

Impact of Grouping Type in Descriptive Collaborative Writings on Iranian EFL Learners' Written Grammatical Accuracy

Hossein Mohammadzadeh, Ph.D. Candidate, Department of English Language Teaching, Tabriz Branch, Islamic Azad University, Tabriz, Iran

Mohammadzadeh255@gmail.com

Nasser Ghafoori*, Assistant Professor, Department of English Language Teaching, Tabriz Branch, Islamic Azad University, Tabriz, Iran

ghafoori.nasser1968@gmail.com

Mahnaz Saeidi, Associate Professor, Department of English Language Teaching, Tabriz Branch, Islamic Azad University, Tabriz, Iran

mnsaeidi@yahoo.com

Abstract

The current study was an attempt to investigate the impact of grouping type on the grammatical accuracy of Iranian EFL learners in collaborative writing. Through administering the Michigan Test of English Language Proficiency, 64 female university students available participated in this study and were assigned to two groups--heterogeneous and homogeneous. The treatment process lasted 12 weeks of collaborative writing and revising, with emphasis on the development of the the participants' written grammatical accuracy. They were pretested and posttested through a descriptive free writing measure. The results revealed that the participants in both groups had higher posttest scores. Using ANCOVA, it was found that there was a differential effect of grouping type on the grammatical accuracy of homogeneous and heterogeneous pairs, indicating that the participants in heterogeneous group had significantly higher posttest scores than those in homogeneous group. Furthermore, although the grammatical accuracy of both groups increased, a significant difference was observed between them, revealing that the participants in heterogeneous group outperformed those in homogeneous group. The findings of this study can be interpreted in terms of the sociocultural theory and Vygotsky's idea of the zone of proximal development.

Keywords: Collaborative writing, grouping type, homogeneous group, heterogeneous group, grammatical accuracy

Introduction

The use of pair and small group work in which learners in groups of two or three write collaboratively in both first (L1) and second (L2) language learning classrooms is one well-liked method of writing instructions (Adams & Hamm, 1996; Doughty & Pica, 1986). Vygotsky stated that "social interaction comes before development; consciousness and cognition are the end result of socialization and social behavior" (Heidar, 2016). The basic idea of collaborative writing was built on this Vygotskian idea of having to collaborate with others through sharing ideas in order for quality learning and growth to happen. (Heidar, 2016).

The use of pair and group work in the L2 classroom is advocated by the social constructivist perspective of learning. Originally rested on the work of Vygotsky (1978), this perspective of learning claims that human development is inherently attracting through social activity. Lightbown and Spada (1999), Doughty and Williams (2013), amongst others, have investigated how opportunities are provided by interaction for learners not only to negotiate the

input message, but, in so doing, to concentrate on its form as well. Studies that compare individual and pair work on writing as well as grammar-focused tasks have proven some advantage for pair-work.

Storch (1999) contrasted pair and individual performance on a variety of grammar-focused exercises such as text reconstruction, multiple-choice and cloze passage and came to a conclusion that pairs performed the exercises more reasonably than the learners who worked individually. Nevertheless, because the same learners carried out the tasks individually and in pairs, the end results might have been affected by the effect of practice. Concerning those conflicts in collaborative writing that students experience while having discussion with their group, some researchers showed that interaction and discussion within participants in groups are influenced by the group selection (Adodo & Agbayewa, 2011; Cady, 2011; Doughty & Pica, 1986; Kuiken & Vedder, 2002; Maftoon & Ghafoori, 2009; Mahenthiran & Rouse, 2000). Kuiken and Vedder (2002) clearly revealed that the formation of carelessly arranged group which they referred to as students' language proficiency dynamic can affect the learner's text quality. It stated indirectly that group and pair selection should be regarded as the first stage to make an effective collaborative writing in the classroom by teachers. They should decide over the best group selection methods carefully prior to the beginning of the class. Thus, the students will be benefited to a large extent from their interaction with the other members. So, another closely connected subject is the level of proficiency of collaborators in collaborative pair and small group works.

Some studies emphasizing the method of group selection including heterogeneous and homogeneous grouping type in different educational areas have been done widely since 1990s to understand the success of these group selection ways (Adodo & Agbayewa, 2011; Cady, 2011; Doughty & Pica, 1986; Kian-sam, 1999; Maftoon & Ghafoori, 2009; Webb, 1989). Later, it has become a debatable problem because each type has its own benefits and disadvantages.

Baer (2003) proposes two most important methods to group learners in CL which are known as homogeneous and heterogeneous groupings. Those who advocated heterogeneous grouping compare the school atmosphere to a practical work atmosphere. They realize that once learners enter the work place they will be required to work with people that have a vast range of aptitudes, ages, abilities, etc. They also regard that the students with lower and average levels will be benefited from peer interaction while the students of higher levels can help their own learning by teaching other students (Kruse, 2011). On the other hand, advocates of homogeneous group method think that homogeneous grouping lets the learners make progress with faster rate without being held by slower learning rate of other learners (Kruse, 2011). In addition, Adodo and Agbayewa (2011) realized that the homogeneous ability level grouping was more powerful than mixed ability grouping. The low- and average ability learners gain benefit academically from homogeneous grouping type than the heterogeneous grouping.

The above-mentioned debatable problem which considers the development (or lack of development) of heterogeneous and homogeneous grouping types, as well as the interaction between the grouping methods and the achievement level (high and low achievers), were the foci of the current study. Therefore, efforts were made to test the following null hypothesis and answer the posed research question:

Q. Does grouping type in collaborative writing have any significant effect on Iranian EFL learners' written grammatical accuracy?

HO. Grouping type (homogeneous vs. heterogeneous group) in collaborative writing does not have any differential impact on Iranian EFL learners' written grammatical accuracy.

Review of Literature

Peer–Peer Interaction and Proficiency Differences

Although Storch's focus was not on proficiency differences, her 2001 study revealed that the pair with the highest proficiency difference (low and upper intermediate) was most collaborative in engaging in the task compared to the other two pairs. Furthermore, the pair with some degree of homogeneity (low and intermediate) was found to be a non-collaborative dominant/dominant pair, which showed less transfer of knowledge and more missed opportunities. Storch therefore suggested that proficiency differences may not be the major reason for a non-collaborative orientation. However, Kowal and Swain (1994) revealed contradictory findings. Their data with grade eight French immersion students documented that in a highly heterogeneous grouping (e.g. upper-middle and low), the stronger student tended to carry out most of the work either because the weaker student was too intimidated to say anything, willing to let the stronger student do the task, or was not allowed to do any of the task whether their opinion was valid or not. Successful scaffolding requires the group members to respect one another's perspectives and trust each other's opinions (Stone, 1993). This may be difficult to achieve when proficiency differences are too large.

Yule and Macdonald (1990) investigated whether different proficiency pairs could work successfully if each member was given appropriate interactive roles. In their task with adult ESL pairs, the more dominant role was to provide map directions and the less dominant role was to identify the directions with a slightly different map. They found that when the lower proficiency member was responsible for the more dominant role, there was more negotiation of meaning and a successful resolution of referential conflicts. Conversely, when the higher proficiency member played the more dominant role, they engaged in little negotiation. The higher proficiency members seemed to ignore their lower proficiency partners' contribution while the lower proficiency partners often assumed a passive role.

Leeser (2004) focused on the impact of learner proficiency on language related episodes (LREs) 2 – instances of collaborative dialogue (Swain, 2001) – in an adult L2 Spanish class. An LRE is defined by Swain and Lapkin (2002) as 'any part of the dialogue where learners talk about the language they produced, and reflect on their language use' (p. 292). Leeser analyzed the frequency, type (i.e. lexical or grammar-based) and outcome of LREs (i.e. problem solved correctly, not solved or solved incorrectly) produced by three different groupings: high-high, high-low and low-low. He found that as the overall proficiency of a pair increases, the learners produce a greater number of LREs, correctly resolve more LREs, and focus more on form than on lexical items. Because the high-low pairs fell between the high-high and low-low peers in their performances, Leeser was led to wonder if the high-proficiency learners actually benefited from their interaction with their low-proficiency partner, and what the basis was of the improved performances of the low proficiency partners (relative to those in the low-low groupings).

Storch (2007), investigated the effects of second language (L2) proficiency differences in pairs and patterns of interaction on L2 learning, making use of both qualitative and quantitative data. She designed the study in such a way that four different *core* participants interacted with higher and lower proficiency *non-core* participants. These learners engaged in a three-stage task involving pair writing, pair comparison (between their original text and a reformulated version of it) and individual writing. The core participants also engaged in a stimulated recall after the task. The researcher analyzed each pair's collaborative dialogue in terms of language-related episodes

and patterns of pair interaction (Storch, 2002a) as well as each learner's individual post-test score. The findings suggested that the patterns of pair interaction greatly influenced the frequency of LREs and post-test performance. When the learners engaged in collaborative patterns of interaction, they were more likely to achieve higher posttest scores regardless of their partner's proficiency level. It seems that proficiency differences do not necessarily affect the nature of peer assistance and L2 learning.

What role did the nature of the interactions within each pair play? These questions, however, cannot be answered from the quantitative analysis of LREs alone. Our literature review revealed that despite its significance, the issue of peer–peer learning between learners of different L2 proficiency levels elicited

few studies in the field of SLA. Moreover, the few studies that addressed this issue examined the interaction of pairs involving different proficiency learners. Such a research design fails to consider how the same learners interact with higher and lower proficiency peers. The present study was conducted in an attempt to provide insight into this under-explored area, using a more appropriate research design.

Recent Studies on Homogeneous and Heterogeneous Grouping in Iranian EFL Context

In the Iranian EFL context, which is the context of the current study, Some relevant studies have been conducted (Maftoon & Ghafoori, 2009; Pishghadam & Ghadiri (2011) Ghanbari & Ghaffar Samar, 2016; Zamani, 2016) . The first one was done by Maftoon and Ghafoori (2009) on the effect of homogeneous (symmetrical) and heterogeneous (asymmetrical) collaborative interaction on the development EFL learners' writing skill. Sixty female students of TEFL participated in the study. The participants were divided into two groups based on their English proficiency test scores. The homogeneous group consisted of 14 participants paired with partners with similar English proficiency test scores, while the heterogeneous group consisted of 16 participants who were paired with partners who had higher test scores. The pairs had interaction and peer collaboration before carrying out three types of writing tasks.

The findings about G-Hom show that the pair work the students had before writing their assignments had a positive effect on their writing ability. The collaboration and interaction the participants had with their partners who had almost equal language proficiency level helped them improve their writing skill. As Storch (2007) suggests, “pair work afforded learners opportunities to pool their linguistic resources and co-construct knowledge about language” (p. 155).

The results for G-Het also show that the participants' writing scores increased significantly from the pretest of writing to all three posttest writing tasks. The collaborative interaction that the participants had with partners who were more proficient helped them to improve their writing skill. The positive effect of collaborative dyadic interaction was seen in the student/writers' higher scores in their posttest individual performances. In sum, their findings showed that although the writing skill of both groups increased significantly as the result of interaction, no significant difference was observed between the two groups. The findings of this study were not completely supported by those by Pishghadam and Ghadiri (2011) who investigated the effect of symmetrical and asymmetrical scaffolding on the reading skill of Iranian EFL learners. The results of this study indicated that the participants in the asymmetrical group which was composed of partners of unequal proficiency outperformed their counterparts in the symmetrical group.

In 2016, Ghanbari and Ghaffar Samar conducted a study to investigate the effect of grouping in terms of language ability and gender of the Iranian EFL learners on their written performance. Two elementary-level male and two elementary-level female intact classes were

selected and randomly assigned to either homogeneous or heterogeneous groups. The treatment sessions lasted 9 weeks of instruction with a special emphasis on developing the skill of writing. Learners were pre-tested and post-tested through a free writing measure. Analyses of variances indicated that learners in homogeneous groups outperformed those in heterogeneous ones, though it did not reach any significance level. Language ability of the learners did not make any difference and only the gender made a significant difference between the groups. A more focused analysis of the five components of writing (i.e. content, organization, grammar, mechanics and vocabulary) revealed the same pattern as that of the learners' whole composition in the two grouping formats. The obtained results were a confirmation of a new strand of research which has questioned the long-dominant heterogeneous grouping in cooperative learning settings to design groups in a way that promotes learners' achievement to the extent possible.

Zamani (2016) conducted a study as an attempt to investigate the impact that homogeneous and heterogeneous groupings of Iranian EFL learners regarding their prior levels had on their writing ability when working cooperatively. Having administered a standardized preliminary English test (PET) and a writing test taken from PET sample tests as a pre-test, 66 high and low proficient learners were assigned into three groups: heterogeneous, homogeneous high, and homogeneous low groups. Following the end of the treatment that took 10 sessions each for 30 min, all groups received a writing test as a post-test. The results demonstrated that learners improved their performance through cooperation, whether working with stronger or weaker peers. However, heterogeneous grouping showed superiority over homogeneous grouping at the low level. Low students in the heterogeneous class made more relative gains than high students in the same class. It must be noted that low students did not improve at the expense of high students. The results revealed that cooperative learning could be especially beneficial for low students.

As the findings of the previously-mentioned empirical studies show, there is no consensus among the researchers on the effect of some variables such as interlocutors' level of proficiency and interaction patterns on the academic achievement of EFL learners. The context of studies becomes even more complex if the teacher is supposed to be one of the interlocutors in interaction configurations or proximal processes.

Methodology

Design

The design of the study was an intact group quasi-experimental one. Thus, out of four available classes, 64 female students voluntarily took part in the study. Since the major goal was to determine the possible impact of homogeneous and heterogeneous grouping types on the learners' grammatical accuracy in their writings, the independent variable consisted of two grouping types groups, i.e. heterogeneous and homogenous pairs, and the dependent variable was the participants' grammatical accuracy in their descriptive writings.

Participants

The participants for this study were selected through convenient sampling, that is, sixty four EFL learners who were studying Teaching English as a Foreign Language at Islamic Azad University, Tabriz Branch, were chosen to take part in this study. The participants were all female students because majority of TEFL students (80%) available to the researchers at the time of the study were females. Another rationale for this was the administration facility and controlling the gender variable. The age range of the participants was 19 to 26. Based on the the results of an English language proficiency test (Michigan English Language Assessment Battery-

-MELAB), the participants were assigned to two experimental groups (two grouping types, i.e. heterogeneous and homogenous pairs).

Instrumentation

Michigan Test

The first instrument in the present study was the Michigan English Language Assessment Battery (MELAB). It was used assign participants to two proficiency levels (low and high learners) and to match them in groups to form heterogeneous and homogenous pairs. In fact, this test was exploited for the purpose of measuring the participants' level of proficiency and making them homogenous. The test contained 40 grammar, 40 vocabulary, and 20 reading comprehension items, i.e. 100 multiple choice items in total. Each correct answer was assigned one point-- totally 100 points.

Writing Topics (Descriptive writing prompts)

The descriptive 4 topics used in this study were adapted from the textbooks that are generally studied in EFL writing courses: *Communicate What You Mean, TOEFL Writing Topics and Model Essays* (2002) to ensure that the topics were parallel. One of the important variables that could affect the results of the research was the task type effect. In order to preclude the task type effect, the researchers deliberately limited the task type to descriptive type of paragraph writing. Moreover, the same 4 topics were introduced as both pretest and posttest writing prompts. and the time interval between pretest and posttest was ten weeks. It seemed to be enough time to preclude the practice effect. It is worth mentioning that before the beginning of the study and before writing about the topics in the treatment process, an informal oral opinion poll was used regarding every topic to make sure students had sufficient topical knowledge about these topics and to elicit sufficient data. Students were allocated 45 minutes to complete each task.

Scoring Scale

To correct the participants' writings at the beginning and at the end of the treatment (after pretest and posttest), the ratio of number of grammatical errors or incorrect grammatical items to the total number of T-units (Wolf-Quintero et al. 1998) was measured. The focus of scoring was on the grammatical accuracy of the texts. The rating was also done by a rater other than the researchers to ensure reliability. Cronbach's alpha was used to detect the degree of inter-rater reliability between the two rating scores for the variable of accuracy.

Procedures (Data collection and Data Analysis)

As it was mentioned above, the participants were divided into two experimental groups according to their language proficiency test scores: Homogenous Group (G-Hom) and Heterogeneous Group (G-Het). In the first stage, the participants were pretested. The interval between the pretest and posttest lasted ten weeks. This almost long period could help in precluding the practice effect. In order to elicit enough data, both in pretest and posttest stages, the participants were required to write a descriptive essay individually, although the they did it collaboratively in the experimentation or treatment process. The intention was to know if collaborative grouping type could help them to develop their written grammatical accuracy.

During the treatment process, the participants in both homogeneous pairs and heterogeneous pairs were required to write and revise their writings collaboratively. Regarding the descriptive topics they were provided, the members of both groups became familiar with them

through an informal oral opinion roll to make sure that they had enough knowledge to write about them. During the informal opinion roll about the topics, the participants were free to use both English as a foreign language and their own L1--Farsi. This way, they were required to write collaboratively on the topics. They had 45 minutes to complete the task. During collaborative writing session, each pair was asked to write two paragraphs in order to give a chance to each participant to be the initiator of one paragraph. To check if both participants in each pair were involved in collaborative writing, their dialogues were recorded.. Then, in the next session the pairs were required to jointly revise their produced writings in connection with the indirect feedback provided by the researchers on their writings. The feedback included some underlined grammatical errors without providing the correct forms. In the revising stage, the members in each pair ignored some of the underlined grammatical errors, because neither of them knew the correct forms. The same procedure (informal opinion roll, collaborative writing, providing indirect feedback, and collaborative revising) was applied to other descriptive topics.

The treatment cycle continued until the pairs in groups worked on all four topics. Because the emphasis of the study was on the grammatical accuracy of the texts, following (Wolfe-Quintero et al,1998), grammatical accuracy was assessed as the ratio of incorrect grammatical forms to the total number of terminal units (T-units) used. Thus, the participants were asked to take part in the collaborative writing task with a partner and after receiving indirect feedback on grammatical accuracy of some structures in their joint writing, they were required to revise their writing task on the basis of the indirect feedback. In fact, they discussed, talked and reflected on the indirect feedback they received from their teacher in order to revise their jointly produced writings in terms of grammatical accuracy. Through the application of collaborative revision technique and the tasks related to it, the participants were actually encouraged to involve in a deeper process of thinking about the accuracy of grammatical forms in their writings.

At the end of the treatment process, the same 4 topics in the pretest writing were presented to the participants and they were required to write another composition on each topic individually. The scores obtained here were considered as posttest scores. The collaborative process of error correction in some cases offered a chance for the participants to make use of each other's ideas and reach a higher level of language knowledge which, in turn, helped them to expand their language knowledge in general and grammatical accuracy knowledge in particular. In other words, during the pair talk collaborative writing and revising, participants shared their ideas and negotiated meaning to prepare themselves for the writing task.

Finally, to answer the posed research question and in order to show the results of pretest-posttest comparisons in homogeneous and heterogonous groups and their improvement after treatment process, analysis of covariance (ANCOVA) was done.

Results

To determine the homogeneity and heterogeneity of the participants, the Michigan Test was given to the participants in intact classes. Table 1 below shows the descriptive statistics of the participants' proficiency scores.

Table 1. *Descriptive Statistics of the Iranian Participants' Proficiency Test Scores*

	N	Minimum	Maximum	Mean	Std. Deviation
Michigan proficiency scores	test64	23.00	85.00	62.30	12.83
Valid N (listwise)	64				

As Table 1 shows, the standard deviation and the mean of the Iranian participants' proficiency test scores were 62.30 and 12.83 ($M= 62.30$, $SD= 12.83$). Thus, the participants whose scores fell within one standard deviation above and one standard deviation below the mean were selected as the homogeneous group and those whose scores fell within two standard deviations above and two standard deviations below the mean were selected as the heterogeneous group.

Written Grammatical Accuracy Pretest Scores

In order to make sure about the homogeneity of the participants regarding their written grammatical accuracy in heterogeneous and homogeneous groups before the treatment, the written grammatical test as a pretest was administered, since the homogeneous group included 16 participants matched with interlocutors with similar test scores in English proficiency, and heterogeneous group included 16 participants paired with partners who had higher test scores. Table 2 below presents the results of descriptive statistics for the pretest scores between heterogeneous and homogeneous groups.

Table 2. Descriptive Statistics of the Participants' Pretest Written Grammatical Scores

	Grouping type	N	Mean	Std. Deviation
Pretest grammatical accuracy scores	Heterogeneous group	32	.70	.38
	Homogeneous group	32	.81	.44

According to Table 2, the mean score of the heterogeneous grouping type was .70 with the standard deviation of .38, but the mean score of the homogeneous group was .81 with the standard deviation of .44.

As for the inter-rater reliability of the participants' written grammatical scores in the pretest, two raters scored the written grammatical accuracy, and Cronbach's alpha was run. Table 3 below demonstrates the results.

Table 3. Inter-rater Reliability of the Participants' Written Grammatical Accuracy Scores in Pretest

		Rater 1	Rater 2
Rater 1 Pretest	Pearson Correlation	1	.991**
	Sig. (2-tailed)		.000
	N	64	64
Rater 2 Pretest	Pearson Correlation	.991**	1
	Sig. (2-tailed)	.000	
	N	64	64

As is illustrated in Table 3, according to Cohen's (1988) guideline, there was a significant and strong correlation between the scores of two scorers since $r = .991$, $p = .000$, implying the high inter-rater consistency between the two raters. Besides, it was necessary to check whether the participants' written grammatical scores in the pretest was normally distributed. So, One-sample kolmogorove-smirnov test was run. Table 4 below specifies the results.

Table 4. *One-Sample Kolmogorov-Smirnov Test for the Participants' Pretest Scores*

		Grammatical accuracy scores in Pretest
N		64
Normal Parameters ^{a,b}	Mean	.75
	Std. Deviation	.41
Most Extreme Differences	Absolute	.093
	Positive	.093
	Negative	-.075
Test Statistic		.093
Asymp. Sig. (2-tailed)		.200 ^{c,d}

As is illustrated in Table 4, the significant value of .200, higher than alpha level ($P = .200 > .05$), indicated that the participants' written grammatical scores in the pretest were normally distributed since the normality test is a prerequisite before the parametric test of Independent samples *t*-test. After the normality assumption was met, the Independent Samples *t*-test was used to see if there was a significant difference between the participants' pretest written grammatical scores in heterogeneous and homogeneous groups. Table 5 below reveals the results of this test.

Table 5. *Independent Samples t-test for the Participants' Pretest Written Grammatical Scores*

		Levene's Test for Equality of Variances		<i>t</i> -test for Equality of Means				95% Confidence Interval of the Difference		
		F	Sig.	T	Df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
Pretest grammatical accuracy scores	Equal variances assumed	.103	.749	1.136	62	.032	.12	.10	-.09	-.32
	Equal variances not assumed			1.136	60.583	.032	.12	.10	-.09	-.32

The results of Table 5, according to Levene's test for equality of variances, the significant value of .749, higher than alpha level, revealed that the equal variances were assumed and the

results of first row should be read. Since $t(62) = 1.136$, $P = .032 < .05$, it was revealed that there existed a significant difference between the participants' pretest written grammatical scores in the heterogeneous and homogeneous group before the treatment.

Results for the Null hypothesis

The null hypothesis stated that grouping type (homogeneous vs. heterogeneous group) in collaborative writing does not have any differential impact on Iranian EFL learners' written grammatical accuracy. So, to see whether the homogeneity of regression slopes was met or not. To yield this purpose, the researcher used test of between-subjects effects. Table 6 below exhibits the test results.

Table 6. Homogeneity of Regression Slopes for the Participants' Posttest Written Grammatical Scores

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Corrected Model	1.823 ^a	3	.608	16.604	.000	.454
Intercept	.129	1	.129	3.529	.065	.056
Grouping types	.104	1	.104	2.842	.097	.045
Pretest grammatical accuracy scores	1.768	1	1.768	48.308	.000	.446
Grouping types * Pretest grammatical accuracy scores	.172	1	.172	4.689	.094	.072
Error	2.196	60	.037			
Total	14.243	64				
Corrected Total	4.019	63				

a. R Squared = .454 (Adjusted R Squared = .426)

As it is indicated in Table 6, since $F=4.689$, $P=.094 > .05$, it was revealed there was not significant interaction between independent variable and intervening variable (groups and pretest written grammatical accuracy scores). So, the assumption was met in terms of homogeneity of regression slopes.

Written Grammatical Accuracy Posttest Scores

In addition, to see whether there was a significant difference between the variances of the participants' written grammatical accuracy scores in the posttest, test of homogeneity of variances was done. Table 7 shows the results.

Table 7. *Test of Homogeneity of Variances between the Participants' Posttest Written Grammatical Scores*

Dependent Variable	Posttest grammatical accuracy scores		
F	df1	df2	Sig.
.049	1	62	.826

As indicated in Table 7, the significant value .82 revealed that there was not a significant difference between the variances of the participants' posttest written grammatical accuracy scores in heterogeneous and homogeneous groups. So, the assumption of homogeneity of variances was met.

After the prerequisite assumptions were met, the posttest was administered to the participants in homogeneous and heterogeneous groups.

Table 8. *Descriptive Statistics of the Participants' Posttest Written Grammatical Scores*

Dependent Variable: Posttest scores			
Grouping types	Mean	Std. Deviation	N
Heterogeneous group	.39	.29	32
Homogeneous group	.41	.22	32
Total	.40	.20	64

As it is noticed in the Table 8, the mean score of heterogeneous group was .39 with the standard deviation of .29 ($M = .39, SD = .29$), whereas the mean score of the homogeneous group was .41 with the standard deviation of .22 ($M = .41, SD = .22$), meaning that the heterogeneous group outperformed the homogeneous group.

Also, after some preliminary analyses for ANCOVA were met, to check whether there was a significant difference between the participants' written grammatical accuracy scores in the posttest, One-way ANOVA between Groups was used. Table 9 reports the results of this ANCOVA test.

Table 9. *Dependent Variable: Posttest grammatical accuracy scores*

Source	Type III Sum of Squares	Df	Mean Square	F	Sig.	Partial Squared	Eta Squared
Corrected Model	1.651 ^a	2	.826	21.275	.000	.411	
Intercept	.150	1	.150	3.878	.053	.060	
Pretest grammatical accuracy scores	1.642	1	1.642	42.305	.000	.410	
Grouping types	.007	1	.007	.193	.000	.203	
Error	2.367	61	.039				
Total	14.243	64					
Corrected Total	4.019	63					

a. R Squared = .411 (Adjusted R Squared = .392)

As represented in Table 9, there was a significant difference between the dependent variable and the independent variable since $p = .000$. Also, the effect size, as mentioned in the partial eta squared column, was .203. Based on Cohen's (1988) guideline, there was a small effect of groups on the written grammatical accuracy scores in the posttest. In addition, 20.3 percent of the variance in the posttest written grammatical accuracy was explained by independent variable. Hence, the first null hypothesis was rejected and the answer to the research question was positive.

Discussion

The study was an effort to investigate the possible effect of homogeneous and heterogeneous groupings of Iranian university EFL learners on their written grammatical accuracy when working collaboratively. The findings showed that grammatical accuracy of both groups improved because of interaction. However, the results revealed that there was an important difference between the performances of heterogeneous pairs compared to the performance of the pairs in homogeneous group type in terms of written grammatical accuracy. In other words, collaborative writing helped the pairs in heterogeneous group to improve their grammatical accuracy more than the pairs in homogeneous group type.

In general, the results of this study can be discussed in two respects. Firstly, the obtained results revealed that in both homogeneous and heterogeneous groups the dyadic activity the learners had before writing and revising their tasks had a positive impact on the written grammatical accuracy. In fact, the positive impact of collaborative paired interaction was noticed in the learner/writers' higher grades in their posttest individual actions. It seems that part of the improvement in both groups can be justified by a feeling of responsibility that develops between the participants in pairs. This result is in line with the findings of the research done by Spector et al. (2016) in a different context. Their research shows that the effectiveness of collaborative writing is dependent on the responsibility that each student takes on.

The results for G-Hom showed that the dyadic work the participants had before writing their tasks had a positive impact on their written grammatical accuracy. The cooperation and interaction the learners had with their teammates who had nearly equal language proficiency level helped them develop their grammatical accuracy. As Storch (2007) proposes, "dyadic work provided learners with chances to share their linguistic knowledge and collaboratively construct knowledge about language" (p. 155). As for G-Het group, the findings of the study also revealed that the participants' written grammatical accuracy scores improved significantly from the pretest to posttest. Actually, the mutual interaction that they had with interlocutors who were more capable helped them to develop in this respect. The positive impact of collaborative pair interaction was noticed in the learner/writers' higher grades in their posttest individual achievements.

Hence, one justification for why there were improvements in the grammatical accuracy of the participants who carried out pair interaction is that they had "chances for learning" (Lantolf, 2000) or "affordances" (Van Lier, 2000). Regarding the heterogeneous group, the affordance and scaffolding came from the more proficient pair, while in the pairs of the homogeneous group the support was collaboratively co-constructed but not necessarily with a more proficient peer. This finding is considered to be in line with what some other researchers have already noticed concerning the advantages of homogeneous dyadic work (Donato, 1994; Kowal & Swain, 1994; Ohta, 2001; Swain & Lapkin, 1998; Villamil & de Guerrero, 1996).

Secondly, the results revealed a significant difference in terms of posttest scores of the participants in favor of those who took part in heterogeneous group. The improvement on the part

of low achievers of heterogeneous group can be justified from a sociocultural perspective, too. Vygotsky (1978) states that from the very start of life, in order to achieve development, a child is in need of interaction with a more capable society member to get help, which has been mentioned to as “Scaffolding”. Ellis (2013) also states that to gain advantage from exchanges and interactions, L2 learners are in need of having communication with somebody who has satisfying amount of proficiency in the same language to make sure that the input is not just at the student’s level, but sometimes, a little bit beyond it. This is what was found in this study, that is, learners with a lower ability of English need to get assistance and feedback from their higher proficient partners.

The progress on the part of the more capable learners in the heterogeneous group can be justified regarding sociocultural perspective as well. Van Lier (2014) claims that though Vygotsky’s work put an emphasis on the cognitive development of children, the theory can be generalizable to all learning and to both symmetrical (i.e. equal-ability) and asymmetrical (i.e. expert-novice) groupings. In this way, learners can learn from the teaching act of others. The teaching act or giving explanations to others may help L2 learners develop their language proficiency and internalize what they learnt before (Allwright, 2014).

On the whole, the results of the current study advocated the hypothesis that paired collaborative writing has a significant impact on developing grammatical accuracy of the EFL learners’ writings; hence, advocating the preceding studies ((Kuiken and Vedder, 2002; Storch and Wigglesworth, 2007). As for the effect of grouping type on grammatical accuracy, the results showed that both homogeneous and heterogeneous grouping types are effective when the participants in pairs were involved in collaborative writing. However, the total results indicate that the heterogeneous grouping type approach is better than the homogeneous grouping one to develop the learners’ written grammatical accuracy. The study, in fact, revealed that the learners in the heterogeneous group continually represented better action than those in homogeneous grouping type. Thus, the results accord with those investigations which advocate the superiority of heterogeneous grouping (Pishghadam and Ghardiri, 2011); Cooper et al., 1990; Johnson et al., 1994; Slavin, 1995; Fauziah and Latief, 2015).

Conclusions

The results of this study shows that when organizing pairs, teachers need to select learners from different proficiency levels. The study also shows that the knowledge or skill needed for scaffolding does not necessarily reside within learners but can be co-constructed by peers through collaboration. To get such successful and effective collaboration, at least one of the participants must exceed a knowledge threshold to provide a correct base from which joint scaffolding can be built. Thus, the finding that pairs with different as well as similar levels of L2 proficiency can take advantage from working and sharing ideas with one another is positive one. It reveals the fact that the act of social mediation comes not only from an expert, including teachers, but also from peers with nearly equal proficiency level, and even from the peers of little proficiency as Watanabe and Swain (2007) revealed in their study on adult ESL students. Thus, teachers should pay more attention to the fact that grouping different proficiency peers can be more provoking to L2 learning. In order that the teachers ease constructive pair work in L2 classrooms, they need to know the basics that incite learners to interact in a particular manner and may be, more significantly, and to understand how to provoke learners to work through collaboration. A number of ideas have been proposed in the related literature considering how instructors can best make “communities of learning” in which collaboration takes place and is highly regarded (for example, Lockhart and Ng, 1995).

References

- Adams, D., & Hamm, M. (1996). *Cooperative learning: Critical thinking and collaboration across the curriculum* (2nd ed.). Springfield: Charles C Thomas.
- Adodo, S. O., & Agbayewa, J. O. (2011). Effect of homogenous and heterogeneous ability grouping class teaching on students interest, attitude and achievement in integrated science. *International Journal of Psychology and Counselling*, 3(3), 48-54.
- Allwright, D. (2014). *Observation in the language classroom*. Abingdon: Routledge.
- Baer, J. (2003). Grouping and Achievement in Cooperative Learning. *College Teaching*, 51(4), 169-175. doi:10.1080/87567550309596434.
- Cady, J. L. (2011). *The Effects of implementing heterogeneous writing groups in a fifth grade classroom*. New Jersey: Rowan University.
- Ghanbari, N., & Ghaffar Samar, R. (2016). Grouping Strategies and Writing Achievement in Cooperative Learning. *Journal of Applied Linguistics and Language Research* 3,(7),85-97.
- Cooper, J., Prescott, S., Cook, L., Smith, L., Mueck, R & Cuseo, J. (1990). *Cooperative learning and college instruction: Effectiveness of student learning teams*. California State University Foundation. Long Beach, CA.
- Donato, R. (1994). *Collective scaffolding in second language learning*. In J. P. Lantolf, & G. Appel. New Jersey: Ablex.
- Doughty, C., & Pica, T. (1986). "Information Gap" Tasks: Do They Facilitate Second Language Acquisition? *TESOL Quarterly*, 20(2), 305-325. doi 3586546/10/10, 2307.
- Doughty, C., & Williams, J. (2013). *Focus on form in classroom second language acquisition*. Cambridge: Cambridge University Press.
- Ellis, N. C. (2013). *Second language acquisition*. In *The Routledge handbook of second language acquisition*. Oxford: Oxford University Press.
- Fauziah, H., & Latief, M. (2015). The effect of working in heterogeneous and homogeneous Pairs on the students' writing skill. *Arab World English Journal (AWEJ) Vol.6. No.2*.
- Heidar, D. M. (2016). ZPD-assisted intervention via web 2.0 and listening comprehension ability. *English for Specific Purposes World*, 17(4), 1–17.
- Johnson, D.W., & Johnson, R.T. (1994). *An overview of cooperative learning*, Baltimore: Brookes Press.
- Kian-sam, H. (1999). Cooperative CBI: The effects of heterogeneous versus homogeneous grouping, student ability and learning accountability on achievement. *Educational Research Journal*, 14(2), 301-313.
- Kowal, M., & Swain, M. (1994). Using collaborative language production tasks to promote students' language awareness. *Language Awareness*, 3(2), 73-93. doi:10.1080/09658416.1994.9 959845.
- Kowal, M. and Swain, M. (1997) From semantic to syntactic processing: How can we promote it in the immersion classroom? In K. Johnson and M. Swain (eds) *Immersion Education: International perspectives* (pp. 284–309). Cambridge: Cambridge University Press.
- Kruse, A. J. (2011). *Proceedings from CIED 500 principle of educational research. The benefits of heterogeneous groupings over homogeneous groupings*. Minneapolis: University of Saint Thomas.
- Kuiken, F., & Vedder, I. (2002). The effect of interaction in acquiring the grammar of a second language. *International Journal of Educational Research*, (3), 343-358. doi:https://doi.org/10.1016/S0883-0355(03)00009-0

Lantolf, J. P. (2000). *Introducing sociocultural theory*. In J. P. Lantolf (Ed.), *Sociocultural theory and second language learning*. Oxford: Oxford University Press.

Leeser, M. J. (2004) Learner proficiency and focus on form during collaborative dialogue. *language Teaching Research* 8, 55–82.

Lightbown, P., & Spada, N. (1999). *How languages are learned* (2nd ed.). Oxford: Oxford University Press.

Lockhart, C., & Ng, P. (1995). Analyzing Talk in ESL Peer Response Groups: Stances, Functions, and Content. *Language Learning*, 45(4), 605-651. doi:10.1111/j.1467-1770.1995.tb00456x

Maftoon, P., & Ghafoori, N. (2009). A comparative study of the effect of homogeneous and heterogeneous collaborative interaction on the development of EFL learners' writing skill. *The Journal of Applied Linguistics*, 2(1), 127-158.

Mahenthiran, S & Rouse, P. J. (2000). The impact of group selection on student performance and satisfaction. *International Journal of Educational Management*, 14(6), 255-264. doi:10.1108/09513540010348043.

Pishghadam, R., & Ghardiri, S. (2011). Symmetrical or asymmetrical scaffolding: Piagetian vs. Vygotskian views to reading comprehension. *Journal of Language and Literacy Education*, 7(1), 49-64.

Slavin, R. E. (1995). *Cooperative learning: Theory, research, and practice*, 2nd Ed., Boston: Allyn & Bacon.

Spector, J. M., Ifenthaler, D., Samspon, D., Yang, L., Mukama, E., Warusavitarana, A., Lokuge Dona, K., Eichhorn, K., Fluck, A., Huang, R., Bridges, S., Lu, J., Ren, Y., Gui, X., Deneen, C. C., San Diego, J., & Gibson, D. C. (2016). Technology enhanced formative assessment for 21st century learning. *Educational Technology & Society*, 19(3), 58–71.

Stone, C.A. (1993) What is missing in the metaphor of scaffolding? In E.A. Forman, N. Minick and C.A. Stone (eds) *Contexts for Learning: Sociocultural Dynamics in Children's Development* (pp. 169–183). Oxford: Oxford University Press.

Storch, N. (1999). Are two heads better than one? Pair work and grammatical accuracy. *System*, 27(3), 363-374. doi:https://doi.org/10.1016/S0346-251X(99)00031-7.

Storch, N. (2001). How collaborative is pair work? ESL tertiary students composing in pairs. *Language Teaching Research*, 5, 29–53.

Storch, N. (2002a). Patterns of interaction in ESL pair work. *Language Learning*, 52, 119–58.

Storch, N. (2007). Investigating the merits of pair work on a text editing task in ESL classes. *Language Teaching Research*, 11, 143-159.

Storch, N., & Wigglesworth, G. (2007). Writing tasks: The effects of collaboration. In M. Garcia Mayo (Ed.), *Investigating tasks in formal language learning* (pp. 157-177). Clevedon, UK: Multilingual Matters.

Swain, M. (2001) Examining dialogue: Another approach to content specification and to validating inferences drawn from test scores. *Language Testing* 18, 319–346.

Swain, M. and Lapkin, S. (2002) Talking it through: Two French immersion learners' response to reformulation. *International Journal of Educational Research* 37, 285

Van Lier, L. (2000). *From input to affordance: Socio-interactive learning from an ecological perspective*. In J. P. Lantolf (Ed.), *Sociocultural theory and second language learning*. Oxford: Oxford University Press.

Van Lier, L. (2014). *Interaction in the language curriculum: Awareness, autonomy and authenticity*. New York, NY: Routledge.

Villamil, O. S., & de Guerrero, M. C. (1996). Peer revision in the L2 classroom: Social-cognitive activities, mediating strategies, and aspects of social behavior. *Journal of Second Language Writing*, 5(1), 51-75. doi:[https://doi.org/10.1016/S1060-3743\(96\)90015-6](https://doi.org/10.1016/S1060-3743(96)90015-6).

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

Watanabe, Y., & Swain, M. (2007). Effects of proficiency differences and patterns of pair interaction on second language learning: Collaborative dialogue between adult ESL learners. *Language Teaching Research*, 11(2), 121-142.

Webb, N. M. (1989). Peer interaction and learning in small groups. *International Journal of Educational Research*, 13(1), 21-39. doi:[https://doi.org/10.1016/0883-0355\(89\)900141](https://doi.org/10.1016/0883-0355(89)900141)

Yule, G. and Macdonald, D. (1990) Resolving referential conflict in L2 interaction: The effect of proficiency and interactive role. *Language Learning* 40, 539–556.

Zamani., M. (2016). Cooperative learning: Homogeneous and heterogeneous grouping of Iranian EFL learners in a writing context. *Cogent Education*,30, 79–123.

