lower Bound	Upper Bound		
Low	312	29.3651	4.28915	.54038	28.2849	30.4453	18.00	38.00
Mid	145	31.1739	6.05019	1.26155	28.5576	33.7902	16.00	40.00
High	201	37.7955	4.25129	.64091	36.5029	39.0880	19.00	40.00
Total	658	32.5385	5.98208	.52466	31.5004	33.5765	16.00	40.00

First, teaching dispositions was examined. Descriptive statistics for the three groups can be seen in Table 1. To see whether these differences are statistically significant, F value was checked (see Table 2).

TABLE 2
F VALUE FOR TEACHING EXPERIENCE REGARDING TEACHING DISPOSITIONS

	Sum of Squares	Df	Mean Square	F	Sig.
Between Groups	1893.241	2	946.621	44.149	.000
Within Groups	2723.067	656	21.441		
Total	4616.308	658			

As demonstrated by table 2, there is significant difference among the three groups concerning teaching dispositions [$F_{(656, 2)} = 44.14$, p<.05]. To trace the accurate place of difference, post-hoc analysis with Tukey was run (see Table 3).

TABLE 3
MULTIPLE COMPARISONS OF THREE AGE GROUPS OF EXPERIENCE AND TEACHING DISPOSITIONS

(I) years		(J) years		Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
							Lower Bound	Upper Bound
dimension2	Low	dimension3	mid	-1.80883	1.12809	.248	-4.4841	.8664
			high	-8.43038[*]	.90975	.000	-10.5879	-6.2729
	Mid	dimension3	low	1.80883	1.12809	.248	-.8664	4.4841
			high	-6.62154[*]	1.19145	.000	-9.4471	-3.7960
	High	dimension3	low	8.43038[*]	.90975	.000	6.2729	10.5879
			mid	6.62154[*]	1.19145	.000	3.7960	9.4471

As shown by table 3, there is significant, meaningful difference between the low experience group and high experience group (mean difference= 8.43, p<.05). As the mean of the high experience (M= 37.79) is higher than that of the low experience (M= 29.36), it can be concluded that by increasing the teaching experience, teaching dispositions increases, too. Moreover, there is a statistically significant difference between the mid experience group and high experience group (mean difference= 6.62, p<.05). As the mean of the high experience (M= 37.79) is higher than that of the mid experience (M= 31.17), it can be concluded that by the increase of experience, teaching dispositions increases, too. As a result it can be inferred that there is a meaningful relationship between subscale of teaching dispositions and Iranian EFL teachers' teaching experience.

Then, teaching knowledge was examined.

TABLE 4
DESCRIPTIVE STATISTICS FOR THREE AGE GROUPS OF EXPERIENCE AND TEACHING KNOWLEDGE

	N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean		Minimum	Maximum
					Lower Bound	Upper Bound		
Low	312	80.7619	9.56086	1.20456	78.3540	83.1698	54.00	104.00
Mid	145	78.4348	10.62512	2.21549	73.8401	83.0294	58.00	101.00
High	201	89.5682	10.33035	1.55736	86.4275	92.7089	66.00	106.00
Total	658	83.3308	10.93277	.95887	81.4336	85.2279	54.00	106.00

Descriptive statistics for the three groups can be seen in Table 4. To see whether these differences are statistically significant, F value was checked (see Table 5).

TABLE 5
F VALUE FOR THREE AGE GROUPS OF EXPERIENCE AND TEACHING KNOWLEDGE

	Sum of Squares	Df	Mean Square	F	Sig.
Between Groups	2678.901	2	1339.450	13.353	.000
Within Groups	12739.876	656	100.314		
Total	15418.777	658			

As inferred from table 5, there is meaningful, positive difference among the three groups regarding teaching knowledge [$F_{(656, 2)} = 13.35$, p<.05]. To find the accurate place of difference, post-hoc analysis with Tukey was run (see Table 6).

TABLE 6
MULTIPLE COMPARISONS OF THREE AGE GROUPS OF EXPERIENCE AND TEACHING KNOWLEDGE

(I) years		(J) years		Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
							Lower Bound	Upper Bound
dimension2	Low	dimension3	Mid	2.32712	2.44003	.607	-3.4594	8.1137
			High	-8.80628*	1.96778	.000	-13.4729	-4.1397
	Mid	dimension3	Low	-2.32712	2.44003	.607	-8.1137	3.4594
			High	-11.13340*	2.57708	.000	-17.2450	-5.0218
	High	dimension3	Low	8.80628*	1.96778	.000	4.1397	13.4729
			Mid	11.13340*	2.57708	.000	5.0218	17.2450

It can be inferred from table 6 that, there is a statistically significant difference between the low experience group and high experience group (mean difference= 8.80, p<.05). As the mean of the high experience (M= 89.56) is higher than that of the low experience (M= 80.76), it can be concluded that by the increase of teaching experience, teaching knowledge increases, too. Moreover, there is a statistically significant difference between the mid experience group and high experience group (mean difference= 11.13, p<.05). As the mean of the high experience (M= 89.56) is higher than that of the mid experience (M= 78.43), it can be concluded that by increasing the experience, teaching knowledge increases, too. Therefore there is a meaningful relationship between subscale of teaching knowledge and Iranian EFL teachers' teaching experience.

Finally, teaching performance was examined.

Descriptive statistics for the three groups can be seen in Table 7. To see whether these differences are statistically significant, F value was checked (see Table 8).

TABLE 7
DESCRIPTIVE STATISTICS FOR THE THREE GROUPS

	N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean		Minimum	Maximum
					Lower Bound	Upper Bound		
Low	312	35.2742	6.02228	.76483	33.7448	36.8036	16.00	54.00
Mid	145	36.9565	5.60421	1.16856	34.5331	39.3800	28.00	50.00
High	201	43.6136	8.49135	1.28012	41.0320	46.1952	20.00	50.00
Total	658	38.4186	7.83292	.68965	37.0540	39.7832	16.00	54.00

Descriptive statistics for the three groups can be seen in Table 7. To see whether these differences are statistically significant, F value was checked (see Table 8).

TABLE 8
F VALUE FOR THREE GROUPS OF EXPERIENCE REGARDING TEACHING PERFORMANCE

	Sum of Squares	Df	Mean Square	F	Sig.
Between Groups	1849.668	2	924.834	19.409	.000
Within Groups	6003.727	656	47.649		
Total	7853.395	658			

Table 8 indicates that, there found an important difference among the three groups with regard to teaching performance [F (656, 2) = 19.40, p<.05]. To place the accurate place of difference, post-hoc analysis with Tukey was run (see Table 9).

TABLE 9
MULTIPLE COMPARISONS OF THREE GROUPS OF EXPERIENCE AND TEACHING PERFORMANCE

(I) years		(J) years		Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
							Lower Bound	Upper Bound
dimension2	Low	dimension3	mid	-1.68233	1.68529	.579	-5.6794	2.3147
			high	-8.33944*	1.36068	.000	-11.5666	-5.1123
	Mid	dimension3	low	1.68233	1.68529	.579	-2.3147	5.6794
			high	-6.65711*	1.77612	.001	-10.8696	-2.4446
	High	dimension3	low	8.33944*	1.36068	.000	5.1123	11.5666
			mid	6.65711*	1.77612	.001	2.4446	10.8696

Table 9 demonstrates that, there is meaningful difference between the low experience group and high experience group (mean difference= 8.33, p<.05). As the mean of the high experience (M= 43.61) is higher than that of the low experience (M= 35.27), it can be concluded that by the increase of teaching experience, teaching performance increases, too. Moreover, there is a statistically significant difference between the mid experience group and high experience group (mean difference= 6.65, p<.05). As the mean of the high experience (M= 43.61) is higher than that of the mid experience (M= 36.95), it can be concluded that by the increase of experience, teaching performance increases, too.

Therefore, it can be inferred that there is a meaningful relationship between subscale of teaching performance and Iranian EFL teachers' teaching experience.

The second question of the study was aimed to find out the relationship between Iranian EFL teachers' assessment literacy and their age. In order to answer this research question, one-way analysis of variance (ANOVA) was used for each of the subscales of assessment literacy. Teachers were grouped into three groups of low (20-29 years old), mid (30-39 years old), and high (above 40) age groups.

First, teaching dispositions which relates to teachers' beliefs in assessment related matters was examined.

TABLE 10
DESCRIPTIVE STATISTICS OF THREE AGE GROUPS

	N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean		Minimum	Maximum
					Lower Bound	Upper Bound		
Low	325	30.2031	4.67407	.58426	29.0356	31.3707	16.00	40.00
Mid	187	30.8846	6.32662	1.24075	28.3292	33.4400	18.00	39.00
High	146	37.3500	4.78539	.75664	35.8196	38.8804	19.00	40.00
Total	658	32.5385	5.98208	.52466	31.5004	33.5765	16.00	40.00

Descriptive statistics for the three groups can be seen in Table 10. As it is presented by the table the population of low age group was more than mid and high age groups. To see whether these differences are statistically significant, F value was checked (see Table 11).

TABLE 11
F VALUE FOR AGE GROUPS REGARDING TEACHER DISPOSITIONS

	Sum of Squares	Df	Mean Square	F	Sig.
Between Groups	1346.194	2	673.097	26.141	.000
Within Groups	3270.113	656	25.749		
Total	4616.308	658			

Table 11 shows, there is meaningful, significant difference among the three groups with regard to teaching dispositions [F (656, 2) = 26.14, p<.05]. To find the accurate place of difference, post-hoc analysis with Tukey was run (see Table 12).

TABLE 12
MEAN DIFFERENCES OF AGE GROUPS

(I) Age		(J) Age	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
						Lower Bound	Upper Bound
dimension2	Low dimension3	mid	-.68149	1.18011	.832	-3.4801	2.1172
		high	-7.14688*	1.02277	.000	-9.5724	-4.7214
	mid dimension3	low	.68149	1.18011	.832	-2.1172	3.4801
		high	-6.46538*	1.27831	.000	-9.4969	-3.4339
	high dimension3	low	7.14688*	1.02277	.000	4.7214	9.5724
		mid	6.46538*	1.27831	.000	3.4339	9.4969

*. The mean difference is significant at the 0.05 level.

Table 12 indicates, there significant difference between the low age group and high age group (mean difference= 7.14, p<.05). As the mean of the high age (M= 37.35) is higher than that of the low age (M= 30.20), it can be concluded that by increasing the age, teaching dispositions increases, too. Moreover, there is a statistically significant difference between the mid age group and high age group (mean difference= 6.46, p<.05). As the mean of the high age (M= 37.35) is higher than that of the mid age (M= 30.88), it can be concluded that by increase of the age, teaching dispositions increases, too. Therefore, it can be inferred from the analysis that there is a significant relationship between Iranian EFL teachers' teaching dispositions and their age.

Then, teaching knowledge that explains what teacher should know about assessment was examined.

TABLE 13
DESCRIPTIVE STATISTICS FOR THE THREE AGE GROUPS

	N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean		Minimum	Maximum
					Lower Bound	Upper Bound		
Low	325	80.6406	8.94670	1.11834	78.4058	82.8754	63.00	103.00
Mid	187	79.5000	11.63873	2.28254	74.7990	84.2010	54.00	104.00
High	146	90.1250	10.49832	1.65993	86.7675	93.4825	66.00	106.00
Total	658	83.3308	10.93277	.95887	81.4336	85.2279	54.00	106.00

Descriptive statistics for the three groups can be seen in Table 13. To see whether these differences are statistically significant, F value was checked (see Table 14).

TABLE 14
F VALUE FOR THREE AGE GROUPS REGARDING TEACHER KNOWLEDGE

	Sum of Squares	Df	Mean Square	F	Sig.
Between Groups	2691.168	2	1345.584	13.427	.000
Within Groups	12727.609	656	100.217		
Total	15418.777	658			

Table 14 demonstrates that, there is meaningful difference among the three groups with regard to teaching dispositions [F (656, 2) = 13.42, p<.05]. To find the precise place of difference, post-hoc analysis with Tukey was run (see Table 15).

TABLE15
MULTIPLE COMPARISONS OF THREE AGE GROUPS AND TEACHING DISPOSITIONS

(I) Age	(J) Age		Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
						Lower Bound	Upper Bound
dimension2	Low	dimension3 mid	1.14063	2.32818	.876	-4.3807	6.6619
		high	-9.48438*	2.01775	.000	-14.2695	-4.6993
	Mid	dimension3 low	-1.14063	2.32818	.876	-6.6619	4.3807
		high	-10.62500*	2.52189	.000	-16.6057	-4.6443
	High	dimension3 low	9.48438*	2.01775	.000	4.6993	14.2695
		mid	10.62500*	2.52189	.000	4.6443	16.6057

Table 15 indicates, there is significant difference between the low age group and high age group (mean difference= 9.48, p<.05). As the mean of the high age (M= 90.12) is higher than that of the low age (M= 80.64), it can be concluded that by increasing the age, teaching knowledge increases, too. Moreover, there is a statistically significant difference between the mid age group and high age group (mean difference= 10.62, p<.05). As the mean of the high age (M= 90.12) is higher than that of the mid age (M= 79.50), it can be concluded that by increasing the age, teaching knowledge increases, too. Therefore, there is a meaningful relationship between subscale of teaching knowledge and Iranian EFL teachers' age.

Finally, teaching performance was examined. Descriptive statistics for the three groups can be seen in Table 12. To see whether these differences are statistically significant, F value was checked (see Table 16).

TABLE 16
STATISTICAL DIFFERENCES FOR THREE AGE GROUPS

	N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean		Minimum	Maximum
					Lower Bound	Upper Bound		
Low	325	36.2540	6.21385	.78287	34.6890	37.8189	16.00	54.00
Mid	187	34.1923	4.81680	.94465	32.2468	36.1379	25.00	44.00
High	146	44.5750	8.17089	1.29193	41.9618	47.1882	20.00	50.00
Total	658	38.4186	7.83292	.68965	37.0540	39.7832	16.00	54.00

Descriptive statistics for different age groups is shown by table.

TABLE 17
F VALUE FOR THREE AGE GROUPS REGARDING TEACHER PERFORMANCE

	Sum of Squares	Df	Mean Square	F	Sig.
Between Groups	2275.645	2	1137.823	25.703	.000
Within Groups	5577.750	656	44.268		
Total	7853.395	658			

Table 17 indicates, there is a meaningful difference among the three groups regarding their teaching performance [F (656, 2) = 25.70, p<.05]. To find the accurate place of difference, post-hoc analysis with Tukey was run. (See table 18).

TABLE 18
MULTIPLE COMPARISONS OF THREE AGE GROUPS AND TEACHING PERFORMANCE

(I) Age	(J) Age		Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
						Lower Bound	Upper Bound
dimension2	low	dimension3 Mid	2.06166	1.55089	.382	-1.6166	5.7400
		High	-8.32103*	1.34512	.000	-11.5113	-5.1308
	mid	dimension3 Low	-2.06166	1.55089	.382	-5.7400	1.6166
		High	-10.38269*	1.67610	.000	-14.3579	-6.4074
	high	dimension3 Low	8.32103*	1.34512	.000	5.1308	11.5113
		Mid	10.38269*	1.67610	.000	6.4074	14.3579

*. The mean difference is significant at the 0.05 level.

Table 18 indicates, there is significant meaningful difference between the low age group and high age group (mean difference= 8.32, p<.05). As the mean of the high age (M= 44.57) is higher than that of the low age (M= 36.25), it can be concluded that by increasing the age, teaching performance increases, too. Moreover, there is a statistically significant difference between the mid age group and high age group (mean difference= 10.38, p<.05). As the mean of the high age (M= 44.57) is higher than that of the mid age (M= 34.19), it can be concluded that by the increase of age, teaching performance increases, too. Therefore, it can be inferred from the analysis that there is a significant relationship between subscale of teaching performance as one of assessment literacy subscales and Iranian EFL teachers' age.

V. CONCLUSION

Having investigated the relationship between EFL teachers' assessment literacy and their teaching experience, it was found that, EFL teachers with more years of teaching experience demonstrated to be more knowledgeable in assessment related matters. In other words, by increasing teaching experience, teachers' assessment literacy increases too. Therefore, the first hypothesis of the study which stated there is no relationship between Iranian EFL teachers' assessment literacy and their teaching experience is rejected.

Having recognized EFL teachers' assessment literacy and their age, it was revealed that there is a positive relationship between Iranian EFL teachers' assessment literacy and their age, therefore the second hypothesis of the research which claimed that, there is no relationship between Iranian EFL teachers' assessment literacy and their age is rejected.

REFERENCES

[1]   Alkharusi, H. (2011). Teachers' classroom assessment skills: Influence of gender, subject area, grade level, teaching experience and in-service assessment training. *Journal of Turkish Science Education*, 8(2), 39-48.
[2]   Bol, L., & Strage, A. (1996). The contradiction between teachers' instructional goals and their assessment practices in high school biology courses. *Science Education*, 80(2), 145-163
[3]   Bol, L., Stephenson, P. L., O'connell, A. A., & Nunnery, J. A. (1998). Influence of experience, grade level, and subject area on teachers' assessment practices. *The Journal of Educational Research*, 91(6), 323-330.
[4]   Brookhart, S. M. (2011). Educational assessment knowledge and skills for teachers. *Educational Measurement: Issues and Practice*, 30(1), 3-12.
[5]   Chen, H. M. (2003). A study of primary school English teachers' beliefs and practices in multiple assessments: A case study in Taipei City. *Unpublished master theses. Taipei: National Taipei Teachers College*.
[6]   Cheng, L., Rogers, T., & Hu, H. (2004). ESL/EFL instructors' classroom assessment practices: Purposes, methods, and procedures. *Language Testing*,21(3), 360-389.
[7]   Mertler, C. A. (2009). Teachers' assessment knowledge and their perceptions of the impact of classroom assessment professional development. *Improving Schools*, 12(2), 101-113.
[8]   Plake, B. S. (1993). Teacher Assessment Literacy: Teachers' Competencies in the Educational Assessment of Students. *Mid-Western Educational Researcher*, 6(1), 21-27.
[9]   Popham, W. J. (2009). Assessment literacy for teachers: faddish or fundamental?. *Theory into practice*, 48(1), 4-11.
[10] Saad, M. R. B. M., Sardareh, S. A., & Ambarwati, E. K. (2013). Iranian Secondary School EFL Teachers' Assessment Beliefs and Roles. *Life Science Journal*, 10(3).
[11] Shohamy, E. (2000). The relationship between language testing and second language acquisition, revisited. *System*, 28(4), 541-553.
[12] Stiggins, R. J., & Conklin, N. F. (1992). In teachers' hands: Investigating the practices of classroom assessment. SUNY Press.
[13] Wiggins, Grant. (1989). "A true test." *Phi Delta Kappan* 70.9 703-713.
[14] Wishon, P. M., Crabtree, K., & Jones, M. E. (1997). Curriculum for the primary years: An integrative approach. Prentice Hall.
[15] Zhang, Z., & Burry-Stock, J. A. (2003). Classroom assessment practices and teachers' self-perceived assessment skills. *Applied Measurement in Education*, 16 (4), 323-342.

Samaneh Zolfaghari, is M.A. in ELT from Azad University of Torbat-e-Heydarieh, Iran. She has been teaching English since 2008. Her areas of interest are language testing, teacher education, and critical discourse analysis.

Hamid Ashraf, Ph.D. in ELT from University of Pune (India) has been a member of faculty at English Department, Islamic Azad University, Torbat-e-Heydarieh, Iran since 1996. He has worked on language learning skills, testing, critical thinking, language learners'' characteristics, and e-learning through doing research, presenting papers in national and international conferences and authoring a book.