

Journal of English language
Teaching and Learning
University of Tabriz
Volume 12, Issue 25, (Spring & Summer 2020)

**Exploring the Effect of Paper-and-pencil vs. Blog JW on L2
Writing in terms of Accuracy, Fluency, Lexical Complexity, and
Syntactic Complexity***

Shohreh Teimournezhad**

PhD Candidate, Department of English Language and Literature, Faculty of
Literature, Alzahra University

Elaheh Sotoudehnama* (Corresponding Author)**

Professor, Department of English Language and Literature, Faculty of Literature,
Alzahra University

S. Susan Marandi****

Associate Professor, Department of English Language and Literature, Faculty of
Literature, Alzahra University

Abstract

In the field of second/foreign language learning, Journal Writing (JW) has been used as a tool for self-expression, acknowledged for its multiple benefits with regard to various aspects of language learning. With the advancements of technology, however, few studies have addressed the potentials of internet-based platforms such as blogs in the practice of JW. The present study, therefore, has attempted to juxtapose the traditional paper-and-pencil mode of JW with Blog JW to explore the potential impacts on L2 learners' writing skill. To this end, four elements of accuracy, fluency, lexical, as well as syntactic complexity were examined in the participants' writing samples through relevant measures. To analyze the data, (M) ANCOVAs and (M) ANOVAs were conducted, the results of which are summarized and presented. The general conclusion of the study was that, JW, in particular, in the form of blogging has the potential of enhancing the quality of language learners' writing.

Keywords: Journal Writing, Blog, Fluency, Accuracy, Lexical Complexity, Syntactic Complexity.

* Received date: 2019/12/07

Accepted date: 2020/05/08

**E-mail: Sh.Teimournezhad@gmail.com

***E-mail: esotoude@alzahra.ac.ir

****E-mail: susanmarandi@alzahra.ac.ir

Introduction

Writing, as an old means of communication and self-expression (Olson, 2009), has always been regarded as an effective way to activate thinking skills (Forsman, 1985). In the arena of foreign language learning, writing, as one of the main four skills, has been proposed to enable students to communicate in the target language with ease, partly because it reduces their fear of making mistakes and being laughed at (Peterson, 1985). In accordance with this line of thought, Journal Writing (JW), as one of the approaches to writing as reflection, has served various pedagogical purposes such as enhancing self-reflection skills and control over the process of writing (Brown, 2004), and improving other language skills (El-Koumy, 1998; Hemmati & Soltanpour, 2012).

On the other hand, technology, as an indispensable part of life, plays an undeniable role in every aspect; thus, many researchers have attempted to tackle the role of technology in educational contexts (e.g. Palalas, 2011; Parvin & Salam, 2015). In this respect, with the emergence and popularity of blogs, many have turned to them as a medium for sharing their personal concerns, as well as using them in educational contexts.

In a similar vein, the present study has attempted to investigate the effect of JW, but this time with a comparative perspective juxtaposing its two forms, i.e. the traditional paper-and-pencil mode vs. JW via blogs, on the participants' writing skills in terms of accuracy, fluency, lexical complexity as well as syntactic complexity.

Literature Review

The benefits of JW have been frequently referred to in the literature, including the chance it provides for self-reflection about learning experiences, or developing creativity, self-expression and personal growth (Brown, 2004; Castellanos, 2008; Cooper, 2006; Marefat, 2002). Writing journals is also believed to enhance meaningful learning in educational contexts (Dewey, 1938), for it can engage students in the learning process in their own way and helps them focus on their personal experiences. Yinger and Clark (1981) argue that writing

reflective journals can be recognized as a powerful learning tool through which one can express ideas in various modes of representation, while the opportunity for reflection and receiving feedback can also facilitate learning. In addition, the free-writing nature of JW, Castellanos (2008) states, can make it a proper alternative for purposes such as problem-posing and problem-solving, reflecting on important lessons, and monitoring the students' classroom performance. Castellanos also refers to the potential of JW in providing the students the chance to write expressively about their preoccupations, without being concerned with being evaluated.

On the other hand, blogs, as an offspring of the Internet, have been gaining great popularity as a personal way of expressing one's mind online since the late 1990s (Armstrong & Retterer, 2008). Zhang (2009) refers to the features of blogs such as the ease of use and access, the security and privacy options available to the users, and the equality of users, all of which make blogs a viable tool to be used in educational fields as a medium for student-teacher communication and provision of learning materials, facilitating collaboration and professional development among instructors.

Many studies have explored journal writing and its benefits. Liao and Wong (2007), for instance, concludes that dialogue JW can improve writing abilities, motivation and reflective skills, while reducing anxiety as well. In an academic writing course for EFL students, for instance, Cahyono (1997) investigates the effect of JW on the students' ability to write English essays. Despite the non-significant results, the students were found to show a positive tendency towards JW and to appreciate the freedom of writing, motivation, and teacher-student communication it provides. Similarly, Gholami Mehrdad (2008) provide evidence for the positive effects of JW on students' writing skills, suggesting that students can reflect on their own progress and gain better insights into their own learning process via JW. Elsewhere, engagement in JW has been reported to enhance writing fluency (VanderMolen, 2011) and students' self-awareness and confidence (Puengpipattrakul, 2009).

With respect to the forms of JW, Bouldin, Holmes, and Fortenberry (2006) highlight the advantages of blog journals over the hardcopy version, such as provision of the exact time and date of each journal entry, which can in turn prevent the students from last-minute assignment submissions; the ease of access to a large number of files; and the spelling-check option. Furthermore, in a literature review regarding the use of blogs in ELT contexts, in particular in writing courses, Zhang (2009) acknowledges that if used properly, blogs can facilitate critical thinking and meaningful learning, provide learning models for students, affect the quality of their writing, and motivate them for purposeful writing. Which aspects of writing are improved, however, were not specified by the author. In addition, using blogs as a means of online diary writing has been also found to increase writing fluency and the tendency to use more low-frequency words as a sign of lexical complexity (Fellner & Apple, 2006). Similarly, blogging has been reported to enhance grammatical accuracy and vocabulary use (Rahmany, Sadeghi, & Faramarzi, 2013).

Although there are numerous studies on the merits of JW or blogs in EFL contexts, none has adopted a comprehensive look at all aspects of writing skill. In addition, no studies have adopted a comparative approach toward traditional vs. modern forms of JW to examine whether or not they have similar impacts. The present study is thus aimed at providing more evidence with respect to the influence of JW, in its traditional mode and via blogs, on various aspects of L2 writing, namely accuracy, fluency, and lexical as well as syntactic complexity .

Accuracy, fluency, and complexity were brought together by Skehan (1998) for the first time as a three-dimensional model of L2 development. *Accuracy* is often linked with the amount of errors in language production, while *fluency* is often attributed to the ease of language. *Lexical complexity*, also known as lexical richness (Lu, 2012), is mainly linked to the learners' communicative skills in both written and oral forms, whereas *syntactic complexity*, also known as linguistic complexity or syntactic maturity (Ortega, 2003), is generally

attributed to the variation and sophistication of grammatical structures (Foster & Skehan, 1996, Ortega, 2003).

Methodology

Participants

The participants of the study were adult engineering students, taking a General English course at a technical university in Tehran, Iran. Their ages were between 17 and 25 years. Their language proficiency levels were evaluated using the paper and pen version of the Quick Placement Test (QPT) (Oxford University Press, 2001), based on which a total of 90 students were selected from among 116 students in three classes, and classified into three groups of equal number, i.e. two experimental and one control groups. The participants included in the final analyses were those who were found to be of lower-intermediate level of language proficiency (that is, level B1 of the Common European Framework of Reference for Languages) (UCLES, 2014). This level was considered appropriate for the purpose of the study as they had an acceptable mastery over the language to express their ideas through writing journals.

Design and Data Collection

The required data was gathered through a pre-test and two post-tests, which were sample writing tasks taken from Cambridge Preliminary English Test (PET). In order to avoid the order effect, the order of treatments, i.e. writing journal entries in paper-and-pencil mode vs. blog-writing, were switched after the first half of the journal entries were collected. The first post-test was administered at this point, and the second one at the end of the course.

Different measures were employed to examine L2 writing factors. The number of words per T-unit and the number of error-free clauses were used as measures of fluency and accuracy, respectively, while lexical and syntactic complexity of the texts were judged by the computational systems, i.e. the *Lexical/Syntactic Complexity Analyzers* (Lu, 2010).

Regarding the treatments, the participants in the experimental groups were primarily introduced to the concept and objectives of journal writing and guided to write their feelings, ideas and assessments of their experiences every other session. In the first group, this was done through the traditional paper-and-pencil mode while the second experimental group wrote their journal entries on a class weblog, on <http://www.edublog.org>. Providing an educational platform, *Edublogs* claims to have been the largest blogging network in the world since 2005, and offers free blogging options to students and instructors.

Once journal entries were written by the participants, the instructor went through them, and provided feedback in the form of comments and questions. In neither group were the writings scored in terms of grammatical accuracy. All in all, each participant handed in 13 journal entries throughout the term, including both handwritten and online journals, each written on different topics and approximately 100-150 words long.

Instruments

L2 Syntactic Complexity Analyzer: Syntactic Complexity Analyzer (Lu, 2010) is a web-based computational system operating on a UNIX-like system, designed to analyze written English language samples using fourteen different indices of syntactic complexity. The fourteen measures included in this system fall into five major categories (Lu, 2010). The first group of measures take into account the length of production units; the second category deals with sentence complexity ratio; the third class assesses subordination including T-unit complexity ratio; the fourth group evaluates coordination in terms of coordinate phrases per clause, coordinate phrases per T-unit, and sentence coordination ratio; and finally the last category take particular structures into consideration. It has to be noted that in these measures, a T-unit is taken as “one main clause plus any subordinate clause or non-clausal structure that is attached to or embedded in it.” (Hunt, 1970, p. 4, as cited in Lu, 2010).

From among the fourteen indices provided by the L2 syntactic analyzer, only MLC (mean length of clauses), C/S (sentence complexity

ratio), and DC/C (dependent clause ratio), as well as T/S (sentence coordination ratio) are considered in this study, as they have been used in the literature more frequently to assess syntactic complexity.

Lexical Complexity Analyzer: a similar web-based system designed to examine written texts in terms of lexical complexity indices is the Lexical Complexity Analyzer (Lu, 2012). Using 25 different metrics of lexical richness constituting three difference sub-categories of lexical density, variation and sophistication, this system was used to run an analysis of the participants' writing samples in terms of lexical complexity. However, given the large number of measures, only three of them were considered for the purpose of this study, each standing for one of the three aspects of lexical density, lexical sophistication, and lexical variation. Following the mainstream literature in this regard, LD (lexical density as measured by the number of lexical words per the whole number of words), LS (lexical sophistication measured as the ratio of the number of sophisticated word types in a text to the total number of word types), and TTR (type-token ratio as a measurement of lexical variation) were used in this study. With respect to the second dimension, i.e. lexical sophistication, it is worth mentioning that Lu (2012) has considered words as sophisticated if they are not among the first 2000 most frequent words introduced by British National Corpus (2001).

Results

In order to analyze the data for accuracy and fluency, mixed-plot ANCOVAs and ANOVAs were run, considering *group* as the between-subject independent variable, *time* as the within subject independent variable measured as pre- and post-tests, and fluency or accuracy measures as dependent variables. In a very similar vein, mixed-plot MANCOVAs and MANOVAs were run for complexity measures and their sub-components. All the analyses were run using SPSS software (version 23).

The study was guided by four research questions, considering four elements to measure L2 proficiency. The questions are thus explored separately to examine the results found for each.

Descriptive statistics and Assumptions

Prior to the main analyses for each question, first the descriptive statistics as well as the required assumptions were checked. Table 1 below represents the means and standard deviation for the various factors examined in each of the questions. It is worth mentioning that experimental group 1 is the one having started with the paper-and-pencil mode, and experimental group 2 represents the one initiating with Blog JW. Moreover, *W* stands for *writing*, and *P1* and *P2* represent post-tests 1 and 2, respectively.

Table 1

Descriptive Statistics of the Three Groups on the Four Aspects of Writing

	Control				Experimental 1. p&p-blog				Experimental2. blog-p&p				
	P1		P2		P1		P2		P1		P2		
	Mean	Sd	Mean	Sd	Mean	Sd	Mean	Sd	Mean	Sd	Mean	Sd	
Accuracy	10.83	2.92	10.50	2.60	11.90	2.21	13.63	2.87	13.50	2.54	16.20	2.42	
Fluency	7.81	2.19	8.32	2.49	8.32	2.81	7.99	2.45	8.82	2.41	9.85	3.05	
L	LD	4.62	.79	4.85	.88	4.59	.42	4.85	.48	4.73	.53	4.54	.60
C	LS	3.48	1.20	3.40	.99	1.88	.85	2.57	1.15	2.62	.88	2.54	.99
	TTR	.54	.11	.54	.13	.56	.08	.57	.07	.54	.07	.56	.07
S	ML	6.77	.877	7.05	1.15	7.02	.93	7.73	1.31	6.99	1.13	7.99	.96
C	C												
	CS	2.11	.53	2.16	.64	1.90	.48	2.13	.56	2.00	.62	2.07	.50
	DC	.28	.13	.29	.09	.29	.12	.38	.17	.29	.09	.33	.11
	C												
	TS	1.36	.17	1.34	.22	1.24	.12	1.39	.25	1.28	.20	1.35	.27

The data in the table above can provide a rough comparison of the two post-tests for the groups. Nevertheless, in order to see if the two groups have shown any significant differences in terms of the dependent variables, the inferential statistics should be observed. Yet, before reporting the findings, the assumptions of normality, equality of covariance, as well as homogeneity were examined for all analyses.

As for the normality of distribution, the relevant tests and graphs were run and inspected. The box plots for each of the dependent variables revealed a normal distribution of the data. Furthermore, the Shapiro-Wilk test of normality did not prove significant for any of the dependent variables. Moreover, the measures of skewness and kurtosis were roughly within the ± 1 range (Tabachnick & Fidell, 2013). Therefore, in all cases, the data was considered to enjoy a normal distribution.

Box's M test was also run to test the homogeneity of covariance matrices and showed a non-significant p value for accuracy, fluency, and lexical complexity. Considering the critical value of $p = .001$ (Tabachnick & Fidell, 2013), it was indicated that the covariance matrices of the dependent variables were equal across groups for these factors. A significant value, however, was reported for syntactic complexity. Therefore, following Tabachnick and Fidell (2013), Pillai's criterion will be used to report the multivariate analyses, as it is more robust to violations of this assumption.

Finally, Levene's test of equality of error variances was also run to check the homogeneity of variances, showing no significant values for accuracy, fluency, and lexical complexity. This indicates that this assumption was also met and the data was considered ready for the analyses. It has to be noted that for syntactic complexity, the results of this test indicated that this assumption was met in most cases ($p > .05$), except for two, namely CSp1, and DCCp2 ($p < .05$). Yet, following Meyers, Gamst, and Guarino (2013), the analyses were undertaken as the data was normally distributed and the cell sizes were equal.

Therefore, four research questions were posed, for which the relevant null hypotheses were formulated and checked. The results of each question are being presented in the following section.

Research question 1: *Is there any significant difference among the two experimental and the control group in terms of the accuracy of their writing?*

The statistics pertaining to this question are presented below in the corresponding tables, followed by necessary explanations. In order to answer the first research question, first the multivariate statistics were explored. According to the results of this part, the F value for the interaction between *time* and *group* has been found to be significant [$F(2, 86) = 12.66, p < .05, \eta^2 = .228$]. It should be noted that *time* refers to the difference between the pre- and post-tests.

Additionally, tests of between-subject variables revealed that the groups have been found to differ significantly over time, [$F(2, 86) = 27.311, p < .05, \eta^2 = .36$]. Nevertheless, as the interaction effect has also proved significant, this main effect has to be treated with caution (Hatch & Lazaraton, 1991). Moreover, it was also observed that part of the variation can be attributed to the differences in the pre-tests, [$F(1, 86) = 30.894, p < .05, \eta^2 = .31$]. This difference, however, had already been expected in the analyses as the participants were only homogenized in terms of their general language proficiency level and not for each element separately. Therefore, it was natural to have differences in terms of pre-tests.

These statistics, however, are not very revealing as each experimental group had undergone a different treatment in each of the two halves of the study. Therefore, in order to see how the groups varied from one another in terms of the two modes of JW, separate ANOVAs along with their post hoc tests were run separately for the two halves of the experiment, comparing the tests prior to and following each treatment in the experimental groups.

As for the first half, the tests resulting from the split plot ANOVA for the pre-test and post-test 1 showed a significant main effect for the within-subjects variable of *time*, [$F(1, 87) = 74.28, p < .05, \eta^2 = .46$]. Moreover, the significance value reported for the interaction between time and group, [$F(2, 87) = 4.90, p < .05, \eta^2 = .10$] indicates that the difference between the two time points was different across groups and the main effects have to be interpreted with respect to this finding (Hatch & Lazaraton, 1991).

Table 2
Multivariate Tests for Accuracy for Writing (first half)

Effect		Value	F	Hypothesis df	Sig.	Partial Eta Squared
Time	Wilks'	.539	74.	1.000	.000	.461
	Lambda		287			
time * group	Wilks'	.899	4.9	2.000	.010	.101
	Lambda		07 ^b			

Tests of between-subjects variable also revealed a significant main effect for *group* [$F(2, 87) = 4.99, p < .05, \eta^2 = .10$]. This means that, the groups have had different performances. However, since the interaction effect has also been found to be significant, this finding is to be dealt with in view of the interaction effect.

Following the main analysis, the post hoc Tukey test was run to pinpoint the differences during the first half of the experiment. As Table 3 reveals, the results have shown a statistically significant difference between the control group and the second experimental group (having undergone the blog writing mode) ($p < .05$). The two experimental groups, however, have not been found to be statistically different ($p > .05$). Nevertheless, the mean differences show that the second experimental group has outperformed both the control and experimental group 1. In other words, in this first period of the study, the participants writing their journals on the weblog have outperformed both the control and the paper-and-pencil JW group. The difference, however, is only significant between the blog-writing group and the control one.

Table 3
Post hoc Test Results for the First Half of the Experiment in terms of Accuracy of Writing

(I) group	(J) group	Mean Difference (I-J)	Sig.
Control	experimental 1.p&p-blog	-.5000	.620

Tukey	experimental2. blog- p&p	-1.6500*	.008
y	experimentaControl	.5000	.620
HSD 1	1. p&p -experimental2. blog- p&p	-1.1500	.086
	blog		
	experimentaControl	1.6500*	.008
	12. blog-experimental 1. p&p -blog	1.1500	.086
	p&p		

The same analysis was run for the second period of the study (from post-test one to post-test two) when the treatments in the two experimental groups were exchanged. The results of the multivariate tests presented in Table 4 provide evidence not only for the significant main effect of *time* [$F(1, 87) = 30.16, p < .05, \eta^2 = .25$], but also for the *interaction* between time and group [$F(2, 87) = 12.92, p < .05, \eta^2 = .22$].

Table 4

Multivariate Tests for Accuracy for Writing (second half)

Effect		Value	F	Hypothesis df	Sig.	Partial Eta Squared
Time	Wilks' Lambda	.743	30.164	1.000	.000	.257
time group	* Wilks' Lambda	.771	12.926	2.000	.000	.229

Tests of between-subjects effects also showed a main effect for *group*, [$F(2, 87) = 24.213, p < .05, \eta^2 = .33$], suggesting that the groups have had statistically significant different performances over the second period of the experiment as well. Nonetheless, as the interaction was also significant, the interpretation of this main effect must be done with regard to the interaction between *time* and *group*.

Therefore, in order to understand where these differences were lying, the post hoc Tukey test was conducted. It can be seen from the results that there are significant differences between the control group

and each of the experimental ones. The mean differences show that both experimental groups have outperformed the control group.

Table 5

Post hoc Test Results for the Second Half of the Experiment in terms of Accuracy of Writing

	(I) group	(J) group	Mean Difference (I-J)	Sig.
Tukey HSD	Control	experimental1. p&p-blog	-2.1000*	.002
		experimental2. blog- p&p	-4.1833*	.000
	experimental1. p&p-blog	Control	2.1000*	.002
		experimental2. blog- p&p	-2.0833*	.002
	experimental2. blog- p&p	Control	4.1833*	.000
		experimental1. p&p-blog	2.0833*	.002

Furthermore, drawing a comparison between the two experimental groups, it can be seen from Table 5 that the two have been found to differ significantly as well. It has to be mentioned, however, that, for the second period, experimental group 1 had experienced blog writing while the participants in the second had written their journal entries via the paper-and-pencil mode. Therefore, it can be observed that, contrary to the first half, the pencil-and-paper group outperformed the blog group during the second half of the experiment. In other words, considering the whole experiment, it can be concluded that, irrespective of the mode of writing, the participants in experimental group 2 have performed better over the second half in terms of the accuracy of their writing when compared with the other groups. The plot represented in Figure 1 gives a schematic picture of the performance of the three groups throughout the experiment in terms of accuracy in their writing.

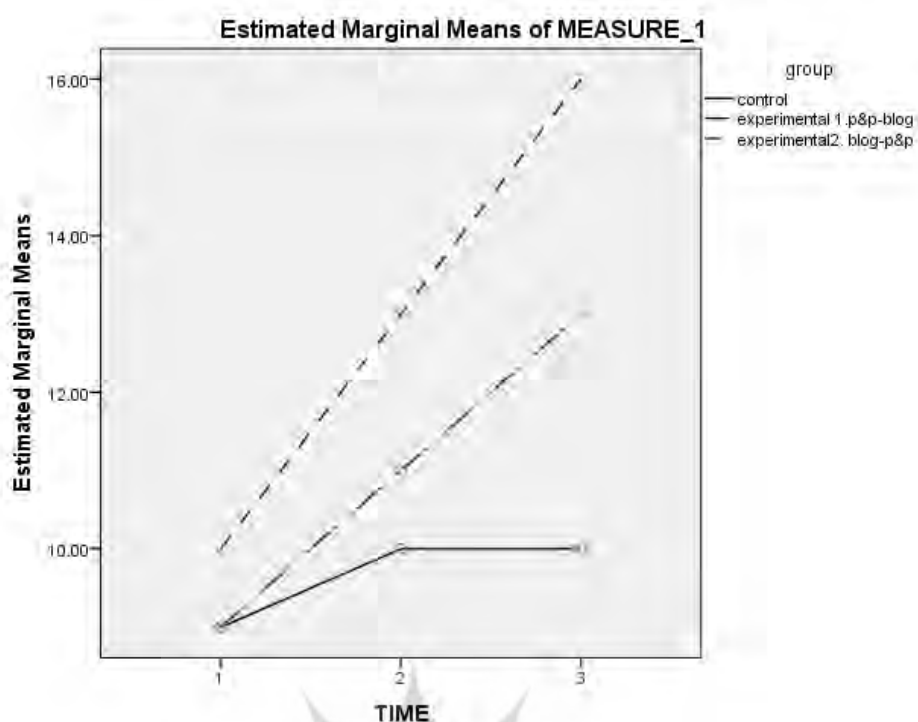


Figure 1. The three groups' performances in terms of accuracy of writing

As the slopes of the lines in the plot illustrate, from the pre-test to post-test one, all the three groups have demonstrated an increase of the means, with the experimental group 2 (blog-writing) showing a more considerable increase than the other two groups, and the experimental group 1, i.e. the paper-and-pencil JW group, outperforming the control group. Regarding the second period, i.e. from post-test one to post-test two, again experimental group 2, in which now paper-and-pencil was used as the mode of writing, shows a slightly greater increase than the other experimental group (blog writing). The control group, however, has had a slight decrease in the mean. The general overview provided in this plot supports the claim that JW in general has had a positive effect on the participants' writing in terms of accuracy. It can be thus said from the findings that the hypothesis stating that the groups are different in terms of accuracy in writing has been confirmed and the null hypothesis can be rejected.

Research question 2: *Is there any significant difference among the two experimental and the control group in terms of the fluency of their writing?*

In response to the second research question, after the assumptions were checked, the multivariate tests were run, revealing a significant result in terms of the interaction between time and group [$F(1, 86) = 7.77, p < .05, \eta^2 = .153$].

Additionally, tests of between-subjects effects indicated statistical differences for the pre-tests [$F(1, 86) = 194.105, p < .05, \eta^2 = .68$] as well as among the groups [$F(2, 86) = 27.609, p < .05, \eta^2 = .36$].

Then separate ANOVAs along with post hoc tests were run to pinpoint the differences between pairs of the tests in each of the two halves of the study.

Taking the first half into account, the multivariate tests summarized in Table 6 below yielded significant results ($p < .05$) for both *time* [$F(1, 86) = 43.35, p < .05, \eta^2 = .33$] and the *interaction* between time and group [$F(2, 86) = 21.03, p < .05, \eta^2 = .32$].

Table 6
Multivariate Tests Fluency in Writing (first half)

Effect		Value F	Hypothesis df	Sig.	Partial Eta Squared
Time	Wilks' Lambda	.667 43.358	1.000	.000	.333
time * group	Wilks' Lambda	.674 21.035	2.000	.000	.326

Nevertheless, tests of between-subjects effects did not show any significant differences among groups, $p > .05$. However, this main effect does not concern us much as the interaction has already been reported to be significant.

In a similar vein, a second ANOVA was run for times 2 and 3. The multivariate tests for *time* [$F(1, 86) = 35.42, p < .05, \eta^2 = .28$]. and the interaction between *time* and *group* [$F(2, 86) = 7.32, p < .05, \eta^2 = .14$]

have been found to be statistically significant, $p < .05$, as illustrated in Table 7.

Table 7

Multivariate Tests Fluency in Writing (second half)

Effect		Value	F	Hypothesis df	Error df	Sig.	Partial Eta Squared
Time	Wilks' Lambda	.711	35.422	1.000	87.000	.000	.289
time group	*Wilks' Lambda	.856	7.327	2.000	87.000	.001	.144

In the same vein, the tests of between-subjects effects indicated no significant differences among groups in terms of the fluency of writing during the second half either.

Nevertheless, although no statistically significant differences have been found among the three groups in terms of fluency, a comparison of means reveals that the two experimental groups showed rather a better performance in comparison with the control group which can be attributed to their engagement in JW, regardless of the mode of writing. The plot provided in Figure 2 illustrates this further.

پروپوزیشن گاہ علوم انسانی و مطالعات فرهنگی
پرتال جامع علوم انسانی

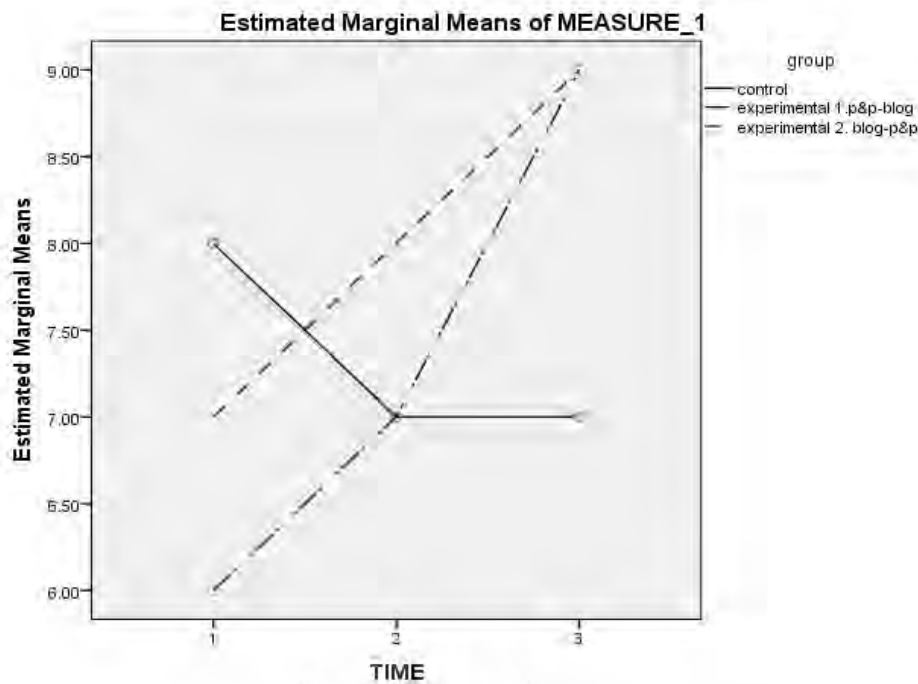


Figure 2. The three groups' performances in terms of fluency of writing

The profile plot depicts the differences between the three groups rather more comprehensibly. From the slopes of the lines, it could be understood that, from the pre-test to post-test one, both experimental groups have shown a better performance than the control group, with the paper-and-pencil group slightly outperforming the others. Of course, as can be viewed, this group has had a lower mean score on the pre-test. The mean of the control group has shown a decrease on time 2. Between times two and three, the control group has remained the same, while the experimental ones have kept doing better, almost at the same level. Therefore, the results indicate that despite the statistically non-significant differences, the experimental groups do tend to outperform the control one.

Research question 3: *Is there any significant difference among the two experimental and the control group in terms of lexical complexity of their writing?*

The third research question dealt with the performance of the participants on pre- and post-tests of writing in terms of *Lexical Complexity*, measured through three indices of Lexical Density (LD), Lexical Sophistication (LS), and Type/Token Ratio (TTR). Initially, the descriptive statistics and assumptions were explored, as presented earlier.

In the next step, the MANCOVAs were run, which indicated significant differences on the pre-test of LD [$F(3,2)=32.46$, $P < .05$, $\eta^2=.543$] and LS [$F(3,82)=38.2$, $p < .05$, $\eta^2=.58$]. This difference, however, had already been predicted. Moreover, the interaction between time and group was also reported to be significant [$F(6, 164)=3.92$, $P < .05$, $\eta^2= .13$].

To interpret the results, separate MANOVAs were conducted for each of the two halves of the experiment in terms of LC, along with their post hoc tests when necessary.

As can be seen in the following table, during the first half of the experiment, there was a significant difference in terms of within-subject factors of *time* [$F(3, 85) = 10.58$, $p < .05$, $\eta^2 = .27$] and the *interaction* [$F(6, 170) = 2.22$, $p < .05$, $\eta^2 = .07$] as well as the between-subjects factor of *group* [$F(6,170) = 8.10$, $p < .05$, $\eta^2 = .22$].

Table 8
Multivariate Tests for LC in Writing (first half)

Effect		Value	F	Hypothesis df	Partial Eta Squared
Between Subjects	Group	Wilks' Lambda .605	8.105	6.000	.000 .222
Within Subjects	Time	Wilks' Lambda .728	10.583	3.000	.000 .272
	time group	*Wilks' Lambda .860	2.223	6.000	.043 .073

The post hoc comparisons presented in Table 9 illustrate that the statistically significant difference lies between the control group and the

two experimental ones in terms of lexical sophistication (LS), with the control group outperforming the experimental ones. This finding, however, has to be interpreted with more care as the interaction effect as well as the covariate of pre-tests might have played a role. On the other hand, the two experimental groups do not appear to vary significantly.

Table 9

Post hoc Comparisons for LC in Writing (first half)

Measure	(I) group	(J) group	Mean Difference		
			(I-J)	Sig.	
LD	Tukey HSD	Control	experimental 1.jw	.1317	.662
			experimental2. Blog	.0183	.992
		experimental1 .jw	Control	-.1317	.662
			experimental2. Blog	-.1133	.736
		experimental2. Blog	Control	-.0183	.992
			experimental 1.jw	.1133	.736
LS	Tukey HSD	Control	experimental 1.jw	1.6583*	.000
			experimental2. Blog	1.1617*	.000
		experimental 1.jw	Control	-1.6583*	.000
			experimental2. Blog	-.4967	.094
		experimental2. Blog	Control	-1.1617*	.000
			experimental 1.jw	.4967	.094
TTR	Tukey HSD	Control	experimental 1.jw	.9176	.461
			experimental2. Blog	.9239	.456
		experimental 1.jw	Control	-.9176	.461
			experimental2. Blog	.0063	1.000
		experimental2. Blog	Control	-.9239	.456
			experimental 1.jw	-.0063	1.000

The data in Table 10 represents multivariate analyses of LC in writing during the second half of the study. It can be viewed that there is a significant p value for the between-subjects variable of *group* [$F(6,170) = 4.71, p < .05, \eta^2 = .14$] as well as the *interaction* between time and group [$F(6,170) = 4.95, p < .05, \eta^2 = .14$].

Table 10

Multivariate Tests for LC in Writing (second half)

Effect		Value	F	Hypothesis df	Sig.	Partial Eta Squared
Between Subjects	Group Wilks' Lambda	.735	4.711	6.000	.000	.143
Within Subjects	Time Wilks' Lambda	.929	2.172	3.000	.097	.071
	time *group Wilks' Lambda	.725	4.954	6.000	.000	.149

Therefore, post hoc test are needed to tell us where the differences lie. Table 11 represents the results of the post hoc Tukey test for the second half of the experiment.

Table 11

Post hoc Test for LC in writing (second half)

Measure		(I) group	(J) group	Mean Difference (I- J)	Sig.
LD	Tukey HSD	Control	experimental1. blog	p&p- .0167	.992
			experimental2. p&p	Blog- .1067	.732
		experimental1. p&p-blog	Control	-0.0167	.992
			experimental2. p&p	Blog- .0900	.801

		Experimental2. Control Blog-p&			-.1067	.732
			experimental1. blog	p&p-	-.0900	.801
LS	Tukey HSD	Control	experimental1. blog	p&p-	1.2200*	.000
			experimental2. p&p	Blog-	.8642*	.002
		experimental1. Control p&p-blog	experimental2. p&p	Blog-	-1.2200* -.3558	.000 .313
		experimental2. Control Blog-p&p	experimental 1. blog	p&p-	-.8642* .3558	.002 .313
TTR	Tukey HSD	Control	experimental 1. blog	p&p-	-.0222	.591
			experimental2. p&p	Blog-	-.0123	.849
		experimental 1. Control p&p-blog	experimental2. p&p	Blog-	.0222 .0098	.591 .901
		experimental2. Control Blog-p&p	experimental 1. blog	p&p-	.0123 -.0098	.849 .901

The data provided in the table indicate that, in the second period of the study too, the control group has shown a better performance than the two experimental ones on LS. Yet, it has to be noted that this difference can be due to the differences on the pre-test, as already indicated by the multivariate analyses presented at the beginning of this section. As for the other components, no significant p values have been reported. In order to understand the results better, the relevant plot will be explored below.

In order to better grasp the results, we will take a glance at the plot provided by SPSS. The plot illustrates the three groups' performances in terms of their mean scores on LS on the pre- and post-tests. It is worth

mentioning that, since the results have not yielded any significant values for the other measures of lexical complexity, only the plot related to LS has been presented here.

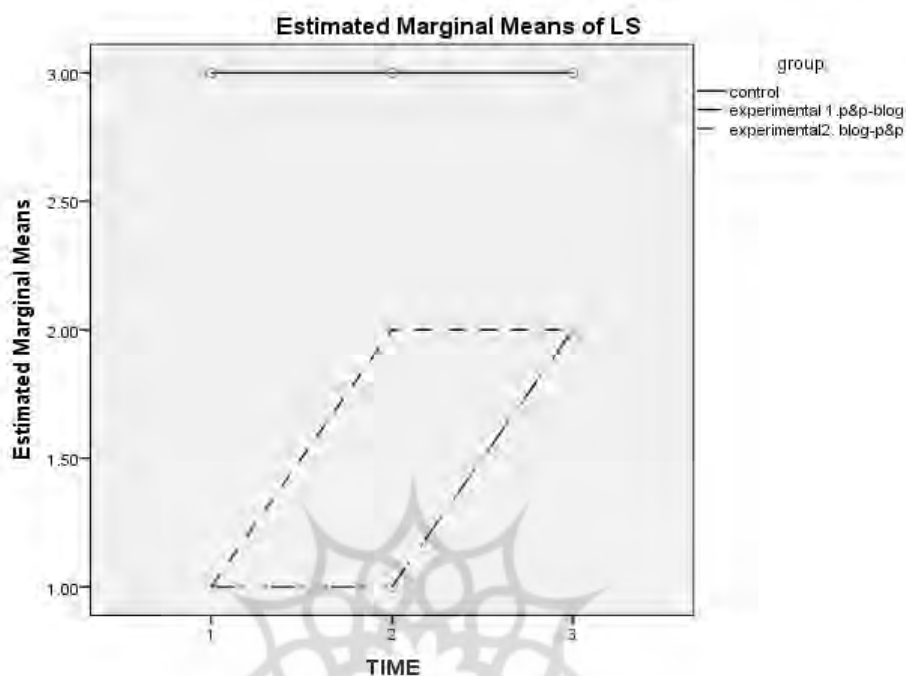


Figure 3. The three groups' performances in terms of lexical sophistication (LS) in writing

In the first half, it can be observed from Figure 3 that, although the control group has apparently had a higher mean score on both the pre-test and the first post-test, the participants' performance in this group has not improved considerably during the first half of the study. In fact, the significant difference found can be attributed to the predicted pre-existing differences among the groups, which is also observable in the figure.

Considering the experimental groups, however, it is obvious that the second experimental group, i.e. the blog writing, has shown an increase in the mean from the pre-test to the first post-test, while the first experimental group, i.e. the paper-and-pencil group has almost

remained at the same level of performance. Therefore, it could be seen that in the first half of the study, the blog-writing group has enhanced in terms of LS and outperformed the other two groups. However, the difference found through the analyses was not significant.

With respect to the groups' performances over the second half in terms of lexical sophistication, it is seen that the control group has obtained a higher mean on both tests than the two experimental ones, which accounts for the significant results in favor of the control group. Nevertheless, no improvement has occurred in the performance of this group from post-test1 to post-test 2. A quick look at the experimental groups in Figure 3, however, reveals that experimental group one (which, in this period, was involved in blog-writing) has improved more noticeably than experimental group 2, which was engaged in paper-and-pencil JW. In fact, the paper-and-pencil group has remained at the same level. This is similar to what had been previously observed for the first half of the experiment. That is to say, the blog-writing group has been found to show a more remarkable raise in the mean scores in both the first and the second halves of the study. However, this observation was not confirmed by the statistical analyses, partly due to the pre-existing differences among the groups.

Research question 4: *Is there any significant difference among the two experimental and the control groups in terms of the syntactic complexity of their writing?*

As for the last part of the analyses regarding the writing skill, the groups were compared in terms of syntactic complexity in their writing skill, measured through four indices of the Mean Length of Clause (MLC), the number of Clauses in proportion to the number of Sentences (C/S), the number of Dependent Clauses relative to the whole number of Clauses (DC/C), and the number of T-units per Sentence (T/S), using the web-based L2 Syntactic Complexity Analyzer (Lu, 2010).

Once the assumptions were checked, a MANCOVA was first conducted, which, as expected, showed differences in terms of the pre-tests, as well as the interaction between the *time* and *group* [$F(8,162) =$

3.85, $p < .05$, $\eta^2 = .16$). However, further analyses are to be explored for a more profound understanding.

In the next step, separate MANOVAs along with their post hoc tests were carried out for each of the two halves of the experiment. The multivariate tests presented in Table 12 report a significant p value in the case of both *time* [$F(4,84) = 3.30$, $p < .05$, $\eta^2 = .13$] and the interaction between *time* and *group* [$F(8,170) = 2.56$, $p < .05$, $\eta^2 = .10$]. The groups, however, did not show significant differences.

Table 12
MANOVA for SC in Writing (first half)

Effect			Value F		Hypothesis <i>df</i>	Sig.	Partial Eta Squared
Between Subjects	Group	Pillai's Trace	.142	1.626	8.000	.121	.071
Within Subjects	Time	Pillai's Trace	.136	3.304	4.000	.014	.136
	time group	*Pillai's Trace	.215	2.561	8.000	.012	.108

Following the same procedure for the second half of the study, another MANOVA was run. It is observable from the following table that, for *time* [$F(4, 84) = 21.70$, $p < .05$, $\eta^2 = .508$] and the interaction between *time* and *group* [$F(8, 170) = 3.263$, $p < .05$, $\eta^2 = .26$], the p values were significant, whereas the groups have not performed differently from one another [$F(8, 170) = 1.5$, $p > .05$, $\eta^2 = .06$].

Table 13
MANOVA for SC in Writing (second half)

<i>Effect</i>			<i>Value</i>	<i>F</i>	<i>Hypothesis</i> <i>s df</i>	<i>Sig.</i>	<i>Partial Eta</i> <i>Squared</i>
Between Subjects	group	Pillai's Trace	.135	1.537	8.000	.148	.067
Within Subjects	Time	Pillai's Trace	.508	21.706	4.000	.000	.508
	time group	*Pillai's Trace	.266	3.266	8.000	.002	.266

In the next step, the post-hoc tests were examined, which did not reveal any differences among the groups in terms of SC in writing. Therefore, it can be deduced from the analyses of this section that the null hypothesis for this section has been confirmed. That is to say, there was no statistically significant difference among the groups with respect to syntactic complexity.

Discussion

The present study attempted to explore the effect of JW in two modes on different aspects of L2 writing, namely accuracy, fluency, lexical complexity, as well as syntactic complexity. It was observed from the statistics that, generally speaking, while the mode of JW did not distinguish between the two experimental groups with regard to accuracy in the first half, the paper-and-pencil group outperformed the other two groups during the second period. The blog-writing group also showed a significant better mean than that of the control group in the both periods. It can be thus said that, regardless of its mode, JW has had a positive impact on the participants' writing in terms of accuracy.

In terms of fluency, however, no statistically significant differences were observed among the three groups. In addition, in the case of complexity, statistically significant differences were reported among the three groups in terms of lexical sophistication as an indicator of lexical complexity. The difference, however, was in favor of the control group, with no difference having been found between the two

experimental ones. Moreover, as for syntactic complexity, the groups were not found to differ significantly on any of the subcomponents.

Therefore, despite relative differences among the means of the experimental groups, which were displayed through the relevant graphs, the findings of the present study do not provide a strong basis for definite conclusions regarding the superiority of one mode over the other. Regardless of its mode, however, JW has been found to be effective in enhancing the participants' writing in terms of accuracy.

Nevertheless, although the differences between the two modes of JW were not significant, it is hard to deny the potentials of blog-writing in the enhancement of educational outcomes, which makes blogs a viable option to be used for instructional purposes whenever possible.

For instance, the generally better results in terms of accuracy, achieved by the second group which had started with blog-writing can be interpreted as a sign of motivation among the participants, that can provide the opportunity for learning and self-corrections, perhaps due to the newness of blogs being used in classrooms. In fact, the participants' early encounter with blogs might have resulted in their taking the assignment more seriously and finding JW more useful than the other groups and thus their outperforming even when the mode of writing shifted to the classic paper-and-pencil mode. Such benefits of JW have also been enumerated in a number of previous studies as well.

Tuan (2010), for instance, has concluded that JW can have positive effects on accuracy in writing as well as motivation among students. Another similar study was conducted by Puengpipattrakul (2009) on the effectiveness of JW on accuracy in writing. Although no statistically significant effect was found in that study, it was decided that JW had provoked a positive view among the participants. With respect to blog writing, Kurnia (2015) has concluded that blogs can serve as an effective medium for writing and can enhance writing accuracy. Rahmany, Sadeghi, and Faramarzi (2013) also considered blogging as a practical tool in learning as it can raise accuracy in writing. The

findings of the present study, however, do not provide sufficient evidence for the superiority of blogs over the pencil-and-paper JW.

Despite the non-significant results having been reported for fluency in this study, it was observed through the comparison of the means that JW had indeed shown a potential to enhance writing fluency, regardless of the mode. Therefore, the effectiveness of JW in this respect cannot be overlooked easily, as it has also been acknowledged by researchers such as VanderMolen (2011) and Tuan (2010).

In fact, what this study has to offer is the equal value of blogs as a more modernized user-friendly mode of writing in contexts where technology is more welcomed by both teachers and language learners. The use of blogs as a viable medium for writing has been acknowledged by others as well. Fellner and Apple (2006), for instance, refer to blogs as a means of online diary writing and suggest that it can positively influence the participants' writing fluency in terms of word counts.

Nevertheless, the fact that blogs offer more advantages than the paper-and-pencil mode was not statistically confirmed in the present work. Yet, from the higher means of the experimental groups, it could be said that JW, regardless of its mode, can be regarded as an influential activity that enhances fluency in writing. This can be due to the students' getting used to writing as they write journals on a regular basis. In addition, since language learners are regularly engaged in the process of expressing their personal thoughts via writing, they develop a command of the language in terms of speed and idea development. This can be one factor behind the increased fluency in their writing, as they can come up with ideas more rapidly. In other words, through practice, they no longer have to waste so much time thinking about the topic and tend to write more fluently.

With respect to lexical complexity, it was found that the control group had generally shown a statistically better performance than the other two groups on all the pre-tests and post-tests. The means of the tests obtained by this group, however, did not show any changes throughout the study. This outperformance could be related to pre-

existing differences in the students' academic majors or any other uncontrolled factors. Moreover, though the statistics did not show a significant difference between the two modes of JW, the comparison of the means illustrated through graphs revealed that blog JW has the potential to affect the participants' writing skill in terms of lexical sophistication as a measure of lexical complexity. This could be explained by the availability of online sources such as web-based dictionaries or easy access to relevant texts on the Internet while writing on a topic. Even Microsoft Office Word can provide synonyms for the words one writes through a mere right-click. In fact, this accessibility and the increasing tendency of the learners to get involved in online practices such as blog JW can be regarded as one of the merits of technology to be appreciated particularly by language educators. In addition, the opportunity of reading the posts written by other peers can in effect expose students to a greater range of vocabulary. Therefore, despite the non-significant findings yielded in the current study, the potential of online practices such as blog-writing cannot be easily overlooked, at least in terms of vocabulary enhancement as one can have access to the Internet or even offline sources to look up the necessary words and easily select or replace the ones he/she wants. Although this might also be the case for paper-and-pencil JW, the easier access to such sources when one is writing on a blog is undeniable.

In this respect, Fellner and Apple's (2006) study also acknowledges the potentials of blogs in the learning process. In their study of the effects of blogs as a medium for online diary writing, these researchers have concluded that tasks such as blog writing can enhance fluency and lexical complexity in the participants' writing as they can improve vocabulary recycling and noticing. The present study attempted to juxtapose the two modes of JW though, to pave the way for a comparative view. In order to provide further evidence in favor of blog JW, however, more specific research is to be conducted to support the claim with more certainty as the statistical procedures in the present study did not directly confirm it. Yet for now, it could be concluded that

blog-writing has the potential to enhance lexical complexity in the language learners' writing skill.

In terms of syntactic complexity, no significant differences were observed among the groups, meaning that this element of L2 was not affected by JW in either of the two modes. This may be attributed to the less formal nature of JW which can explain the lower tendency to use complicated and more advanced structures. Moreover, the rarity of syntactically complex sentences may be attributed to the language proficiency level of the students.

Yet, another possible reason might be related to the style of writing required on the pre- and post-tests, as the participants were asked to write stories rather than, an essay, for instance. Naturally, the type of language used for storytelling is not supposed to be very complex and formal, which might have been the reason behind the participants' reluctance to use complex structures in their writing. However, the scarcity of literature in this regard has to be taken into consideration before drawing definite conclusions. Such a finding or even the opposite has not yet been reported in the literature in this regard, suggesting that the effect of JW on the syntactic complexity of the learners' writing skills has not been well explored by researchers. In one study, Kurnia (2015) has investigated the role of blogs in enhancing accuracy and grammatical complexity in the participants' writing, and has reported that blogs can be used as a medium to enhance the students' accuracy and grammatical complexity. That study, however, focused on writing narrative texts and did not exactly resemble the present study. Therefore, it is highly recommended that more studies be conducted to decide whether or not JW can leave an impact on language learners' writing in terms of syntactic complexity.

Conclusion and Implications

The current study was aimed at exploring the effects of JW in two forms, namely the traditional paper-and-pencil mode vs. JW on a blog, on the participants' writing skill in terms of accuracy, fluency, lexical

complexity as well as syntactic complexity. The statistical findings revealed that, JW in general had a positive impact on writing accuracy. It was also observed through juxtapositions of the means that fluency, as well as lexical complexity could be enhanced through blog-writing. Syntactic complexity, however, was not affected by JW in either of its modes. Thus, although there were occasional differences among the means of the groups and the relative outperformance of the blog-writing group, the findings of the current work do not provide strong evidence in favor of blog-writing and its superiority over paper-and-pencil mode. Nonetheless, the potentials of blogs in facilitating the writing process due to accessibility and its motivational effects should not be neglected.

As a result, future researchers are suggested to embark on studies on the role of JW in various forms in the enhancement of different aspects of language learning skills. Moreover, with the advancement and popularity of modernization, potentialities of new technological platforms and devices are to be investigated, not only in terms of their implications in improving writing skill, but with regard to all aspects of language learning.

The results of the present work can provide insights to researchers and students in educational fields, language teachers and educators as well as language learners on the advantages accompanying JW in both its traditional and modern forms, and can pave the way for future research in this area. In fact, the comparative approach adopted in this study can provide evidence that blog JW has the potentiality of replacing the traditional mode of writing and is even likely to provide more benefits to language learners as well. Materials developers can also take advantage of the results in having a general overview of JW and its implementation as a useful language learning activity.

References

- Armstrong, K., & Retterer, O. (2008). Blogging as L2 writing: A case study. *AACE Journal*, 16(3), 233-251.

- Bouldin, A. S., Holmes, E. R., & Fortenberry, M. L. (2006). "Blogging" about course concepts: using technology for reflective journaling in a communications class. *Am J Pharm Educ.* 70-84.
- Brown, H. D. (2004). *Language assessment: Principles and classroom practice*. White Plains, NY: Pearson Education.
- Cahyono, B.Y. (1997). Effectiveness of journal writing in supporting skills in writing English essay. *Journal of Education*, 4(1), 310-318.
- Castellanos, J. (2008). Journal writing and its benefits in an upper intermediate EFL class. *Profile*, 9(1), 111-128.
- Common European Framework of Reference for Languages (UCLES, 2014).
- Dewey, J. (1938). *Experience and education*. New York, NY: Simon & Schuster.
- Cooper, J.W. (2006). Journal writing in career and technical education: A tool to promote critical thinking skills. *Journal of Industrial Teacher Education*, 43(2), 71-81.
- El-Koumy, A. S. A. (1998). Effect of dialogue journal writing on EFL students' speaking skill. Retrieved from ERIC database.
- Fellner, T., & Apple, M. (2006). Developing writing fluency and lexical complexity with blogs. *The JALT CALL Journal*, 2(1), 15-26. Retrieved from http://journal.jaltcall.org/articles/2_1_Fellner.pdf.
- Forsman, S. (1985). Writing to learn means learning to think. In A. R. Gere (Ed.), *Roots in the sawdust: Writing to learn across the disciplines* (pp. 162-174). Urbana, IL: National Council of Teachers of English.
- Foster, P. & Skehan, P. (1996). The influence of planning and task-type on second language performance. *Studies in Second Language Acquisition*, 18, 299-324.
- Gholami Mehrdad, A. (2008). The effect of dialogue journal writing on EFL students' writing skill. *The Journal of applied linguistics*, 1(1), 34-44.
- Hemmati, F., & Soltanpour, F. (2012). A comparison of the effects of reflective learning portfolios and dialogue journal writing on Iranian EFL learners' accuracy in writing performance. *English Language Teaching*, 5(11), 16-28.

- Kurnia, I. (2015). *The effectiveness of blog toward students' accuracy and complexity in writing narrative texts*. Syarif Hidayatullah State Islamic University, Jakarta.
- Liao, M. & Wong, C. (2007). Effects of dialogue journals on L2 students' writing fluency, reflections, anxiety, and motivation. *Reflections on English Language Teaching*, 9(2), 139-170.
- Lu, X. (2010). Automatic analysis of syntactic complexity in second language writing. *International Journal of Corpus Linguistics*, 15(4), 474-496.
- Lu, X. (2012). The relationship of lexical richness to the quality of ESL learners' oral narratives. *The Modern Language Journal*, 96(2), 190-208.
- Marefat, F. (2002). The impact of diary analysis on teaching/learning writing. *RELC Journal*, 33(10), 101-21.
- Meyers, L.S., Gamst, G., & Guarino, A.J. (2013). *Applied multivariate research: Design and interpretation*. SAGE Publications, CA: Thousand Oaks.
- Olson, D. R. (2009). The history of writing. In R. Beard, D. Myhill, J. Riley, & M. Nystrand (Eds.), *Sage handbook of writing development*. (pp. 6-16). London, UK: Sage.
- Ortega, L. (2003). Syntactic complexity measures and their relationship to L2 proficiency: A research synthesis of college-level L2 writing. *Applied Linguistics*, 24(4), 492-518.
- Palalas, A. (2011). ESP for busy college students: Is the blend of in-class, online & mobile learning the answer? *IALLT Journal of Language Learning Technologies*, 41(1), 108-136.
- Parvin, R. H., & Salam, S. F. (2015). The effectiveness of using technology in English language classrooms in government primary schools in Bangladesh. *FIRE Forum for International Research in Education*, 2(1), 47-59.
- Peterson, D. (1985). Writing to learn German. In A. R. Gere (Ed.), *Roots in the sawdust: Writing to learn across the disciplines* (pp. 46- 59). Urbana, IL: National Council of Teachers of English.

- Puengpipattrakul, W. (2009). The use of journals to develop grammatical accuracy in writing. *MANUSYA: Journal of Humanities Regular*, 12(2), 90-108.
- Rahmany, R., Sadeghi, B. & Faramarzi, S. (2013). The effect of blogging on vocabulary enhancement and structural accuracy in an EFL context. *Theory and Practice in Language Studies*, 3(7), 1288-1298.
- Skehan, P. (1998). *A cognitive approach to language learning*. Oxford: Oxford University Press.
- Tabachnick, B. G., & Fidell, L. S. (2013). *Using multivariate statistics* (6th ed.). Needham, MA: Allyn & Bacon.
- Tuan, L. T. (2010). Enhancing EFL learners' writing skill via journal writing. *English Language Teaching*, 3(3), 81-88.
- VanderMolen, M. A. (2011). *Does the use of dialogue journals affect the writing fluency of low-literacy adult Somali students?* (Unpublished master's thesis). Hamline University, Minnesota.
- Yinger, R., & Clark, M. (1981). *Reflective journal writing: theory and practice*. East Lansing, MI: Michigan State University. Institute for Research on Teaching.
- Zhang, D. (2009). The application of blog in English writing. *Journal of Cambridge Studies*, 4(1). Retrieved from <http://journal.acs-cam.org.uk/data/archive/2009/200901-article8.pdf>.