

The Comparative Effect of Portfolio and Summative Assessments on EFL Learners' Writing Ability, Anxiety, and Autonomy

Mania Nosratinia

Islamic Azad University, Central Tehran Branch, Iran

Farahnaz Abdi

Islamic Azad University, Central Tehran Branch, Iran

Abstract—The present study was an attempt to systematically compare the effect of portfolio and summative assessment on writing ability, anxiety, and autonomy of English as a Foreign Language (EFL) learners. The participants were 70 male and female intermediate level EFL learners, between 19 and 35 ($M_{age} = 27$), who were non-randomly selected from among 90 EFL learners through employing a piloted sample of the Preliminary English Test (PET). They were randomly assigned into two experimental groups of 35, named *portfolio assessment writing* and *summative assessment writing*. Prior to the treatment phase, the participants filled out the English versions of Foreign Language Classroom Anxiety Scale (Horwitz, Horwitz, & Cope, 1986) and Zhang and Li's (2004) Learner Autonomy Questionnaire. The portfolio group was instructed based on the Classroom Portfolio Model, by Hamp-Lyons and Condon (2000), whereas in the summative assessment group, the common traditional summative assessment approach was implemented. After the treatment phase, both experimental groups were given another writing section of the PET test and the same anxiety and autonomy questionnaires as the posttests. The analysis of the test scores using two independent-samples *t*-tests and an analysis of covariance (ANCOVA) revealed that the summative assessment group had a significantly higher post-treatment level of anxiety. Furthermore, the level of post-treatment autonomy in the portfolio assessment group was significantly higher. It was also concluded that there was a non-significant difference between the impact of portfolio and summative assessment on writing scores when controlling for the impact of pretest scores.

Index Terms—anxiety, autonomy, portfolio assessment, summative assessment, writing ability

I. INTRODUCTION

Contemporaneous with the social constructivism theory of learning which favors the active construction of language competence through a social and experiential process (Ashton-Hay, 2006; Zaker, 2016), the English Language Teaching (ELT) practice now favors learners' active and steady engagement in the process of learning (Aliweh, 2011; Zaker, 2015). Consequently, learners are now placed in the center of the learning process (Collins & O'Brien, 2003). Moreover, English as a Foreign Language (EFL) learners are expected to actively use the language with the goal of internalizing it more effectively (Nosratinia & Zaker, 2013). This new trend has influenced all aspects of ELT practice, ranging all the way from pedagogical techniques to assessment and evaluation.

Portfolio assessment, as a representative of this new trend, is a process-based approach which is believed to be highly influenced by the abovementioned developments in the ELT domain (Brown & Hudson, 1998). Portfolio assessment has been defined as a collection of evidence able to demonstrate learners' progress and achievement through providing a more inclusive picture (Plaza, Draugalis, Slack, Skrepnek, & Sauer, 2007). In fact, it is believed that through integrating assessment and instruction, portfolio assessment can provide considerable developments in language learning (Poehner, 2008). Furthermore, many have stated that such a learner-centered approach to assessment can provide significant developments in learners' mental, cognitive, and metacognitive characteristics (Aliweh, 2011). On the other hand, the summative assessment approaches rely on testing the final product of learners, while ignoring learners' role in the process of learning (Moya & Melley, 1994), providing EFL teachers with an incomplete picture of learners' Second Language (L2) development.

Systematic studies are required for confirming the advantage of portfolio assessment approach over the traditional summative approach which cannot provide the immediate and contextualized feedback helpful for teachers and students during the learning process (Kozulin & Garb, 2004). Moreover, it is of utmost importance to inspect the way portfolio assessment affects EFL learners' internal factors (e.g. anxiety, autonomy, and creativity) which are believed to play a major role in the process of learning (Fahim & Zaker, 2014). Hoping to address the abovementioned concerns, the present study attempted to systematically compare the impact of summative (traditional) and portfolio assessment on

both language skills and learners' mental qualities. However, in order to enhance the levels of validity and generalizability of the findings, the study narrowed its focus on dealing with writing ability, anxiety, and autonomy.

Among different language skills, writing is considered a major skill, representing one's general language proficiency (Onozawa, 2010). This skill is a significant requirement for EFL learners (Lightbown & Spada, 2013) as well as a "basic communication skill and a unique asset in the process of learning a second language" (Chastain, 1988, p. 244). However, mastering writing has always been a challenge for EFL learners. The existence of this challenge is not only attributed to learners' degree of linguistic ability, but it is also highly dependent on the degree of focus on the learning process and the feedback the learners are provided with during the learning process (Chih, 2008). As stated earlier, portfolio assessment has among its principles the focus on process and providing feedback (Plaza et al., 2007). Furthermore, major components of writing, i.e. generating ideas, drafting, redrafting, and editing, are not sufficiently assessed in a one-shot attempt of traditional summative assessment (Babae & Tikoduadua, 2013).

Anxiety is the other variable of concern in this study whose reaction to a portfolio assessment was to be assessed. It is believed that anxiety is a crucial factor in L2 learning success (O' Donnell, Reeve, & Smith, 2012). In the ELT domain, anxiety is defined as "the feeling of tension and apprehension specially associated with second language contexts, including speaking, listening, and writing" (MacIntyre & Gardner, 1991, p. 284). Although some studies in the ELT domain have predicted a positive association between L2 learning and anxiety (Pimsleur, Ludwig, & Morrison, 1996), numerous studies have reported a negative correlation between anxiety and L2 development (Brown, 2007; Saito & Samimy, 1996). More specifically, it has been stated that learners' inadequacy in writing skill mostly stems from anxiety (Kırmızı & Dağdeviren Kırmızı, 2015). On the other hand, considering the pedagogical concerns, it has been stated that portfolio assessment can significantly contribute to lowering EFL learners' anxiety (Barootchi & Keshavarz, 2002).

Finally, autonomy, the third dependent variable of the study, is regarded a major and influential factor in mastering language skills (Nosratinia & Zaker, 2015). As stated by Dickinson (1995), autonomy is "a situation in which the learner is totally responsible for all the decision concerned with his/her learning and the implementation of those decisions" (p. 11). According to Nosratinia and Zaker (2014), "current EFL pedagogical trends seem to primarily focus on a student-centered methodology in which learner autonomy is given a great value" (p. 1). In other words, EFL learners "are now given a meaningful role in pedagogic decision making by being treated as active and autonomous players (Nosratinia & Zaker, 2014, p. 1). Accordingly, learner autonomy is now considered a major concern when planning the language teaching programs and determining the assessment approach (Bell, 2003; Nosratinia & Zaker, 2015).

The abovementioned concerns were reflected through the following research questions:

Research Question 1: *Is there a significant difference between the impact of portfolio and summative assessment in terms of EFL learners' anxiety level?*

Research Question 2: *Is there a significant difference between the impact of portfolio and summative assessment in terms of EFL learners' autonomy level?*

Research Question 3: *Is there a significant difference between the impact of portfolio and summative assessment in terms of EFL learners' writing ability?*

II. METHOD

Participants

The participants of this study were 70 male and female EFL learners with the age range of 19 to 35 ($M_{age} = 27$) who studied EFL at intermediate level in Tehran Institute of Technology, Tehran, Iran. These participants were non-randomly selected and homogenized from among 90 EFL learners through employing The Preliminary English Test (PET). Furthermore, before administrating the PET test, a group of 30 EFL learners with the same characteristics as the target group participated in the piloting of PET test.

Besides one of the researchers, as the teacher and one of the raters, another trained rater who was an M.A. holder in TEFL with nine years of teaching experience attended the scoring of the writing sections of the PET test.

Instrumentation

In order to fulfill the purpose of the study, the following instruments were utilized:

1. The Preliminary English Test
2. The Rating Scale of PET Writing
3. The Foreign Language Classroom Anxiety Scale
4. The Autonomy Questionnaire
5. The Portfolio Assessment Model
6. The Course Textbook
7. Six Compositions
8. The ESL Composition Profile

The Preliminary English Test (PET)

In order to homogenize the participants in terms of language proficiency, the researchers administered a version of the PET test, adapted from the book *PET Practice Test* (Quintana, 2008). Due to some practicality issues, e.g. the

institutional rules, only the reading and writing sections of the test were administered. The reading section consists of five parts with 35 reading comprehension questions. The writing section consists of three parts with 8 questions. The allocated time for these two sections (i.e. reading and writing) is 1 hour and 30 minutes. Furthermore, in order to estimate participants' post-treatment writing ability, the writing section of another PET test was administered.

The Rating Scale of PET Writing

The analytic writing scale developed by Cambridge under the name of General Mark Schemes for Writing was employed to rate the writing section of the PET in this study. It includes a scale of 0-5 based on content, organization, cohesion, coherence, format, range, mechanical accuracy, word choice, dictation, and sentence structures.

The Foreign Language Classroom Anxiety Scale

The Foreign Language Classroom Anxiety Scale (FLCAS) was developed by Horwitz, Horwitz, and Cope (1986). The scale has 33 items scored on a 5-point Likert scale, ranging from *strongly agree* to *strongly disagree*. The allocated time for completion is forty minutes, and the possible range of scores is 33-165. Horwitz et al. (1986) have reported the reliability index of .93 for the scale; however, the reliability index of FLCAS in this study was estimated to be 0.70 using the Cronbach's alpha coefficient.

The Autonomy Questionnaire

The Autonomy Questionnaire was developed by Zhang and Li (2004). The questionnaire has two parts. The first part contains 11 items and second 10, totally 21 items. The first 11 Likert scale items have five options, ranging from *never* to *always*. The second part is in multiple-choice format, and the participants are expected to choose the closer answer to their beliefs and their attitudes or ideas, from A (1 point) to E (5 points). The participants are required to respond in 30 minutes, and the maximum possible score of the questionnaire is 100. The questionnaire reportedly has high content validity and high reliability (Dafei, 2007). The reliability of this instrument in this study was estimated to be 0.76 using the Cronbach's alpha coefficient.

The Portfolio Assessment Model

The Hamp-Lyons and Condon's (2000) portfolio assessment model was used in this study. It is based on the Classroom Portfolio Model and consists of three procedures: collection, selection and reflection as shown in Figure 3.1.

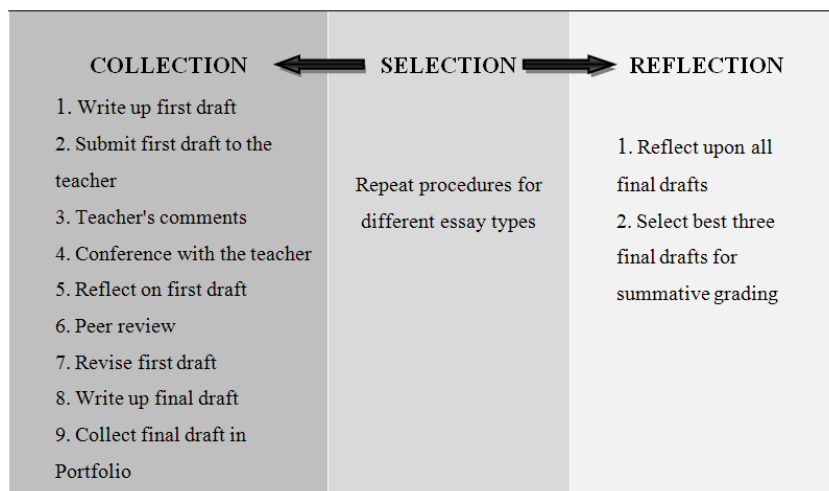


Figure 1: Classroom portfolio model designed by Hamp-Lyons and Condon (2000)

The Course Textbook

The main textbook employed in both of the experimental groups during the instruction was *English Result Intermediate* by Annie McDonald and Mark Hancock (2010), published by Oxford University Press. It consists of 12 units, each including five lessons and one review. In this study, four units of the book (units five to eight) were covered. The titles were Law and Order, Encounters, Appearance, and Communication.

Six Compositions

The participants were asked to write six compositions during the treatment sessions. They had 40 minutes to write about the following predetermined topics:

1. Describe one of the best memories of your school days.
2. Describe one of your favorite films that you have watched or the book that you have read before.
3. Describe the characteristics of a good neighbor.
4. When you go shopping what processes do you follow as your shopping habit? Describe them.
5. Where do you want to go for your holidays? Explain about your accommodation and problems that you may encounter with?
6. What is the most important decision in your lifetime and the degree of hope that you wish to achieve it.

The topics of the compositions were the same across the two experimental groups. The compositions consisted of 100 to 150 words and had to be written in descriptive voice; they needed to have three basic parts: introduction, body paragraphs, and conclusion.

The ESL Composition Profile

In order to score the six abovementioned compositions objectively, the ESL Composition Profile rating scale by Jacobs, Zinkgraf, Wormouth, Hartfiel, and Hughey (1981) was employed. It is an analytic scoring scale and consists of five subcategories of content, organization, vocabulary, language use, and mechanics. The highest possible total score is 100.

Procedure

The researchers followed certain steps, stated in a chronological order in this section. Prior to the treatment, piloting the PET test was the very first step for implementing this study. The test (reading and writing parts only) was administered to 30 non-participating candidates who had almost the same characteristics as the target group. The three characteristics of individual items (Item Facility, Item Discrimination, and Choice Distribution) were calculated, and no malfunctioning items were found. Also, the Cronbach alpha formula was employed for calculating the reliability index.

The writing part of the PET was rated according to the rating scale by the two raters who had already discussed and reviewed the rating procedure. Later on, the inter-rater reliability was calculated which indicated the existence of an acceptable consistency between the two raters. Consequently, the final writing score of each participant was the average of the two scores provided by the raters.

The already piloted PET was administered to 90 non-randomly selected male and female EFL learners. Thence, 70 EFL learners whose scores fell between one standard deviation above and below the mean were chosen as the homogeneous participants. They were randomly assigned into two experimental groups of 35, named *portfolio assessment writing* and *summative assessment writing*. Furthermore, to ensure the pre-treatment homogeneity in terms of writing ability, the scores of the writing section of the PET were analyzed. Furthermore, in the first session, the participants in two groups completed the Anxiety and Autonomy questionnaires, making it possible to check their pre-treatment anxiety and autonomy degrees.

The two groups were instructed by the same teacher, using the textbook. The teacher tried to teach the relevant grammatical points and the essential vocabularies with special focus on the writing skill. The only difference lay in using portfolios in the portfolio writing group. During the course, the participants were asked to write six compositions, and they had 40 minutes for each topic, one topic each session. The compositions had to have 100 to 150 words, written in descriptive voice. The participants were taught how to write a composition including introduction, body paragraphs, and conclusion. The compositions were rated according to the Composition Profile by the teacher. The course consisted of 12 sessions of 3 hours spanning over a period of five weeks.

The Portfolio Writing Group

In the first session, this group was provided with the explanation of the nature, purpose, and the design of the portfolio and procedures of creating a portfolio based on the Classroom Portfolio Model recommended by Hamp-Lyons and Condon (2000). The six compositions (explained in the previous section) were collected, and, using the ESL Composition Profile, the teacher provided notes and comments. In the following section, the papers were then returned to the participants, and they were expected to read the comments and consult their teacher in a one-to-one conference after the class.

The participants were asked to reflect on their own compositions and evaluate them. Also, they were asked to review their peers' writings in groups of three and to give their comments. By means of these comments, i.e., the instructor comments, peers' comments, and their own reflections, the participants gained information about their strengths and weaknesses in their writings and revised them. Afterwards, the participants wrote the final draft, and collected them in their portfolios. At the end of the term, the participants were asked to choose three of their writings for the final evaluation, and the portfolio score of the students was the average of scores on those three final drafts (Lam & Lee, 2008). It should be mentioned that, in order to encourage the participants to actively participate in the study, they were told that their writings were course requirements, impacting their final grades.

The Summative Assessment Writing Group

In this group, the traditional summative assessment was implemented. The participants were asked to write on the six given topics, similar to the other group, and after collecting the compositions, they were rated according to the ESL Composition Profile. In addition, the teacher wrote her notes and comments for each student; however, contrary to the portfolio assessment group, in this group the participants were not asked to reflect on, redraft, and revise their writings, and the evaluation of their writings was summative.

The Post-Treatment Phase

At the end of the treatment phase, all of the participants sat for the writing posttest. Simultaneously, the Anxiety questionnaire and the Autonomy questionnaire were re-administered to the participants so that it was possible to estimate the post-treatment degrees of these two constructs.

III. RESULTS

This study had one independent variable, the assessment technique, with two levels, portfolio and summative assessment. There were three dependent variables, namely writing ability, anxiety, and autonomy. Also, the language proficiency of the participants, intermediate level, was considered the control variable. In order to answer the research questions of this quasi-experimental study, certain analyses were carried out which are reported in this section.

The Pre-Treatment Analyses

Initially, the reading and writing sections of the PET were administered to the pilot group ($n = 30$), and items went through an NRT item analysis procedure, including item facility, item discrimination, and choice distribution. The results showed that all of the items, in the reading section of the test, exhibited acceptable IF, ID, and CD indices. Accordingly, all the test items were employed for the participant selection phase. Furthermore, the internal consistency of the PET in the piloting phase was estimated through using Cronbach's alpha coefficient (.792). In order to make sure that inter-rater reliability index for the raters regarding the writing section of the PET is acceptable, Pearson's product-moment correlation coefficient was run between the two sets of writing scores for the piloting administration whose results indicated a high level of inter-rater reliability between the two raters, $r = .94$, $n = 30$, $p < .01$.

The piloted PET test was administered among 90 individuals. The obtained descriptive statistics indicated that the skewness ratio value (-1.76) fell within the range of -1.96 and +1.96, supporting the normality of the distribution of the scores (Tabachnick & Fidell, 2007). Afterwards, in order to select the participants of the study, the researchers selected those individuals whose PET scores fell within the range of +1 SD and -1 SD above and below the mean score (38.2 to 53.4). Following this procedure resulted in keeping 70 individuals as the homogenous participants of the study; they were randomly assigned to two experimental groups.

Prior to administering the treatments to the two experimental groups, a pre-treatment data set was created using the scores the participants took on the three instruments of the study. Using this data set, it was attempted to make sure that the participants are homogeneous regarding their pre-treatment anxiety, autonomy, and writing ability. The obtained scores were analyzed through running three independent samples t -tests. The obtained results indicated that:

- There was no significant difference in anxiety scores for portfolio assessment ($M = 95.54$, $SD = 8.31$) and summative assessment ($M = 98$, $SD = 11.99$) groups ($t(60.5) = -.99$, $p = .32$, two-tailed). The magnitude of the differences in the means (mean difference = -2.46, 95% CI : -7.39 to 2.48) was very small (eta squared = 0.014).
- There was no significant difference in autonomy scores for portfolio assessment ($M = 65.86$, $SD = 6.95$) and summative assessment ($M = 67$, $SD = 7.95$) groups ($t(68) = -.64$, $p = .524$, two-tailed). The magnitude of the differences in the means (mean difference = -1.143, 95% CI : -4.704 to 2.418) was very small (eta squared = 0.005).
- There was a significant difference in writing scores for portfolio assessment ($M = 19.314$, $SD = 3.5729$) and summative assessment ($M = 21.486$, $SD = 2.9117$) groups ($t(68) = -2.787$, $p = .007$, two-tailed). The magnitude of the differences in the means (mean difference = -2.1714, 95% CI : -3.7261 to -.6168) was medium to large (eta squared = 0.102).

Based on the obtained results in this section, it was concluded that the participants in the two experimental groups demonstrated the same quality regarding their anxiety and autonomy levels; however, it was observed that there is a significant pre-treatment difference between the writing scores of the participants in the two experimental groups. As a result, pertinent statistical analysis, taking into account this initial difference, had to be implemented. This point will be discussed when answering the third research question.

The Post-Treatment Analyses: Checking the Assumptions

Considering the nature of the data and research questions in the present study, two parametric statistical tests were employed, i.e. independent-samples t -test and an Analysis of Covariance (ANCOVA) test. There are a number of assumptions which apply to all parametric tests. The status of these general assumptions is checked in this section; however, the test-specific assumptions are dealt with before answering the research questions in the following sections. The abovementioned general assumptions, as stated by Tabachnick and Fidell (2007), are listed and checked hereunder:

- **The dependent variable should be measured at the interval or ratio level.** Considering characteristics of the instruments used in this study (see instruments), this assumption was met.
- **Random sampling** is another assumption, favored in experimental studies. The initial 90 participants were selected non-randomly, and out of this number, 70 homogenous EFL learners were randomly assigned to the two experimental groups. Therefore, this assumption was partially met.
- **The observations should be independent.** Having the participants performing on the tests independently, this assumption was met.
- **The normality assumption** (dealt with in the following section)

Checking the Assumption of Normality

In order to inspect the normality of the data, descriptive statistics of the two experimental groups regarding their post-treatment levels of anxiety, autonomy, and writing ability were obtained. The obtained results indicated that the distribution of scores for the participants' post-treatment levels of anxiety, autonomy, and writing ability was normal as all skewness ratio and kurtosis ratio values fell within the range of -1.96 and +1.96, supporting the normality of distribution for the scores (Tabachnick & Fidell, 2007).

Answering the Research Questions

In order to accomplish the purpose of this study, three research questions were framed. Based on the design of the study and the characteristics of the variables, the researchers initially opted for running three independent samples *t*-tests. However, as the initial writing-wise homogeneity of the participants was not confirmed, the last question was to be answered through running an ANCOVA test.

The First Research Question

The first intention of the study was to systematically investigate whether there is a significant difference between the impact of portfolio and summative assessment in terms of EFL learners' anxiety level. An independent-samples *t*-test was run in order to answer this research question. As reported in Table 1, the assumption of homogeneity of variances was not met (*Levene's F* = 4.4, *p* = .04). Therefore, the values presented in the second line of the table were used.

TABLE 1:
INDEPENDENT SAMPLES *T*-TEST FOR ANXIETY POSTTEST SCORES

		Levene's Test for Equality of Variances		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
Anxiety	Equal variances assumed	4.402	.040	-5.357	68	.000	-14.171	2.646	-19.451	-8.892
Posttest	Equal variances not assumed			-5.357	60.809	.000	-14.171	2.646	-19.462	-8.881

Table 2 presents the descriptive statistics pertinent to posttest anxiety scores categorized according to the received treatments.

TABLE 2:
DESCRIPTIVE STATISTICS FOR ANXIETY POSTTEST SCORES

Treatment Type	N	Mean	SD	Std. Error Mean
Anxiety Posttest				
Portfolio Assessment	35	79.40	8.965	1.515
Summative Assessment	35	93.57	12.830	2.169

As reported in Table 1, there was a significant difference between anxiety scores for portfolio assessment ($M = 79.4$, $SD = 8.96$) and summative assessment ($M = 93.57$, $SD = 12.83$) groups ($t(61) = -5.357$, $p = .0005$, two-tailed). The magnitude of the differences in the means (mean difference = -14.171, 95% *CI*: -19.462 to -8.881) was very large (eta squared = 0.296).

The Second Research Question

The second intention of the study was to systematically investigate whether there is a significant difference between the impact of portfolio and summative assessment in terms of EFL learners' autonomy level. An independent-samples *t*-test was run in order to answer this research question. As reported in Table 3, the assumption of homogeneity of variances was met (*Levene's F* = 1.47, $p = .23$). Therefore, the values presented in the first line of the table were used.

TABLE 3:
INDEPENDENT SAMPLES *T*-TEST FOR AUTONOMY POSTTEST SCORES

		Levene's Test for Equality of Variances		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
Autonom	Equal variances assumed	1.469	.230	3.776	68	.000	7.400	1.960	3.489	11.311
y Posttest	Equal variances not assumed			3.776	65.91	.000	7.400	1.960	3.487	11.313

Table 4 presents the descriptive statistics pertinent to posttest autonomy scores categorized according to the received treatments.

TABLE 4:
DESCRIPTIVE STATISTICS FOR AUTONOMY POSTTEST SCORES

Treatment Type	N	Mean	Std. Deviation	Std. Error Mean
Autonomy Posttest				
Portfolio Assessment	35	77.03	7.434	1.257
Summative Assessment	35	69.63	8.898	1.504

As reported in Table 3, there was a significant difference between autonomy scores for portfolio assessment ($M = 77$, $SD = 7.43$) and summative assessment ($M = 69.6$, $SD = 8.9$) groups ($t(68) = 3.776$, $p = .0005$, two-tailed). The magnitude of the differences in the means (mean difference = 7.4, 95% *CI*: 3.489 to 11.311) was large (eta squared = 0.173).

The Third Research Question

The third intention of the study was to systematically investigate whether there is a significant difference between the impact of portfolio and summative assessment in terms of EFL learners' writing ability level.

Checking the Assumptions of the ANCOVA Test

Before running the ANCOVA test, it was needed to check three major assumptions. These assumptions, according to Tabachnick and Fidell (2007), are:

- Linearity
- Homogeneity of regression slopes
- Equality of variance

In order to check the assumption of linearity, the following scatterplot was created.

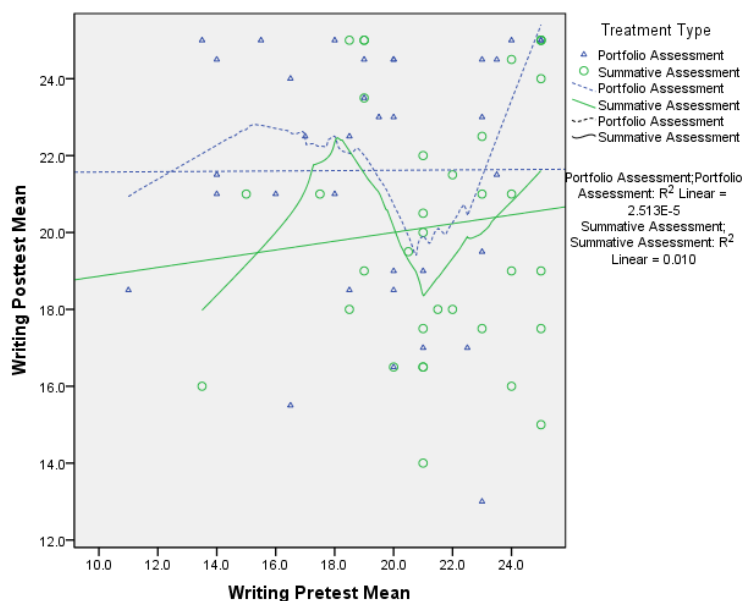


Figure 2: The ANCOVA linearity scatterplot for portfolio and summative assessment on writing

As illustrated in Figure 2, there exists a clearly linear relationship between the dependent variable and the covariates for both of the categories, and there is no indication of a curvilinear relationship. Therefore, the assumption of linearity was not violated. This procedure was followed by checking the homogeneity of regression slopes. The pertinent desired result would confirm that there is no interaction between the posttest scores and the two categories. Table 5 reports the result of the pertinent between-subjects effects.

TABLE 5:
THE TEST OF BETWEEN-SUBJECTS EFFECTS FOR THE TREATMENTS AND WRITING POSTTEST SCORES

Dependent Variable: Writing Posttest Mean					
Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	40.182 ^a	3	13.394	1.197	.318
Intercept	611.766	1	611.766	54.691	.000
Treatment	5.740	1	5.740	.513	.476
Writing Pretest Mean	2.433	1	2.433	.217	.643
Treatment * WritingPretestMean	2.070	1	2.070	.185	.668
Error	738.265	66	11.186		
Total	31334.250	70			
Corrected Total	778.446	69			

a. R Squared = .052 (Adjusted R Squared = .009)

Through inspecting Table 5, it can be noticed that the Sig. value for the interaction between treatment type and vocabulary posttests (.668) is safely above the .05 cut-off value. Therefore, the interaction is not statistically significant, indicating that the assumption is met. Finally, the assumption of equality of variance was checked through running the Levene's Test of Equality of Error Variances (Table 6).

TABLE 6:
LEVENE'S TEST OF EQUALITY OF ERROR VARIANCES FOR THE TREATMENTS AND WRITING POSTTEST SCORES^a

Dependent Variable: Writing Posttest Mean			
F	df1	df2	Sig.
.071	1	68	.790

Tests the null hypothesis that the error variance of the dependent variable is equal across groups.
a. Design: Intercept + Writing Pretest Mean + Treatment

As reported in Table 6, the assumption of equality of variance is not violated as the Sig. value is much larger than the .05 cut-off value (*Levene's F* (1, 68) = .071, $p = .79$). This indicates that the variances are desirably equal. Having the three assumptions checked and met, the researchers could legitimately opt for the ANCOVA test.

Running the ANCOVA Test

After checking the preliminary assumptions, the ANCOVA test was run in order to answer the third research question. The main results of the test are presented in Table 7. This test will indicate whether the two treatments are significantly different in terms of affecting writing (the posttest scores when controlling for the impact of pretest scores).

TABLE 7:
THE ANCOVA TEST RESULTS FOR THE TREATMENTS AND PARTICIPANTS' WRITING

Tests of Between-Subjects Effects
Dependent Variable: Writing Posttest Mean

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Corrected Model	38.112 ^a	2	19.056	1.725	.186	.049
Intercept	671.303	1	671.303	60.753	.000	.476
Writing Pretest Mean	1.679	1	1.679	.152	.698	.002
Treatment	37.615	1	37.615	3.404	.069	.048
Error	740.335	67	11.050			
Total	31334.250	70				
Corrected Total	778.446	69				

a. R Squared = .049 (Adjusted R Squared = .021)

As reported in Table 7, after adjusting the posttest scores, there was not a significant difference between the two intervention groups on the writing posttest scores (writing development), $F(1, 67) = 3.4, p = .069$, partial eta squared = .048 representing a small-to-medium effect size. This indicates that 4.8 percent of the writing posttest scores is explained by the type of treatment. It was also concluded that there is not a significant relationship between the pretest scores and the posttest scores of writing while controlling for the type of treatment. In fact, pretest scores only explain .2 percent of the variance in the posttest scores of writing. Finally, Table 8 presents the adjusted means on writing posttest scores for each treatment group. Here, the effect of the pretest scores has been statistically removed.

TABLE 8:
THE ADJUSTED MARGINAL MEANS ON WRITING POSTTEST SCORES

Treatment Type
Dependent Variable: Writing Posttest Mean

Treatment Type	Mean	Std. Error	95% Confidence Interval	
			Lower Bound	Upper Bound
Portfolio Assessment	21.667 ^a	.578	20.514	22.820
Summative Assessment	20.119 ^a	.578	18.966	21.272

a. Covariates appearing in the model are evaluated at the following values: Writing Pretest Mean = 20.400.

The main obtained values of the study are included in Figure 3 in order to get a holistic picture of the impact of the treatments on participants' anxiety, autonomy, and their writing ability.

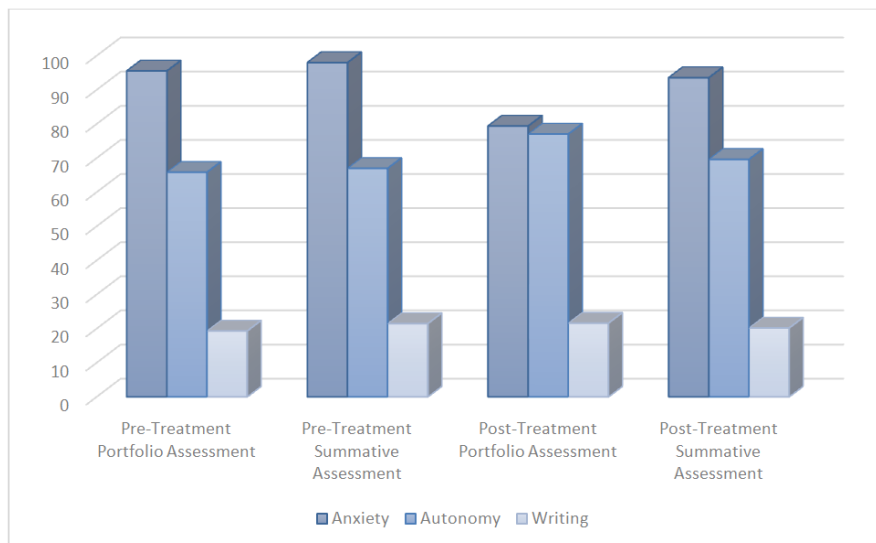


Figure 3: Pretest and posttest scores in the two experimental groups

IV. DISCUSSION

Addressing the first research question of this study through running an independent samples *t*-test revealed that the summative assessment group had a significantly higher post-treatment level of anxiety ($t(61) = -5.357, p = .0005$, two-tailed). This result was further supported by observing a very large effect size (eta squared = 0.296). Considering the pre-treatment homogeneity of participants' anxiety levels, the significantly higher level of anxiety in the summative assessment group indicates the considerable and statistically-supported role of portfolio assessment in diminishing the level of anxiety among EFL learners. This supports the argument proposed by Barootchi and Keshavarz (2002). More importantly, this finding confirms the findings of Ozturk and Cecen's (2007) experimental study which reported a significant role for employing portfolios in helping students to overcome their anxiety.

The second research question attempted to systematically compare the impact of portfolio and summative assessment on another important factor in EFL learning, autonomy (Nosratinia & Zaker, 2014, 2015). The results of an independent samples *t*-test indicated that level of post-treatment autonomy in the portfolio assessment group is significantly higher, compared to the summative assessment group ($t(68) = 3.776, p = .0005$, two-tailed). This result was further supported by the existence of a large effect size (eta squared = 0.173). Considering the pre-treatment homogeneity of participants' autonomy levels, the significantly higher level of post-treatment autonomy in the portfolio assessment group indicates the considerable and statistically-supported role of portfolio assessment in enhancing the level of autonomy among EFL learners. This result confirms the findings of Wen, Tsai, Lin, and Chuang's (2004) study which reported a significant development in learners' metacognitive awareness and autonomy as a result of implementing portfolio assessment. This finding is also in line with the findings of Koyuncu (2006) which reported an improvement in young learners' autonomy through employing portfolio assessment.

Finally, the third/last research question of the study was an attempt to systematically compare the impact of portfolio and summative assessment on EFL learners' writing ability. Comparing the impact of portfolio and summative assessments on L2 writing makes it possible to inspect the direct impact of portfolio assessment on EFL learners' language skills, compared to the first and second research questions which focused on learning-related factors. Considering the pre-treatment heterogeneity of participants' writing ability, the third research question was answered through running an ANCOVA test after checking the pertinent assumptions. Results indicated that there was not a significant difference between the impact of portfolio and summative assessment on writing scores when controlling for the impact of pretest scores ($F(1, 67) = 3.4, p = .069$, partial eta squared = .048 representing a small-to-medium effect size).

Although the results of the ANCOVA test reported a non-significant difference between the two intervention groups, inspecting the table of adjusted marginal means (Table 8) indicated that the adjusted mean on writing posttest scores for the portfolio assessment group was higher than the summative assessment group. Moreover, the partial eta squared (0.048) indicated that, despite being statistically non-significant, 4.8 percent of the writing posttest scores is explained by the type of treatment. This finding provides partial support for Fahed Al-Serhani's (2007) study which reported a significant development in EFL learners' writing as a result of employing portfolio assessment. The same level of support is also provided for Qinghua's (2010) study which reported a development in EFL learners' writing accuracy and coherence when applying portfolio assessment. Results also partially confirmed the findings of Tabatabaei and Assefi's (2010) study in an Iranian EFL context. The last point to make is that the significance or non-significance of the difference between two sets of scores is dependent on the features of the sample, e.g. sample size, mean scores, and variation (Tabachnick & Fidell, 2007). Therefore, a larger sample size could turn this non-significant difference into a significant difference.

V. CONCLUSION

This study was mainly motivated by the desire to highlight the significance of assessment in language programs and study the way the new developments in this area can facilitate and promote language learning. There is now a common consensus among language educators on the significance of the role played by assessment in teaching activities and outcomes (Kozulin & Garb, 2004). However, simultaneous with the developments in the teaching techniques and theories (Zaker, 2015, 2016), the summative assessment approach which adopts a time-bound and narrow approach to assessment has been under serious attack (Garrison & Ehringhaus, 2007). Moreover, it has been stated that the traditional summative assessment is unable to provide the immediate and contextualized feedback helpful for teachers and students during the learning process (Kozulin & Garb, 2004).

As stated earlier, there has been a global shift from traditional teacher-centered approaches to the student-centered approaches in language teaching (Nosratinia & Zaker, 2013, 2014) in which learners voice their opinions on the content, activities, materials, and pace of learning (Collins & O'Brien, 2003; Richards, 2005). As a result, new developments in the area of language assessment have been introduced. Portfolio assessment, as a representative of this new trend, has been defined as a collection of evidence able to demonstrate learners' progress and achievement through providing a more inclusive picture (Plaza et al., 2007). Despite all the arguments stated above, there seems to be a dire need to systematically study the advantages of employing these new assessment approaches, e.g. portfolio, in the area of language teaching and in different EFL contexts.

Therefore, this quasi-experimental study was designed to include anxiety, autonomy, and writing as the dependent variables of the study while comparing two groups, one receiving summative and the other one receiving portfolio assessment. The results of the study led the researchers to the conclusion that

- Portfolio assessment diminishes the level of anxiety among EFL learners.
- Portfolio assessment opens the road for EFL learners' autonomy.
- Portfolio assessment can develop EFL learners' writing.

The important point to mention is that portfolio assessment seems to have a reasonable potential for directly contributing to the process of language learning. Therefore, integrating alternative assessment techniques, especially portfolio, in the TEFL practice seems to be logical and justified attempt. The findings of the present study highlight the significance of adopting alternative and process-oriented approaches to language assessment in EFL contexts. Doing so will enable EFL teachers to enhance the level of learner involvement in the learning process and employ the assessment process as a motivating and informative tool simultaneously. In order for EFL teachers to employ portfolio assessments in the best possible way, they are recommended to:

- a) focus learners' attention on learning processes,
- b) value the revision processes,
- c) attempt to enhance learners' both intrinsic and extrinsic motivation,
- d) involve the learners in the process of teaching and assessment,
- e) establish a friendly atmosphere for a better communication in the class, and
- f) try to familiarize the learners with the advantages of active engagement in the process of learning.

As highlighted in the constructivist theory of learning, learners' active involvement in the process of learning is a must (Ashton-Hay, 2006; Zaker, 2016). Therefore, they should be cognizant of the fact that without playing an active role in an EFL classroom, no learning will take place. Moreover, they should welcome an assessment process which requires them to actively participate all through the period of learning. In an EFL classroom where portfolio assessment is employed, learners constantly receive comments and feedbacks on their performance. They should attempt to critically assess themselves and, through studying the comments, attempt to correct the errors they have made. It is also recommended to exchange the comments with other classmates so that they can engage in a collaborative process which can facilitate language learning.

Based on the findings of the present study, EFL syllabus designers are encouraged to prepare EFL materials in a way that learners are given the chance to engage in a self-assessment process at different points during the language course. EFL syllabus designers should also value the significance of cooperation and interaction in the process of learning, which are the tools for a fruitful engagement in the portfolio assessment process. EFL syllabi should be designed in a way that learners are exposed to a variety of tasks and numerous opportunities to give and receive comments as stated in the principles of portfolio assessment. Finally, EFL syllabi can explicitly familiarize the learners with the principles of portfolio assessment and the role learners are expected to play in it.

Considering the design of this study and its focus, the characteristics of the learners, and the peculiarities of the study, a limited number of recommendations are presented here, hoping that other researchers would find them interesting enough to pursue in the future.

1. It is suggested to compare the impact of portfolio assessment with other types of alternative assessment on language skills.
2. It is suggested to inspect the impact of portfolio assessment on other language skills and personal factors.
3. This study was conducted among EFL learners, between the ages of 19 and 34. The same study could be conducted among other age groups.
4. It is suggested to replicate this study in a way that the numbers of male and female participants are equal. Therefore, gender might not act as an intervening variable.
5. This study can be replicated employing some qualitative instruments to increase the validity and reliability of the results and interpretations.
6. It is suggested to replicate this study in a way that the treatment phase takes a longer time, with higher number of sessions and different types of writing tasks and activities. This way, it would be possible to inspect whether portfolio is able to significantly improve learners' writing performance.
7. This study can be replicated with participants with other levels of language proficiency.

REFERENCES

- [1] Aliweh, A. M. (2011). The Effect of electronic portfolios on promoting Egyptian EFL college students' writing competence and autonomy. *Asian EFL Journal*, 13(2), 90-132.
- [2] Ashton-Hay, S. (2006). Constructivism and powerful learning environments: Create your own! Paper presented at the 9th International English Language Teaching Convention: The Fusion of Theory and Practice, Middle Eastern Technical University, Ankara, Turkey. Retrieved May 22, 2015, from <http://eprints.qut.edu.au/17285/1/17285.pdf>.
- [3] Babae, M., & Tikoduadua, M. (2013). E-portfolios: A new trend in formative writing assessment. *International Journal of Modern Education Forum*, 2(2), 49-56.
- [4] Barootchi, N., & Keshavarz, M. H. (2002). Assessment of achievement through portfolios and teacher-made tests. *Educational Research*, 44(3), 279-288.

- [5] Bell, D. M. (2003). Method and postmethod: Are they really so incompatible? *TESOL Quarterly*, 37(2), 325-336.
- [6] Brown, H. D. (2007). Principles of language learning and teaching (5th ed.). White Plains, NY: Pearson Education Inc.
- [7] Brown, J. D., & Hudson, T. (1998). The alternatives in language assessment. *TESOL Quarterly*, 32(4), 653-675.
- [8] Chastain, K. (1998). Developing second-language skills theory and practice. Florida: Harcourt Brace Jovanovich, Inc.
- [9] Chih, H. K. (2008). Designing an online writing system: Learning with support. *ELC Journal*, 39(3), 285-299.
- [10] Collins, J. W., & O'Brien, N. P. (Eds.). (2003). *Greenwood Dictionary of Education*. Westport, CT: Greenwood.
- [11] Dafei, D. (2007). An exploration of the relationship between learner autonomy and English proficiency. *Asian EFL Journal*, 24(4), 1-23.
- [12] Fahed Al-Serhani, W. (2007). The effect of portfolio assessment on the writing performance of EFL secondary school students in Saudi Arabia (Unpublished master's thesis). Taibah University, Saudi Arabia.
- [13] Fahim, M., & Zaker, A. (2014). EFL learners' creativity and critical thinking: Are they associated? *Humanising Language Teaching*, 16(3). Retrieved October 2, 2015, from <http://www.hltag.co.uk/jun14/mart01.htm>.
- [14] Garrison, C., & Ehringhaus, M., (2007). Formative and summative assessments in the classroom. NIMSA's Annual Conference and Exhibit, November 10-12, Louisville, KY. Retrieved March 4, 2015, from <http://www.amle.org/Publications/WebExclusive/Assessment/tabid/1120/Default.aspx>.
- [15] Hamp-Lyons, L., & Condon, W. (2000). Assessing the portfolio: Principles for practice, theory and research. Cresskill, NJ: Hampton Press.
- [16] Horwitz, E. K., Horwitz, M. B., & Cope, J. (1986). Foreign language classroom anxiety. *The Modern Language Journal*, 70(2), 125-132.
- [17] Jacobs, H. L., Zinkgraf, S.A., Wormouth, D.R., Hartfiel, V. F., & Hughey, J. B. (1981). Testing ESL composition: A practical approach. Rowley, MA: Newbury House.
- [18] Kırmızı, Ö., & Dağdeviren Kırmızı, G. (2015). An investigation of L2 learners' writing self-efficacy, writing anxiety and its causes at higher education in Turkey. *International Journal of Higher Education*, 4(2), 57-66.
- [19] Koyuncu, S. (2006). The effect of the European language portfolio on learner autonomy for young learners (Unpublished master's thesis). Çukurova University, Turkey.
- [20] Kozulin, A., & Garb, E. (2004). Dynamic assessment of literacy: English as a third language. *European Journal of Psychology of Education*, 19(1), 65-77. doi:10.1007/BF03173237.
- [21] Lightbown, P., & Spada, N. (2013). How languages are learned (4th ed.). Oxford: Oxford University Press.
- [22] MacIntyre, P. D., & Gardner, R. C. (1991). Methods and results in the study of anxiety and language learning: A review of the literature. *Language Learning*, 41(1), 85-117.
- [23] Moya, S. S., & O'Malley, J. M. (1994). A portfolio assessment model for ESL. *The Journal of Educational Issues of Language Minority Students*, 13(1), 13-36.
- [24] Nosratinia, M., & Zaker, A. (2013, August). Autonomous learning and critical thinking: Inspecting the association among EFL learners. Paper presented at the First National Conference on Teaching English, Literature, and Translation, Shiraz University, Shiraz, Iran. Retrieved September 6, 2015, from http://www.civilica.com/Paper-TELT01-TELT01_226.html.
- [25] Nosratinia, M., & Zaker, A. (2014). Metacognitive attributes and liberated progress: The association among second language learners' critical thinking, creativity, and autonomy. *SAGE Open*, 4(3), 1-10. doi: 10.1177/2158244014547178.
- [26] Nosratinia, M., & Zaker, A. (2015). Boosting autonomous foreign language learning: Scrutinizing the role of creativity, critical thinking, and vocabulary learning strategies. *International Journal of Applied Linguistics and English Literature*, 4(4), 86-97. doi: 10.7575/aiac.ijael.v4n4p.86.
- [27] O'Donnell, A., Reeve, J., & Smith, J. (2012). *Educational psychology: Reflection for action* (3rd ed.). Hoboken, NJ: Wiley.
- [28] Onozawa, C. (2010). A study of the process writing approach: A suggestion for an eclectic writing approach. *Proceedings of Kyoai Gakuen College, Japan*, 10, 153-163.
- [29] Ozturk, H. & Cecen, S. (2007). The effect of portfolio keeping on writing anxiety of EFL students. *Journal of Language and Linguistic Studies*, 3, 2, 218-236.
- [30] Pimsleur, P., Ludwig, M., & Morrison, A. L. (1962). Student factors in foreign language learning. *Modern Language Journal*, 46(4), 160-170.
- [31] Plaza, C. M., Draugalis, J. R., Slack, M. K., Skrepnek, G. H. & Sauer, K. A. (2007). Use of reflective portfolios in health sciences education. *American Journal of Pharmaceutical Education*, 71(2), 34-55.
- [32] Poehner, M. E. (2008). Dynamic assessment: A Vygotskian approach to understanding and promoting L2 development. New York, NY: Springer.
- [33] 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.
- [34] Quintana, J. (2008). PET practice tests. Oxford: Oxford University Press.
- [35] Richards, J. C. (2005). Communicative language teaching today. SEAMEO Regional Language Centre.
- [36] Saito, Y., & Samimy, K. K. (1996). Foreign language anxiety and language performance: A study of learner anxiety in beginning, intermediate, and advanced-level college students of Japanese. *Foreign Language Annals*, 29(2), 239-249.
- [37] Tabachnick, B. G., & Fidell, L. S. (2007). Using multivariate statistics. Boston, MA: Pearson Education, Inc.
- [38] Tabatabaei, O., & Assefi, F. (2012). The effect of portfolio assessment technique on writing performance of EFL learners. *English Language Teaching*, 5(5), 138-147.
- [39] Wen, M. L., Tsai, C. C., Lin, H. M., & Chuang, S. C. (2004). Cognitive-metacognitive and content-technical aspects of constructivist Internet-based learning environments: A LISREL analysis. *Computers & Education*, 43(3), 237-248.
- [40] Zaker, A. (2015). EFL learners' language learning strategies and autonomous learning: Which one is a better predictor of L2 skills? *Journal of Applied Linguistics-Dubai* 1.1, 27-39.
- [41] Zaker, A. (2016). Social constructivism and metacognition in an EFL context: Inspecting the contribution of critical thinking to EFL learners' social intelligence. *Humanising Language Teaching* 18.6. Retrieved December 27, 2016, from www.hltag.co.uk/dec16/index.htm.

- [42] Zhang, L. X. & Li, X. X. (2004). A comparative study on learner autonomy between Chinese students and west European students. *Foreign Language World*, 4, 15-23.

Mania Nosratinia is Assistant Professor in TEFL at Islamic Azad University, Central Tehran Branch. She has been teaching at undergraduate and graduate levels in the areas of English language teaching methodology, language testing, and research methodology in language studies. She has published in national and international academic journals and presented in several national and international seminars.

Farahnaz Abdi holds an MA in TEFL from Islamic Azad University, Central Tehran Branch and is an English teacher at a number of language schools. Her main areas of research interest include portfolio assessment and second language writing.