Journal of English language
Teaching and Learning
University of Tabriz
Volume 11, Issue 23, (Spring & Summer 2019)

A Comparative Study of Writing Assessment Using Activity Theory-Based Assessment Model (ATBAM) and a Traditional Approach*

Sayyad Mohammad Alavi (PhD)**
Associate Professor, University of Tehran
Shiva Kaivanpanah (PhD)***
Assistant Professor, University of Tehran
Fatemeh Danesh*** **(corresponding author)**
PhD Candidate, University of Tehran, Alborz Campus

Abstract

Assessment of writing skill is generally believed to be judged by a rater subjectively and qualitatively or by using analytic scoring rubrics which can potentially result in somehow not very reliable assessment. It seems that an evaluation of writing based on a model can result in a valid and reliable writing assessment. To achieve such an objective, this study firstly aimed to develop an assessment model based on Activity Theory (AT), i. e., Activity Theory-Based Assessment Model (ATBAM), and then to employ it in the assessment of writing performances of Iranian language learners in a private language college. And finally, to achieve the concurrent validity of ATBAM, its results were compared with those of a traditional approach. Three groups of participants took part in this study: a group of upper intermediate English learners (N=29) who submitted one writing sample per week in four successive weeks, teachers (N=6) who provided learners with feedback and assigned holistic scores and course supervisors (N=2) who reassessed the writing samples on the basis of an analytic rubric of writing assessment. The results showed that using ATBAM in writing assessment results in an exploration of not only learners but also teachers engagement in the development of learners writing ability. The role of teachers feedback and teachers and learners social interaction in the development of learners writing ability could potentially provide comprehensive, fair, reliable, and valid scores in this model.

Keywords: Activity Theory (AT), Activity Theory-Based Assessment Model (ATBAM), Writing Skill Assessment, Teacher s Feedback

*Received: 2019-01-08 Accepted: 2019-04-24

**E-mail: smalavi@ut.ac.ir
***E-mail: shkaivan@ut.ac.ir
****E-mail: t_danesh@ut.ac.ir; t_danesh@daneshlanguage.com

Introduction

It is generally believed that language teaching, learning and testing principles are to be interrelated. Examining these principles and their relationship with one another has been extensively researched to provide guidelines to be practiced in English language pedagogy. This very interrelatedness, dates back to 1930s, has also mutually influenced on each other in pedagogical conditions. In other words, the incorporation of teaching into testing procedures implies that specific approaches to testing and assessment can also be traced to certain eras of dominant language teaching and learning practices. However, it seems that a gap is felt in language testing and assessment procedures when a theory is employed in teaching a language as a foreign language.

The emergence of communicative competence model in 1980s (Canale and Swain, 1980) led to a paradigm shift in language teaching and assessment. The advantages of such a shift over the previous testing approaches are its qualitative mode of assessment and the use of criterion-referenced testing in classroom assessment. Moreover, the assessment of communicative competence directed language teachers to incorporate social factors in their classroom contexts to improve learners' communicative competence. It seems that an integration of social factors to assess learners' communicative competence requires language testers to think about alternative ways of language assessment.

Alternative assessment claims that a link between assessment and instruction needs to be established by 'feedback loop' which allows instructors to continuously monitor and modify instruction based on what is already assessed. In other words, such an alternative assessment can potentially provide instructors with informative information (Alderson & Banerjee, 2001) and motivate language learners to take more responsibility for their own learning (Hamayan, 1995). Feedback as an important feature in improving the learning of the students through assessment can build confidence in students and motivate them (O Farrell, 2016). Hattie and Timperley (2007) suggest

ways of enhancing the effectiveness of feedback in classroom. The relationship between feedback and learning and learners intended performances is documented in many studies (Kluger & DeNisi, 1996; Mory, 2004; and Narciss & Huth, 2004) and the importance of feedback as a significant factor in motivating learning is also documented (Lepper & Chabay, 1985; Narciss & Huth, 2004).

Language testing and assessment circles have witnessed different forms of alternative assessment such as Dynamic Assessment (DA), Authentic Assessment, Performance-based Assessment, Portfolio Assessment, etc. (Barnard Bachelor, 2017) which have been recently developed. However, most teachers still prefer to use the old and traditional approach to testing and assessment. Any alternative assessment claims to perform specific assessment purposes. For example, DA, suggested as an alternative to traditional assessment methods, (Lantolf & Poehner, 2004; Poehner, 2007; Mattarima & Hamdan , 2011) is regarded as a means designed to bring out the potential of learning and to improve the learning effectiveness by providing more opportunities for learners to interact with each other and with teachers (Wang, 2010). In DA, teaching is a part of assessment itself and it supplies very useful information for developing interventions. Performing DA, as Lidz (1991) suggests, can help teachers or examiners discover how the students learn or can be best instructed by looking at their learning processes which are believed to be influenced by learners cognitive, affective, social, and cultural factors.

The impact of sociocultural theories on EFL contexts has been widely researched. However, there seems to be scarcity of research in testing and assessment models in which social dimensions of assessment in classroom contexts are observed. These issues appeared as the tenets of Activity Theory (AT) which is employed as a descriptive meta-theory derived from Vygotsky s(1978) social development idea (Nardi, 1996). Vygotsky (1978) states that in AT, human activity is directed toward an object and mediated by tools and artifacts that are constituted in a social context. Karanasios, Allen and

Finnegan (2015) describe AT system as the interaction between the human agent (the subject) and the world (the object) mediated by tools and artifacts. Chen, Sharman, Rao, and Upadhyaya (2013) believe that the subject in the AT is the active element, an individual or a group, of the process. It considers the whole activity system beyond just one actor or user and it includes cultural and technical mediation of human activity and artifacts in use. According to Jonassen and Rohrer-Murphy (1999), AT is not just the kind of activity to be examined but it looks at any activity in a context in terms of the person engagement, the intended goals to be achieved, the resulted objects or products, and the delineated rules and norms. Since AT claims to provide a useful framework in which the processes of changes in teaching and learning processes can be investigated, it can be argued that AT can offer a new assessment model.

There are three generations AT built on different notions. The first generation drew from Vygotsky (1978) notion of mediation focusing only on individuals and composed of subject, object and actions. This mediation in the second generation AT is on the other components of an activity system and its relationship with them (Engestrom, 1999).The second generation was built on Leont ve snotion of activity system in which human activity doesn happen in isolation but rather it takes place in a community governed by certain rules and divisions of labor (Engestrom 1987). By adding the elements of community, rules and division of labour, this model aims to represent the social/collective elements in an activity system. So, AT can be considered as a framework or a lens for studying developmental processes of human learning in two levels of individual and social interlinked simultaneously. Since the outcomes of such an activity system are prescribed, Engestrom (1990) articulated the need for a third generation AT. This third generation involves at least two interacting activity systems with a shared object. According to Engeström (1999), the third generation AT focuses on the joint activity, not individual activity, as the unit of analysis in which mediated activity can modify not only the subject, but also the

environment. Yamagata-Lynch (2010) states that the third generation AT was developed for the analysis of activity systems when the researchers take participatory and interventionist roles in the participants activity.

This study firstly aimed to develop an assessment model based on the second and third generation AT. To achieve such an objective, this study, more specifically, aimed to suggest Activity Theory-Based Assessment Model (ATBAM) as an assessment model and to employ the model in the assessment of writing performances of Iranian language learners in a private language college. Subjective scoring of writing performances of learners (Pearson, 2004) and using different scoring rubrics cannot solve the problem of subjectivity in the assessment of writing in order to get to a more valid score (Barkaoui, 2010). Therefore, the results of ATBAM were compared with those of a traditional approach to writing assessment in order to prove the validity of ATBAM. In this model all the features of ATBAM are evaluated based on the writing samples of the learners in a classroom context; that is, the contribution between teachers, learners and course supervisors in scoring the writing assignments, the feedback of the teachers to learners writing samples, the rules governed in this process and finally, the community of the assessment.

Method

This study is intended to suggest ATBAM as a model of assessment in assessing learners writing skill ability and to compare the results of ATBAM with those of a traditional approach to writing assessment. To achieve such an objective, both qualitative and quantitative longitudinal researches with a mixed-method quasi-experimental approach were employed. The following sections describe the participants, instruments, scoring scheme, and procedures of this study.

Participants

Three groups of participants, i. e., language learners, language teachers, and course supervisors, took part in this study. The Iranian language learners (N=29), who registered for an upper intermediate

course, were studying in an English language learning program in the private English college, Danesh Language Centre, in Tehran. The learners were all adult learners who have studied English for several years in this college. They participated in twenty 105-minutes sessions for seven weeks. The second group consisted of six experienced language teacher who were interviewed by the researcher. The third group consisted of the course supervisors (N=2) who were particularly instructed to supervise, observe, and score and rescore the writing sample assignments of language learners.

Instruments

Data collection instruments were considered as mediating artifacts in ATBAM. These instruments dealt with the properties of learners writing samples and teachers feedback. The second generation of AT was considered as the basis for the intended writing assessment model of this study. As Engestrom (1987) suggests, in the second generation AT, as presented in Figure 1, each participant activity does not happen in isolation but rather it takes place in a community governed by certain rules and divisions of labor.

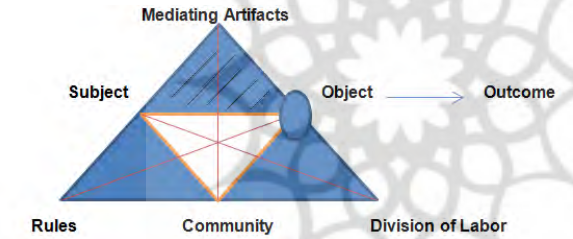

Figure 1. Second Generation Activity Theory Model

Scoring Scheme

In order to achieve the purpose of this study, all features of ATBAM were needed to be scored. Because of the diversity of the data collected for different features of ATBAM, different scoring procedures were used. In division of labor, three interval scores were measured based on the contribution among teachers, learners and course supervisors. First the writing samples of the learners were corrected and scored (0-8) by teachers. Then these sample writings

were rescored (0-5) by supervisors based on writing assessment rubrics (González, Trejo & Roux, 2017) presented in appendix 1. Finally, in the division of labor between teachers and supervisors, a score (1-4) was given to each teacher s total feedback by supervisors based on teacher s feedback rubrics (Wylie & Lyon, 2016) presented in appendix 2. Also, two sets of ordinal scores were measured by teachers for both rules and regulations and community based on Likert scale. Rules and regulations were scored based on teachers observation of learners assessment processes and community was scored based on the amount of teacher-learner contribution in the assessment process. Figure 2 represent the relationship between different features of ATBAM in assigning scores and getting to a final outcome score.

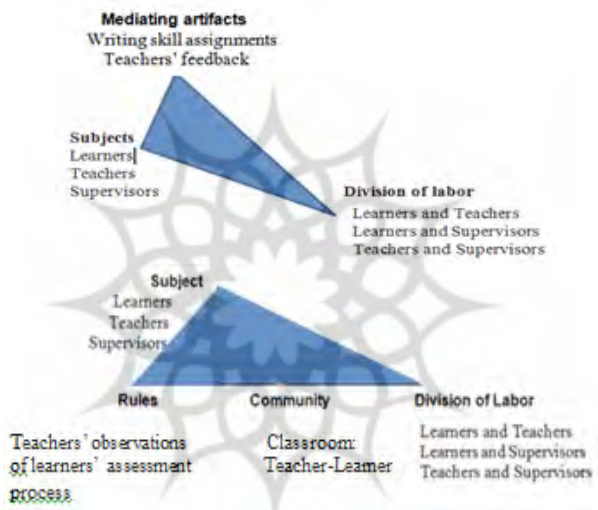

Figure 2. The Collaboration among Features of ATBAM

Table 1 represents features of ATBAM and the scores assigned to each of them. The average of all the scores gained in ATBAM is evaluated to get to a final score as an outcome score of ATBAM for each learner s writing performance.

Table 1
The Scoring Scheme Based on the Collaboration among Features 0f ATBAM

ATBAM Features		**Object**	**Score**	**Outcome Score**
Subjects in Division of Labor	**Learner and Teacher**	-Score based on subjective assessment of writing. -Teacher s feedback.	Interval	The average of scores in all aspects of AT
	Learner and Supervisor	Score based on writing assessment rubrics.	Interval	
	Teacher and Supervisor	Score based on teacher s feedback rubrics.	Interval	
Rules and Regulations		Score based on teachers observation of assessment rules.	Ordinal	
Community		Score based on teacher-learner contribution.	Ordinal	

Data Collection and Analysis

As mentioned earlier, the basis of ATBAM is the second generation AT in which assessment procedures and the final results were dealt with more features rather than any other traditional assessment procedures in EFL context. The features are teachers, learners and course supervisors (subjects), learners writing assessment (objects), learners writing samples and teachers feedbacks (mediating artifacts), teachers observation of learners assessment procedures (Rules and regulations), teachers , learners , and supervisors responsibility (Division of labour), teacher-learner contribution in classroom context (Community) and learners writing performance and improvement, and very safe and reliable scores to learners writing samples (outcomes).

The description of features of ATBAM, as a writing assessment model on the basis of second generation AT, are presented in table 2 adopted from Mwanza & Engestrom (2003). The features of ATBAM are also represented in Figure 3 as a writing assessment model based on second generation AT triangle.

Table 2
Features of ATBAM as a Writing Assessment Model (Adopted from Mwanza & Engestrom, 2003)

Components	**Descriptions**	**Examples**
Activity	What sort of Testing Activity is taking place?	Dynamic Assessment
Object/Motive	Why is the activity taking place?	Learners writing development
Subjects	Who is involved in carrying out the activity?	Learners, teachers, and supervisors
Tools	By what means are the subjects performing the activity?	Learners Writing samples, and teachers feedback
Rules and regulations	Are there any cultural norms, rules or regulations governing the performance of the activity?	Regulations, for writing assessment processes of learners, observed by teachers
Division of labour	Who are responsible for what, when carrying out activity and how are these roles organized?	All participants: Learners, teachers, and supervisors
Community	What is the environment in which this activity is being carried out?	Classroom: Teacher-Learner contribution
Outcomes	What is desired outcome from carrying out this activity?	Learners writing performance and improvement, and More reliable and safe score to learners writing performances

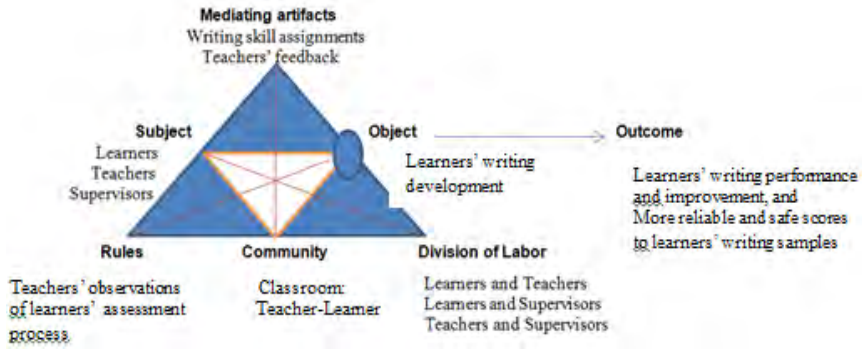

Figure 3. Features of ATBAM as a Writing Assessment Model

The data was collected during one term of English language learning program consisting of 20 sessions held in 7 weeks. The learners attended to their English classes three times a week, each time 105 minutes. Data collection started as of week three and was followed for one month (four successive weeks). As of week three up to week 7, the learners were asked to write one writing sample per week. Learners writing samples were corrected by their teachers who provided them with feedback on their mistakes each week. After four weeks, teachers assigned a holistic score (0-8) to each learner s writing samples subjectively based on their subjective correction during the course. Then at the end of the course, the course supervisors not only rescored the learners writing samples and assigned a score (0-5) based on a writing assessment rubric but also evaluated and scored (1-4) teachers feedbacks based on the descriptive teacher s feedback rubric. These three sets of interval scores were represented as the division of labor feature in ATBAM. The other two sets of scores were also assigned by teachers to rules and regulations, and community features. Scores to rules and regulations were based on teachers observation of learners assessment process and scores to community were assigned based on the amount of contribution between teacher and learner in the classroom.

Results

The results of the assessment of learners writing samples in four successive weeks are analyzed. More specifically, reliability, descriptive, correlational and regression analyses were performed to compare scores of learners writing performances based on ATBAM with the scores based on a traditional approach to assessment.

Reliability Analysis

In order to estimate the consistency of scoring, the writing samples of learners on the basis of the five features of analytical scoring rubric of writing (González, Trejo & Roux, 2017), a reliability analysis was performed. The results in Table 3 show that the course supervisors scoring is consistent.

Table 3

Reliability Co-Efficiency of Features of Writing

Feature of Writing	**No. of items**	**Reliability Co-efficient**
Content	4	.69
Organization	4	.77
Use of Language	4	.76
Use of Vocabulary	4	.67
Mechanics and Spelling	4	.84

Descriptive analysis

In order to have a general picture of the data in this study, a descriptive analysis was performed. As Table 4 shows, the mean scores of learners in the second and fourth samples are low. The homogeneity of scores in the first sample is low while that of the third sample is high. The reason for the low standard deviation in the first sample is that the participants were not much familiar with the scoring criteria. The highest standard deviation in the third sample may be due to the writing improvement of some learners.

Table 4
Descriptive Statistics of Four Samples of Writing

Scores to Writing Samples	N	Minimum	Maximum	Mean	Std. Deviation
First	28	13.00	24.00	19.03	2.97
Second	28	12.00	23.00	18.50	3.30
Third	26	11.00	24.00	19.26	3.81
Forth	26	12.00	23.00	18.53	3.02

The descriptive statistics of writing features in the four writing samples of learners appear in Table 5, indicating that these features were developed almost similarly in the four writing samples.

Table 5
Descriptive Statistics of Total Scores of Features of Writing Rubrics

Features of Writing	N	Minimum	Maximum	Mean	Std. Deviation
Content	23	11.00	20.00	15.91	2.17
Organization	23	11.00	20.00	15.47	2.59
Use of Language	23	9.00	19.00	15.04	2.75
Use of Vocabulary	23	11.00	18.00	15.04	2.20
Mechanics and Spelling	23	11.00	20.00	15.47	2.85

As Table 6 shows, the learners' highest mean score is in Rules and Regulations. This indicates that the teachers were satisfied with the learners participation in learning and assessment of their writing. The lowest mean score is for community feature of AT that shows teachers scores for their collaboration with learners in the class. Table 6 also shows that the scores of learners on division of labor are more homogenous. A comparison of the total scores of ATBAM and a Traditional approach show that total score of learners writing samples in ATBAM has got higher mean than division of labor in which there are a traditional approach to assessment and teachers feedback.

Table 6
Descriptive Statistics of Features of ATBAM

Scores to Features of ATBAM	N	Minimum	Maximum	Mean	Std. Deviation
Division of Labor	23	53.33	97.33	81.47	12.26
Rules and Regulations	29	33.33	100.00	88.50	20.46
Community	29	33.33	100.00	78.16	28.55
Total Score to Writing Samples	23	57.33	99.11	86.57	12.525

Correlational Analyses

Table 7 shows correlation between learners writing scores by teachers and supervisors and their scores based on ATBAM. The results show that there is a positive and moderate correlation between ATBAM based scores and the total scores of supervisors on learners writing. There is a low and positive correlation between ATBAM scores and the total score of teachers. Also, there is a low and negative correlation between supervisors and teachers total scores.

Table 7
Correlation between Learners' Writing Scores and their ATBAM Based Scores

		Supervisors Total Score	Teachers total Score
ATBAM Score	Pearson Correlation	.48*	.13
	Sig. (2-tailed)	.01	.55
	N	23	23
Supervisors Total Score	Pearson Correlation	1	-.20
	Sig. (2-tailed)		.35
	N	23	23

Table 8 shows the correlational analysis of learners writing scores based on their teachers and supervisors. The table shows that there is a low and negative correlation between teachers scores and

supervisors scores. Writing one has got positive and moderate correlation with writing two, three, and four. Writing two has got positive and strong correlation with writing three and four. Writing three has got positive and moderate correlation with writing four.

Table 8

Correlational Analysis of Learners' Four Writing samples' Scores Based on Supervisors and Teachers

		Supervisors Scores to Four Writing Samples			
		First Writing Samples	Second Writing Samples	Third Writing Samples	Fourth Writing Samples
Teachers total Score	Pearson Correlation	-.231	-.020	-.235	.001
	Sig. (2-tailed)	.236	.918	.248	.996
	N	28	28	26	26
Supervisors Score to First Writing Samples	Pearson Correlation		.567**	.485*	.554**
	Sig. (2-tailed)		.002	.012	.004
	N		27	26	25
Supervisors Score to Second Writing Samples	Pearson Correlation			.773**	.817**
	Sig. (2-tailed)			.000	.000
	N			25	26
Supervisors Score to Third Writing Samples	Pearson Correlation				.637**
	Sig. (2-tailed)				.001
	N				23

Development of Writing Features across Writing Samples

In order to find the extent to which the writing features improved in the four sample writings of learners, a multiple regression using Enter

method was employed. The results clearly showed that the writing features were improved more in the third and the fourth samples. The regression analyses, Table 9, showed that organization, and mechanics and spelling have more contribution to writing one. While, in writing four the content and organization, in writing three Use of language, and in writing two, use of vocabulary and mechanics have more predictive power.

Table 9

Contribution of Writing Practices to Development of Writing Features

Writing Features	Constant	Regression coefficients			
		Writing One	Writing Two	Writing Three	Writing Four
Content	2.16	0.03	0.18	0.12	0.37
Organization	-2.090	0.29	0.19	0.01	0.42
Use of Language	-0.8	0.18	0.13	0.39	0.1
Use of Vocabulary	0.65	0.18	0.33	0.18	0.03
Mechanics and Spelling	-2.84	0.29	0.3	0.12	0.22

Table 10 shows a multiple regression analyses of teachers feedback, learners rules and regulations, and teacher-learner community contribution to the writing scores of learners. In order to understand the extent to which providing feedback contribute to the writing scores of learners, a multiple regression analyses was performed. The results show that teachers feedback to writing sample 3 has more predictive power to supervisors score to writing sample four. Also, learners scores to rules and regulations, and community in writing sample three have more predictive power to supervisors score to writing sample four.

Table 10

Contribution of Teachers' Feedback, Learners' Rules and Regulations, and Teacher-Learner Community to Learners' Writing Scores

Scores	Constant	Teachers Total Score	Supervisors Scores to Writing Samples			
			One	Two	Three	Four
Teachers Feedback	-22.83	.04	.28	.75	-.74	.96
Learners Rules and Regulations	51.47	.15	.35	-.18	-.07	.3
Teacher-Leaner Community	14.57	-.17	.23	.14	-.33	1.07

Table 11 shows the contribution of teachers and supervisors scores to ATBAM based scores. The multiple regression analysis shows that the learners writing score based on supervisors is a better predictor of the scores the leaners received from their teacher. The contribution of teachers and supervisors scores on writing performances of learners to ATBAM based scores is also appeared in Equation (1).

Table 11

Contribution of Teachers' and Supervisors' Scores on Writing Performances of Learners To ATBAM Based Scores

Total Scores	Unstandardized Coefficients		Standardized Coefficients	t	Sig.
	B	Std. Error	Beta		
(Constant)	19.87	25.24		.78	.44
Teachers	.20	.16	.23	1.24	.22
Supervisors	.61	.22	.53	2.79	.01

a. Dependent Variable: Total ATBAM score

Equation (1)

Total ATBAM Score=19.87+0.2 *learner writing score by teacher
+0.61*learner writing score by supervisor

Table 12 shows the contribution of writing features to ATBAM based scores. The results show that the total score to Mechanics and

Spelling and use of vocabulary are the first and the second predictors of scores in ATBAM. Equation (2) present the formula based on the information in Table 12.

Table 12

Contribution of Writing Features' Scores to ATBAM Based Scores

	Unstandardized Coefficients		Standardized Coefficients		
Total Scores	B	Std. Error	Beta	T	Sig.
(Constant)	43.26	16.84		2.56	.02
Content	.08	.49	.07	.16	.87
Organization	-.01	.38	-.01	-.04	.96
Use of Language	-.51	.32	-.56	-1.56	.13
Use of Vocabulary	.42	.63	.37	.66	.51
Mechanics and Spelling	.57	.28	.66	2.02	.05

a. Dependent Variable: Total ATBAM Score

Equation (2)

ATBAM based scores of learners =43.26+0.08 Total Content score in four writing samples

-0.01Total Organization Score in four writing samples

-0.51Total Language Use Score in four writing samples

+0.42 Total Vocabulary Use Score in four writing samples

+0.57 Total Mechanics and Selling Score in four writing samples

Discussion of the Results

The objective that this study intended to achieve was to develop an assessment model based on AT (ATBAM) and to employ such a model in the assessment of writing samples of learners. The results of ATBAM were compared with the results of a traditional approach to writing assessment; that is, subjective scoring by teachers themselves or based on an assessment rubric without any useful feedback to learners writing samples. The results showed that the role of teachers feedback and teachers and learners social interaction cannot be ignored in learners writing improvement. It seems that jut scoring the learners writing by using scoring rubrics cannot lead teachers and

researchers to very valid and accurate scores. The main purpose of feedback in this study was to reduce the gap between learners performance and the goal. This is in line with Hattie and Timperley (2007) who suggests the ways of enhancing the effectiveness of feedback in classroom. As a result, the effectiveness of teachers feedback on the writing of the learners results in learners writing skill improvement. The homogeneity of scores and the writing improvement of learners in writing sample 3 in week 5 show this fact. Comparing to week 1, this improvement in week 5 may be because of the learners familiarity with scoring criteria and their teachers feedback. Also all features of writing samples in scoring rubric, which was used by supervisors scoring of the writing samples of learners as a part of division of labor in ATBAM, were developed almost similarly in four writing samples of the learners. This can also due to the effectiveness of feedback.

Literature extensively documented the relationship between feedback and learning and learners intended performances. The following sources show such a relationship (Kluger & DeNisi, 1996; Mory, 2004; and Narciss & Huth, 2004). And some other studies have done on the importance of feedback as a significant factor in motivating learning (Lepper & Chabay, 1985; Narciss & Huth, 2004). O Farrell (2016) refers to the characteristics of a successful feedback as to build confidence in the students, to motivate students, to improve their learning, to provide students with performance improvement information, to correct errors, and to identify strengths and weaknesses.

The multiple regressions also showed the importance of feedback in learners writing improvement. Writing features of scoring rubric were developed across writing samples of learners. These features were improved more in the third and fourth writing samples of learners. The results also showed that the feedback was focused more on the Use of Vocabulary, and Mechanics and Spelling features in writing one, Use of Language feature in writing two, Content and Organization in writing three. These feedbacks in each writing sample

affected the learners better performance in such features in the following writing samples.

Descriptive statistics of features of ATBAM revealed that teachers were satisfied with learners performance in assessment process by observing that the learners respected to rules and regulations of assessment. The mean score of community feature was lower than the other features. The reason is that although teachers feedback and collaboration resulted in better improvement of learners writing development, some teachers in some classes didn have good social collaboration with their learners and provided them with not enough feedback which resulted in less improvement and lower score in ATBAM. Also, higher mean score of total score in ATBAM rather than a traditional assessment shows the importance of social factors which can be observed in Rules and Regulations, and Community as well as teachers feedback.

Comparing the results of learners writing scores in ATBAM with those of a traditional approach included in ATBAM; that is, supervisors scores based on a rubric and those of teachers subjectively, revealed that the correlation between the scores of ATBAM and those of supervisors was better than such a correlation between the scores of ATBAM and teachers scores. Also very low and negative correlation between teachers scores and supervisors scores revealed the fact that assessing based on a rubric can results better than just scoring subjectively. Literature shows that not only subjective scoring of raters (Pearson, 2004) but also using different scoring rubrics cannot solve the problem of subjectivity in order to get to a more accurate score (Barkaoui, 2010). Using ATBAM which includes all the features of social interactive context and even traditional approach of assessment in itself seems very practical solving this problem of subjectivity of writing assessment scoring.

Conclusions and Assessment Implications

The results of this study showed that the use of ATBAM as a social and interactive assessment model developed to assess learners writing performances seems very potential due to its inclusion of different

assessment features needed in any assessment context. ATBAM includes not only traditional approach to assessment, i.e., teachers subjective assessment of learners writing performances and assigning a holistic score subjectively, and teachers or assessors use of writing assessment rubrics and assigning a score subjectively (Pearson, 2004), but also communicative approach to assessment in which social and interactive factors and whoever and whatever in the assessment context must be evaluated. This developed assessment model imply that it is not just learners that need to be assessed but even teachers feedback to learners writing samples must be evaluated. The reason is that If the learners are provided with not sufficient feedback, they cannot improve in their further writing samples. And this is the duty of just a good teacher to provide the learners with their needed feedback. As a result, teacher s feedback evaluation in assessment process which has been neglected in traditional approach to language assessment must be done in order to get to a fairer score for each learner.

Also comparing the results of writing assessment of learners in ATBAM with those of a traditional assessment including in ATBAM show that ATBAM is more capable to assess the writing skills of individual learners and engage them more in social and interactive contexts that could potentially provide comprehensive, fair, reliable, and valid scores of their writing performances. Using ATBAM in writing assessment also results in an exploration of learners engagement in the development of their writing ability, teacher and learners interaction in the development of written social interaction which have been neglected in the traditional assessment of writing.

Indicating the potentiality and effectiveness of ATBAM designed for the purpose of this study theoretically and experimentally may lead EFL language teachers, course supervisors and researchers to substitute and use this model of an integrated, collaborative, interactive and multidimensional language assessment to assess learners writing skill in English as a Foreign Language (EFL) context.

Using this model may solve all the previous problems of raters, and scoring rubrics in order to get to a more accurate score.

References

Alderson, J. C. & Banerjee, J. (2001). 'Language testing and assessment (Part 1): state-of-the-art review'. *Language Teaching, 34*(4), 213 -36.

Barkaoui, K. (2010). Variability in ESL essay rating processes: The role of the rating scale and rater experience. *Language Assessment Quarterly,7*(1), 54-74.

Barnard Bachelor, R. (2017). Alternative Assessments and Student Perceptions in the World Language Classroom. *Online Submission*, *6*, 30-44.

Canale, M., & Swain, M. (1980). Theoretical bases of communicative approaches to second language teaching and testing. *Applied Linguistics*, *1*, 1-47.

Chen, R., Sharman, R., Rao, H. R., & Upadhyaya, S. J. (2013). *"Data Model Development for Fire Related.*

Engeström, Y. (1987). *Learning by expanding: An activity-theoretical approach to developmental research.* Helsinki: Orienta-Konsultit.

Engeström, Y. (1990). *When is a tool? Multiple meanings of artifacts in human activity*. In Y. Engeström, Learning, working and imagining: Tweleve studies in activity theory. Helsinki: Orienta-Konsultit.

Engeström, Y. (1999) *Innovative learning in work teams: analysing cycles of knowledge creation in practice*, in: Y. ENGESTRÖM et al (Eds.) Perspectives on Activity Theory, (Cambridge, Cambridge University Press), 377-406.

González, E. F., Trejo, N. P. & Roux, R. (2017). Assessing EFL university students writing: a study of score reliability. *Revista Electrónica de Investigación Educativa, 19*(2), 91-103. https://doi.org/10.24320/redie.2017.19.2.928.

Hamayan, E.V. (1995). Approaches to alternative assessment. *"Annual Review of Applied Linguistics," 15*, 212-226.

Hattie, J. and Timperley. H. (2007). The Power of feedback. *Review of Educational Research*, 77, 81-112.

Jonassen, D. H., & Rohrer-Murphy, L. (1999). Activity theory as a framework for designing constructivist learning environments. *Educational Technology Research and Development, 47*, 61˚79.

Karanasios, S., Allen, D., & Finnegan, P. (2015). "Information Systems Journal Special Issue On: Activity Theory in Information Systems Research," *Information Systems Journal* (25:3), pp. 309-313.

Kluger, A., & DeNisi, A. (1996). The effects of feedback interventions on performance: A historical review, a meta-analysis, and a preliminary feedback intervention theory. *Psychological Bulletin, 2*, 254 284. doi:10.1037/0033-2909.119.2.254

Lantolf, J. P., & Poehner, M. E. (2004). *Dynamic assessment in the*

language classroom (CALPER Professional Development Document CPDD-0411). University Park, PA: The Pennsylvania State University, Center for Advanced Language Proficiency Education and Research.

Lepper, M. R., & Chabay, R. W. (1985). Intrinsic motivation and instruction: Conflicting views on the role of motivational processes in computer-based education. *Educational Psychologist, 20*(4), 217˚230.

Lidz, Carol S. (1991). *Practitioner's guide to dynamic assessment.* New York, NY: Guilford Press.

Mattarima, K. & Hamdan, A. R. (2011). Understanding students learning strategies as an input context to design English classroom activities . *International Journal of Psychological Studies, 3* (2), 238-248.

Mory, E. H. (2004). *Feedback research review*. In D. Jonassen (Ed.), Handbook of research on educational communications and technology (pp. 745˚783). Mahwah, NJ: Erlbaum Associates.

Mwanza, D. and Engestrom, Y. (2003). *Pedagogical Adeptness in the Design of E-learning Environments: Experiences from the Lab@Future Project.* In: Proceedings of World Conference on E-Learning in Corporate, Government, Healthcare, and Higher Education 2003, 7-11 Nov 2003, Phoenix. USA.

Narciss, S., & Huth, K. (2004). *How to design informative tutoring feedback for multimedia learning.* In H. M. Niegemann, D. Leutner, & R. Brunken (Ed.), Instructional design for multimedia learning (pp. 181-195). Munster, New York: Waxmann.

Nardi, B. (1996). *Activity theory and human-computer interaction.* In B. Nardi (Ed.), Contexts and consciousness. Activity theory and human-computer interaction (pp. 7-16). The MIT Press: Cambridge.

O Farrell, C. (2016). *Enhancing Student Learning through Assessment.*

Pearson, P.C. (2004). *Controversies in second language writing: Dilemmas and decisions in research and instruction.* The University of Michigan Press.

Poehner, M. E. (2007). Beyond the Test: L2 Dynamic Assessment and the Transcendence of Mediated Learning. *Modern Language Journal, 91*, 323-340.

Vygotsky, L. (1978). *Mind in society: The development of higher psychological processes* (Ed. By M. Cole, V. John-Steiner, S. Scribner, & E. Souberman). Cambridge, MA: Harvard University Press.

Wang, T. H. (2010). Web-based dynamic assessment: taking assessment as teaching and learning strategy for improving students e-Learning effectiveness. *Computers & Education, 54*(4), 1157˚1166.

Wylie, C. & Lyon, C. (2016). *Using the Formative Assessment Rubrics, Reflection and Observation Tools to Support Professional Reflection on Practice (Revised)*. Formative Assessment for Students and Teachers State Collaborative on Assessment and Student Standards of the Council of Chief State School Officers. Washington, DC.

Yamagata-Lynch, L. C. (2010). "*Understanding Cultural Historical Activity Theory,*" in Activity Systems Analysis Methods. Springer, pp. 13-26.

Appendix 1

Scoring analytical rubric for writing assessment adopted from González, Trejo & Roux (2017)

Score	Content	Organization	Use of Language	Use of Vocabulary	Mechanics and Spelling.
5	Text shows knowledge of the topic and gives details or examples to support main ideas. Text fully corresponds to task requirements. Communication is effective.	Organizational skills are present in the text making flow and coherence of ideas smooth. Main ideas and structure of text are easily found and logically sequenced.	Text makes use and maintains use of complex language structures effectively. There are no errors of idioms, collocations and grammar in general. Facility in use of language is apparent.	Demonstrates sophisticated and broad use of vocabulary. Effective and appropriate use of idiomatic expressions and colloquialisms; shows awareness of connotations and their meaning.	Writing shows mastery of punctuation and spelling conventions. Capitalization and paragraphing errors and typos are not found.
4	Task is answered in its majority but information may be redundant or unnecessary. Some detail is given. Sufficient development of main ideas. Some gaps may be found among information.	Adequately organized with the use of organizational patterns and connectors but sequencing of information is incomplete. Connection of main ideas may be lost but meaning is still understood.	Grammatical accuracy consistently maintained. Few errors of idioms, collocations and grammar in general. Complex sentences present minor errors.	Demonstrates sophisticated use of vocabulary. Good command of idiomatic expressions and colloquialisms. Minor errors in vocabulary use.	Writing shows occasional errors of punctuation and spelling conventions. Capitalization and paragraphing errors and typos are occasionally found.
3	Task is addressed adequately but information may be missing. Some details are used to support the main idea. Shows some knowledge of the main topic and limited development of main ideas.	Some organizational skills are present. Use of cohesive devices makes text clear and understood. Occasional deficiencies can lead to "jumpiness" among information.	Some grammatical "slips" may be found. Grammatical errors such as verb tense, verb agreement, number, word order, articles, pronouns, and prepositions are found but they do not lead to misunderstanding. Context given in text allows for interpretation of meaning.	Vocabulary accuracy is high though occasional errors may be found. Adequate and appropriate word/idiom choice and use. Some incorrect word choice does occur without impeding communication.	Writing shows few errors of punctuation and spelling conventions. Few capitalization and paragraphing errors and typos are found.
2	Task reveals little relevance to the topic. Major gaps in information are found and insufficient details to support main ideas are given. Inappropriate information. Pointless repetition of information.	Small pieces of text are linked with basic connectors. Unsatisfactory cohesion may cause most but not all of the information to seem sloppy and non-fluent.	Frequent grammatical inaccuracies found. Frequent and basic errors of tense, agreement, number, word order, articles, pronouns, and prepositions are found. Understanding of ideas is seldom confusing.	Sufficient control of elementary vocabulary to express basic ideas. Repetition of vocabulary is frequent. Frequent misuse of word form use, word/idiom choice and use, making communication confusing.	Writing shows frequent errors of punctuation and spelling conventions. Capitalization and paragraphing errors and typos are frequently found. Meaning may be confusing.
1	Task presents limited relevance to main topic. Inadequate development of topic. Details are not given.	Groups of words connected with simple connectors such as "and", "but" or "because". Cohesion is almost absent. Connection among ideas is difficult to find making information confusing or misleading.	Almost all or most of the basic grammatical constructions are inaccurate. Major issues in simple sentences. Errors of negation, agreement, number, word order, articles, pronouns, and prepositions frequently found. Understanding of information difficult.	Text has little knowledge of English vocabulary, idioms and word forms. Language sufficient for coping with simple survival needs. Information is basically translated. Inappropriate choice of word forms.	Almost all spelling is inaccurate and the text shows an ignorance of punctuation conventions . Text is dominated by capitalization and paragraphing errors and typos. Meaning is obscured.
0	Task does not reveal development topic. Totally inadequate answer to task. No details are given. Content insufficient to assess.	Cohesion is totally absent. Writing is fragmented making communication impossible to obtain. Lack of structure in information leads to absence of organization. Content insufficient to assess.	All language use is inaccurate. Meaning obscured. Content insufficient to assess.	No apparent vocabulary use or vocabulary comprehension is present in text. Content insufficient to assess.	All spelling is inaccurate and the text shows an ignorance of punctuation conventions . Text is dominated by capitalization and paragraphing errors and typos. Meaning is obscured. Content insufficient to assess.

Appendix 2

Rubric of descriptive teachers feedback adopted from Wylie & Lyon (2016)

1 Beginning	2 Developing	3 Progressing	4 Extending
The teacher provides evaluative feedback on a specific piece of work (e.g., a score, grade, or other summative feedback). *OR* Feedback seems disconnected to the intended learning goals. ☐	The teacher provides descriptive feedback on a specific piece of work that supports the learning goals and/or reflects the criteria for success. ☐	The teacher provides descriptive feedback on a specific piece of work that supports the learning goals and/or reflects the criteria for success. ☐	The teacher provides descriptive feedback on a specific piece of work that supports the learning goals and/or reflects the criteria for success. ☐
Corrective feedback does all the thinking for the students; subsequent student actions consists solely of following directions. ☐	Corrective feedback sometimes does all the thinking for the students; other times it appropriately scaffolds the next steps that students are to take. ☐	Corrective feedback appropriately scaffolds the next steps students are to take, pointing out one or more areas to work on, followed by a suggestion, reminder, or question to elicit further learning from the students. ☐	Corrective feedback appropriately scaffolds the next steps students are to take, pointing out one or more areas to work on, followed by a suggestion, reminder, or question to elicit further learning from the students. ☐
The teacher does not have a systematic approach for providing feedback to most or all students. ☐	It is unclear whether the teacher has a systematic approach for providing feedback to most or all students. ☐	It is unclear whether the teacher has a systematic approach for providing feedback to most or all students. ☐	It is clear that the teacher has a systematic approach for providing feedback to most or all students. ☐
There is no opportunity for students to review the feedback, ask questions in order to internalize the feedback, or apply the feedback to their work in meaningful ways. ☐	There is little or no opportunity for students to review the feedback, ask questions in order to internalize the feedback, or apply the feedback to their work in meaningful ways. ☐	Students are provided with limited structures and supports (e.g., limited time is provided or students are confused about the process) to review the feedback, ask questions in order to internalize the feedback, or apply the feedback to their work in meaningful ways. ☐	Students are provided with ample structures and supports (e.g., time, feedback structures, etc.) to review the feedback, ask questions in order to internalize the feedback, or apply the feedback to their work in meaningful ways. ☐