The Impact of Portfolio Assessment on the Accuracy and Complexity of TEFL Postgraduate Students' Writing

Zohreh Seifoori[1]

Abstract

This quasi-experimental study examined the impact of portfolio assessment (PA) on the accuracy and complexity of postgraduate TEFL students' writing. It was hypothesized that engaging learners in the process of assessing their performance can focus their attention on formal and textual features and thereby promote their writing. 40 intermediate TEFL postgraduate students, recruited from a population of fifty taking the "Advanced Writing Course" at Islamic Azad University, Tabriz Branch, participated in the study and were randomly assigned as the control and experimental groups after their homogeneity was assessed based on a Preliminary English Test (PET) and a writing test. Both groups received process-oriented instruction based on the same teaching materials for 12 sessions and underwent a process-oriented teaching approach complemented by interactive feedback, peer-editing, and teacher's feedback on their writing samples. The experimental group was additionally engaged in weekly PA of their wiring. The paired samples t-test analysis of their writing post-test revealed that the experimental group surpassed the control group and produced more accurate and complex texts. The results accentuate the significance of engaging postgraduate TEFL students not only in the learning process but also in the process of evaluating their own progress over time and offer pedagogical implications.

Key Words: accuracy, complexity, interactive feedback, portfolio assessment, writing

۱- Department of English Language, Tabriz Branch, Islamic Azad University, Tabriz, Iran

تأثیر ارزشیابی پوشه‌ایی بر صحت دستوری و پیچیدگی نوشتاری دانشجویان کارشناسی‌ارشد آموزش زبان‌انگلیسی

چکیده

تحقیق شبه تجربی حاضر به بررسی تأثیر ارزشیابی پوشه ایی بر صحت دستوری و پیچیدگی نوشتاری دانشجویان کارشناسی ارشد آموزش زبان پرداخته است و بر این فرضیه تحقیق استوار می‌باشد که شرکت در فرآیند ارزشیابی عملکرد خود توجه زبان آموزان را نسبت به ویژگی‌های صوری و متنی کلام نوشتاری جلب نموده و موجب پیشرفت آنها می‌شود. آزمودنی‌های حاضر در این پژوهش شامل ۴۰ دانشجوی مقطع کارشناسی‌ارشد آموزش زبان انگلیسی از جمعیت پنجاه نفری دانشجویانی انتخاب شدند که درس نگارش پیشرفته را اخذ نموده و پس از تأیید همگنی در مهارت‌های درک مطلب و نوشتار به صورت تصادفی به دو گروه تحقیق و شاهد تعیین شدند. هر دو گروه به مدت ۱۲ جلسه بر اساس مطالب آموزشی یکسان و روش تدریس فرآیند-محور توام با بازخورد تعاملی، خود ویراستاری و بازخورد معلم آموزش دریافت نمودند. گروه تحقیق به صورت هفتگی در ارزشیابی پوشه‌ایی تکالیف نوشتاری خود نیز شرکت نمودند. تحلیل آماری داده‌های پس‌آزمون نشانگر پیشرفت معنی‌دار گروه تحقیق در صحت دستوری و پیچیدگی نوشتاری بود. نتایج پژوهش ضرورت شرکت دادن دانشجویان کارشناسی‌ارشد آموزش انگلیسی در فرآیند آموزش و ارزشیابی پیشرفت آموزشی خود را مورد تأکید قرار می‌دهد و کاربردهای آموزشی بسیاری دارد.

واژگان کلیدی: ارزشیابی پوشه‌ایی، بازخورد تعاملی، پیچیدگی، صحت دستوری، نوشتار

1. Introduction

Writing is a complex communicative process of converting thoughts to written language which usually starts off as a chance to practice vocabulary and grammar in pursuit of linguistic accuracy, but soon evolves into an opportunity to develop organizational skills with a focus on meaning. The complexity involved in forming original thoughts and organizing them, on the one hand, and the grammatical knowledge and organizational skill required for coherently transforming those thoughts to written language, on the other, has made learning of writing as an academic communicative skill a daunting task for many language learners in ESL and EFL contexts. Many language teachers and learners find it difficult to advance in writing by leaps and bounds partly owing to the inadequacy of the prerequisite sub-skills required for writing, as suggested by Wolf (2000, as cited in Topuz, 2004). Actually, teachers in EFL contexts hardly accomplish the objective of addressing multiple features of writing and thus highlight the linguistic component. For the same reason, EFL learners mostly deal with writing as a

support technique to reinforce handwriting, grammar and vocabulary, and rarely succeed in developing a communicative command of this productive skill in its own right.

The focus on formal features of the language and overlook of meaning and compositional skills in writing and assessment of writing characterized the traditional product-oriented approach that dominated the under-graduate, graduate, and post-graduate levels of many educational systems. However, the pedagogical pendulum started to sway gently and gradually away from the product to the process of constructing meaning and culminated in a more balanced learner-oriented, if not learner-centered, pedagogy even in EFL contexts where the paramount significance of engaging learners in the process of writing has now been widely acknowledged. The reformist process-oriented approach is totally in tandem with Progressive Educational Philosophy (PEP) (Clarke, 1987) and purports to extricate learners from the tyranny of teachers by teaching them how to manage the process of writing and how to evaluate their performance.

The new trend called into question the static nature of evaluative techniques as well. Alternative assessment techniques (AAT) like self and peer-assessment, journal writing, and portfolio assessment (PA) were thus introduced to realign the traditional summative teacher-led uni-directional evaluation (Belanoff & Dickson, 1991). A portfolio has been defined as a purposeful collection of students' work reflecting their gradual progress, weaknesses, and achievements in various areas (Genesee & Upshur, 1996). Based on the type of collected data, PA might be used to serve a wide range of pedagogical purposes in second language learning contexts of all kinds (Shohamy & Walton, 1992). According to Genesee and Upshur (1996), such analysis and evaluation might focus on some specific aspect of the language, such as writing, or have a broad focus that includes examples of all aspects of language development. Yet, in EFL contexts this technique, as proposed by Mousavi (1999), represents an approach most widely employed to evaluation of writing where teachers provide suggestions on how to revise the product and comment on the individual's progress in writing.

Unlike the traditional evaluation techniques like tests which offered a unidirectional flow of feedback from the teacher to the learner, PA, as postulated by Song and August (2002), enables learners to participate in the evaluation of their own progress over time and take responsibility for their own learning. Of course, the interactive nature of PA escalates the complexity and dynamicity of the evaluation process which, as posed by Karrol (1998, cited in Reid, 2000), originates from the difficulty and complexity of theoretically defining writing in the first place and the

challenge of simultaneously controlling various factors involved in writing and assessing. Through this dynamic process student learning is monitored while students themselves are involved in making decisions about the degree to which their performance matches their ability. For this very reason, assessment can be regarded as a formative attempt rather than summative (Bachmann, 1990), and as postulated by Spolsky (1992), it can rightly be regarded as curriculum-driven as any other formative assessment because it shadows the curriculum and provides feedback to students and teachers. Coombe, Folse, and Hubley (2007) highlighted the compatibility of PA with process-oriented approach and its capability to engage the learners in self-assessment of multiple pieces of writing produced over time. Likewise, O'Malley and Valdez Pierce (1996) advocated this ongoing self-evaluation of writing samples as a motivating technique that encourages learners to reflect on their work.

1.   Review of the Related Literature

Empirical studies have addressed various aspects of the teaching and learning processes in EFL/ESL contexts in an attempt to find ways of facilitating the development of this skill for learners. One such fertile soil for research has been the use of portfolios with two major lines of research. The first line is mainly concerned with the effect of portfolio assessment on learners' achievement in writing (Qinghua, 2010; Barootchi & Keshavaraz, 2002, Elahinia, 2004; Nezakatgoo, 2011; Song & August, 2002), the second with learners' reflections, comments, and attitudes toward portfolio assessment (Starck, 1999; Kear, Coffman, McKenna, & Ambrosio, 2000; Spencer, 1999). More recently, PA has been merged with merged with individual multiple intelligences (Faravani & Atai, 2015) to promote EFL learners' reflective thinking skills. The present study is of the first type and it would be enlightening to review some empirical studies focused on the impact of portfolio assessment on language learning in some EFL contexts.

In a longitudinal study, Song and August (2002) investigated the effect of portfolio assessment on two groups of advanced ESL learners' composition writing. They found that the success rate of the PA group was higher than that of the control group not only in writing but also in the college exit exam. More recently, Qinghua (2010) conducted a quasi-experimental study to investigate the impact of PA on writing development of 34 EFL Chinese learners, aging from 18 to 2, in two sophomore English major classes of the same size, gender distribution, and writing proficiency. The findings revealed significant differences between the two groups in terms of accuracy and coherence with the PA group surpassing their counterparts in the control group.

In the Iranian context, Barootchi and Keshavaraz (2002) examined the effect of PA on 60 Iranian female high-school sophomores' achievement of course objectives and their feeling of responsibility towards monitoring their progress. The analysis of the scores obtained from the teacher-made achievement test and a satisfaction questionnaire verified the significant effect of PA. The portfolio assessment scores correlated significantly with those of the teacher-made achievement test, and high inter-rater reliability was also achieved. They, thus, suggested PA as a promising testing and teaching tool which can be deployed in conjunction with teacher-made tests to promote educational outcomes and learner engagement.

Elahinia (2004) investigated the effect of portfolio assessment on writing achievement of 34 male and female Iranian EFL learners majoring English Translation. Analysis of their scores in a final essay writing exam displayed the positive effect of PA on the participants' writing performance and attitudes. The effect of PA of writing was likewise explored by Tabatabaei and Assefi (2012) who reported significant impacts on focus, vocabulary, organization, conventions, and vocabulary of 40 male and female EFL learners at upper-intermediate level of proficiency. Still in another study, Nezakatgoo (2011) suggested that evaluating learners' work through portfolio-assessment could impact students' test performance.

Apart from the impact of PA on the writing skill of EFL learners, however, more innovative application of this alternative assessment technique uncover its wide-ranging applicability in language pedagogy and far-reaching impacts on various features of the learning process. Oakley, Pegrum, and Johnston (2014), explored the impact of Wi-Fi-based e-portfolios into a Master of Teaching program at university level in the form of personal learning environments to accentuate reflective practice. While discussing the issue from different perspectives, they proposed that the e-portfolio helped the pre-service teachers acknowledge the value of reflection and the use of a guiding structure and achieved a medium level of reflection during the first year of study.

In the context of Iran, Jafarigohar and Mortazavi (2013) examined different types of reflective individual, peer collaborative, and teacher collaborative journal writing on 60 upper intermediate female English learners' self-regulated learning as measured through Self-Regulated Learning Scale (ASRL-S). The findings revealed that journal writing contributed significantly to the self-regulatory skills of the participants and that sharing the journal with peers and teachers could boost the impact. Self-assessment was also found effective in impacting EFL learners' goal orientation (Baleghizadeh & Masoun, 2014).

In another study, Faravani and Atai (2015) merged dialogic-based PA with individual multiple intelligences to promote reflective thinking skills of EFL participants. This innovative application of PA proved to escalate the growth of higher order thinking skills of forty EFL participants and the impact was found to be greater among intellectually homogeneous participants in the experimental group.

Very few studies, to my best knowledge, have yet delved into the impact of PA on the accuracy and complexity of postgraduate TEFL students' writing. It is normally assumed that these learners have already mastered conventions of the target writing system and should concern more with conceptual content of their courses and concentrate on higher level facets like content, organization, coherence, and other features of academic writing. A cursory examination of doctoral dissertations and master theses, however, is sufficient to uncover the full scale of the problems postgraduate students still have with accuracy and complexity, which are among the basic requirements in academic writing, despite years of schooling. Their writing is highly restricted in terms of grammatical and lexical accuracy, reflects traces of L1 influence, and is far from standard natural written discourse. Most importantly, the excessive simplicity of the expressions employed hampers the exchange of ripeness and complexity of the intellectual content required at this level. Such inadequacies are evident in the written assignments produced even by advanced learners who have already passed various English grammar, vocabulary enrichment, and composition writing courses through process-oriented instruction in writing and various types of feedback. Surprisingly, they know most of the rules of writing and even spend time generating ideas and editing their first drafts but still fail to approximate native-like standards of writing.

Despite the prevalence of process-orientation in teaching writing to graduate and postgraduate students, they are rarely engaged in the evaluation process of evaluating the products they produce owing to various executive constraints. It is my conviction that although any serious treatment of the problem calls for radical improvements in the teaching conditions, it might still be viable to help postgraduate TEFL students attend to the accuracy and complexity of writing through engagement in PA. Hence, this quasi-experimental study was designed to examine the effect of involving Iranian TEFL learners in PA on the development of their writing skill and the following research question was formulated:

- **Does PA impact the accuracy and complexity of postgraduate TEFL students in a process-oriented writing course?**

2.   Method

3.1. Participants

The participants in the current quasi-experimental study included 40 male and female Iranian post-graduate TEFL students at Islamic Azad University, Tabriz Branch who were within the age range of 24-36. All participants were taking the "Advanced Writing Course" and were attending two intact classes once a week during a fifteen-session semester. Most of them had learned English as a third language with Turkish and Farsi as their first and second languages. The classes were randomly assigned as the experimental group and the control group, each including mostly female participants. Initial homogeneity of the groups was assessed using a Preliminary English Test (PET) and a writing test.

3.2. Instruments

Two instruments were used to collect the research data. First, a modified 40-item version of PET was administered to verify the homogeneity of the participants in reading comprehension (35 items) and basic grammar (15 items). In addition, a writing test was administered with the topic of "Rural Life and Urban Life" to see if the participants were homogeneous in terms of the accuracy and complexity of their writing. The same topic was assigned as the post-test and the practice effect was controlled in two ways. Firstly, I had noted the participants that the purpose of the initial writing was to delineate their entry level in writing and verify the compatibility of the syllabus and that they would not expect any kind of feedback. Further, I administered the post-test with a 12-week interval at the end of the treatment during which the participants were working on eight different writing genres and assignments. To relieve the burden on the participants and alleviate their stress, both writing tests were administered during the class time as class work.

3.3. Materials and Procedure

The course book selected as the teaching materials was "Advanced Writing" (Birjandi, Alavi & Mosalanejad, 2004) which comprises twelve chapters, four on preliminary information about paragraph writing and eight on various genres of numeration, chronology, process, description, definition, cause and effect, comparison/contrast and argumentation. The eight genres were presented in the class during the 12-session treatment and the participants were weekly asked to write compositions based on each of these genres after they were presented in class.

As righty argued by many experts in testing evaluation, a major concern

in PA is maintaining the fairness of the evaluation which might be threatened owing to the growing subjectivity on the part of the teacher. In order to minimize the inevitable impact of this factor, I followed O'Malley and Valdez Pierce (1996) suggestion and based the evaluation on a common rubric that described numerical points. This rubric was focused on four major categories of unity, organization, accuracy, and complexity each with varying levels of specification and four levels representing a scale of 1, the weakest performance, to 4 which represented near native-like performance (See Appendix). Descriptions of the levels were based on a modified version of the writing scale in Jacobs, Zingraft, Wormuth, Harfield, and Hughey (1981).

In order to direct the participants' self and peer-assessment attempts, I introduced the rubric as a vivid evaluation criterion during the first session when its content was fully and interactively discussed. One sample paragraph was also evaluated based on the rubric to help learners understand how their writings would be quantified and evaluated. They were recommended to employ it in self and peer-assessment of their writings.

During the course, the participants in both groups were first introduced to the basic concepts of the writing genre in question and followed the recognition to production order of learning based on the model paragraphs presented in the book. Their attention was drawn towards structural and organizational characteristics of each paragraph type and finally they were demanded to produce the first draft of a sample paragraph on one of the topics suggested in the book. The writing process would start in class and was to be completed and self-edited at home.

The following session, the teacher would display a sample of the students' writings on the projector screen asking all of the participants to read and edit the text in 15 minutes while the text was being evaluated and revised interactively in class. The teachers' role during this interactive feedback session was to draw the participants' attention to defective features through some general questions. It was assumed that the text along with the teachers' scaffolding questions would function as object-regulation and other-regulation in the zone of proximal development and finally lead learners to self-regulation (Lantolf, 2000).

Then, the participants in both groups were required to peer-edit one of their classmate's writings in the class leaving some comments and notes for them to observe. The next step was for the writers to subsequently revise their writings based on the feedback offered by their peers. The third drafts were collected for final teacher assessment.

In the experimental group, the participants had to revise their texts once more based on the teachers' comments and to collect in a portfolio of four samples of their selected writings, along with the three revisions for each. Meanwhile, weekly meetings were arranged with the teacher to randomly assess and negotiate recently added content of the portfolios. Owing to the wide-range of grammatical and organizational problems the participants had in writing, the assessment focus in portfolio sessions was inevitably restricted to grammatical accuracy and complexity of the teaching points peculiar to each writing genre.

No portfolio meetings were, however, arranged for the participants in the control group nor did they trace their gradual development through collecting and comparing versions of their work and evaluating various features of writing. The last session was allocated to collect the final sample of the participants' writing in the form of a long paragraph on the same topic as the pre-test.

3.4. Measures

Grammatical accuracy might be measured as the ratio of error-free terminal units (t-units) or in terms of inaccuracy as the ratio of errors per t-unit (Larsen-Freeman, 2006). A t-unit is defined as each independent utterance providing referential or pragmatic meaning (Foster & Skehan, 1999) and may be made up of one simple independent finite clause or an independent finite clause plus one or more dependent finite or non-finite clauses. In this study, accuracy was quantified as the percentage of error-free clauses in overall writing (Ellis & Yuan, 2004; Foster & Skehan, 1999; Tavakoli & Skehan, 2005; Yuan & Ellis, 2003).

Following Foster and Skehan (1999), complexity was also measured as the ratio of subordinate clauses to the overall t-units produced. These two methods supplanted the levels of the EFL composition profile to address the reliability concern posed by Song and August (2002) who underscored the increased subjectivity as a serious problem in PA. It was assumed that these quantification methods would render more objective measures of accuracy and complexity to estimate the extent to which the treatment was effective.

Moreover, two independent scorers rated the participants' pre-test and post-test writings to ensure the reliability of the accuracy and complexity measures. The inter-orator reliability of the score sets was further found to be .78 and .83 for accuracy and .86 and .91 for complexity measures of the pre-test and post-test scores, respectively.

3.   Results and Discussion
4.1. Results

The research data obtained from the pre-tests were analyzed to assess the initial homogeneity of the groups while the post-test scores were analyzed to find out the impact of PA on the dependent research variables.

3.1.1. The Pre-test

To ascertain the groups' initial homogeneity in reading, grammar, and accuracy and complexity of writing, I first estimated the descriptive statistics of the data obtained from the pre-test scores, as shown in Table 1.

Table 1. Descriptive Statistics of the Groups' PET Scores and Pre-writing Accuracy and Complexity Measures

	Groups	N	Mean	Std. Deviation	Std. Error Mean
PET	Con. G.	20	23.40	6.30	1.40
	Exp. G	20	26.15	5.16	1.15
Pre-test Accuracy	Con. G.	20	.41	.075	.01
	Exp. G	20	.39	.09	.02
Pre-test Complexity	Con. G.	20	.20	.04	.01
	Exp. G	20	.23	.05	.01

Table 1 shows difference in the groups' pre-test mean scores; hence, to verify the significance of the apparent differences, I ran three independent samples t-tests, the results of which are displayed in Table 2.

Table 2. Independent Samples t-test of the Groups' PET Scores and Pre-writing Accuracy and Complexity Measures

		Levene's Test		t-test for Equality of Means					95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
PET	Equal Variance assumed	.702	.407	-1.50	38	.139	-2.75	1.82	-6.43	.938
	Equal Variance not assumed			-1.50	36.58	.140	-2.75	1.82	-6.44	.942
Acc.	Equal Variance assumed	2.21	.145	.87	38	.388	.023	.02	-.03	.077
	Equal Variance not assumed			.87	36.22	.388	.023	.02	-.03	.078
Com.	Equal Variance assumed	.30	.585	-1.65	38	.106	-.027	.01	-.06	.006
	Equal Variance not assumed			-1.65	37.59	.106	-.027	.01	-.06	.006

As indicated in Table 2, the Levene's Test verified the equality of variances of the score sets from the two groups on the PET ($t =.407>.05$), the accuracy measures ($t =.145>.05$), and the complexity scores ($t=.585>.05$). As it is evident, the differences between the groups did not reach significance level in their reading comprehension and grammar, as measured by PET, $t = .14>.05$, in accuracy, $t =.38>.05$, nor in complexity, $t =.10>.05$.

3.1.2. The Impact of Portfolio Assessment

Having verified the initial homogeneity of the experimental and control groups, it was safe to analyze the post-test scores to answer the research question. Hence, the descriptive statistics of the groups' post-test accuracy and complexity measures were estimated to explore the impact of the PA on the accuracy and complexity of the participants' writing. Table 3 presents the results.

Table 3. Descriptive Statistics of the Groups' Post-test Accuracy and Complexity Measures

	Groups	N	Mean	Std. Deviation	Std. Error Mean
Post-test Accuracy	CG	20	.54	.22	.05
	EG	20	.74	.19	.04
Post-test Complexity	CG	20	.60	.28	.06
	EG	20	.86	.14	.03

Table 3 displays increases in the accuracy measures from .41 on the pre-test, as indicated in Table 1, to .54 in the control group (CG), and from .39 on the pre-test to .74 on the post-test in the experimental group (EG). Mere inspection is sufficient to indicate a substantial raise in the proportion of

accurate forms to the overall number of t-units produced by both groups that might be attributed to the impact of the process-oriented writing and the effectiveness of interactive feedback and learner engagement in self and peer-assessment activities. A similar pattern of growth was also detected in complexity from the means of .20 and .23 for the control and experimental groups to .60 and .86, respectively.

Yet, to verify the significance of the difference that PA could have made and to answer the research question, it was necessary to submit the research data to another set of independent samples t-test. The results are depicted in Table 4.

Table 4. Independent Samples t-test Analysis of the Groups' Post-test Accuracy and Complexity Measures

		Levene's Test		t-test for Equality of Means					95% Confidence Interval	
		F	Sig.	t	df.	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
Acc.	Equal Variance assumed	.24	.62	3.04	38	.004	.20	.06	.06	.33
	Equal Variance not assumed			3.04	37.17	.004	.20	.06	.06	.33
Com.	Equal Variance assumed	6.43	.01	-3.60	38	.001	-.25	.07	-.39	-.11
	Equal Variance not assumed			-3.60	28.30	.001	-.25	.07	-.40	-.11

Table 4 supports the equality of variances of the accuracy measures (t =.62> .05) but not of the complexity scores (t =.01 <.05). However, the difference between the control and experimental groups reached significance level with regard to accuracy (t =.004 <.05) and complexity (t =.001 <.05).

Following, Cohen (1988), I decided to calculate eta square through the following formula,

$$\text{Eta squared} = \frac{t^2}{t^2 + (N12 + N2\text{-}2)}$$

Replacing with appropriate values from the Tables 2 and 3, the estimated values were found to be 0.19 for accuracy and .025 for complexity. According to Cohen, any value of .01 would be interpreted as small effect, .06 as moderate effect, and values larger than .14 would be interpreted as large effect. Hence, the estimated values were both higher than .14

representing a large effect size. Therefore, the research question is answered positively: engaging Iranian TEFL learners in PA of their own writing enhance the accuracy and fluency of their writing.

4.2.   Discussion

The findings from the present inquiry verified superior performance of the experimental group who underwent PA compared to the control group in terms of accuracy and complexity of writing. The findings lend credence to previous studies which underscored the supremacy of PA over teacher-made tests in enhancing the achievement of course objectives (Barootchi & Keshavaraz, 2002), and in promoting writing and attitude of the participants (Elahinia, 2004). The findings are also compatible with those who reported positive effects from portfolio-based writing assessment on the accuracy and coherence (Qinghua, 2010), who detected the impact of self-assessment on EFL learners' goal-orientation (Baleghizadeh & Masoun, 2014), who suggested that evaluating learners' work through portfolio-based assessment could improve their writing on final examinations (Nezakatgoo, 2011), and who reported significant impact of PA on the growth of focus, vocabulary, organization, conventions and vocabulary, at upper-intermediate level of proficiency (Tabatabaei & Assefi, 2012).

The findings might be explicated in terms of deeply-rooted Socio-Cultural Theory (SCT) (Lantolf, 2000; Swain & Lapkin, 1998; Vygotsky, 1987) according to which higher forms of human mental activity are construed as mediated processes which hinge on collaborative exploitation of symbolic tools to lead the learner from object-regulation through other-regulation to self-regulation (Lantolf, 2000). The domain where this process can most productively take place, according to Vygotsky (1978), is the Zone of Proximal Development (ZPD) or "the child's developmental level as determined by independent problem solving and the higher level of potential development as determined through problem solving under adult guidance in collaboration with more capable peers" (p. 85).

The assembled portfolios in the case of postgraduate TEFL students represented their ZPDs and the collaborative interaction they engaged in during the portfolio sessions served to trigger interest, simplify the task and maintain pursuit of the goal which was accuracy and complexity to finally help them mark discrepancies between their writings and the target norms (Wood, Bruner, & Ross, 1976). The participants' superior performance in terms of accuracy and complexity might be substantiated with regard to the scaffolding support they received from their teacher which seems to have assisted them in appropriating the accuracy and complexity of writing and escalated their capability of regulating their own writing.

5. Conclusion

The results emerging from the current enquiry allude to two major conclusions. Firstly, to learn how to write, postgraduate TEFL students ought to actively participate in the process of learning the textual organization and structural building blocks, generating and organizing relevant ideas, transforming them into written language and editing their own and their peers' writing. Such effective cognitive engagement prior to and during the writing process and attending to the feedback they receive viably provides the requirements for taking in instructional input, noticing deficiencies in their initial attempts (Swain & Lapkin, 1998) and intriguing them into learning.

The burgeoning impact of process-oriented writing with interactive feedback might be escalated even more if the instruction is complemented by involving the learners in the portfolio assessment of their own writing. The impact might be associated, as suggested by experts (Genesee & Upshur, 1996; Shohamy & Walton, 1992) and researchers (Faravani & Atai, 2015; Jafarigohar & Mortazavi, 2013), with the sense of ownership portfolios can give the learners and the opportunities they offer for creative and critical thinking. The reflective component is assumed to provide learners the chance to reflect upon their learning, assess their strengths and weaknesses, and become more aware of their learning and writing processes and strategies.

The results of the present inquiry emphasize a number of implications for in-service teachers and professors. Since portfolios can give the learners a sense of ownership of their work and the evaluation process, it is highly recommended the teachers incorporate this process-oriented evaluative technique to personally engage learners in self-evaluation of their work. Weekly out-of-class individual or group portfolio or e-portfolio sessions might be scheduled to help students incrementally attend to the lesson objectives and draw students' attention to their points of weakness and strength. One significant factor to be controlled, however, is the size of class. PA works efficiently in small classes; hence, the institutional policy should allow small writing classes of maximally 15 if any significant improvement is to be observed in postgraduate TEFL students' writing through implementation of process-oriented evaluation techniques. Albeit such techniques might be criticized as time-consuming (Ediger, 2000; Fenwick & Parsons, 1999; Juniewicz, 2003) and costly (Ediger, 2000; Gomez, 2000) evaluation techniques, the benefits they promisefor the learners, as postulated by Fenwick and Parsons (1999), far outweigh the

expenses if adequate time and resources are allotted and teachers and learners are trained to deploy portfolios of various types,.

Moreover, the use of PA, as reflected in more recently explored spheres of education (Baleghizadeh & Masoun, 2014; Faravani & Atai, 2015; Oakley, Pegrum, & Johnston, 2014), might be extended to other areas of language learning such as creative and critical thinking through various collaborative and cooperative feedback sessions. Of course, neither higher order thinking and reflective skills nor participants' multiple intelligences were taken into consideration in the present study. Thus, replication of the study with supplementary instruments and methodology to accommodate these two variables might offer illuminating results.

References:

Bachman, L. F. (1990). *Fundamental considerations in language testing*. Oxford: Oxford University press.

Baleghizadeh, S., & Masoun, A. (2014). The effect of self-assessment on EFL learners' goal orientation. *Iranian Journal of Applied Linguistics,* 17(1), 25-48.

Barootchi, N., & Keshavarz, M. H. (2002). Assessment of achievement through portfolio assessment and teacher-made tests. *Educational Research*, 44(3), 279-288.

Belanoff, P., & Dickson, M. (1991). *Portfolios: Process and product*. Portsmouth, NH: Goynton/Cook Publishers Inc.

Clarke, J. L. (1987). *Curriculum renewal in foreign language learning*. Oxford: Oxford University Press.

Cohen, J. W. (1988). *Statistical power analysis for the behavioral science* (2nd Ed.). Hillsdale, NJ: Lawrence Erlbaum Associates.

Coombe, C., Folse, K., & Hubley, N. (2007). *Assessing English Language Learners*. Michigan: Michigan Publishing.

Ediger, M. (2000). *Assessment with portfolio and rubric use*. (ERIC Document Reproduction Service No. ED440127). Retrieved from ERIC database: http://www.eric.ed.gov/:q=Assessment with portfolio and rubric use.

Ellis, R., & Yuan, F. (2004). The effects of planning on fluency, complexity and accuracy in second language narrative writing. Studies *in Second Language Acquisition,* 26, 59-84.

Faravani, A. & Atai, M. R. (2015). Merging multiple intelligence with dialogic-based portfolio assessment to expedite Iranian EFL learners' higher order thinking skills. *The Journal of Teaching Language Skills, 6*(4), 19-44.

Foster, P., & Skehan, P. (1999). The influence of planning and focus of planning on task-based performance. *Language Teaching Research, 3,* 215-247.

Gardner, H. (1999). *Intelligence reframed: Multiple intelligences for the 21st century*. New York: Basic Books.

Genesee, F., and Upshur, J. (1996). *Classroom–based evaluation in second language education*. Cambridge: Cambridge University Press.

Gomez, E. (2000). *Assessment portfolios: Including English language learners in large-scale assessments*. Retrieved from ERIC database: http://www. eric.ed.gov/?q=Assessment portfolios%3A+IncludingEnglish language learners in large scale assessment

Jacobs, H. L., Zingraf, S. A., Wormuth, D. R., Hartfield, V. F., & Hughey, J. B. (1981). *Testing ESL composition: A practical approach*. Rowley, Mass: Newbury House.

Jafarigohar, M., & Mortazavi, M. (2013). The effects of different types of reflective journal writing on learners' self-regulated learning. *Iranian Journal of Applied Linguistics*, 16(1), 59-78.

Kear, D. J., Coffman, G. A., McKenna, M. C., & Ambrosio, A. L. (2000). Measuring attitude toward writing: A new tool for teachers. *The Reading Teacher*, 54 (1), 10-23.

Lantolf, J. (2000). Second language learning as a mediated process. *Language Teaching*, 33, 79-96.

Larsen-Freeman, D. (2006). The emergence of complexity, fluency, and accuracy in the oral and written production of five Chinese learners of English. *Applied Linguistics*, 27(4), 590-619.

Mousavi, A. (1999). *A dictionary of language testing*. Teharn: Rahnam publication.

Nezakatgoo, B. (2011). Portfolio as a viable alternative in writing assessment. *Journal of Langue Teaching and Research*. 2(4), 747-756.

O'Malley, J. M., & Valdez Pierce, L. (1996). *Authentic assessment for English language learners: Practical approaches for teachers*. Reading, MA: Addison-Wesley Publishing Company.

Paulson, F. L., Paulson, P. R., & Meyer C. A. (1991). What makes a portfolio a portfolio? *Educational Leadership*, 48 (5), 60-63.

Polio, C. G. (2001). Research methodology in second language writing research: The case of text-based studies. In T. Silva & P. K. Matsuda (Eds.), *Second language writing* (pp. 91-116). Mahwah, NJ: Lawrence Erlbaum.

Polio, C., Fleck, C., & Leder, N. (1998). If only I had more time: ESL learners' changes inlinguistic accuracy on essay revisions. *Journal of Second Language Writing*, 7, 43-68.

Phillips, D. (1996). *Longman preparation course for the TOEFL test* (2nd Ed.). New York: Longman.

Qinghua, L. (2010). The impact of portfolio-based writing assessment on EFL writing development of Chinese learners. *Chinese Journal of Applied Linguistics*, 33(2), 103-116.

Reid, J. (2000). Advanced EAP writing and curriculum design: What do we need to know? In R. Carter, & D. Nunan (Eds.), *The Cambridge guide to teaching English to speakers of other languages* (pp.28–34). Cambridge: Cambridge University Press.

Shohamy, E., & Walton, R. (1992). *Language assessment for feedback: Testing and other strategies*. Dubuque, IA: Kendall/Hunt Publishing Co.

Song, B., & August, B. (2002). Using portfolios to assess the writing of ESL students: Apowerful alternative? *Journal of Second Language Writing*, 11(1), 49-72.

Spencer, D. M. (1999). *An exploration of portfolio assessment and its influence on children'snwriting* (Unpublished master's thesis). Regina University, Canada.

Spolsky, B. (1992). Diagnostic testing revisited. In Shohamy, E. & Walton, R.A., (Eds.), *Language assessment and feedback: Testing and other strategies* (pp. 29-39).Dubuque, IA: Kendall/Hunt Publishing Co.

Starck, T. L. (1999). *Student portfolios: Impact on writing skills and attitudes* (Unpublished doctoral dissertation). Memphis University, Tennessee.

Swain, M., & Lapkin, S. (1998). Interaction and second language learning: Two adolescent French immersion students working together. *Modern Language Journal*, *82*, 320-337.

Tabatabaei, O., & Assefi, F. (2012). The effect of portfolio assessment technique on writing performance of EFL learners. *English Language Teaching*, 5(5), 138-147.

Tavakoli, P., & Skehan, P. (2005). Strategic planning, task structure and performance testing. In R. Ellis (Ed.), *Planning and task performance in a second language* (pp.239-277). Amsterdam: Benjamins.

Topuz, E. (2004). *The effects of two different goal setting processes on students' attitudes towards writing and towards a writing course* (Unpublished master's thesis). Bilkent University, Ankara, Turkey.

Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological Processes*. Cambridge, MA: Harvard University Press.

Wang, Y. H., & Liao, H. C. (2008). The application of learning portfolio assessment for students in the technological and vocational education system. *Asian EFL Journal*, 10(2), 132-154.

Wolfe-Quintero, K., Inagaki, S. & Kim, H. Y., (1998). *Second language development in writing: Measures of fluency, accuracy and complexity*. Honolulu, HI: National Foreign Language Resource Center.

Wood, D., Bruner, J., & Ross, G. (1976). The role of tutoring in problem solving. *Journal of Child Psychology and Psychiatry***,** 17, 89-100.

Yuan, F., & Ellis, R. (2003). The effects of pre-task planning and on-line planning on luency, complexity and accuracy in L2 oral production. *Applied Lingusitics,* 24(1), 1-27.

Appendix: EFL Composition Profile

Student:		Date: Composition Topic:
Features	**Score**	**Criteria**
Unity	4 3 2 1	**Very good to excellent**: All the sentences are related to one single topic and controlling idea. **Good to average**: most of the sentences support the topic and the controlling idea with very few irrelevant sentences. **Fair to poor**: A topic is evident though the controlling idea is not adequately specified and not uniformly supported by sentences. **Very poor**: the topic is not adequately presented and there is not controlling ideas and relevant major and minor support sentences.
Organization	4 3 2 1	**Excellent to very good:** ideas are clearly stated and supported coherently and cohesively in a well-organized and logical way. **Good to average:** the text is choppy, loosely organized but main ideas stand out with limited support logical but incomplete sequencing. **Fair to poor:** the text is non-fluent and ideas are disconnected or confused. It lacks logical sequencing and development. **Very poor:** the text does not communicate; it lacks organization, coherence, cohesion, and development.
Accuracy	4 3 2 1	**Excellent to very good:** accurate with very few errors of agreement, tense, number, word order, article, pronouns and prepositions. **Good to average:** effective with some errors of agreement, tense, number, word order, article, pronouns and prepositions. **Fair to poor:** major problems in simple structures related to agreement, tense, number, word order, article, pronouns and prepositions. **Very poor:** full of various grammatical problems that hamper communication and reflect no mastery of conventions.
Complexity	4 3 2 1	**Excellent to very good:** effective complex communication using nominal, adjectival, and adverbial clauses appropriately and correctly. **Good to average:** complex communication has started using nominal, adjectival or adverbial clauses though with some problems. **Fair to poor:** very few uses of correct and appropriate nominal, adjectival, and adverbial clauses. **Very poor:** far from complex communication. Very few, if any, use of complex structures.
Total Score		Comments