Collaborative and Individual Writing: Effects on Accuracy and Fluency Development

MEHDI SARKHOSH
Urmia University

SHIVA NAJAFI
Urmia Azad University

ABSTRACT: The present study investigates the differential effects of collaborative vs. individual writing approaches on the development of fluency and accuracy among male and female EFL learners. The study is unprecedented in terms of investigating the effect of these two approaches on fluency and accuracy development in the long-run through a delayed posttest. Additionally, the study examines the development of fluency and accuracy over seven successive sessions longitudinally. The written outputs, produced by the participants, were scored for fluency and accuracy according to the method applied by Wigglesworth and Storch (2009). The Data, analyzed through MANCOVA and ANOVA analyses, revealed that collaborative writing led to more fluent texts for both males and females in the short and the long run. Also, the collaboratively written compositions were more accurate in comparison to those written individually for males and females in the short and long run.
Keywords: Sociocultural Theory, Collaborative Writing, Fluency, Accuracy

Escritura colaborativa e individual: Efectos en el desarrollo de la precisión y fluidez

RESUMEN: Este estudio investiga los efectos diferenciales de la redacción colaborativa versus la versión individual y sus acercamientos del desarrollo de la precisión y fluencia entre los aprendices varones y mujeres EFL. El estudio no tiene precedente alguno en término de investigación de los efectos de estos dos acercamientos en un largo manejo a través de post prueba (posttest) dilatada. Además, el estudio comprueba el desarrollo de la precisión y fluencia en siete sesiones sucesivas longitudinalmente. Las producciones escritas hechas por los participantes fueron apuntadas para la fluencia y precisión en concordancia con el método aplicado por Wigglesworth y Storch (2009). Este dato analizado a través de la Mancova y Anova y sus análisis revela que la redacción colaborativa dirigida a más textos fluentes para ambos, varones y mujeres, en un manejo largo y un manejo corto. También las composiciones escritas eran más precisas en comparación con aquellas escritas individualmente para los aprendices varones y mujeres en el manejo largo y el corto.
Palabras clave: Teoría sociocultural, redacción colaborativa, fluencia, precisión

1. INTRODUCTION

With the advent of the communicative approach, collaborative learning (CL) such as pair and group work grew more popular. As a consequence, a considerable number of studies tried to investigate the effectiveness of CL situations on L2 learning (e.g., Dobao, 2012; Storch,

2005; Swain & Lapkin, 1998, among others). CL, grounded in Vygotsky's Socio-Cultural Theory (SCT), is a social activity in which the development of learners' knowledge hinges upon the interaction with others (Vygotsky, 1978). In this regard, among language skills, the writing skill appears to be the skill at an advantage (Villamil & de Guerrero, 1996, 1998) since the needed collaboration between the novice (the student) and the expert (either the classmate or the teacher) can easily shape in the classroom. Collaborative writing (CW) is defined as "the production of a single text by co-authors or group authors" (Ede & Lunsford, 1985: 14), and has been paid much attention to since early 1970's since Bruffee (1973) contended that students were able to produce better work when they worked collaboratively. Therefore, it seems that the amount and type of scaffolding or peer feedback can be crucial (Aljaafreh & Lantolf, 1994). In the same vein, Farrah (2012) reported that peer feedback is welcomed by learners and it can provide the opportunity for social interaction.

Researchers have reported a number of advantages for the use of collaboration in wiring. According to Reid (1993), CW can assist learners get well motivated; it also helps students to be more flexible about risk-taking. Foster (1998) notes that CW can improve students' interaction in the classroom. Further, Gousseva-Goodwin (2000) found that learners who wrote collaboratively received higher grades in comparison with those who wrote individually. Other studies have also divulged that the implementation of CW improves learners' production and leads to meaningful revisions (see Storch, 2011).

Although some researchers have tried to investigate the effects of CW on fluency and accuracy (e.g., Storch, 2002; Wigglesworth & Storch, 2009, among others), the long-term effects of collaboration on students' writing have not been amply explored. No study has investigated the long-term effects of CW on the development of accuracy and fluency among EFL learners. In other words, no study has explored the effectiveness of CW through a delayed posttest. Furthermore, no researcher has investigated the development of both accuracy and fluency in writing over several consecutive sessions. The present study attempted to contribute to the existing literature on collaborative vs. individual writing by probing into the effects of CW on both accuracy and fluency development of male and female learners in both the short and the long-run.

2. Review of Literature

CL, which is defined as "a form of indirect teaching in which the teacher sets the problem and organizes students to work it out collaboratively" (Bruffee, 1984: 418), is one of the key characteristics of communicative language teaching (Neumann & McDonough, 2014). Many researchers have tried to investigate the possible effects of collaboration on second language learning (Elola & Oskoz, 2010; Louth, McAllister & McAllister, 1993; Storch, 2001, 2002, 2013). CL is based on Vygotsky's SCT, in which learning is defined as a social activity and learners' knowledge development depends on the interaction with others (Vygotsky, 1978). Accordingly, social context is the core of communication and learning, and social interactions have a significant effect on learning (Vygotsky, 1962). As a consequence, SLA researchers have attempted to investigate second language learning from a sociocultural perspective (see e.g., Donato & Mccormick, 1994).

The zone of proximal development (ZPD), as one of the cornerstones of SCT, suggests that there is a difference between an individual's performance while he/she acts individually, and when he/she is supported by more knowledgeable or capable peers (Vygotsky, 1978). From Vygotsky's point of view, "what the child is able to do in collaboration today he[/she] will be able to do independently tomorrow" (Vygotsky, 1978: 211). The concept of ZPD is closely linked to the notion of scaffolding, which suggests that a knowledgeable person like a teacher or an adult can help a learner to finish a task that he is not able to accomplish by him/herself (Bruner, 1986). According to Donato (1994), scaffolding is a framework for peer interaction and individuals can scaffold each other. Since each learner has different areas of strengths and weaknesses, scaffolding can be used as an invaluable resource for learners while performing at the ZPD level (Ohta, 2001).

The evidence from SCT reveals that both instructors and students can benefit from collaborative techniques. Alghasab, Hardman and Handley (2019), in a Wiki-mediated CW instruction, found the positive and significant role of teacher in promoting CW and stressed that dialogic engagement between teacher and students impacted positively on CW process. Block (1996) suggests that language classrooms are essentially social events. Therefore, students can gain more learning opportunities through collaborative activities. CW, as a way of developing second language competence, is based on SCT. Storch (2005) indicated that during a CW, task writers decide how to express their own ideas in the target language so that they can construct structures to express these ideas. Although writing has been usually deemed as an individual activity (Wigglesworth & Storch, 2009), it is possible to create a situation in which learners may benefit from social interactions in writing classes through collaborative or group activities. The product of CW is a document that learners produce through meaningful interactions, shared decision-making and responsibility for the document (Allen, Atkinson, Morgan, Moore, & Snow, 1987).

A number of benefits have been reported for collaborative activities (see Storch, 2013). For example, learners suggest that collaborative learning situations can provide more learning opportunities (Storch, 2005). Moreover, it has been reported that knowledge can be co-constructed through collaborative activities (Swain & Lapkin, 1998). Most of the students who have had the experience of working in groups and pairs suggest that there are more ideas and creativity during collaborative working (Dobao & Blum, 2013), and collaboration during a writing task can lead the students to grammatical accuracy and better use of vocabulary. In the same vein, Saunders (1989) showed that collaboration between learners could help them to complete the task successfully during co-writing activities. Bikowski and Vithanage (2016) examined the impact of repeated in-class web-based CW tasks on students' (L2) writing scores. Final findings revealed that the participants valued the collaborative in-class writing tasks and those in the individual writing group wished they had been in-class collaborative web-based writing group.

Regarding collaborative vs. individual writing, some researchers have tried to compare the performance of students when they accomplish a writing task collaboratively with those who write individually. Kuiken and Vedder (2002), in their study, found no significant difference between the performance of the students who wrote collaboratively and those who wrote individually. On the contrary, the majority of studies on individual vs. CW, have found more gains from CW. The nature of interactions during CW activities was investigated by analyzing Language Related Episodes (LREs), which are considered as the instances of

collaborative dialogues and are defined as ''any part of a dialogue where the students talk about the language they are producing, question their language use, or correct themselves or others'' (Swain & Lapkin, 1998: 328). The results of this study showed that LREs provided opportunities for language learning, which means that knowledge can be co-constructed through collaborative dialogues. Dobao (2012) investigated LREs between the learners who were working in pairs and in small groups during their negotiations throughout the CW task. The findings revealed that both the students working in pairs and in groups focused on language relatively often; groups made more LREs and more correctly resolved LREs than pairs. Chen and Yu (2019), as well, showed that discrepancies in the attitudes of learners toward CW influenced their patterns of interaction and LREs.

Patterns of interaction during collaborative activities were investigated by Storch (2001, 2002). Storch (2001) found that the quality of the final product is affected by the level of collaboration. In another study, Storch (2002) reported that collaborative work had better results as the learners preferred to work in groups.

Recently, researchers have focused on understanding what effects CW situations may have on the development of fluency, accuracy and complexity. For example, Dobao (2012) investigated the effects of the number of participants in a writing task on fluency, accuracy and complexity. According to the results, groups wrote more accurate compositions in comparison to pairs and individuals; however, the participants in the individual group wrote longer texts in comparison to those in pairs and groups. The performance of individuals, pairs and groups was similar in terms of complexity. In a study on the impact of collaborative planning on writing quality, Kang and Lee (2019) studied the impact of collaborative planning on L2 writing and found it to be effective in enhancing writing fluency and complexity but not in improving accuracy. In another study, Dobao and Blum (2013) investigated the differences between the writing performance of the students who worked in pairs, in groups of four, and individually. Although the students benefited from collaborative working both in pairs and in groups of four, they had more opportunities for exchanging ideas on knowledge in groups of four since the group had more members and each member had different areas of strength.

Storch (2005) revealed that although pairs took more time to complete the task, they produced shorter texts in comparison to individuals. The texts produced by pairs were grammatically more accurate and also more complex. Pair dialogues were examined for planning, writing and revision phases. The analysis showed that the planning phase was very brief and most of the time was spent on the writing one, and the revision phase was also very short. Similar results were found by Wigglesworth and Storch (2009). They investigated the differences between collaborative and individual writing approaches in terms of fluency, accuracy and complexity. The findings showed that there were not any significant differences between the collaborative and the individual groups in terms of fluency and complexity. However, the participants in the collaborative group produced more accurate pieces in comparison to those in the individual group. Similarly, Jafari and Nejad-Ansari (2012) found that the students who worked collaboratively produced more accurate texts in comparison to those who worked individually. In line with Wigglesworth and Storch (2009), Biria and Jafari (2013) reported that collaboration does not affect fluency significantly. On the other hand, Soleimani, Modirkhamaneh and Sadeghi (2015) reported that the performance of the students in the collaborative group was significantly better than those in the individual group in terms of fluency and accuracy but not complexity.

Some researchers have also reported the benefits of collaborative prewriting. Neumann and McDonough (2014) found that, having been given a collaborative prewriting task, English L2 university students produced paragraphs which received higher analytic ratings compared to those who experienced individual prewriting. McDonough, De Vleeschauwer, & Crawford (2018) reported that Thai EFL students did not show any significant differences for syntactic complexity measures but produced problem/solution paragraphs, which were more accurate than paragraphs written by individual planners.

The literature above reveals that no study has investigated the long term effects of CW on the development of fluency and accuracy through a delayed posttest. Since the majority of studies reviewed above have reported on the positive effects of CW in writing, investigating the long term effects of CW, which is missing in the literature, seems warranted and propitious. The present study attempted to fill this gap by studying the effects of CW on accuracy and fluency development of males and females after several successive writing sessions. For this purpose, the following four research questions were formulated:

1. Is there a difference between collaborative and individual writing approaches in terms of their immediate effect on the accuracy development of male and female learners?
2. Is there a difference between collaborative and individual writing approaches in terms of their delayed effect on the accuracy development of male and female learners?
3. Is there a difference between collaborative and individual writing approaches in terms of their immediate effect on the fluency development of male and female learners?
4. Is there a difference between collaborative and individual writing approaches in terms of their delayed effect on the fluency development of male and female learners?

3. Method

3.1. Design and Participants

The design of the present study is an experimental pretest, immediate, delayed posttest design, which takes place over seven successive sessions and renders this study a longitudinal one. There are two independent variables: a) writing approach at two levels of individual or collaborative, b) sex at two levels of male or female; therefore, the study enjoys a (2*2) factorial design. There are two dependent variables: a) accuracy, b) fluency, which are investigated among male and female participants in the short and long run. The participants in the present study included 120 students at the elementary level of proficiency in Iran Language Institute (a popular private language school in Urmia, Iran). According to the institutes' policies, learners have to take part in a placement test in order to be placed in the appropriate level, based on their language proficiency. In addition, at the end of each term, learners take an achievement test in order to be eligible to pass to the next level. Therefore, the learners in each level are almost at the same level of proficiency. The age range of the participants was between 12 and 14 (mean = 12.5). Although the participants knew Persian (the official language of the country) pretty well, their mother tongue was Turkish since they lived in Urmia (a Turkish speaking city in the north west of Iran).

3.2. Procedure

Four classes (two male and two female classes) with thirty students in each were selected randomly from among elementary classes. In one male and one female class, the students were supposed to work collaboratively with their self-selected pairs.; however, one male and one female class were asked to write individually. The experiment lasted for seven sessions. The participants were cautioned against absenteeism and were promised course credit in return for their attendance throughout the term, which paid off and no one missed any of the seven sessions. The researchers selected topics (see Appendix), and the participants were supposed to write descriptive compositions based on these topics. Learners were asked to write at least a 100 to 150 word composition each session. The writing time was determined according to IELTS writing task one, which is 20 minutes for a 150 word composition. The first session was considered as the pretest in which the teacher researchers asked the participants to write a composition on the given topic. All the participants were asked to write the pretest compositions individually. These papers were then analyzed to make sure that the participants were homogenized in terms of their writing skill. Moreover, the results of the pretest were used to be compared with those of the posttests in order to investigate the effects of collaborative and individual writing approaches. The compositions were taken home, were corrected and returned to the participants in the second session. In the second session, the participants were given some time to reflect on their errors and the corresponding corrections. The feedback on the grammatical errors made by the learners was unfocused and targeted all types of grammatical errors except for spelling errors. This process of error correction and feedback was followed after each writing task during the present study. In the second session, the students in the collaborative groups were asked to write collaboratively with their pairs and those in the individual groups were supposed to write the compositions by themselves. This resumed for six sessions both in the collaborative and the individual groups. Session six was the immediate posttest session after which there was a one month interval. After a month, the students in the collaborative and individual groups were asked to write the seventh composition, which was considered as the delayed posttest.

3.3. Scoring compositions for accuracy and fluency

The data was analyzed for accuracy and fluency according to the method applied by Wigglesworth and Storch (2009). Fluency was measured based on, a) the average number of words, b) the average number of clauses (dependent clauses and independent clauses), and c) the average number of T-units. The compositions were typed in MS Word 2013 in order to count the average number of words and calculate the length of the compositions. Then, the average number of clauses (dependent clauses and independent clauses) and T-units were calculated. Accuracy was measured by, a) the global units expressed in terms of the proportion of error-free T-units of all T-units (EFT/T) and, b) error-free clauses of all clauses (EFC/C).

3.4. Data analysis

The SPSS software, version 21 was employed to analyze the data. Before embarking on the analyses, the assumptions of Multivariate Analysis of Covariance (MANCOVA) including

normality of variables across groups, homogeneity of regression slopes, and homogeneity of variance and covariate matrix were tested. These assumptions were tested for each MAN-COVA analysis. However, due to the large number of research questions and considerable amount of analyses in the present study, some of the tests are very briefly reported below and the descriptive statistics and some other less necessary statistics and the related tables are excluded.

4. Results

Due to massive data analysis employed in the present study and in order to save space, normality tests and some other analyses are not presented in tables and we sufficed to report them briefly. Before each MANCOVA test, Kolmogorov-Smirnov Test was run in order to investigate the normality of the scores. Regarding fluency, for both male and female groups, the scores for average number of words, average number of clauses and average number of t-units, at pretest, immediate and delayed posttests were normally distributed ($p > 0.05$). Regarding accuracy, for both male and female groups, the scores for the number of correct clauses and number of correct t-units at pretest, immediate and delayed posttests were normally distributed ($p > 0.05$). Therefore, parametric statistics were utilized throughout the following analyses.

Before each MANCOVA test, the homogeneity of regression slopes test was run in order to investigate the relationship between the covariates and the dependent variable for each group and for each analysis. This test showed that there were no interactions between the covariate and the treatment or the experimental manipulation ($p > 0.05$).

Before each MANCOVA test, the homogeneity of variance and covariate matrix test was run in order to investigate the homogeneity of variance and covariance matrix across groups. Throughout the following analyses, the F statistic of Box's M Test was significant at 0.05 ($p < 0.001$) alpha level. According to this significant result, it is concluded that the variance and covariance matrix of research scores are not homogeneous across groups. According to Tabachnick & Fidell (2007), Pillai's criterion must be used instead of Wilks' lambda in order to evaluate multivariate significance.

Throughout the following analyses, MANCOVA was utilized in order to investigate the differences between collaborative and individual writing approaches in terms of their immediate and delayed effects on fluency and accuracy development of male and female learners. Statistical test of MANCOVA was bilateral with Type I error set at 5%. The partial eta-squared (η_p^2) was used in order to measure the effect size, considering that a partial eta-squared of 0.01 was small, 0.04 moderate, and 0.1 large (Huberty, 2002).

4.1. Results for fluency

4.1.1. Fluency at the immediate posttest:

Table 4.1. Result of multivariate test for the fluency of male and female groups at the immediate posttest.

TEST	MULTIVARIATE TEST (PILLAI'S TRACE)	F	Df1	Df2	p	η_p^2
Females	0.6	27.55	3	53	0.001	0.61
Males	0.51	18.99	3	53	0.001	0.52

After controlling for the pretest scores, the F statistic of group membership is significant for both male (F=18.99, $p<0.005$) and female groups (F=27.55, $p<0.005$). This finding shows that there are significant differences between the collaborative and the individual male and the collaborative and the individual female groups at the immediate posttest in terms of fluency. Partial eta-squared, for group membership, is 0.61 for female groups and 0.52 for male groups, which shows that the differences between the collaborative and the individual groups in fluency development in the statistical population, are large and significant. This shows that CW yields more fluent texts at the immediate posttest in both male and female groups. ANCOVA was used in order to examine the groups' differences in the components of fluency (average number of words, clauses and t-units).

Table 4.2. Result of ANCOVA to compare the female groups and male groups in terms of the components of fluency at the immediate posttest.

TEST	VARIABLE	SUM OF SQUARES (GROUP)	SUM OF SQUARES (ERROR)	MEAN SQUARE (GROUP)	MEAN SQUARE (ERROR)	F	p	η_p^2
	Words	2627.25	49246.6	2627.25	895.39	2.93	0.09	0.05
Females	Clauses	2608.54	3920.07	2608.54	71.27	36.59	0.001	0.40
	t-units	1400.49	1119.31	1400.49	20.35	68.81	0.001	0.56
	Words	23407.3	37880.6	23407.3	688.73	33.98	0.001	0.38
Males	Clauses	1475.79	2475.42	1475.79	45.01	32.79	0.001	0.37
	t-units	823.81	773.85	823.81	14.07	58.55	0.001	0.52

According to ANCOVA, the F statistic of group membership is significant in female groups for the average number of clauses (F=36.59, $p<0.005$) and t-units (F=68.81, $p<0.005$), and in male groups for the average number of words (F=33.98, $p<0.005$), clauses (F=32.79, $p<0.005$) and t-units (F=58.55, $p<0.005$). However, the F statistic of group membership, for

the average number of words (F=2.93, p>0.05) is not significant in female groups. Partial eta-squared for clauses and t-units is 0.40 and 0.56 for females. In male groups, partial eta-squared for words, clauses and t-units is 0.38, 0.37 and 0.52, respectively. This result indicates that there is a significant difference between the collaborative and individual groups in terms of the average number of clauses and t-units in female groups and in terms of the average number of words, clauses and t-units in male groups.

4.1.2. Fluency at the delayed posttest:

Table 4.3. Result of multivariate test for the fluency of female groups and male groups at the delayed posttest.

TEST	MULTIVARIATE TEST (PILLAI'S TRACE)	F	Df1	Df2	p	η_p^2
Females	0.65	32.80	3	53	0.001	0.65
Males	0.59	25.65	3	53	0.001	0.5

After controlling for the pretest scores, the F statistic of group membership is significant for both female groups (F=32.80, p<0.005) and male groups (F=25.65, p<0.005). This finding shows that there are significant differences between the collaborative and the individual male and the collaborative and the individual female groups at the delayed posttest in terms of fluency. Partial eta-squared, for group membership, is 0.65 for female groups and 0.59 for male groups, which shows that the differences between the collaborative and the individual groups in fluency development in the statistical population, are large and significant. Therefore, CW yields more fluent texts at the delayed posttest in both the male and female groups. ANCOVA was used in order to examine the groups' differences in the components of fluency development.

Table 4.4. Result of ANCOVA to compare the female groups and male groups in terms of the components of fluency at the delayed posttest.

TEST	VARIABLE	SUM OF SQUARES (GROUP)	SUM OF SQUARES (ERROR)	MEAN SQUARE (GROUP)	MEAN SQUARE (ERROR)	F	p	η_p^2
	Words	1535.74	52518.17	1535.74	954.87	1.60	0.21	0.03
Females	Clauses	2387.40	2387.40	2387.40	75.97	31.42	0.001	0.36
	t-units	1559.09	1559.09	1559.09	21.08	73.96	0.001	0.57
	Words	149.42	934028.12	149.42	16982.32	0.01	0.92	0
Males	Clauses	1358.48	2278.89	1358.48	41.43	32.78	0.001	0.37
	t-units	1072.03	756.57	1072.03	13.75	77.93	0.001	0.59

After controlling for the pretest scores, the F statistic of group membership is significant in female groups for the average number of clauses (F=31.42, $p<0.005$) and t-units (F=73.96, $p<0.005$). Similarly, in male groups, the F statistic of group membership is significant for the average number of clauses (F=32.78, $p<0.005$) and t-units (F=77.93, $p<0.005$). However, the F statistic of group membership, for the average number of words is not significant in the female groups (F=1.60, $p>0.05$) and male groups (F=0.01, $p>0.05$). Partial eta-squared for clauses and t-units is 0.36 and 0.57 for females, and 0.37 and 0.59 for males, respectively, which indicates that there is a significant difference between the collaborative and the individual groups in terms of the average number of clauses and t-units in female and male groups at the delayed posttest.

4.2. Results for accuracy

4.2.1. Accuracy at the immediate posttest:

After controlling for the pretest scores, the F statistic of group membership is significant for both female groups (F=28.58, $p<0.005$), and male groups (F=28.33, $p<0.005$).

Table 4.5. Result of multivariate test for the accuracy of female groups and male groups at the immediate posttest.

Test	Multivariate test (Pillai's Trace)	F	Df1	Df2	p	η_p^2
Females	0.51	28.58	2	55	0.001	0.51
Males	0.50	28.33	2	55	0.001	0.51

This finding shows that there are significant differences between the collaborative and the individual male and the collaborative and the individual female groups at the immediate posttest in terms of accuracy. Partial eta-squared, for group membership, is 0.51 for both the female groups and male groups, which shows that the differences between the collaborative and the individual groups in accuracy development in the statistical population, are large and significant. Therefore, CW gives more accurate texts at the immediate posttest in both the male and female groups. ANCOVA was used in order to examine the groups' differences in the components of accuracy.

Table 4.6. Result of ANCOVA to compare the female groups and male groups in terms of the components of accuracy at the immediate posttest.

TEST	VARIABLE	SUM OF SQUARES (GROUP)	SUM OF SQUARES (ERROR)	MEAN SQUARE (GROUP)	MEAN SQUARE (ERROR)	F	p	η_p^2
Females	Correct Clauses	2168.69	2185.01	2168.69	39.01	55.58	0.001	0.49
	Correct t-units	1101.81	1083.55	1101.81	19.24	56.94	0.001	0.50
Males	Correct Clauses	1808.52	1789.26	1808.52	31.95	56.60	0.001	0.50
	Correct t-units	686.41	711.07	686.41	12.69	54.05	0.001	0.49

The F statistic of group membership is significant in female groups for correct clauses (F=55.58, $p<0.005$) and correct t-units (F=56.94, $p<0.005$). Similarly, in male groups, the F statistic of group membership is significant for correct clauses (F=56.60, $p<0.005$), and correct t-units (F=54.05, $p<0.005$). Partial eta-squared for correct clauses and correct t-units is 0.49 and 0.50 for females, and 0.50 and 0.49 for males, respectively. This result indicates that there is a significant difference between the collaborative and the individual groups in terms of the proportion of correct clauses and correct t-units to the total number of clauses and t-units in the female and male groups.

4.2.2. Accuracy at the delayed posttest:

Table 4.7. Result of multivariate test for the accuracy of female groups and male groups at the delayed posttest.

TEST	MULTIVARIATE TEST (PILLAI'S TRACE)	F	Df1	Df2	p	η_p^2
Females	0.59	39.03	2	55	0.001	0.59
Males	0.59	39.73	2	55	0.001	0.59

The F statistic of group membership is significant for both female groups (F=39.03, $p<0.005$) and male groups (F=39.73, $p<0.005$). This divulges that there are significant differences between the collaborative and the individual male and the collaborative and the individual female groups at the delayed posttest in terms of accuracy. Partial eta-squared, for group membership, is 0.59 both for female groups and male groups, which shows that the differences between the collaborative and the individual groups in accuracy development in the statistical population are large and significant. This reveals that CW produces more accurate texts at the delayed posttest in both male and female groups. ANCOVA was used in order to examine the groups' differences in the components of accuracy.

Table 4.8. Result of ANCOVA to compare the female groups and male groups in terms of the components of accuracy at the delayed posttest.

TEST	VARIABLE	SUM OF SQUARES (GROUP)	SUM OF SQUARES (ERROR)	MEAN SQUARE (GROUP)	MEAN SQUARE (ERROR)	F	p	η_p^2
Females	Correct Clauses	3048.90	2726.29	3048.90	48.68	62.63	0.001	0.53
	Correct t-units	1448.56	1030.44	1448.56	18.40	78.72	0.001	0.58
Males	Correct Clauses	2618.17	1903.32	2618.17	33.98	77.03	0.001	0.58
	Correct t-units	1002.68	761.39	1002.68	13.59	73.74	0.001	0.57

The F statistic of group membership is significant in female groups for correct clause (F=62.63, $p<0.005$) and correct t-units (F=78.72, $p<0.005$). Moreover, the F statistic of group membership is significant in male groups for correct clauses (F=77.03, $p<0.005$) and correct t-units (F=73.74, $p<0.005$). Partial eta-squared for correct clauses and correct t-units is 0.53 and 0.58 for females, and 0.58 and 0.57 for males, respectively, which tells that there is a significant difference between the collaborative and the individual groups in terms of the proportion of correct clauses and correct t-units to the total number of clauses and t-units in female and male groups at the delayed posttest.

5. DISCUSSION

The objective of the present study was to investigate the effects of collaborative and individual writing approaches on the development of writing fluency and accuracy for males and females. The study, particularly, intended to investigate the long-term effects of these two approaches on the development of writing fluency and accuracy through a delayed posttest. Analyzing the results of the immediate posttest scores revealed that the performance of the female learners in the CW group was significantly better than the performance of those in the individual writing group in terms of fluency. Moreover, data analysis disclosed that the texts written by the female collaborative group were more fluent in comparison to those written by the female individual group at the delayed posttest. In terms of accuracy, the results showed that the writing pieces of the collaborative female group were more accurate than the texts written by the individual group. Accuracy at the delayed posttest in the female collaborative group was also significantly better in comparison to that in the individual group. Data analysis in the male group showed that the performance of male learners in the collaborative group was significantly better than those in the individual group in terms of fluency at the immediate posttest. Moreover, the texts written by the male collaborative group were more fluent in comparison to those written by the individual group at the delayed

posttest. In terms of accuracy, it was found that writing pieces of the collaborative male group were more accurate than the texts written by the individual group. Accuracy at the delayed posttest in the male collaborative group was also significantly better than that in the individual group. Further, based on partial eta squared of MANCOVA, in the collaborative groups, females had better performance in comparison to males in terms of fluency both at the immediate and delayed posttests. However, the performance of males and females was equal in terms of accuracy.

The findings of the present study can be discussed in light of Vygotsky's (1987) SCT, which suggests that learning is a social activity and students can learn through interaction with others. During the process of collaboration, students can scaffold each other in order to reach higher mental development. This aspect of SCT can positively affect writing accuracy in as much as students can share knowledge and ideas in order to select the best grammatical structures. Moreover, students can provide useful grammatical feedback for their pairs. Collaborative activities can also be worthwhile in terms of fluency development as students can negotiate and scaffold each other in order to have a better vocabulary use since using first language as one of the features of CW (Fung, 2010) can help students ask their peers the meaning of the words they do not know.

The results of this study revealed that CW can be beneficial for the students at the elementary level of proficiency. This finding is complementary to Storch (1998) and Williams (2001), who found that collaborative activities are more effective when they are applied at advanced levels. The value of collaborative work has been corroborated in the previous studies (e.g., Louth *et. al.*, 1993; Soleimani *et. al.*, 2015; Storch, 2005). This study also provides strong evidence on the effectiveness of CW, both in the short and long run. The results of the present study are consistent with the studies that reported the students who write collaboratively can produce more grammatically accurate texts in comparison to those who write individually (Storch, 1999, 2005; Storch & Wigglesworth, 2007; Wigglesworth & Storch, 2009). In addition, in line with several recent studies (e.g. Dobao, 2012; Soleimani, *et al.*, 2015), this study confirms that the CW approach can positively improve fluency in writing.

The findings of the present study are, however, in contrast to Storch (2005), Storch and Wigglesworth (2009) and Biria and Jafari (2013), who reported collaboration does not affect writing fluency. The contrast between these studies and the current study, only as regards fluency development, may be due to the small sample size of the previous studies. Moreover, the previous studies investigated the effects of collaboration on writing fluency in a single research session. It seems that since the present study investigated the effects of CW after several successive sessions, it led to more fluent writing pieces.

Since this study is unprecedented in terms of investigating the long term effects of CW, it seems that more research is needed to shed light on the effectiveness of CW in terms of its delayed effects. Our findings confirm the findings of most of the studies on CW and the idea that CW can be used as an effective approach in order to improve students' writing skill in terms of both fluency and accuracy. CW has proved effective in developing writing skill in L2; it demands reflective thinking, helps learners to focus on grammatical accuracy, lexis and discourse, and it encourages sharing of knowledge about the language (DiCamilla & Anton, 1997; Donato, 1994; Hirvela, 1999; Storch, 2002; Swain & Lapkin, 1998). Accordingly, it behooves teachers who, due to lack of time or other reasons, are not willing to apply CW in their classes, to reconsider CW and the amount of time they devote

to this activity. It seems that, in this way, students can become more autonomous language learners as they exchange knowledge, skills, and strategies. Growing cognizant of the advantages of collaboration, more capable students can guide and help the weaker partners through scaffolding and encourage them to develop their independence and responsibility to construct knowledge on their own.

6. Conclusion

The present study investigated the impact of individual versus CW instruction on the development of accuracy and fluency in the short and the long run. The final findings divulged that both male and female students who were in the CW groups, yielded more accurate and fluent texts both in the immediate and delayed posttests. Succinctly answering the research questions of the study, the researchers of the present study concede that CW tasks were more effective than individual ones in that both males and females' accuracy and fluency in the collaborative groups improved significantly both in the short and the long run.

One salient contribution of this study was the investigation of the impact of CW tasks in the long run, which was measured through delayed posttests. The results were suggestive of the positive effect of CW tasks even in the long run on both accuracy and fluency development. Since the findings regarding the positive impact of CW on the improvement of fluency and accuracy in writing outputs are consistent with the majority of the previous studies (Dobao & Blum, 2013; Storch, 2013; Bikowski & Vithanage, 2016; Alghasab *et al*., 2019; Kang & Lee, 2019), the present authors highly recommend that English teachers incorporate CW activities and tasks into their writing classes in order to improve student writers' written outputs fluency and accuracy-wise

The implications for writing instructors are significant. It should be mentioned that implementing CW is far from easy for most language teachers as this approach is almost novel and most EFL teachers were trained in the traditional educational systems, which were not conducive to collaborative activities. Writing instructors who wish to see student writers who produce accurate and fluent texts should invest in CW activities, which based on the findings of this study and the vast majority of other studies reviewed before, yield better writers both accuracy and fluency-wise. Those teachers who want to have superior individual writers, too, are recommended to apply CW tasks in their classes since we believe, in line with Elola and Oskoz (2010), that CW is not to surrogate individual writing, rather, it scaffolds student writers to further improve their own individual writing. Therefore, it behooves writing instructors to realize that the ultimate goal is to enhance student writers' individual writing and have autonomous individual writers in the end of the day.

Finally, it should be noted that, the present study was limited to elementary level learners; therefore, it is suggested that future studies investigate the effects of CW in the long run on various proficiency levels. Additionally, since in the present study, the development of fluency and accuracy was investigated, investigating the effects of CW on the development of complexity, in the long run, is a propitious suggestion and will be a further contribution to the literature.

7. Rᴇғᴇʀᴇɴᴄᴇs

Alghasab , M., Hardman, J., & Handley, Z. (2019). Teacher-student interaction on wikis: Fostering collaborative learning and writing. *Learning, Culture and Social Interaction*, 21: 10-20.

Aljaafreh, A. & Lantolf, J. P. (1994). Negative feedback as regulation and second language learning in the Zone of Proximal Development. *The Modern Language Journal,* 78, 4, 465-483.

Allen, N., Atkinson, D., Morgan, M., Moore, T., and Snow, C. (1987). What experienced collaborators say about CW, *Journal of Business and Technical Communication,* 1(2), 70-90.

Bikowski, D. & Vithanage, R. (2016). Effects of web-based collaborative writing on individual writing development". *Language Learning and Technology*, 20 (1), 79-99.

Biria, R. & Jafari, S. (2013). The impact of CW on the writing fluency of Iranian EFL learners. *Journal of Language Teaching and Research,* 4 (1),164-175.

Block, D. (1996). Not so fast: Some thoughts on theory culling, relativism, accepted findings and the heart and soul of SLA. *Applied Linguistics*, 17(1), 63-83.

Bruffee, K. A. (1973). Collaborative Learning: Some practical models. *College English,* 34 (5), 634-643.

Bruffee, K. A. (1984). Collaborative learning and the conversation of mankind. *College English,* 46 (7), 635-652.

Bruner, J. (1986). *Actual minds, possible worlds.* Harvard University Press.

Chen, W. & Yu, Sh. (2019). A longitudinal case study of changes in students' attitudes, participation, and learning in collaborative writing. *System*, 82: 83-96.

DiCamilla, F. J. & Anton, M. (1997). Repetition in the Collaborative Discourse of L2 Learners: A Vygotskian Perspective. *Canadian Modern Language Review,* 53 (4), 609-33.

Dobao, A. F. (2012). CW tasks in the L2 classroom: Comparing group, pair, and individual work. *Journal of Second Language Writing*, 21 (1), 40-58.

Dobao, A. F. and Blum, A. (2013). CW in pairs and small groups: Learners' attitudes and perceptions. *System*, 41(2), 365-378.

Donato, R. (1994). Collective scaffolding in second language learning. In J. P. Lantolf, & G. Appel (Eds.) *Vygotskian approaches to second language research* (pp. 33-56). Norwood, NJ: Ablex.

Donato, R. & McCormick, D. (1994). A sociocultural perspective on language learning strategies: The role of mediation. *The Modern Language Journal*, 78 (4), 453-464.

Ede, L. & Lunsford, A. (1985). Let them write- together. *English Quarterly,* 18 (4), 119-127.

Elola, I. & Oskoz, A. (2010). Collaborative writing: Fostering foreign language and writing conventions development. *Language Learning & Technology*, 14 (3), 51-71.

Farrah, M. (2012). The impact of peer feedback on improving the writing skills among Hebron university students. *An-Najah Univ. J. Res. (Humanities),* 26 (1), 179-210.

Foster, P. (1998). A classroom perspective on the negotiation of meaning. *Applied Linguistics,* 19 (1), 1-23.

Fung, Y. M. (2010). CW features. *RELC Journal*, 41 (1), 18-30.

Gousseva-Goodwin, J. V. (2000). *CW assignments and on-line Discussions in An advanced ESL composition class* (Unpublished doctoral dissertation). University of Arizona, USA.

Hirvela, A. (1999). CW instruction and communities of readers and writers. *TESOL Journal*, 8 (2), 7-12.

Huberty, C. J. (2002). A history of effect size indices. *Educational and Psychological Measurement*, 62 (2), 227-240.

Jafari, N. & Nejad Ansari, D. (2012). The effect of collaboration on Iranian EFL learners' writing accuracy. *International Education Studies,* 5, (2), 125-131.

Kang, S. & Lee, J. H. (2019). Are two heads always better than one? The effects of collaborative planning on L2 writing in relation to task complexity. *Journal of Second Language Writing*, 45, 61-72.

Kuiken, F. & Vedder, I. (2002). The effect of interaction in acquiring the grammar of a second language. *International Journal of Educational Research,* 37(3-4): 343-358.

McDonough, K., De Vleeschauwer, J. & Crawford, W. (2018). Comparing the quality of collaborative writing, collaborative prewriting, and individual texts in a Thai EFL context. *System, 74*: 109–120.

Neumann, H. & McDonough, K. (2014). Exploring student interaction during collaborative prewriting discussions and its relationship to L2 writing. *Journal of Second Language Writing*, 27: 84-104.

Ohta, A. S. (2001). *Second language acquisition processes in the classroom*. Mahwah, NJ: Lawrence Erlbaum.

Reid, J. M. (1993). *Teaching ESL writing.* NJ: Prentice Hall Regents.

Saunders, W. M. (1989). CW tasks and peer interaction. *International Journal of Educational Research,* 13 (1), 101-112.

Soleimani, M., Modirkhamene, S. & Sadeghi, K. (2015). Peer-mediated vs. individual writing: Measuring fluency, complexity, and accuracy in writing. *Innovation in Language Learning and Teaching*, 9(3), 1-15.

Storch, N. (1999). Are two heads better than one? Pair work and grammatical accuracy. *System,* 27(3), 363-374.

Storch, N. (2001). How collaborative is pair work? ESL tertiary students composing in pairs. *Language Teaching Research,* 5(1), 29-53.

Storch, N. (2002). Patterns of interaction in ESL pair work. *Language Learning,* 52 (1), 119-158.

Storch, N. (2005). CW: Product, process, and students' reflections. *Journal of Second Language Writing,* 14 (3), 153–173.

Storch, N. (2011). CW in L2 Contexts: Processes, Outcomes, and Future Directions. *Annual Review of Applied Linguistics,* 31(1), 275-288.

Storch, N. (2013). *Collaborative writing in L2 classrooms*. Bristol, UK: Multilingual Matters.

Swain, M. & Lapkin, S. (1998). Interaction and second language learning: Two adolescent French immersion students working together. *Modern Language Journal*, 82 (3), 320-337.

Tabachnick, B. G. & Fidell, L. S. (2007). Multivariate analysis of variance and covariance. *Using Multivariate Statistics*, 3, 402-407.

Villamil, O. S. & de Guerrero, M. C. M. (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.

Villamil, O. S. & de Guerrero, M. C. M. (1998). Assessing the impact of peer revision in L2 writing. *Applied Linguistics*, 19(4), 491–514.

Vygotsky, L. (1962). *Thought and language.* New York: The MIT Press, Massachusetts Institute of Technology and John Wiley & Sons, Inc.

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

Williams, J. (2001). The Effectiveness of Spontaneous Attention to Form. *System*, 29(3), 325–340.

Wigglesworth, G. & Storch, N. (2009). Pair versus individual writing: Effects on fluency, complexity and accuracy. *Language Testing*, 26(3), 445-466.A