



## **Enhancing English Vocabulary Retention Among Iranian Trainee Teachers Through Literature-Based Interactive Input-Output Tasks**

**Sara Salehpour<sup>1</sup>, Biok Behnam<sup>2,\*</sup>, and Zohreh Seifoori<sup>3</sup>**

<sup>1</sup>*PhD Candidate in TEFL, Tabriz Branch, Islamic Azad University, Tabriz, Iran  
Email: s.salehpour2017@yahoo.com*

<sup>2</sup>*Corresponding author: Associate Professor in TEFL, Department of English, Faculty of Literature and Foreign Languages, Tabriz Branch, Islamic Azad University, Tabriz, Iran  
Email: behnam\_biok@yahoo.com ORCID: 0000-0002-8936-0701*

<sup>3</sup>*Associate Professor in TEFL, Department of English, Faculty of Literature and Foreign Languages, Science and Research Branch, Islamic Azad University, Tehran, Iran  
Email: zseifoori2005@yahoo.com ORCID: 0000-0002-4296-8226*

### **Abstract**

To surmount the obstacles a deficient productive vocabulary retrieval places in the way of EFL teachers, the current study sought to ascertain whether or not involving trainee teachers in interactive input-output activities enhance their vocabulary retention. To this end, a convenience sample including 49 Iranian EFL trainee teachers were recruited to take part in a quasi-experimental pretest-posttest-delayed posttest study. Having been grouped into three comparison groups, the participants were exposed to the same literary texts; however, the differential treatment of the study entailed three different interaction modes (individual, collaborative, and collaborative-cooperative) and two types of input / output processing (non-reciprocal and reciprocal). A repeated measure analysis of covariance (RM ANCOVA) was performed on the participants' achievements in the pre- post- and delayed posttest measures and the results revealed that the two groups involved in literature-based interactive (collaborative and collaborative-cooperative) reciprocal input-output activities showed significantly higher levels of vocabulary retention compared to the group exposed to literature-based individual non-reciprocal input-output tasks. Additionally, contributing to significantly higher levels of long-term retention, the collaborative-cooperative mode of interaction was found to be more effective than the collaborative one. The findings corroborated the need for including literature-based interactive input-output tasks in EFL teacher training curriculum.

*Keywords:* input-output tasks, interactive vocabulary learning, literature-based approach, teacher training course, vocabulary retention

### ARTICLE INFO

Research Article

Received: Sunday, May 16, 2021

Accepted: Thursday, April 19, 2022

Published: Sunday, May 1, 2022

Available Online: Thursday, April 19, 2022

DOI: <http://dx.doi.org/10.22049/jalda.2022.27233.1315>



© The Author(s)

## **Introduction**

Gaining unrestricted access to a rich internally-developed lexical database is undoubtedly a burning ambition of everyone interested in acquiring a good command of a second / foreign language (L2 / FL) knowledge. Such a fierce desire for vocabulary learning, as corroborated theoretically (Cameron, 2001; Lewis, 2000; McCarthy, 1990), may be attributed to the contributory role of vocabulary knowledge in serving a broad range of communicative needs of language users. Functioning as the building blocks of both language comprehension (i.e., reading and listening) and production (i.e., speaking and listening) skills (Chen & Chun, 2008), an in-depth vocabulary knowledge not only serves the learners' general need for language comprehension and use, but also could profit teachers to retain their position of power and authority in today's communicative classrooms.

The prominence of vocabulary in language acquisition and its leading role in facilitating an effective communication have compelled those in charge of language teaching programs to place a special emphasis on vocabulary development tasks / activities. Despite its central position in language teaching / learning curriculums, vocabulary has always been presumed to be a problematic domain the bulk of communication breakdowns could be attributed to (Cook, 2013). To reach a clear realization of the reason behind such deficiency, the discrimination made by Nation (2001) between receptive and productive vocabulary knowledge should be taken into account. Relying upon Nation's (2001) classification, to effectively navigate an L2 / FL-mediated communication, along with adequate receptive vocabulary knowledge to comprehend the interlocutor's message, there is a need for high levels of productive (expressive) vocabulary knowledge to convey the intended message.

As claimed by Min and Hsu (2008), an absolute majority of language users can easily retain a high proportion of newly-learned receptive vocabulary for a relatively long period of time; however, their productive vocabulary knowledge falls into decline within a limited time span. The disequilibrium between receptive and productive vocabulary knowledge, therefore, could account for the majority of communication breakdowns experienced by language users. A workable vocabulary enhancement system, therefore, needs to strike a balance between productive and receptive vocabulary knowledge. The necessity of adopting such a balanced system seems to be of paramount importance in teacher training and apprenticeship programs intended to help trainee teachers to become academically well-versed in language teaching.

Although gaining native-like competence in English vocabulary, as a potentially endless area of language learning, seems to be a faint possibility owing to the lack of authentic material exposure in EFL contexts like Iran, English teachers are in desperate need of an advanced knowledge of productive vocabulary to retain their dominant position in language learning classrooms. Having the chance to deal with many EFL trainee teachers, the authors of the current paper came to the conclusion that trainee teachers involved in academic TEFL training programs, as

future EFL teachers, in spite of possessing a wide repertoire of English vocabulary, could only put a smattering of this knowledge into actual practice while speaking or writing. Relying on a multiplicity of theoretical frameworks (e.g., Krashen's (1977) input hypothesis, Swain's (1985) comprehensible output hypothesis, Rosenblatt's (1978) transactional theory of literary works, and Long's (1981) interaction hypothesis), the current quasi-experimental study set out to explore whether going through a three-tenet cycle of vocabulary learning including iterative exposure to authentic language input, interactive word-processing, and comprehensible language output production could profit Iranian EFL trainee teachers to improve in terms of vocabulary retention.

## **Literature Review**

### **Vocabulary Retention**

The process of vocabulary retrieval, known as "vocabulary retention", is referred to as the ability to retrieve a word from memory after a specific time span (Richards & Schmidt, 2002). This ability, according to Richards and Schmidt (2002), is largely a function of teaching quality, materials quality, and learner motivation. As claimed by Gairns and Redman (1986), although the constant repetition of unknown vocabulary items enables learners to commit new words, and the meaning thereof, to their short-term memory, vocabulary retention, known generally as the ability to effectively retrieve proper vocabulary items from long-term memory, is in need of more elaborate strategies. A large body of research, therefore, has been conducted to delve into the best strategies for committing new words to both long-term and short-term memory. Given the findings of these studies, L2 / FL learners can take advantage of critical reading strategies (Khabiri & Pakzad, 2012), T-coding methods (Sadeghi Beniss & Ehsani Moghadam, 2016) semantic mapping (Nilfroushan, 2012), word list methods (Baleghizadeh & Ashoori, 2010), and mnemonic strategies (Kayaalt, 2018) to boost vocabulary retention. Notwithstanding the fundamental differences between the strategies enumerated above, there is a great deal of common ground between them on the significance of input processing quality in vocabulary retention.

### **Comprehensible Input**

Input, generally defined by Carroll (2001) as the communicatively intended language data either heard or read by language learners, has been attached special significance by many scholars in the field of L2 / FL acquisition (e.g., Ellis, 2007; Ellis & Wulff, 2015; VanPatten, 2004). Nonetheless, they are not unanimous in attaching the same prominence to the role it plays in promoting language acquisition. Such dissension between second language acquisition (SLA) pioneers is easily traceable in different theories that underpin different input-based L2 / FL instruction such as Krashen's (1982) monitor theory, VanPatten's (1996; 2004) model of input processing, the theory of emergentism (Ellis, 2007), and the interaction hypothesis (Gass 2003; Gass & Mackey, 2007). Nonetheless, all the theories enumerated above provided a fertile ground for a major reform of language education system which, according to Ellis (2012), "involves the manipulation of the input that learners are exposed to or are required to process" (p. 285).

What accounts for both L1 and L2 acquisition from Krashen's (1977) point of view is an innate mental structure called language acquisition device (LAD) which is put into action by the sufficient quantity of comprehensible input ( $i+1$ ). Comprehensible input, according to Krashen (1982) is "the only causative variable for SLA" (p. 57) and includes language structures that are beyond learners' current level of competence. Although Krashen's (1977) input hypothesis aroused considerable controversy over the adequacy of comprehensible input per se to develop L2 / FL language knowledge, it was helpful in turning the spotlight on the prominence of input in SLA. Such prominence has triggered numerous studies in the investigation of issues related to the effective type of language learning input. One branch of investigation concentrated on the use of literary texts which can serve as authentic language input for L2 / FL teaching purposes.

Literary texts, as the kind of language which closely reflects the language used by native speakers in real life situations, are rich and natural sources of target language use which basically concern genuine feelings of the writers (Bobkina & Romero, 2014). The richness and authenticity of these texts may account for their superiority over designed (controlled) texts written to allow learners cultivate a particular language skill (Norland & Pruet-Said, 2006). Additionally, literary texts serve as a means of art presentation which, at the same time, allow language use. Such dual function of literary texts would be another reason why designed texts are mainly superseded by literary ones.

### **Comprehensible Output**

Notwithstanding the broad consensus on the significance of comprehensible input, as a vital ingredient in SLA, the input-based instructional methods of L2 / FL teaching fueled controversies among the scholars of the field. One of the persuasive arguments about the input hypothesis has been put forward by White (1987) who believed that the hypothesis underestimates the importance of comprehension difficulties or input incomprehensibility in enhancing the process of SLA through the negative feedback they provide. Swain (1985) was another pioneer who provided empirical evidence that the one-dimensional view on input fails to promote SLA, showing the contributory role of meaningful language production, referred to as *output*, in improving the accurate use of language. In Swain's (1985) view, "comprehensible output is a necessary mechanism of acquisition independent of the role of comprehensible input" (p. 252). This view, labeled as the comprehensible output hypothesis, is characterized by the emphasis placed on language production and compensates for the drawbacks of comprehension-only approaches including the lack of automaticity in language use, overemphasizing semantic processing, underestimating syntactic processing, and incomplete interlanguage systems (Zhang, 2011).

Regarding the specific domain of L2 / FL vocabulary, as pinpointed by Shirzad et al. (2017), most of the instructional methods draw upon one or both of the two deep-rooted hypotheses about input / output consideration: Krashen's (1977) input hypothesis and Swain's (1985) output hypothesis. Nonetheless, there is no empirical evidence whether or not the language pedagogies underpinned by each or

both of the two theories took account of the reciprocal relationship between input and output.

### **Interactive (Collaborative / Cooperative) Language Teaching**

In an attempt to compensate for the inadequacy of comprehensible input, Long (1981) proposed the interaction hypothesis, stating that participation in conversation with native speakers, which is made possible through the modification of interaction, is the necessary and sufficient condition for second language acquisition. The interaction hypothesis attracted the attention of a huge number of scholars in the field of English language teaching (ELT). For instance, Ellis (1991) posited that Long's interaction hypothesis gives prominence to the role of interaction, while not denying nor downgrading the significance of comprehensible input in SLA. The significance of interaction was also underscored by Gass and Mackey (2007) who stated that through interaction, learners' selective attention is focused on problematic features. In other words, through interaction, learners may initially realize whether the way they convey their message differs from the way a native speaker puts it and, as a result, notice that they failed to convey what they wish to get across.

Based on a detailed review of the literature on interactive language teaching / learning, the scientific community appears deeply split on the issue of operationalizing interaction in language learning classrooms as either teacher-learner interaction (e.g., Adams, 2007) or learner-learner interaction (e.g., Gass & Alvarez Torres, 2005). Regardless of the type of interaction, any process through which learners can take advantage of scaffolding provided by other individuals in the classroom is known as an instance of interactive language learning, known also as either collaborative or cooperative language learning. Although a long hard look at the related literature testifies to the interchangeable use of the two terms *collaboration* and *cooperation* to refer to a learner-learner interaction while language learning, the former (collaboration) seems to be a better label for a teacher-learner interaction, owing to its precise definition.

### **Empirical Background to the Study**

In spite of the abundance of empirical data highlighting the significant role of input enhancement methods (e.g., Bisson et al., 2013; Bulan & Kasapoglu, 2021; Gass & Mackey, 2007) output processing methods (e.g., Gass & Selinker, 2008; Hashemi & Kassaian, 2011; Kwon, 2006; Soleimani & Mahmoudabadi, 2014), interactive classroom techniques (e.g., Dalođu & Duzan, 2010; Dobao, 2014; Hoa & Trang, 2020; Motaie et al., 2018; Rezaei Gashti, 2021; Shokouhi & Pishkar, 2015), and the input-output cycle (e.g., Benati, 2017; Kaivanpanah et al., 2020; Pei & Lin, 2020; Shirzad et al., 2017) in FL / L2 vocabulary learning and retention, no study, to the best of the authors' knowledge, has explored the cumulative impact of these factors on English vocabulary retention among trainee teachers. As a novel scientific endeavor, therefore, the current study explored the practicality of enhancing trainee teachers' vocabulary retention through involving them in literature-based interactive input-output tasks. To deeply delve into the domain under investigation, an

interactive task was operationalized as a task entailing either teacher-learner (cooperation) or teacher-learner and learner-learner interaction (collaboration-cooperation). Accordingly, the following research questions were formulated:

1. Does the involvement of Iranian EFL trainee teachers in literature-based collaborative input-output tasks significantly enhance their vocabulary retention?
2. Does the involvement of Iranian EFL trainee teachers in literature-based collaborative-cooperative input-output tasks significantly enhance their vocabulary retention?
3. Is there any significant difference between literature-based collaborative input-output and collaborative-cooperative input-output tasks in enhancing Iranian EFL trainee teachers' vocabulary retention?

## **Method**

### **Participants**

A convenience sample including 49 (28 female and 21 male) Iranian trainee teachers, with an age range from 19 to 37, constituted the participant sample of the study. They were all involved in an English teaching program in either a university (English teaching field) or an institution (teacher training course). The homogeneity of the participating teachers in terms of general English proficiency was ensured by administering a paper-based test of English as a foreign language (PBT TOEFL) prior to the main study. Based on the results, none of the TOEFL scores fell beyond one standard deviation greater or lower than the average one. Taking their performance on PBT TOEFL, the participants were then divided into three homogeneous groups. To this end, the first male and female participants with the highest scores were assigned to one group and the two next pairs with the highest scores were assigned to the other two groups. This process continued till the ones with the lowest scores were assigned to the three groups as well. The groups were then randomly named as Group 1, Group 2, and Group 3.

### **Design**

The current study was a quasi-experimental research with a pretest / posttest (both immediate and delayed posttest) comparison group design. Benefiting from such design, the study was intended to explore whether or not being involved in literature-based interactive input-output vocabulary tasks (the independent variable) have any effects on vocabulary retention (the dependent variable) of Iranian trainee teachers. Based on the between-group design of the study, three comparison groups were involved. The three groups were asked to work on the same instructional materials (literary texts) but under different interaction modes (individual, collaborative, and collaborative-cooperative) and receiving different types of input / output processing (non-reciprocal processing vs. reciprocal processing).

## **Instructional Materials**

The literary texts decided on as the chief instructional materials were extracted from the book *Harry Potter and the Sorcerer's Stone* by Rowling (1997). This imaginary-theme novel includes 14 chapters revolving around various unusual events in the everyday life of a young wizard named Harry Potter. Aside from its fantasy genre – which provided ample room for exposure to more abstract words – what accounted for the selection of the novel was that it chronicles the events in daily life of a school-age hero (Harry Potter), thereby reflecting a context-appropriate authentic language use beneficial to the teacher participants mainly motivated to make an effective communication with school-age students.

## **Testing Instruments**

### ***PBT TOEFL***

To ensure the homogeneity of the participating individuals in terms of general English proficiency, a PBT TOEFL practice test was administered prior to the course. The test, extracted from an actual TOEFL corpus (i.e., *Longman Preparation Course for the TOEFL Test*), also served as a scale whereby the participants were divided into three homogeneous groups. Although the original version of the practice test encompassed four sections including listening, structure and written expression, reading comprehension, and writing, the writing section was excluded in the current study owing to some practical constraints.

### ***Course-Based Vocabulary Measures***

Relying upon the book, *A Resource Guide to use with Harry Potter and the Sorcerer's Stone*, three equivalent course-based measures were developed to gauge the participants' lexical knowledge of the target content before, immediately after, and one month after receiving the study treatment. The resource guide served as the basis for the overall structure of the items; however, the target words / expressions were mainly replaced by more complicated content-derived words / phrase / expressions owing to the participants' proficiency level. The rationale behind using equivalent versions was to eliminate the risk of content familiarity (practice effect). Each of the three counterbalanced measures – entitled the pretest, posttest, and delayed posttest – comprised of 40 items. The comparability of the measures in terms of lexical difficulty was ensured consulting a word frequency index, namely Brown Corpus.

The reliability of the instruments was pilot tested through both internal consistency and inter-measure equivalence techniques. To this end, prior to the main study, the three researcher-devised measures were administered to a group of trainee teachers who enjoyed characteristics (e.g., age, gender, proficiency level, etc.) similar to those of the main participants. The statistical analysis of the pilot data, using Cronbach's Alpha ( $\alpha$ ) formula, showed that the three measures enjoyed acceptable degrees of internal consistency. Additionally, the estimated correlation coefficients between every possible pair of measures implied the equivalence of the measures. The content validity of the instruments was established through expert appraisal.

## **Data Collection Procedure**

To collect the data required to address the research questions, the three groups took part in a 15-session literature-based course. Before receiving the study treatment, however, all the participants were given the vocabulary pretest to be assessed in terms of initial knowledge of the target words. Having been pretested in terms of the target vocabulary, the participants in all the study groups received over 20 hours of content-relevant instruction. To cover the whole instructional content, each instructional session was allocated to one particular chapter. As a preliminary stage in training, the participants in the three groups were asked to read the target chapter prior to each session and prepare a summary of it. They were invited to use a dictionary to check the meaning of the new words and expressions they face throughout the text.

The instructor commenced each instructional session with randomly asking the participants to summarize the story. To involve the whole class in active learning, the instructor interrupted the participants and randomly chose another participant to resume the summary. The process continued until nearly everyone contributed to the summary of the target chapter. Despite the similarities between the three groups in benefiting from literary texts as the main input and putting special emphasis on summary as the core output, they differed in terms of the kinds of instruction and interaction whereby the input and output were processed.

In Group 1, the process of summarizing was followed by a conventional vocabulary instruction in which summary was presumed to be a motivator enabling learners to concentrate on the key words and phrases throughout an assigned text. Accordingly, the instructor proceeded with the instruction and asked a number of comprehension-check questions to reach a detailed account of how the story unfolds. Without any particular focus on the new vocabulary items, the questions were intended to measure the participants' ability to perceive and recollect what they read in the text. Throughout summarizing or answering the comprehension check questions, the instructor disregarded the incidentally-encountered new word / phrases. Nonetheless, when any questions arose about the problematic points in the text, the instructor intervened and explicitly taught the vocabulary to clarify the problem. In sum, in Group 1, every individual participant was in charge of his / her own vocabulary learning and no reciprocal relationship was found between the input (the literary texts) and the output (the chapter summaries). Accordingly, the instructor's corrective feedback on the target vocabulary was provided only when the need arose.

On the other hand, in Group 2, the process of summarizing the target chapter was followed by a collaborative input-output instruction which promoted an active interaction between the instructor and every individual participant while working on the summaries. Playing a dual role in progressing the instruction, the summaries acted as not only the content-oriented output but also a comprehensible input for further processing in terms of the target vocabulary. To this end, once the summary of the target chapter was fully rounded off, the instructor asked a number of vocabulary reinforcement and comprehension check questions to check not only the



participants' overall comprehension of the story but also the extent to which they had fully grasped the new words / phrases. To satisfy such a dual purpose, the participants were asked to answer the questions benefiting from the words / phrases / expressions / collocations used throughout the target chapter. Asking the questions continued until all the events and new words / expressions / phrases in the chapter were marked. In cases that any participant failed to answer the question, another one was asked and when the whole class failed to provide the word, the instructor provided a clue as to how they can spot the word. Having elicited a particular target word, the instructor embarked on expanding it, introducing synonyms, and collocations. The instructor took advantage of implicit corrective feedback (e.g., recast) to correct errors made by the participants.

The training phase in Group 3 bore a remarkable resemblance to that of the Group 2; however, the active interaction between the instructor and every individual learner (collaboration) was accompanied by a learner-learner interaction (cooperation). Accordingly, the whole instructional procedure subsequent to summary telling was carried out in a collaborative and cooperative fashion in Group 3. Such interaction mode entailed an active involvement of the instructor and the whole class. To this end, once almost everyone contributed to the summary of the target chapter, the instructor went over the chapter from the very beginning and drew the participants' attention to the new words, phrases, expressions, and collocations asking them to paraphrase some sentences and compile as many synonymous words and phrases as possible. Working in groups or pairs under the instructor's expert guidance, the participants in Group 3 then disintegrated the text into smaller parts and phrases. Accordingly, the texts were classified under some distinctive labels such as *fear*, *surprise*, *happiness*, *anger* and so on and so forth. Additionally, the participants benefited from the collaborative construction of a variety of frames representing each of the labels. The frames were mainly shaped using the chunks extracted from the story. The participants of the study were required to practice the frames in cooperation with their partner / team-mates so as to internalize and personalize the chunks. In addition to the instructor's corrective feedback, the participants in Group 3 benefited from peer-correction and recasts to correct their mistakes.

Immediately after the last training session, the posttest was administered to all the participants so as to measure any changes in their knowledge of the target vocabulary as a result of receiving the differential treatment of the study. In order to ascertain the longer-term effects of the treatment, the participants were also asked to take the delayed posttest one month after the last training session.

### **Data Analysis**

Statistical package for social sciences (SPSS, version 22) was employed to analyze the data quantitatively and the level of significance (alpha) was set at .05. The analytical procedure of the study entailed a descriptive analysis of the pretest, posttest, and delayed posttest scores in the three study groups. Subsequently, an RM ANCOVA was carried out to address the three research questions.

## Results

### Descriptive Statistics

Table 1 displays the central tendency (mean) and dispersion measures estimated based on the three study groups' performance on the pretest, posttest, and delayed posttest.

**Table 1**

*Descriptive Statistics of the Participants' Scores on the Vocabulary Measures*

Group	Variable	N	Min	Max	Mean	SD
Group 1	Pretest Scores	16	8	19	13.13	2.58
	Posttest Scores	16	21	35	29.63	3.70
	Delayed Posttest Scores	16	19	33	26.88	3.67
Group 2	Pretest Scores	16	8	15	11.44	2.39
	Posttest Scores	16	27	36	31.38	2.82
	Delayed Posttest Scores	16	23	32	27.69	3.03
Group 3	Pretest Scores	17	10	17	12.24	1.98
	Posttest Scores	17	28	37	32.35	2.76
	Delayed Posttest Scores	17	28	37	32.24	2.86

As depicted in Table 1, the performance of Group 1 on the pretest ( $M = 13.13$ ,  $SD = 2.58$ ) was, on average, better than those of the other two groups (Group 2:  $M = 11.44$ ,  $SD = 2.39$ ; Group 3:  $M = 12.24$ ,  $SD = 1.98$ ). Such a marked difference between the three groups in terms of initial knowledge of the target vocabulary implied the necessity of including the pretest scores, as the covariate variable, in the analysis. Regarding the posttest, the participants in Group 3 ( $M = 32.35$ ,  $SD = 2.76$ ) outperformed their counterparts in Group 2 ( $M = 31.38$ ,  $SD = 2.82$ ) and Group 1 ( $M = 29.63$ ,  $SD = 3.70$ ). Nonetheless, the amount of difference between Groups 1 and 3 was remarkably greater than that of Groups 2 and 3. Taking the three groups' performance on the delayed posttest, the participants in Group 3 outperformed ( $M = 32.24$ ,  $SD = 2.86$ ) their counterparts in Group 2 ( $M = 27.69$ ,  $SD = 3.03$ ) and Group 1 ( $M = 26.88$ ,  $SD = 3.67$ ).

A pairwise comparison of the participants' pretest and posttest scores in every study group revealed a substantial improvement; however, the greatest amount of increase belonged to Group 3. Furthermore, in comparison with the participants' performance on the posttest measure, all the three groups gained, on average, lower scores on the delayed posttest measure.

### Inferential Statistics

To address the research questions which explored the main as well as comparative impact of collaborative input-output and collaborative-cooperative input-output on the participants' vocabulary retention, an RM ANCOVA was conducted, regarding the achievements in the posttest and delayed posttest as two

measures of a variable called *elapsed time*. Before performing the main analysis, however, the preliminary assumptions underlying a RM ANCOVA (i.e., normality of residuals, homogeneity of variances, linearity, homogeneity of regression slopes, and equality of covariance matrices) were checked and the results were found to be satisfactory (see the Appendix). Table 2 displays the results drawn from tests of within-subject effects.

**Table 2**  
*Results Drawn from Tests of Within-Subject Effects*

Source	Type III Sum of Squares	Df	Mean Square	F	Sig.	Partial Eta Squared
Time	2.623	1	2.623	7.229	.010	.138
Time * Pretest Scores	.128	1	.128	.353	.555	.008
Time * Group	55.003	2	27.501	75.804	.000	.771
Error(Time)	16.326	45	.363			

As shown in Table 2, the elapsed time (as measured by the posttest and delayed posttest) had a significant influence on vocabulary achievements,  $F(1, 45) = 7.229, p < .05$ , partial  $\eta^2 = .138$ . In simpler terms, there was a significant difference in the vocabulary achievements of the participants in the posttest and delayed posttest measures. It was also found that the time elapsed between the posttest measure and the delayed one explained 13.8% of the variance in the participants' vocabulary achievements.

Additionally, the results in Table 2 indicated a non-significant interaction between the elapsed time factor and the pretest scores,  $F(1, 45) = .353, p = .555$ . In other words, the difference between the learners' performance on the posttest and delayed posttest was not dependent on their initial differences in terms of vocabulary achievements.

Based on the results in Table 2, there was also a statistically significant interaction between the elapsed time and the group factor on the vocabulary achievements,  $F(2, 45) = 75.804, p < .001$ , partial  $\eta^2 = .771$ . This significant interaction implied that the changes in the participants' vocabulary achievements from the immediate posttest to the delayed one was significantly different among the three study groups. This significant difference could account for 77.1% of the variance in the vocabulary achievements in the two measures.

In addition to the aforementioned significant within-subject differences, there was a significant difference between the three groups in terms of the post-intervention achievements of the participants in the two measures administered following the treatment,  $F(2, 45) = 19.875, p < .001$ , partial  $\eta^2 = .469$  (see Table 3). Accordingly, 46.9% of the variance in the post-intervention scores could be attributed to such significant difference among the three groups. Additionally, the

significant main effect of the pretest scores on the average post-intervention scores,  $F(1, 45) = 49.014, p < .001$ , partial  $\eta^2 = .521$ , affirmed the idea of including the participants' initial differences, as the covariate, in the analysis.

**Table 3**  
*Results Drawn from Tests of Between-Subject Effects*

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Intercept	1041.705	1	1041.705	107.788	.000	.705
Pretest Scores	473.689	1	473.689	49.014	.000	.521
Group	384.160	2	192.080	19.875	.000	.469
Error	434.898	45	9.664			

Table 4 displays the breakdown of the marginal means estimated based on the post-intervention scores for the different study groups after detaching the impact of pre-existing between-group differences (covariate effect).

**Table 4**  
*Marginal Means of the Vocabulary Post-Intervention Scores*

Group	Mean	Std. Error	95% Confidence Interval	
			Lower Bound	Upper Bound
Group 1	27.411	.562	26.278	28.544
Group 2	30.339	.562	29.208	31.470
Group 3	32.294	.533	31.220	33.368

The comparison of the adjusted means of the post-intervention scores, as displayed in Table 4, revealed that the participants in Group 3 outperformed their counterparts in Group 2 and Group 1.

Table 5 provided adequate support for believing that the significant between-subject difference in terms of the post-intervention vocabulary achievements was rooted in significant differences between every pair of groups.

**Table 5**  
*Pair-wise Comparison of the Marginal Post-Intervention Scores*

(I) Group	(J) Group	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval for Difference	
					Lower Bound	Upper Bound
Group 1	Group 2	--2.928	.812	.002	-4.947	-.909
	Group 3	-4.883	.776	.000	-6.812	-2.954
Group 2	Group 1	2.928	.812	.002	.909	4.947
	Group 3	-1.955	.774	.045	-3.879	-.031
Group 3	Group 1	4.883	.776	.000	2.954	6.812
	Group 2	1.955	.774	.045	.031	3.879

As shown in Table 5, there was a significant difference between Group 1 and Group 2 in terms of their participants' performance on the two post-intervention measures after controlling the initial differences between them,  $p < .05$ . This result led to the rejection of the first null hypothesis of the study (i.e., literature-based collaborative input-output instruction does not significantly affect vocabulary retention). Additionally, given the significant difference between Group 3 and Group 1 in the post-intervention marginal mean scores ( $p < .05$ ), the second null hypothesis of the study (i.e., literature-based collaborative-cooperative input-output instruction does not significantly affect vocabulary retention) was rejected as well. The case was similar with respect to the difference between Group 2 and Group 3, taking their differences at the outset of the study into account,  $p < .05$ . Consequently, the third null hypothesis (i.e., literature-based collaborative input-output and collaborative-cooperative input-output instruction do not affect vocabulary retention differently) was rejected as well.

### **Discussion**

As its primary area of inquiry, the study explored the main impact of literature-based collaborative input-output and literature-based collaborative-cooperative input-output tasks on the trainee teachers' vocabulary retention. The results drawn from the descriptive and inferential analysis of the data indicated that both collaborative and collaborative-cooperative mode of a literature-based interactive input-output instruction could affect the participants' ability to commit newly-learned words / phrases to their long-term memories. The efficacy of classroom interaction, variously labeled as collaborative learning, in accelerating vocabulary retention has been widely confirmed by research (e.g., Daloğlu & Duzan, 2010; Motaei et al., 2018; Niu & Helms-Park, 2014; Shokouhi & Pishkar, 2015; Soleimani & Mahmoudabadi, 2014). To provide a logical explanation for the effectiveness of classroom interaction, as a distinctive feature of the differential treatment administered to Groups 2 and 3, a variety of speculations could be made. The most noteworthy explanations would be a) the active participation of the participants in the input- and output-based tasks, b) ideal opportunities for word processing so as to gain a deeper knowledge of the target words / phrases, c) the provision of teacher and peer scaffolding, and d) due consideration for contextual vocabulary learning.

The fact that the participants in the two interactive groups could retain the target words / phrases better than the individual group may provide further evidence for the effectiveness of scaffolding provided through either teacher-learner or learner-learner interaction. Having been assisted in working on a variety of interactive vocabulary-processing tasks, the interactive groups participants were more likely to gain a deeper understanding of the target vocabulary items in comparison with those in individual group (Group 1) who were left on their own to tease out the meaning of the unknown vocabulary. To endorse such a claim, one can refer to Hunt and Beglar's (1998) view that in the absence of appropriate assistance, learners are very prone to wrongly guess the meaning of the unknown words from the context. Such an unscaffolded path to learning, according to Hulstijn (1992), might yield the retention of fossilized incorrectly-inferred meaning. Accordingly, the significance of scaffolding caused by interactive output processing was deemed

to be twofold: first, it may have led the participants toward deep learning of the unknown words / phrases and second, it may have prevented them from digging out the wrongly-inferred meanings from the context.

Attributing the significant impact of literature-based interactive input-output tasks on vocabulary retention to the reciprocal relationship between input and output, the findings are compatible with the findings of a couple of previously-conducted studies (i.e., Rott et al., 2002; Shirzad et al., 2017) showing the contribution of an input-output cycle to better lexical acquisition and retention. To justify the significance of a proper emphasis on the reciprocal relationship between input and output, the researchers of these studies referred to the ample room provided for word processing. Relying upon such speculation, the use of chunking, grouping semantically-related words, and framing, not only served the purpose of processing input and turning as much input to intake but also help the participants reap the benefits of output functions such as noticing the gap, hypothesis testing, and consciousness-raising.

Taking advantage of the reciprocal input-output tasks, the participants of the two interactive groups benefitted from plenty of opportunity to work on the target vocabulary both prior to and throughout the class time. Unlike the individual non-reciprocal input-output tasks implemented in Group 1, which led the participants to minimally concentrate on the target vocabulary, the interactive reciprocal input-output treatment of the study demanded great concentration on new words / phrases throughout the instructional stages including receiving corrective feedback on the summaries, answering the comprehension questions, scanning the input for the answers to the comprehension questions, and receiving instruction on the deep meaning of the target vocabulary items (i.e., synonym, antonym, collocation, pronunciation, grammatical behavior). The greater exposure to as well as the longer time spent on the target vocabulary items might have opened up a golden opportunity for the participants of the two interactive groups to gain an in-depth knowledge of vocabulary. Better retention of new vocabulary items could be the probable outcome of such a thorough knowledge of the target words. This explanation owed its justification to the previous research studies (e.g., Daloğu & Duzan, 2010) which revealed that the amount and quality of attention to various aspects of words significantly affect the retention quality.

As its secondary aim, the study examined the comparative impact of the two interaction modes while working on reciprocal input-output tasks. The comparative inferential statistics revealed a significant difference in terms of vocabulary retention between the two groups involved in collaborative and collaborative-cooperative input-output tasks. Based on the results, the participants of the collaborative-cooperative group who benefited from both teacher-learner and learner-learner classroom interaction showed significantly higher achievements in the two post-intervention measures (the immediate and delayed posttest) compared to their counterparts in the collaborative group who were only involved in teacher-learner interaction. Corroborating the claim made by Gairns and Redman (1986) that a cooperative learning setting can yield longer retention spans, the finding provided empirical evidence for the significant association between cooperation and word

retention. A possible explanation for the finding could be the learners' constant involvement in the elaboration of new concepts while interacting with their partner / group mates. Such cooperative learning activities may have improved the amount and the quality of attention that the collaborative-cooperative groups participants paid to various aspects of the target words / phrases, thereby encouraging longer retention periods.

As a matter of fact, the pair / group input-output activities (i.e., grouping, chunking, and framing) the collaborative-cooperative experimental group were actively involved in were basically a number of either rehearsal (grouping, chunking, framing) or contextualizing (internalization of the frames) strategies which belong respectively to the macro categories of memory and activation in Gu and Johnson's (1996) classification of vocabulary learning strategies. According to the classification developers, these two categories (i.e., memory and activation) are directly in charge of vocabulary retention, whereas the two other categories (i.e., cognitive and metacognitive) handle the processes contributing to vocabulary learning. Going interactively (in pairs or groups and under the direct guidance of the instructor) through three successive stages including grouping the target words under different general labels, shaping a body of language based on a grammar frame dressed with different sorts of lexical chunks extracted from both the input (the assigned text) and output (the summaries), and promoting the use of words in different contexts (internalization), the participants exposed to collaborative-cooperative tasks were much more likely to have the chance for retaining the target vocabulary after an overall elapsed time of one month.

Additionally, taking advantage of memory- and activation-based techniques to focus their efforts on the critical analysis of the source text, the participants of the collaborative-cooperative group may have been provided with an ideal opportunity to undertake a systematic contextual review of the target vocabulary. This differential feature could account for the significantly higher levels of retention among those who benefited from the literature-based collaborative-cooperative input-output tasks. The endorsement of such claim lies in Pimsleur's (1967) graduated-interval recall hypothesis which states that learners' knowledge of newly-learned vocabulary will rapidly fade, in the absence of an effective reviewing process.

In sum, the efficacy of the literature-based interactive input-output tasks in enhancing vocabulary retention among the participating trainee teachers seems reasonable believing that every single component (i.e., comprehensible input, scaffolding, and interactive input- and output-based activities) may act as a link in the chain of interaction, as suggested by interaction hypothesis (Gass & Mackey, 2015). Owing to the apparent lack of empirical evidence on the efficacy of the multi-faceted tasks implemented in the current study, however, further research is needed to add credibility to the findings.

### **Conclusion**

The findings of the current study offered the view that trainee teachers' exposure to literary texts would act as a trigger point for the enrichment of their repertoire of productive vocabulary. To hit the target, which is the active retention of

this sort of vocabulary in real communication, however, there is a real need to an instructional mechanism whereby every piece of comprehended literary text (the intake) stands the chance of becoming the basis for language production (output) and being processed in terms of target vocabulary items (the new input). Nonetheless, such a mechanism seizes up in the absence of effective interaction which could provide scaffolding required to forge a reciprocal link between output and input. Although the incorporation of literature-based interactive reciprocal input-output tasks into the mainstream teacher training pedagogy seems to have the potential to bridge the gap in EFL teachers' receptive and productive vocabulary knowledge, its contribution to longer retention span may be maximized benefiting from a full pattern of classroom interaction involving both collaboration and cooperation.

Taking advantage of an optimized form of scaffolding while being involved in a variety of input- and output-based activities, trainee teachers may be offered a systematic approach to enlarge their repertoire of authentically-used productive (active) vocabulary. Building a lexicon that approximates those of native speakers, EFL teachers would wield their authority so as to propel their learners into an effective language learning. Having the first-hand experience of vocabulary building benefiting from literature-based collaborative-cooperative input-out instruction, the trainee teachers may be persuaded to adopt the approach for their learners, enjoying both intention and practical experience required to go through such an elaborate instructional method. To satisfy this purpose, syllabus designers are recommended to search English literature for literary novels, readers, and short stories that suit learners of different proficiency levels.

Owing to the fact that the study was performed on a limited-size convenience sample of Iranian trainee teachers, the replication of the study on a larger and / or more varied sample may cast more light on the findings. To generalize the effectiveness of the specific treatment of the study in enhancing EFL vocabulary retrieval, there is a need for further research on other groups of English users (EFL teachers and students) with different language proficiency levels, age range, and learning styles. Researchers motivated to expand upon the findings are recommended to explore the difficulties of implementing literature-based interactive input-output tasks in EFL teaching / learning contexts so as to ascertain the practical ways of overcoming the many obstacles to an effective implementation of these multi-dimensional tasks.

## **References**

- Adams, R. (2007). Do second language learners benefit from interacting with each other? In A. Mackey (Ed.), *Conversational interaction in second language acquisition* (pp. 29–51). Oxford: Oxford University Press.
- Baleghzadeh, S., & Ashoori, A. (2011). The impact of two instructional techniques on EFL learners' vocabulary knowledge: Flash cards versus word lists. *Mextesol Journal*, 35(2), 1–9.



- Bulan, N., & Kasapoglu, K. (2021). An investigation of the effect of TPRS on vocabulary acquisition among third graders. *Journal of Language and Linguistic Studies*, 17(1), 645–662. <https://doi.org/10.17263/jlls.903530>
- Benati, A. (2017). The role of input and output tasks in grammar instruction: Theoretical, empirical and pedagogical considerations. *Studies in Second Language Learning and Teaching*, 7(3), 377–396. <https://doi.org/10.14746/ssllt.2017.7.3.2>
- Bisson, M. J., van Heuven, W. J. B., Conklin, K., & Tunney, R. J. (2013). Incidental acquisition of foreign language vocabulary through brief multi-modal exposure. *PLOS ONE*, 8(4), 1–7. <https://doi.org/10.1371/journal.pone.0060912>
- Bobkina, J., & Romero, E. D. (2014). The use of literature and literary texts in the EFL classroom; between consensus and controversy. *International Journal of Applied Linguistics and English Literature*, 3(2), 248–260. <https://doi.org/10.7575/aiac.ijalel.v.3n.2p.248>
- Cameron, L. (2001). *Teaching languages to young learners*. Cambridge University Press.
- Carroll, S. E. (2001). *Input and evidence: The raw material of second language acquisition*. John Benjamins.
- Chen, M., & Chung, J. (2008). Personalized mobile English vocabulary learning system based on item-response theory and learning memory cycle. *Journal of Computer and Education*, 51(2), 624–645. <https://doi.org/10.1016/j.compedu.2007.06.011>
- Cook, V. (2013). *Second language learning and language teaching* (Rev. ed.). Routledge.
- Daloğlu, A. & Duzan, K. C. (2010). Cooperative learning and vocabulary retention. *ORTESOL Journal*, 28, 15–21.
- Dobao, A. M. F. (2014). Vocabulary learning in collaborative tasks: A comparison of pair and small group work. *Language Teaching Research*, 18(4), 497–520. <https://doi.org/10.1177/1362168813519730>
- Ellis, N. (2007). The associative-cognitive CREED. In B. VanPatten & J. Williams (Eds.), *Theories in second language acquisition: An introduction* (pp. 77–95). Erlbaum.
- Ellis, N. C., & Wulff, S. (2015). Second language acquisition. In E. Dąbrowska & D. Divjak (Eds.), *Handbook of cognitive linguistics* (pp. 409–431). DeGruyter Mouton.
- Ellis, R. (1991, April). *The interaction hypothesis: A critical evaluation*. [Paper presentation]. The Regional Language Center Seminar, Singapore.
- Ellis, R. (2012). *Language teaching research and language pedagogy*. John Wiley & Sons. <https://doi.org/10.1002/9781118271643>

- Gairns, R., & Redman, S. (1986). *Working with words: A guide to teaching and learning vocabulary*. CUP.
- Gass, S. M. (2003). Input and interaction. In C. Doughty & M. Long (Eds.), *The handbook of second language acquisition* (pp. 224–255). Blackwell.
- Gass, S. M., & Alvarez Torres, M. J. (2005). Attention when? An investigation of the ordering effect of input and interaction. *Studies in Second Language Acquisition*, 27(1), 1–31. <https://doi.org/10.1017/S0272263105050011>
- Gass, S. M., & Selinker, L. (2008). *Second language acquisition: An introductory course*. Taylor & Francis.
- Gass, S. M., & Mackey, A. (2007). Input, interaction, and output in second language acquisition. In B. VanPatten & J. Williams (Eds.), *Theories in second language acquisition: An introduction* (pp. 175–199). Erlbaum.
- Gass, S. M., & Mackey, A. (2015). Input, interaction and output in second language acquisition. In B. VanPatten & J. Williams (Eds.), *Theories in second language acquisition: An introduction* (pp.180–206). Routledge.
- Gu, Y., & Johnson, R. K. (1996). Vocabulary learning strategies and language learning outcomes. *Language Learning*, 46(4), 643–679. <https://doi.org/10.1111/j.1467-1770.1996.tb01355.x>
- Hashemi, S., & Kassaian, Z. (2011). Effects of learner interaction, receptive and productive learning tasks on vocabulary acquisition: An Iranian case. *Procedia Social and Behavioral Sciences*, 15, 2165–2171. <https://doi.org/10.1016/j.sbspro.2011.04.073>
- Hoang, T. M., & Trang, T. T. T. (2020). Effect of the interactive whiteboard on vocabulary achievement, vocabulary retention and learning attitudes. *Anatolian Journal of Education*, 5(2), 173–186. <https://doi.org/10.29333/aje.2020.5215a>
- Hulstijn, J. H. (1992). Retention of inferred and given word meanings: Experiments in incidental learning. In P. J. L. Arnaud, & H. Béjoint (Eds.), *Vocabulary and applied linguistics* (pp. 113–125). Macmillan.
- Hunt, A., & Beglar, D. (1998). Current research and practice in teaching vocabulary. *JALT*, 22(1), 827–841. <https://doi.org/10.1017/CBO9780511667190.036>
- Kaivanpanah, Sh., Alavi, S. M., & Ravandpour, A. (2020). The effect of input-based and output-based tasks with different and identical involvement loads on Iranian EFL learners' incidental vocabulary learning. *Cogent Psychology*, 7(1), 1–16. <https://doi.org/10.1080/23311908.2020.1731223>
- Kayaalt, M. (2018). Mnemonic technique: An effective vocabulary teaching method to plurilingual students. *Modern Journal of Language Teaching Methods (MJLTM)*, 8(5), 388–400. <https://doi.org/10.26655/mjltm.2018.5.5>

- Khabiri, M. & Pakzad, M. (2012). The effect of teaching critical reading strategies on EFL learners' vocabulary retention. *The Journal of Teaching Language Skills (JTLS)*, 4(1), 73-106. <https://doi.org/10.22099/jtls.2012.325>
- Krashen, S. (1977). Some issues relating to the monitor model. In H. Brown, C. Yorio, & R. Crymes (Eds.). *Teaching and learning English as a second language: Trends in research and practice* (pp. 144–158). TESOL.
- Krashen, S. (1982). *Second language acquisition and second language learning*. Pergamon.
- Kwon, S. (2006). *Roles of output and task design on second language vocabulary acquisition*. [Unpublished doctoral dissertation]. University of Florida, USA. Retrieved from [http://etd.fcla.edu/UF/UFE0014501/kwon\\_s.pdf](http://etd.fcla.edu/UF/UFE0014501/kwon_s.pdf)
- Lewis, M. (2000). *Teaching collocation: Further developments in the lexical approach*. Language Teaching Publications.
- Long, M. H. (1981). Input, interaction, and second language acquisition. In H. Winitz (Ed.) *Native language and foreign language acquisition* (p. 379). Annals of the New York Academy of Sciences.
- McCarthy, M. (1990). *Vocabulary*. Oxford: Oxford University Press.
- Min, H. T. & Hsu, W. S. (2008). The impact of supplemental reading on vocabulary acquisition and retention with EFL learners in Taiwan. *Journal of National Taiwan Normal University*, 53(1), 83–115.
- Motaei, B., Ahangari, S., & Hadidi Tamjid, N. (2018). Impact of interaction and output modality on the vocabulary learning and retention of Iranian EFL learners. *International Journal of Foreign Language Teaching and Research*, 7(25), 65–85.
- Nation, I. S. P. (2001). *Learning vocabulary in another language*. Cambridge: Cambridge University Press.
- Nilforoushan, S. (2012). The effect of teaching vocabulary through semantic mapping on EFL learners' awareness of the affective dimensions of deep vocabulary knowledge. *English Language Teaching*, 5(10), 164-172. <https://doi.org/10.5539/elt.v5n10p164>
- Niu, R., & Helms-Park, R. (2014). Interaction, modality, and word engagement as factors in lexical learning in a Chinese context. *Language Teaching Research*, 18, 345–372. <https://doi.org/10.1177/1362168813510383>
- Norland, D. & Pruetz-Said, T. (2006). *A kaleidoscope of models and strategies for teaching English to speakers of other languages*. Libraries Unlimited.
- Pei, A., & Lin, Y. (2020). A study on the input and output of vocabulary teaching based on noticing theory. *Studies in English Language Teaching*, 8(2), 123-135. <https://doi.org/10.22158/selt.v8n2p123>

- Pimsleur, P. (1967). A memory schedule. *Modern Language Journal*, 51(2), 73–75. <https://doi.org/10.2307/321812>
- Rezaei Gashti, Z. (2021). The impact of storytelling and cooperative learning on Iranian EFL learners' vocabulary achievement. *International Journal of Linguistics, Literature and Translation*, 4(5), 63–76. <https://doi.org/10.32996/ijllt.2021.4.5.5>
- Richards, J., & Schmidt, R. (2002). *Longman dictionary of language teaching and applied linguistics*. Pearson Education.
- Rosenblatt, L. (1978). *The reader, the text, the poem: The transactional theory of the literary work*. Southern Illinois University Press.
- Rott, S., Williams, J., & Cameron, R. (2002). The effect of multiple-choice L1 glosses and input-output cycles on lexical acquisition and retention. *Language Teaching Research*, 6(3), 183–222. <https://doi.org/10.1191/1362168802lr1080a>
- Rowling, J. K. (1997). *Harry Potter and the sorcerer's stone*. Bloomsbury.
- Sadeghi Beniss, A. R., & Ehsani Moghadam, T. (2016). Foreign language vocabulary retention: Investigating the role of T coding method in comparison with rote rehearsal learning. *International Journal of Humanities and Cultural Studies*, 3(1), 2328–2343. <https://civilica.com/doc/491895>
- Shirzad, M., Eslami Rasekh, A., & Dabaghi, A. (2017). The effects of input and output tasks on the learning and retention of EAP vocabulary. *Theory and Practice in Language Studies*, 7(2), 145–152. <http://doi.org/10.17507/tpls.0702.09>
- Shokouhi, A., & Pishkar, K. (2015). Collaborative method and vocabulary retention of teenage EFL learners. *Theory and Practice in Language Studies*, 5(11), 2395–2401. <https://doi.org/10.17507/tpls.0511.26>
- Soleimani, H. & Mahmoudabadi, Z. (2014). The impact of interactive output tasks on develop vocabulary knowledge of Iranian EFL learners. *Iranian Journal of Applied Linguistics (IJAL)*, 17 (2), 93–113.
- Swain, M. (1985). Communicative competence: Some roles of comprehensible input and comprehensible output in its development. In S. Gass & C. Madden (Eds.), *Input in second language acquisition* (pp. 235–256). Newbury House.
- VanPatten, B. (1996). *Input processing and grammar instruction: Theory and research*. Ablex.
- VanPatten, B. (2004). *Processing instruction: Theory, research, and commentary*. Lawrence Erlbaum.
- White, L (1987). Against comprehensible input: The input hypothesis and the development of L2 competence. *Applied linguistics*, 8, 95–110. <https://doi.org/10.1093/applin/8.2.95>
- Zhang, Y. (2011). Content-based instruction in grammar teaching of English majors. *Journal of Taiyuan Normal University*, 10(5), 155–157.

## Appendix

### Results Related to Assumption Testing

**Table A1**

*Results of Normality Testing for Unstandardized Residuals of the Posttest Scores*

Variable	Kolmogorov-Smirnov <sup>a</sup>			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	Df	Sig.
Residuals for Posttest Scores	.113	49	.152	.982	49	.639
Residuals for Delayed Posttest Scores	.085	49	.200	.984	49	.656

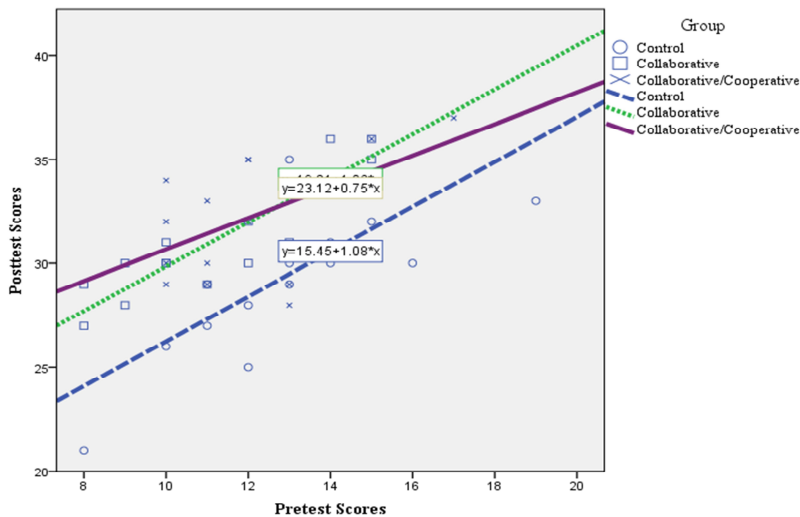
**Table A2**

*Results of Levene's Test on the Posttest and Delayed Posttest Scores*

Variable	F	df1	df2	Sig.
Residuals for Posttest Scores	1.770	2	46	.182
Residuals for Delayed Posttest Scores	2.162	2	46	.127

**Figure A1**

*The Line Chart Representing the Linear Relationship Between the Pretest and Delayed Posttest Scores*



**Table A3**

*RM ANCOVA Results for the Significance of Interaction Between the Covariate and Independent Dependent Variables*

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Intercept	1149.881	1	1149.881	114.390	.000
Group	26.171	2	13.086	1.302	.283
Pretest Scores	414.711	1	414.711	41.255	.000
Group * Pretest Scores	7.492	2	3.746	.373	.691
Error	432.249	43	10.052		

**Table A4**

*Test of Equality of Covariance Matrices on the Posttest and Delayed Posttest Scores*

Variable	Box's <i>M</i>	<i>F</i>	<i>df1</i>	<i>df2</i>	<i>Sig.</i>
Post-intervention Scores	11.042	1.724	6	51542.228	.111

### Authors' Biographies

---



**Sara Salehpour** is a PhD candidate at Islamic Azad University Tabriz branch. She has several years of experience in teaching English and tutoring Advanced English learners and English teachers in Advanced Vocabulary, Grammar, Speaking, and Writing. She has published research articles in national and international journals. Her research interests include Teaching Methodology and Language Learning through Literature.



**Biook Behnam** is an Associate Professor of Applied Linguistics at Azarbaijan Shahid Madani University. He has widely presented papers on national and international conferences in North America, Australia, Europe, China, India, and South East. He has acted, for around ten years, as the Editor-in-Chief of the *Journal of Applied Linguistics* published by Tabriz IAU, and is on the editorial board of some national and international academic journals.



**Zohreh Seifoori** is an Associate Professor of TEFL at Islamic Azad University, Tabriz Branch, and at Science and Research Branch, where she is a researcher, educator, and teacher trainer. She has published research articles in renowned national and international journals and her research interests include Teacher Education, Learner Autonomy, and Teaching Methodology.

---