

Traditional Versus Alternative Writing Assessment

Ali Forutan, Ph.D. Candidate,
English Department, Farhangian University of Isfahan, Isfahan, Iran
forutanali@yahoo.com

Abstract

The present study intended to evaluate EFL students' performance on a traditional method of writing assessment in comparison to an alternative assessment technique by taking into account the effects of teacher and peer feedback. Moreover, the attitudes of the participants regarding assessment procedures and activities were focused in this study. To this end, thirty male and female language learners taking Advanced Writing course at university level were chosen as the participants of the study. First, all the participants were given a traditional writing test. Then, they were provided with a video and a reading passage related to the writing topic. Additionally, they received teacher and peer comments and feedback on their essays. As a further step, a questionnaire was distributed to collect information about the participants' opinions on the provided treatments. Analysis of the results indicated that in general most EFL learners had a better performance in alternative assessment mainly on the components of content, organization and style in comparison to traditional one. Also, the participants showed a positive attitude in general about alternative assessment through teacher and peer feedback, chiefly towards teacher comments.

Keywords: alternative assessment, feedback, analytic rating scale, traditional assessment, feedback

Introduction

Writing has been an essential area in applied linguistics for more than half a century and is a domain of dynamic academic investigation and discussion. Its multifarious, versatile character appears continuously to elude sufficient description and clarification, and various types of research have been called for to make clear both how writing works and how it should be taught and tested. One factor is, of course, the overarching implication of writing in our lives, its multiple roles in social, professional and academic contexts and the importance it has in determining our life chances. The various purposes of writing, then, the increased complexity of its contexts of use and the different backgrounds and needs of those wishing to learn it, all push the study of writing into wider frameworks of analysis and understanding (Hyland, 2002).

Furthermore writing has also turn out to be more significant as principles of CLT. It refers to teaching a foreign language as a system of communication not as an object under investigation in EFL and ESL contexts. The conventional perspective in language teaching that writing acts chiefly to help and strengthen patterns of spoken language application, grammatical structures, and words, is being replace by the concept that writing in another language is a valuable activity in and of its own (Weigle, 2002). Since the introduction of writing cognitive model by Flower and Hayes (1981) a significant quantity of investigation has been conducted that clarifies the recurring character of the writing process (Bereiter and Scardamalia 1987; Berninger et al. 1996; Fayol 1991; Raimes 1987 and Zamel 1983). Research suggests that topical knowledge and revision are integral parts of text development and that good writing is a function of such factors (Bosher 1998; Wallace et al. 1998 and Weigle 2002). These results led to the process-oriented approach to writing and had impacts on lots of academic contexts at the educational and evaluation stages.

As Susser (1994) argued, the term process can be defined in three ways: (1) the act of writing, (2) the emphasis of writing instruction on process writing pedagogies, and (3) theories of writing. When writing timed impromptu essays, students go through a process in its first usage; writing, composing, or transcribing. The term process in this study is used in its second sense, where process signifies a process of discovery in which ideas are produced and not just transcribed as writers think through and organize their ideas before writing and revising their drafts. This second sense of process is also related to an expanded theoretical understanding of the construct of writing that is more social and less bound by the ability of an individual, as many traditional theories of writing suggest. There is a difference between process-oriented English as a Second- or Foreign Language (ESL/EFL) writing instruction and product-oriented assessment. Although the methodology for teaching writing in ESL/EFL classes has shifted towards process-oriented approaches over the last two decades, the assessment of writing in such settings has continued to focus on written products such as timed impromptu essays.

Background of the study

Defining the skill that is intended to be assessed is a critical starting point in designing a test, and the definition of writing ability for a particular context depends in large measure on such considerations as the particular group of second/foreign- language writers and the kind of task that these learners are probably to involve in (Weigle, 2002). To present such a definition, reference is made to what Hyland (2002, p.5) has outlined as general approaches to writing.

Table 2.1 General approaches to writing (Hyland 2002, p.5)

The first approach focuses on the products of writing by examining texts in various ways, either through their formal surface elements or their discourse structure.

The second approach, loosely divided into Expressivist, Cognitivist and Situated strands, focuses on the writer and describes writing in terms of the process used to create texts.

The third approach emphasizes the role that readers play in writing, adding a social dimension to writing research by elaborating how writers engage with an audience in creating coherent texts.

Since the present study is basically concerned with the second approach, the main contours of three positions which together have contributed to the tremendously influential process writing movement will be sketched:

- *The first focuses on the personal creativity of the individual writer
- *The second on the cognitive process of writing
- *The third on the writer's immediate context

Founded on the work of Elbow (1998), Murray (1985) and others, this view persuades writers to discover their own insight to construct writing that is new and natural. There is a fundamental supposition that thinking emerges before writing and that the free expression of thoughts leads to enhancement of self-awareness and cognitive development. Writing development and personal development are seen as symbiotically interwoven to the extent that 'good therapy and composition aim at clear thinking, effective relating, and satisfying self-expression' (Moffett, 1982, p.235). At the heart of this model is the view that writing is a 'non-linear, exploratory and generative process whereby writers discover and reformulate their ideas as they attempt to approximate meaning' (Zamel, 1983, p. 165). Following Emig's (1983, cited in Hyland, 2002) description of composing as recursive, not as an interrupted

task, a great deal of research has revealed the complexity of planning and editing activities, the influence of task and the value of examining what writers do as a source of data.

A knowledge-telling model indicates the reality that beginners plan and revise less in comparison to advanced writers and they mainly concentrate on creating content out of the internal sources. Novices' major purpose is merely to express what they can remember based on the homework, the genre or the subject. A knowledge-transforming model suggests how skilled writers use the writing task to analyze problems and set goals. These writers are able to reflect on the complexities of the task and resolve problems of content, form, audience, style, organization, and so on within a content space and a rhetorical space, so that there is continuous interaction between developing knowledge and text. Knowledge transforming thus involves actively reworking thoughts so that in the process not only text, but also ideas, may be changed. This model provides valuable psychological insights into writing activity and helps explain the difficulties often experienced by skilled writers because of task complexity and lack of topical knowledge. It also helps account for reflective thought in writing, and this has obvious implications for teaching and testing. It also put emphasis on the significance of feedback and revision in the process of creating content and expression. It is unclear from this model; however, how writers really neither make the cognitive transition to a knowledge-transforming model, nor is it spelt out what occurs in the intervening stages and whether the process is the same for all learners.

In contrast to the cognitivist view, writing as a situated act gives greater emphasis to the actual performance of writing in a particular context. Writing is seen as a social act that can merely take place within a particular condition. It is therefore influenced both by the personal attitudes and social experiences that the writer brings to writing and the impact of particular political and institutional contexts in which it takes place. Research here is going to go beyond the feasible workings of writers' minds and into the substantial and pragmatic milieus in which writing happens. This perspective refutes the myth of the scattered creator and sets out to explain how context cues cognition (Flower, 1989). The emphasis is actually placed on a notion of context as the situation of expression (Nystrand, 1987). (Flower (1989) elaborates this as the effects of prior knowledge, assumptions and expectations together with features of the writing environment which selectively tap knowledge and trigger specific processes.

In choosing or designing a writing test, the logical place to begin is by considering what we plan to use the test for. In other words, why are we interested in testing writing ability- what is our purpose? Bachman and Palmer (1996) discuss two main purposes for language tests, of which writing tests can be considered to be a subtest. The primary purpose is to make inferences about language ability, and the secondary purpose is to come to conclusion based on those inferences. Such inferences are subsequently utilized as information for making various decisions at an individual, educational setting, or curriculum level such as proficiency, diagnosis, and evaluation of attainment. Inferences about general language proficiency can be used to make decisions such as entrance to educational courses, assignment into various levels of language classes, exclusion from specific assignments, or choosing for a special occupation. Inferences about diagnosis- that is, the strong and weak points of students, are used mostly by teachers to adapt their teaching to satisfy their learners' requirements. Deductions about improvement or the degree to which learners attained definite educational objectives are made use of to make judgment about ranking and promotion at the individual level, and adjustment of teaching at the level of classroom. Moreover inferences regarding attainment are used on state or national levels to make assessment about syllabus and financial support for educational plans.

Backman and Palmmer (1996, p.17) maintain that 'the most important consideration in designing and developing a language test is the use for which it is intended, so that the most important quality of a test is its usefulness.' They define test usefulness in terms of six

qualities, including: reliability, construct validity, authenticity, interactiveness, impact, and practicality. While these qualities are all important, it must be emphasized that it is virtually impossible to maximize all of them. In particular, practicality, or the amount of available resources, is a limiting factor, and requires prioritization among the other qualities. Therefore, instead of attempting to maximize each quality of usefulness, the task of the test developer is to establish a suitable balance through the qualities for definite conditions.

An assessment task in which examinees essentially generate a sample of composition, often referred to as a direct test of writing, is probably the most common method of testing writing in both first and second language contexts (Hamp-Lyons, 1990). Hamp-Lyons (1991 a, cited in Weigle, 2002)) offers some features of a supposed direct test of writing. Test takers should write as a minimum one section of uninterrupted text (at least 100 words); writers are provided with a set of directions or prompt and they have significant leeway in replying the prompt; each text is read by one, and in general two or more, skilled scorers; assessments are based on a general benchmark, for instance one or two rating scales. A couple of extra features of the majority of these writing tests are the fact that samples are written in a restricted period of time, normally within thirty minutes to two hours, and that the subject matter is unfamiliar to test takers before running the test.

The term direct is commonly used to contrast this type of test with so-called indirect tests of writing—most often, multiple-choice tests of grammar and usage. However, the term direct is somewhat problematic, since any test is at best an indirect indicator of an underlying ability (Messick, 1994). Therefore, the term traditional or timed impromptu writing test would be used throughout this study to refer to this approach to assessing writing. This account of a traditional writing test makes it different from so-called indirect tests of writing and portfolio assessment, or the assessment of a number of texts written in non testing circumstances during an extensive period of time.

Traditional tests enjoy a firm inquiry foundation and have appeared broadly conventional as a method for judging writing in various conditions. According to Cohen (1994) assessments are categorized according to three key applications: research, instructional, and administrative. In nearly all of the conditions that these sorts of tests have been investigated administrative tests which are placement, general proficiency assessments, and so forth have been used. Traditional or timed impromptu writing tests are regularly applied in research as well, for instance, to gather data of creative language use for investigations of second-language development in different situations. Another key component of a writing assessment after the issues of the development and trialing of tasks is the way written product is going to be scored. The scoring methods are very important since the score is eventually a tangible criterion in decisions making and deductions about test takers. A score in a writing assessment is the outcome of an interaction that involves not merely the test taker and the test, but the test taker, the prompt or task, the written text itself, the rater(s) and the rating scale (Hamp-Lyons, 1990; Mcnamara, 1996). Weigle (2002) argues that, of these elements, two signify fundamental considerations in scoring: defining the rating scale, and ensuring that raters use the scale appropriately and consistently.

Among the primary decisions to be made in choosing a system of scoring is what sort of rating scale is needed: it means should a single score be assigned to each composition, or every text needs to be assessed on several various factors? During the last three decades, this matter has been the focus of a lot of studies and debates. Three major categories of rating scales are talked about in the literature of writing. They are primary trait scales, holistic scales, and analytic scales. These three types of scales can be characterized by two typical aspects: Firstly, whether the scale is planned to be peculiar to a single writing assignment or applicable to a group of tasks and the second one is whether a single score or several scores are assigned to each writing (Weigle, 2002). Weigle (2002) notes that the idea supporting

primary trait scoring is that it is essential to recognize how well learners are able to write in a limited identified type of discourse such as explanation. Through this procedure of scoring, the rating scale is described according to a particular writing task and papers are evaluated in accordance with the level of achievement with which the learner has performed the task. Hayes et al., (2000) describe holistic scoring as the dedicating of a single mark to a draft with respect to the general impression of the text. According to Hughes (1989) , “This kind of scoring has the advantage of being very rapid” (p. 86). In a usual holistic scoring procedure, each text is examined rapidly and after that assessed according to a rating scale that plans the scoring principles. A scoring rubric differentiates between holistic scoring and the former, low reliable one, general impression marking, in which scoring standards are not clearly expressed. In analytic scoring, compositions are evaluated on numerous facets of writing instead of an assigned single score. Relevant to the intention of the assessment essays may be judged on characteristics like organization, unity, language, syntax, or mechanics. Therefore analytic scoring schemes present further comprehensive information regarding a writer's capability in various aspects of writing and are consequently better than holistic schemes in the view of a lot of writing experts (Weigle, 2002). Weigle considers the most important shortcoming of analytic scoring the longer time it takes than holistic scoring, since evaluators are required to make more than one decision for every script. An additional problem with some analytic scoring schemes is that, if scores on the different scales are combined to make a composite score, a good deal of the information provided by the analytic scale is lost.

It seems to the researcher, at least based on common sense and experience, that writing teachers of Iranian universities are least informed and familiar, if at all, with the current writing assessment approaches and methods in ESL/EFL contexts. This is in line with Hoing (1997), “Obviously, many teachers and lecturers are not aware of the fact that there is such a wide variety of assessment scenarios and applied criteria (p. 29).” In short, it can be a possible claim that the dominant trend for assessing writing in academic settings in Iran is far behind the modern ones practiced in distinguished universities throughout the world. One piece of evidence can be the frequent negative feedbacks teachers seem to receive from the students about the final exams of writing in every semester.

In fact, the kind of writing test which is carried out at Iran universities as final exams is an obvious example of traditional tests. One of the problems with the traditional essay test is that the type of writing task that the examinees are asked to do has no place in their theory or real-life contexts. Pointing to a new form for writing evaluation, Camp (1993a) argues, ‘... a further limitation becomes apparent when we match our current concepts of writing with the construct implied in and the information provided by traditional formats for writing assessment’ (p.58). Theory suggests that writing in general is a cognitively demanding task, and that improving written texts always require topical knowledge, reflection, feedback, revisions and, naturally, extended time to make them, which the traditional essay test does not allow. Therefore, the potential of process-oriented writing assessment is rarely studied empirically in EFL assessment circumstances in terms of the level of revision, scores, and complete text analysis. The present study intended to show how the notion of process-oriented writing for writing assessment is going to be operationalized in its classroom writing assessment context. To this end, the following research questions were formulated:

Research Questions

1. Does the evaluation of the writing ability of the same examinees remain the same between the first drafts, which are similar to the traditional timed impromptu tests, and the final drafts which are turned in after going through the processes of discussion, feedback, reflection and revision?

2. How do the test takers perceive the new type of writing assessment through teacher and peer feedback?

Method

Participants

A group of 60 EFL students were selected. They were all Persian native speakers aged between 22 and 25 who English Teaching majors were taking Advanced Writing course at Azad University. Since the purpose of the study was to make a comparison between the same students' first and final drafts to detect any change in the quality of written texts (essays) as a result of administering the new assessment method, the participants all took a TOEFL test so that a homogeneous group including 30 participants was finally selected. In order to make sure that the members of the group are homogeneous in terms of language proficiency in general, the researcher applied a one-way ANOVA to the students' scores on the TOEFL test.

Instrumentation

The following instruments were employed in this study:

- *The standard proficiency test (TOEFL)
- * A video related to the topic the students are supposed to write on
- *A reading text on the same topic of essay that the students were to develop.
- *Guided questions which are developed by the researcher to enhance the process of feedback and discussion.
- *Guidelines to prepare participants for peer response
- *A questionnaire on the impact of the new writing assessment model on the performance of the test takers.

Procedures

In order to select a homogeneous group of 30 out of 60 male and female students taking advanced writing course at Islamic Azad University, a grammar and vocabulary part of a TOEFL was administered. At the first phase of the study the test takers were given a topic on smoking to write on within a time limit of 35 minutes. This step was exactly the same as timed impromptu essays which are currently in use at Iran universities for assessing students' writing ability. Then the participants tried the new enhanced writing assessment model. Two notable features of this kind of enhanced classroom writing assessment are: (a) *extended time* for writing, which is one of the characteristics of real-world type of writing, and (b) refined facilitative activities such as brainstorming, discussion, the provision of topical knowledge, feedback, and revision. In this phase of the study the teacher presented and defined the topic. Next the students watched a video on the topic in order to provide them with topical knowledge and make them ready to give feedback and comments on their friends' essays and have interaction with teacher and classmates. They also read a text on the same topic of essay that the students were to develop and in pair discussed the content, grammar, organization and style of the text.

In groups of three or four, the students brainstormed answers to guided questions proposed by the researcher about the topic. Later the researcher formed groups of three, and the students took turns reading each essay and writing comments and suggestion on the peer review sheets. To familiarize students with peer review, a class activity where the teacher asked the students to respond to a sample paragraph based on a peer review sheet was introduced to the students. Then they talked about each other composition and responses. When the students had peer feedback activities, the teacher offered suggestions and they had

interactive discussions. After this activity the teacher presented oral feedback to the class. Next he gave written feedback on their compositions and their peers' comments.

Finally the students wrote and revised their final essays based on the topical knowledge, group discussion, peer comments and teacher feedback. After collecting the papers the questionnaire which was adapted from the one used in Tisu&Ng's research (2000, pp.154-156) research on the impact of the new enhanced writing assessment model on the performance of the test takers were distributed. For the present study, use was made of Analytic Scale for Rating Composition Tasks developed by Brown and Bailey (1984 cited in Brown, 2004) and the essays written in the first and second phases of the study were scored based on this rating scale by the researcher and a second trained rater (the total score =50). Then the pair-wise correlations was calculated and showed that inter-rater reliabilities between the researcher and second rater, $r = .93$, which indicates that the level of inter-rater reliabilities is high.

Results

Table1 indicates the descriptive statistics for comparing traditional and alternative assessments. It shows that the mean scores of participants in alternative assessment are higher than their mean scores in traditional one.

Table 1. Descriptive Statistics for Comparing Traditional and Alternative Assessment

Group	N	Mean	SD	SEM
Traditional	30	34.25	4.449	.812
Alternative	30	39.95	3.241	.592

Table 2 displays the results of matched-paired t-test for comparing traditional and alternative assessments. This table represents that the difference between learners' performance in two types of performance is significant.

Table 2. The Results of Matched-Paired T-test for Comparing Traditional and Alternative

T	Df	Sig.	Mean Difference
-6.230	29	.000	-5.70

The results of the analysis of the subcomponents of the writing skill including content, organization, grammar, style and mechanics for two types of assessments are given in the following tables.

Table 3. Descriptive Statistics for Content

Group	N	Mean	SD	SEM
Traditional	30	7.03	1.033	.189
Alternative	30	8.47	.615	.112

Table 4. The Results of Matched-Paired T-test for Content

T	Df	Sig.	Mean Difference
-6.696	29	.000	-1.43

Table 5. Descriptive Statistics for Organization

Group	N	Mean	SD	SEM
Traditional	30	6.08	1.076	.196
Alternative	30	7.65	.873	.159

Table 6. The Results of Matched-Paired T-test for Organization

t	Df	Sig.	Mean Difference
-5.568	29	.000	-1.57

Table 7. Descriptive Statistics for Style

Group	N	Mean	SD	SEM
Traditional	30	6.37	1.058	.193
Alternative	30	8.70	.702	.128

Table 8. The Results of Matched-Paired T-test for Style

t	Df	Sig.	Mean Difference
-11.140	29	.000	-2.33

According to the Tables 3, 5 and 7 the participants had higher means in the area of content, organization and style in alternative assessment. Also Tables 4, 6 and 8 indicate that the differences are significant.

Table 9. Descriptive Statistics for Grammar

Group	N	Mean	SD	SEM
Traditional	30	7.92	1.091	.199
Alternative	30	8.12	1.031	.188

Table 10. The Results of Matched-Paired T-test for Grammar

t	Df	Sig.	Mean Difference
-.886	29	.383	-.20

Table 11. Descriptive Statistics for Mechanics

Group	N	Mean	SD	SEM
Traditional	30	6.83	1.109	.203
Alternative	30	6.98	.782	.143

Table 12. The Results of Matched-Paired T-test for Mechanics

t	Df	Sig.	Mean Difference
-.699	29	.490	-.15

However Tables 9 and 11 display that in the subcomponents of grammar and mechanics in two types of assessments the students had approximately the same mean. Also Tables 10 and 12 indicate that the differences are not significant.

Table 13. Descriptive Statistics for Type of Feedback

Feedback	N	Mean	SD	Min	Max
Teacher Comment	30	20.93	1.230	19	23
Peer writing	30	18.30	1.745	15	21
Writing revision	30	19.23	2.063	15	23
General View	30	18.07	2.067	13	21
Peer comment	30	13.97	1.564	12	17

Table 13 is the results of the analysis of the student reactions to different sections of questionnaire. This questionnaire contains 30 items and is divided in to 5 parts. 6 items is about students 'general view of teacher comments,7 items about students 'general view of writing revision, 6 items about students 'general view of reading peer writing, 6 items about students 'general view of reading peer comments and finally 5 items is about students 'general view of alternative assessments. The means cores of the subcomponents of the questionnaire show that in general students had a positive attitude toward new type of assessment. To support the results, a one-way ANOVA was run to cross compare the different sections of the questionnaire. The results are represented in Table14.

Table 14. Results of the One-way ANOVA for Types of Feedback

Source	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	793.133	4	198.283	63.839	.000
Within Groups	450.367	145	3.106		
Total	1243.500	149			

As the table displays, the difference between the means is significant.

To trace the significant difference among the parts of the questionnaire, a post-hoc Scheffe was run and its results are represented in table 15 below.

Table 15. The Results of the Scheffe Post hoc Test

	Sections	Mean Difference	Sig.
Teacher comment	Peer writing	2.63*	.000
	Writing revision	1.70*	.009
	General view	2.87*	.000
	Peer comment	6.97*	.000
Reading peer writing	Teacher comment	-2.63*	.000
	Writing revision	-.93	.383
	General view	.23	.992
	peer comment	4.33*	.000
Writing revision	Teacher comment	-1.70*	.009
	Peer writing	.93	.383
	General view	1.17	.167
	Peer comment	5.27*	.000

General view	Teacher comment	-2.87*	.000
	Peer writing	-.23	.992
	Writing revision	-1.17	.167
	Peer comment	4.10*	.000
Peer comment	Teacher comment	-6.97*	.000
	Peer writing	-4.33*	.000
	Writing revision	-5.27*	.000
	General view	-4.10*	.000

A comparison between opinion about teacher comments and other subcomponents of the questionnaire shows that it had significant difference with all of them, which means the students had the most positive attitudes toward teacher comments. By comparing reading peers' writing view with other subcomponents of the questionnaire it turned out that it had only significant difference with teacher comments and peer comments. In fact, students had less positive attitudes toward reading peers' writing than teacher comments and more positive attitudes than peer comments. The results of comparing attitude about writing revision and other parts, show that that it had only significant difference with teacher comment. Table 15 represents that students' general view of alternative assessments had a significant difference with teacher comments and peer comments, that is, they had more positive attitude toward general view of alternative assessments than peer comments. Finally a comparison between peer comments and other subcomponents revealed that it had a significant difference with the rest subparts which means the participants had the least positive attitude toward peer comments than other items of the questionnaire.

Discussion

Concerning the first phase of the study, the results indicate that participants generally improved in their writing ability in alternative assessment especially in the area of content, organization and style. Providing learners with topical knowledge and a reading text similar to their essay topic can be considered as a crucial factor to have a better performance in terms of content in their second essays. Reading their peers' compositions, teacher feedback and peer feedback can give them a hint to change their organization and style, although, it cannot be claimed that this improvement is permanent and requires a delayed test to check whether it is permanent or temporary. The results of the present study is in the contrast with the finding of Hirose (2008) who employed peer feedback in essay writing with university students which they didn't have improvement in their writing. However in the present study in addition to peer feedback the students had teacher comments and perhaps the students benefit more from teacher feedback as the main source of knowledge in the classroom than their classmates. In terms of grammar and mechanics the participants didn't have a significant improvement in their second writing. It seems if this process is applied in writing classes for a long period of time the result will be different because they need more time to improve all areas of their writing. According to Liu (2008) through teacher direct correction and indirect correction, by indicating that there is an error without correction, students were able to self-edit their texts. However direct correction improved writers' errors in the immediate draft, it was not permanent in the next situations. Indirect feedback helped them decrease some morphological errors. Therefore offering corrective feedback on writers' compositions is not a suitable procedure to improve students' accuracy or grammatical forms in writing.

Regarding the second phase of the study, the findings indicated that all the participants in general had positive attitudes toward alternative assessment. In fact almost all university

students experience traditional assessments or a sort of product oriented assessment in their writing classes and they suffer from writing under the pressure of time and lack of sufficient content knowledge about the topic. Accordingly they advocate having more time, interaction and negotiation with their instructors and classmates. The result of the present study is in line with Cohen (1994) who claimed that Taiwanese learners had a strong positive attitude in alternative assessments that in other writing classes they asked for the same activities. The participants in this study also had the maximum positive attitudes regarding teacher comments and the least about peer comments. One interpretation is that the students consider teachers as the main source of knowledge and power in the educational settings and they believe that teachers are infallible in terms of accuracy, understanding and information. Therefore they accept their teachers' feedback undoubtedly and it will have influence on their subsequent writing. Also students may consider giving comments on their classmate writing a difficult task, because they are not confident enough to evaluate their classmates' essays and they do not accept the peers' opinions unquestionably because they underestimate their general proficiency. According to (Min, 2003; Nelson & Murphy, 1993) peer suggestions were generally considered as casual, disappointing assessment, ineffective and superficial. The result of the study is in line with the findings of Sengupta (1998) as well. She referred to the classic view that learners have regarding the role of teacher and students in educational setting. She added that learners' imagination of teacher is the supreme authority of the classroom, therefore they believe that peer assessment and group learning is worthless and time consuming. Another important finding of this study is the students' agreement on writing revision. The explanation is that by providing topical information and given sufficient time the condition is no more stressful and the learners are more confident in putting their ideas on the paper.

Conclusions

As the results in the previous section revealed, through two different types of writing assessments and a questionnaire and the results of data analysis the researcher has been able to conclude that in general most EFL learners in the present study had better performance in alternative assessment mainly on the components of content, organization and style in comparison to traditional one. However there was no significant difference in their performance regarding grammar and mechanics component. Furthermore the majority of the participants had a positive view toward feedback exchange and interaction in contribution to their writing improvement although they considered the impact of peers' comments the least, comparing to teacher feedback, revision, and their general view and reading peers' essays, on their final drafts. Students' attitude toward the comments of teacher in this study was higher than other subcomponents of the questionnaire.

Pedagogical implications

The conclusions drawn from the study have implications for EFL teachers and learners. The outcome of the present paper indicates that overall students prefer an alternative assessment which is a sort of process writing in contrast to the traditional product oriented one. Teachers should consider that the essential steps in alternative assessment for instance, topical knowledge, revision, feedback and discussion, are the requirements of integration of all skills and a step toward communicative language learning and teaching. Therefore through this procedure learners can be encouraged to expand their four skills in writing courses. In this new method of writing learners also feel a new role and their participation in different activities makes them more active, responsible, motivated and interested in writing which was a complicated and tedious task for them. In addition learners should be aware of the fact that learning is the product of involvement, cooperation, feedback and interaction. In short, in

writing classes assessments must be justifiable, based on recent theories and methods, highly dependable and credible.

References

- Bachman, L. F., & Palmer, A. S. (1996). *Language testing practice*. Oxford: Oxford University Press.
- Bereiter, C. & Scardamalia, M. (1987). *The psychology of written composition*. Hillsdale, NJ: Erlbaum.
- Berninger, V., Whitaker, D., Feng, Y., Swanson, H.L., & Abbott, R.D. (1996). Assessment of planning, translating, and revising in junior high writers. *Journal of School Psychology*, 34, 23-52.
- Bosher, S. (1998). The composing processes of three southeast Asian writers at the post-secondary level: an exploratory study. *Journal of Second Writing*, 7 (2), 205-33.
- Brown, H. D. (2004). *Language assessment: Principles and classroom practices*. New York: Longman.
- Camp, R. (1993). The place of portfolios in our changing views of writing assessment. In R. E. Bennett, & W.C. Ward (eds.). *Construction versus choice in cognitive measurement: Issues in constructed response, performance testing, and portfolio assessment* (pp. 183-212). Hillsdale, NJ: Lawrence Erlbaum.
- Cohen, A. D. (1994). *Assessing language ability in the classroom*. (2nd ed.). Boston, MA: Heinle and Heinle.
- Elbow, P. (1998). *Writing with power: techniques for mastering the writing process*. New York and Oxford: Oxford University Press.
- Emig, J. (1971). *The composing process of twelfth graders*. Urbana, IL: National Council of Teachers of English.
- Fayol, M. (1991). From sentence production to text production: Investigating fundamental processes. *European Journal of Psychology of Education*, 6(2), 101-119.
- Flower, L., & Hayes, J. (1981) A cognitive process theory of writing. *College Composition and Communication*, 32(1), 365-87.
- Flower, L. (1989). Cognition, context and theory building. *College Composition and Communication*, 40, 282-311.
- Hamp-Lyons, L. (1990). Second language writing: Assessment issues. In B. Kroll (ed.), *Second language writing: Research insights for the classroom*. New York: Cambridge University Press.
- Hayes, J. R. (1996). A new framework for understanding cognition and affect in writing. In C. M. Levy and Ransdell (eds.), *The science of writing*. NJ: Lawrence Erlbaum Associates.
- Hoing, H. (1997). Positions, power and practice: Functionalist approaches and translation quality assessment, in C. Schaffner (Ed.), *Translation and quality* (pp. 6-34).
- Hughes, A. (1989). *Testing for language teachers*. Cambridge: Cambridge University Press.
- Hyland, K. (2002). *Teaching and researching writing*. Hong Kong: Longman.
- Hirose, K. (2008). University students' perceptions of peer feedback: An exploratory study. *Chubuchiku eigokyouikugakkai* (The Chubu English Language Education Society) *Kiyo*, 37, 299-306
- Liu, Y. (2007). The Effects of Error Feedback in Second Language Writing. *Arizona Working Papers in SLA & Teaching*, Vol. 15, 65-79
- Min, H. T. (2003). Why peer comments fail. *English Teaching and Learning*, 27, 85-

- Mcnamara, T. F. (1996). *Measuring second language performance*. London and New York: Longman
- Messick, S. (1994). The interplay of evidence and consequences in the validation of performance assessments. *Educational Researcher*, 23 (2), 13-23.
- Moffett, J. (1982). Writing, inner speech and mediation. *College English*, 44, 231-44.
- Murray, D. (1985). *A writer teaches writing*, 2nd edn. Boston, MA: Houghton Mifflin.
- Nelson, G. L, & Murphy, J. M. (1993). Peer response groups: Do L2 writers use peer comments in writing their drafts? *TESOL Quarterly*, 27, 135-142.
- Nystrand, M. (1987). The role of context in written communication. In R.Horowitz, and S. J. Samuels (eds), *Comprehending oral and written language* (pp. 197-214). San Diego, CA: Academic Press.
- Rames, A. (1987). Language proficiency, writing ability and composing strategies. *Language Learning*, 37, 439-68.
- Sengupta, S. (1998). Peer-evaluation: 'I am not the teacher'. *EFL Journal*, 52/ pp.19-26
- Susser, B. (1994). Process approaches in ESL/EFL writing instruction. *Journal of Second Language Writing*, 3, 31-47.
- Wallace, M.J. (1998). *Action research for language teachers*. Cambridge: Cambridge University Press.
- Weigle, S. C. (2002). *Assessing Writing*. Cambridge: Cambridge University Press.
- Zamel, V. (1983). The composing process of advanced ESL students: six-case-studies. *TESOL QUARTERLY*, 17(2), 165-87.