



Contributory Role of Collaborative Assessment in Improving Critical Thinking and Writing

Mansoor Fahim

Department of Foreign Languages and Literature, Allameh Tabataba'i University, Southern Allameh st., Tehran, Iran
E-mail: dr.mfahim@yahoo.com

Mowla Miri (Corresponding author)

Department of Foreign Languages and Literature, Allameh Tabataba'i University, Southern Allameh st., Tehran, Iran
E-mail: molamiri84@gmail.com

Yaghoub Najafi

Semnan University, Iran
E-mail: yaghoubnajafi@yahoo.com

Received: 20-08-2013

Accepted: 24-09-2013

Published: 01-01-2014

doi:10.7575/aiac.ijalel.v.3n.1p.1

URL: <http://dx.doi.org/10.7575/aiac.ijalel.v.3n.1p.1>

Abstract

Instilling critical thinking skills in language learners' mind and enhancing second language writing are gaining momentum in the field of English language teaching. Though some approaches to meeting these objectives have been proposed, review of the literature revealed that the contributory role of collaborative assessment in fostering critical thinking and second language writing has been overlooked to date. Thus, this study was conducted to delve into this issue. To this aim, two intact intermediate EFL classes in Iran, each of which contains 18 learners, were included in the study; they were randomly assigned to Student-Student (S-S) collaborative assessment and Teacher-Student (T-S) collaborative assessment groups. After receiving six sessions of treatment, both groups wrote an essay on an IELTS topic. Results of the study indicated that collaborative assessment, regardless of its type, has the potential to foster critical thinking and writing proficiency. Further, it came to light that S-S collaborative assessment group significantly outperformed the T-S collaborative assessment group in terms of gains in critical thinking and writing proficiency. In light of the findings, language teachers are suggested to involve learners in collaborative assessment processes; further, some suggestions are offered at the end to show some fruitful avenues for further research.

Keywords: collaborative assessment, critical thinking, second language writing

1. Introduction

Language assessment can play a monumental role in fostering language learning along with its primary role i.e. measuring learning outcomes. Some researchers are rightly of the opinion that assessment has not grabbed a worthy attention as a learning and developmental tool rather it has been primarily deployed as an evaluative instrument (e.g. Chau, 2005; Gardner, 1999). Gardner points out that "rather than being imposed externally at odd times during the year assessment ought to become part of the natural learning environment. As much as possible it should occur 'on the fly', as a part of natural engagement of one in learning situation" (p. 102). Further, he opines that it is more likely to introduce assessment explicitly at initial stages; however, as some time passes by, it is better to let the wealth of assessment "occur naturally on the part of student and teacher, with little need for explicit recognition or labeling on anyone's part" (p. 102). One type of assessment having the potential to assess and foster learning in general and language learning in particular is collaborative assessment.

Collaborative assessment (CA) highlights mutual goals (working towards a mutually acceptable assessment grade), dynamic exchange of information (presenting, defending and elaborating views on the grade by tutor and student) and role interdependence (emphasizing individual accountability for meaningful exchange to take place) as key characteristics of collaboration. (Chau, 2005, p. 27)

Further, instilling critical thinking in learners' minds has been at the top of educational agenda for a long time, but recently, foreign language teaching research has practically recognized its critical importance in language learning and has thought of different ways to introduce into language classrooms. For instance, Kabilan (2000) proposed involving in pedagogy of the questioning as a fruitful approach; in a similar line, Zainuddin and Moore (2003) argued that engaging in a structured controversial dialogue technique is conducive to critical thinking development. In summary, the bottom line of the bulk of research highlights that teachers should primarily concentrate on activities entailing

forming relationships, compare and contrast, classify, evaluate, rank, identifying right from wrong, facts from opinion, cause and effect, bias, to give reasons for causes, to foresee consequences, making inferences and summary, generalizations, interpret, identifying main, supporting and detailed ideas as well as making decisions and solving problems. (Mahyuddin et al, 2004, pp. 24-5)

Considering the enumerated features of helpful activities into account, one can draw the conclusion the collaborative assessment has the potential to pave the ground for involving learners in activities provoking the employment of critical thinking skills. Additionally, convincing body of research has offered considerable evidence for the positive role of alternative assessment in developing second language writing (Cheng & Warren, 2005), but the share of collaborative assessment in enhancing writing skills has been not addressed up to now. Thus, this study was set up to delve into the potential of collaborative assessment in fostering critical thinking and second language writing in Iranian context.

2. Review of the related literature

2.1 Alternative assessment

Brown and Abeywickrama (2010) assert that though some criticisms are leveled against alternative in assessment on the grounds that they demand a substantial amount of time and cost on the part of all stakeholders in language learning process. They are worth the time and cost since they can offer markedly greater washback and authenticity. Brown and Hudson (1998) put forward that alternatives in assessment are beneficent because they are of the potential to make learners use language creatively in real-world situations. Further, they assist learning by taking both products and processes of learning into account while traditional testing is primarily concerned with learning products. Additionally, alternatives in assessment direct learners' attention to their weak and strong points by doing so, in fact, learners' higher-level thinking skills will be targeted. Finally, it is worth mentioning that Brown and Abeywickrama hold that alternatives in assessment deepen learners' insight into assessment criteria and raters' characteristics.

2.2 Collaborative assessment

Gardner asserts that "all assessment should be undertaken primarily to aid students. It is incumbent upon the assessor to provide feedback to the students that will be helpful at the present time- identifying areas of strength as well as the weakness ..." (p. 103). As a category of alternative assessment, collaborative assessment has attracted that attention of very few SLA studies. Collaborative assessment aims to involve learners to reach a consented score with those who have scored their performances. In fact, the learners learn to self-evaluate their own performance and compare it with an external rater's evaluation, whether a peer or teacher and when they find a difference between these two kinds of evaluation, they seek opportunities to discuss it and reach mutual agreement. It is held that involving in such evaluative process can make a positive contribution to language learning (Chau, 2005). In the same way, Dickinson claimed for the successful role of collaborative assessment where learners find opportunities to mull over the criteria and ponder on their own performance (1996 as cited in Chau, 2005). Even though, collaborative assessment seems a fecund ground for critical thinking and second language writing research, very few studies have addressed this issue up until now.

2.3 Second Language writing and assessment

Second language writing is a complicated skill in which learners should orchestrate different cognitive and linguistic resources to come up with a coherent text. Which individual differences mainly account for variety in writing processes and the quality of the written product have been less investigated by SLA researchers (Kormos, 2012). Additionally, Kormos points out that second language writing is an activity which takes a long time and demands concentration and determination. To meet learners' writing objectives, recently some researchers have argued for the role of assessment in fostering writing skill (e.g. Massa, 1997). Massa points out that there

"is a need to redefine the objectives of writing assessment, moving it from a punitive, gatekeeping tool that measures deficits, to a facilitative tool that informs novice academic writers of the characteristics of clear expression of thought, informs teachers of students' potential, and informs the classroom curriculum". (87)

2.4 Critical thinking

Critical thinking is defined as "the ability to analyze facts, generate and organize ideas, defend opinions, make comparisons, draw inferences, evaluate arguments and solve problems" (Chance, 1986, p. 6). In another attempt to offer a definition pertinent to academic milieu, Gieve (1998) points out that learners should "examine the reasons for their actions, their beliefs, and their knowledge claims, requiring them to defend themselves and question themselves, their peers, their teachers, experts, and authoritative texts, both in class and in writing" (p. 126). Facione (1990) condenses all common themes of such attempts to come up with the following definition of critical thinking.

We understand critical thinking to be purposeful, self-regulatory judgment which results in interpretation, analysis, evaluation, and inference, as well as explanation of the evidential, conceptual, methodological, criteriological, or contextual considerations upon which that judgment is based. (1990, p. 2)

Fahim and Pezeshki (2012), similarly, put forward that critical thinking requires reasoning and thorough reflection and rejects blind acceptance of others' opinions.

2.5 Critical thinkers' characteristics

Those marked as critical thinkers move with reason, Siegel (1990) points out. They do not take things for granted; they make a pause and cast a healthy doubt on what they are going to do (Sofa, 2004). Additionally, Paul and Elder (2006)

assert that critical thinkers meet the following criteria when they reason: clarity, relevance, logicalness, accuracy, depth, significance, precision, breadth, and fairness. Hopper (2003) states that "a critical thinker is constantly asking questions, trying to distinguish between fact and opinion. Not about memorizing, but analyzing all sides of an issue to find more in the situation than the obvious and makes assertions built on sound logic and solid evidence" (p.37). In one of the best endeavours in this regard, Ennis (1993, p. 179) mentions that critical thinkers examine the authenticity and reliability of information sources, pinpoint the reasons and tenets of arguments and evaluate their soundness, formulate their opinions and defend them logically; further, they try to be open to reasons and shed light on issues by posing pertinent and shrew questions. In a similar line of inquiry, Wood (2002) contends that critical thinkers are those who are keen to probe all perspectives open-mindedly; they welcome dangers and opposing ideas since they are on a quest for truth.

2.6 EFL and critical thinking

Integrating critical thinking skills into language learning programs has been supported on some grounds recently. Shirkhani and Fahim (2011) highlight that learners having a higher degree of critical thinking can manage more successfully to evaluate and monitor their own work. Additionally, previous research shows that there is a high correlation between critical thinking skills and language learning and proficiency (Liaw, 2007). Arend (2009) states that "although critical thinking is a goal of ... education, instructors may not often be using the most effective methods to encourage students to use critical thinking strategies" (p. 2). Concerning language teaching, sowing and fostering the seeds of critical thinking in language learners' minds has been a peripheral objective for language teaching methodologies even for communicative language teaching (Kabilan, 2000; Pica, 2000). Thus, it incumbent upon teachers, as pointed out by Lipman (2003), to put more store on making critical thinking as an integral part of language learning and assessment.

2.8 Collaborative assessment and critical thinking

Yuretich (2004) articulates that classroom activities which involve learners in posing and reflecting on questions, discussing over the questions in groups as well as rationalizing and explicating their answers to their peers push learners to move toward more advanced levels of Bloom's taxonomy. Likewise, Smith (1990) holds that engaging learners in activities requiring arguments and challenges in a respectful atmosphere can be conducive to cultivating critical thinking. It can be induced that collaborative assessment has the potential to engage learners in meaningful arguments with either their peers or teachers. The learners should analytically evaluate their own or their peers' writings so as to rationally score or rescore texts. It is apparent that collaborative assessment demands identifying and focusing on a point of conflict, analyzing the evidence for pinpointing the reasons behind the conflict by resorting to inductive and deductive logic. The traces of these factors can be identified in critical thinkers' characteristics. Pithers and Soden (2000) assert that critical thinking entails a number of abilities such as identifying a problem and the assumptions on which it is based, focusing on the problem, analyzing, understanding and making use of inferences, inductive and deductive logic, and judging the validity and reliability of assumptions and sources of data. Thus, it seems that involvement in collaborative assessment has the potential to set a fertile ground for cultivating critical thinking.

2.9 Statement of the problem and research questions

Only a few studies (e.g. Chau, 2005) have addressed the potential of collaborative assessment for fostering language development in general and writing proficiency in particular. Additionally, as aptly raised by Shirkhani and Fahim (2011), the introduction of critical thinking into EFL curriculum is in its infancy and still growing. Needless to say, it is incumbent on EFL teachers to use the materials and activities which implant and foster critical thinking in EFL learners. Shirkhani and Fahim opine that teachers should assist learners to grow critical thinking via teaching material and classroom activities. Classroom activities should have the potential to set the ground for learners to express their ideas, reflect on their learning outcome and processes, and get involved in continuous interaction and communication with their classmates, teachers, and others both within the borders of the classroom and out of them (Kusaka & Robertson, n.d.). One of the possible classroom activities speculated to make a contribution to cultivating and raising critical thinking is collaborative assessment, whether teacher-student or student-student collaborative assessment. Additionally, second language writing has been unanimously considered as a highly complicated and frustrating skill, since language learners should manage to orchestrate a wide range of skills such grammar, vocabulary, mechanics, and content simultaneously to come up with a coherent piece of writing. In spite of its salient importance and frustrating complexity, Zhang (2008) holds that instructing second language writing "has been an overlooked and underrepresented aspect of L2 writing research" (p. 133). Thus, this study was set up to cast some empirical light on the effect of collaborative assessment on developing writing proficiency and critical thinking skills in Iranian context. More specifically, the study addressed the following questions:

1. Does collaborative student-student assessment lead to any significant change in critical thinking of Iranian EFL intermediate learners?
2. Does involving in collaborative teacher-student assessment bring about any significant gains in critical thinking of Iranian EFL intermediate learners?
3. Is there any difference in the effect that student-student and teacher-student collaborative assessments impose on the gains in critical thinking of Iranian EFL intermediate learners?
4. Does collaborative student-student assessment bring about any gains in writing proficiency of Iranian EFL intermediate learners?

5. Does collaborative student-student assessment give rise to any gains in writing proficiency of Iranian EFL intermediate learners?
6. Is there any difference in the effect that student-student and teacher-student collaborative assessments impose on the gains in writing proficiency of Iranian EFL intermediate learners?

3. Method

3.1 Participants and context

This study was carried out in Iran where despite all compelling evidence for effectiveness of alternative assessment, the most prevalent approaches to evaluation are administered teacher-led which this case permeates through both formative and summative assessment at all education levels. Two intact classes of Iranian EFL students at intermediate level of language proficiency took part in this study. They enrolled in an English institute, but in two different branches, at the same time. They have covered the same materials with the same teacher over the past two years so it was supposed that they are somehow homogeneous. To gain more assurance, their achievement scores for the previous terms were consulted; it came to light that their means were not significantly different. The classes were randomly assigned as student-student collaborative assessment group (N = 18) and teacher-student collaborative assessment group (N = 18). It is worthy to point out that their age ranged from 15 to 18 years old.

3.2 Tasks

Six IELTS writing tasks from the general module were selected and given to three groups. These tasks were chosen on the account that IELTS writing tasks enjoy acceptable degrees of validity and reliability (Charge & Taylor, 1997). Second, the participants were prospective IELTS test takers; therefore, doing these tasks was in line with their objectives. It was speculated that working on these tasks encourage the participants to take the tasks more seriously and involve themselves in collaborative assessment.

3.3 Tests

Two writing tasks from the general module were considered as pre- and post-tests. As the pre-test both classes were asked to write an essay containing at least 150 words in 30 minutes. Both of them wrote their essays in class time without receiving any help from other peers or students. After the treatment, both groups sat for the post-test lasting for 30 minutes in class time. They were required to write an essay on the given task entailing at least 150 words. Test topics are as follows:

Pre-test: Education is recognized as vital to the future of any society in today's world. Governments throughout the world should make education obligatory for all children between the ages of 5 and 15. To what extent do you agree or disagree with this statement?

Post-test: The idea of going overseas for university study is an exciting prospect for many people. But while it may offer some advantages, it is probably better to stay home because of the difficulties a student inevitably encounters living and studying in a different culture. To what extent do you agree or disagree with this statement? Give reasons for your answer.

3.4 Analytical scoring checklist

To help learners evaluate writing tasks more consistently and rationally, the scale developed by Jacob et al. (1981) was adopted in the current study. On the basis of the scale, the writing tasks are scored on five dimensions each of which bears different points: 'content' (30 points), 'language use' (25 points), 'organization' (20 points), 'vocabulary' (20 points), and 'mechanics' (5 points). The analytical scoring was selected to serve the objective of the study on the account that they could "provide more detailed information about a test-takers performance in different aspects of writing" (Weigle, 2002, p.114); it is argued that analytical scoring facilitates rater training i.e. novice raters can more easily figure out how to give appropriate weight to different aspects of writing task; additionally, "analytical scoring is particularly useful for second-language learners, who are more likely show a marked or uneven profile across different aspects of writing (Weigle, p. 120). Another grounds on which the analytical scoring scale was opted for is providing the possibility for making more decisions on the part of raters; thus, the probability of provoking controversies over the scores between the first and second raters could be increased. In other words, holistic scoring allows learners to make only one decision while analytical scoring allows the learners to make five decisions on five dimensions of writings; thus, every decision can give rise to a cognitive conflict which are conducive to critical thinking and help the learners to gain a deeper insight into writing processes.

3.5. Critical thinking test

The Watson Glaser Critical Thinking Appraisal (WGCTA) was used in this study to gauge learners' critical thinking both prior and after the treatment. More specifically, form A of this test including 'Inferences', 'Recognition of Assumptions', 'Deduction', 'Interpretation' and 'Evaluation of Arguments', each category entailing 16 items with two to five alternatives, was employed. The test is not subject-specific and takes one hour to be answered. Watson and Glaser (1980) report that its test-retest reliability is acceptably high ($r = 0.81$) and, similarly, the reliability coefficient of its Persian translation estimated by Cronbach's Alpha was very high ($\alpha = 0.85$). This appraisal was adopted to fulfill the objectives of the current study on the ensuing grounds: first, "it has construct validity in that its five sections correspond closely to the primary skills identified by major theorists in the area" (Floyd, 2011, p. 292). Second, a good

number of studies have provided convincing evidence for its considerable validity (e.g. Gadzella et al., 2006; Hashemi & Zabih, 2012).

3.6 Procedures

Once having received the participants' consent to take part in the study, the Critical Thinking Appraisal was administered to both classes to answer in 60 minutes. Additionally, they were asked, as the writing pre-test, to write an essay containing at least 150. It is worthy to add that both classes sat for the pre-test in class without receiving any help from others. Later on, both classes were instructed how to evaluate writing according to an analytical scale since it is unanimously held that evaluation gets much fruitful for the students who know how to offer and use evaluative feedback (Berg, 1999; Min, 2006). For example, Stanley (1992) pointed out that the trained learners were more successful than untrained learners in giving precise responses and they were more ready to respond the feedback they received in revision phase. In each of the collaborative classes, the second researcher distributed and explained the checklist to students; then, a prepared sample of writing scored according to the analytical scoring checklist was given to both groups. The teacher elaborated on the reasons behind the given scores in detail. To hone their assessing skill, another piece of writing was given to the learners; upon the learners' finishing scoring, the teacher wrote the scores he has allotted to different aspects of the writing on the board. Then, he asked the learners to compare both sets of scores and discuss the points of difference. They all were given time to put forward their arguments and defend their own scoring. The teacher tried to push them toward casting healthy doubt on both sets of scores by posing questions. In fact, the teacher endeavored to assist the learners to perceive all steps of collaborative writing assessment since it requires to be "modeled, explicitly taught, and controlled" (Reid, 1993, p. 157). Students were given another two pieces of writing scored analytical according to the scale to score and offer their reasons for assigning their scores. If a difference was found between the learners' scores and those of the teacher, the learners had time in the following session to negotiate on the point of controversy with the teacher and defend their choices in class.

Teacher-Student (T-S) collaborative assessment class was asked to write six pieces of writing which were initially scored by teacher on the basis of the analytical scale. Later on, the learners were asked to score their own writing according to the same scale and pinpoint where they disagree with teacher's scores. Later on, each student was given three minutes to discuss the points of difference and defend his own position against teacher's scores. They both worked in collaboration to settle differences and develop mutual intersubjectivity and agreement on the final score. In the cases that their attempts failed to reach mutual agreement, the learner was given the opportunity to sleep on it for a while and discuss it the other following session or send his/her arguments via email.

For Student-Student (S-S) collaborative assessment group, the learners were supposed to write on the same six topics. Students' pieces of writing were collected and distributed to their peers. Each student was asked to score his peer's writing on the basis of the analytical scale given to them. They were cautioned to be fair and objectives since the papers would be reviewed by the teacher and in the case of flagrant carelessness or bias they would be discredited. In the next phase, each learner appraised his own writing on the basis of the scale and took on the points of difference in scoring. Finally, on two settings of four minutes learners given the opportunity to discuss the points of difference and reach a compromise. However, if their attempts failed to negotiate a compromise and mutual understanding, they were given the opportunity to discuss it for two minutes in the following session.

After the study, a semi-structured interview was held with both classes to delve into their reactions to the use of collaborative assessment and their opinions about the merits and demerits of collaborative assessment.

4. Results and discussions

4.1 Results for the first and second questions

The first two questions were raised to explore whether collaborative assessment can increase gains in critical thinking. In order to pinpoint whether both types of collaborative assessment could impose some effects on developing critical thinking skills, two paired-samples t-tests were conducted on the results. As viewed in Table 1, descriptive statistics for both t-tests revealed that collaborative assessment has made a change in the gains in critical thinking for both groups.

Table 1. Descriptive statistics for the effect of collaborative assessment types on CT

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	CT Pre-t of S-S	29	18	5.88	1.38
	CT Post-t of S-S	41.50	18	6.27	1.47
Pair 2	CT Pre-t of T-S	28.66	18	5.56	1.31
	CT Post-t of T-S	33.50	18	6.46	1.52

Additionally, the inferential statistics unveiled that both types of collaborative assessment gave rise to significant gains in critical thinking. The first paired-samples t-test was run to evaluate the effect of S-S collaborative assessment on critical thinking. There was a statistically significant increase in critical thinking from Time 1 ($M = 29$, $SD = 5.88$) to

Time 2 ($M = 41$, $SD = 6.27$), $t(17) = -20.61$, $p < .05$ (two-tailed). The eta squared statistic (.96) indicated a large effect size. The second paired-samples t-test showed that T-S collaborative assessment led to significant increase in gains in critical thinking from Time 1 ($M = 28.66$, $SD = 5.56$) to Time 2 ($M = 33.50$, $SD = 6.46$), $t(17) = -6.40$, $p < .05$ (two-tailed). Further, the eta squared statistic (.71) showed a large effect size.

Table 2. The effect of collaborative assessment types on gains in CT

Pair		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference				
					Lower	Upper			
Pair 1	CT Pre-t of S-S - CT Post-t of S-S	-12.50	2.57	.60	-13.77	-11.22	-20.61	17	.00
Pair 2	CT Pre-t of T-S - CT Post-t of T-S	-4.83	3.20	.75	-6.42	-3.24	-6.40	17	.00

The results of the paired-samples t-test revealed that S-S collaborative assessment has successfully managed to enhance critical thinking. This finding, in fact, gave more credence to the collaborative learning and assessment claim's that cooperation has the potential to instill and improve critical thinking on the grounds that learners find opportunities to freely involve in reflective evaluation (Saito, 2008). The findings, further, can be explained on Rezaie et al.'s (2011) claim that involving learners in opportunities demanding self- and peer-assessment push them to grow and practice critical thinking skills; "Students should be given a chance to assess not only themselves but also others to enhance their critical thinking ability" (p. 775). Another possible reason behind the findings could be the fact that collaborative assessment sets the scene for self- and other evaluation on the basis of sound criteria and logically sufficient reasons supported with evidence. In fact, each of the learners had the chance to gather evidence and establish an evaluative position about his peers or himself and be ready to defend his own position against probable arguments; therefore, the learner got involved in the components of critical thinking such as inferencing and evaluation. Additionally, collaborative assessment offered them opportunities to argue for or against an evaluative position. They were either question poser or replier. Collaborative assessment cultivated the ground for identifying the problems, analyzing the problem by resorting to inductive and deductive reasoning and working toward a mutual objective. That is, learners in both groups reviewed their writing in retrospect and compared and contrasted their own scores with those of the teacher or their peers; whenever they identified a distinct difference, they formulated it as a problem to be analyzed and elucidated by accruing data from different part of the text or the general picture of writing. These processes were cyclic and learners went through them in treatment time. They hone the skill to cast a healthy doubt on self- and other evaluation and do not take evaluation for granted. They grew the habit to ask questions and to be questioned about their evaluative judgments.

Another plausible explanation behind the efficacy of collaborative assessment in enhancing critical thinking is its potential to provoke a discussion or debate among students. This genuine debate works as a strong impetus for cultivating and raising critical thinking skills since it is controversial and pertinent to class and syllabus, learners know about the points of discussion in advance so they can reflect upon it and formulate their arguments and defenses on the basis of solid evidence. The enumerated characteristics of discussions resulting from engaging in collaborative assessment are recognized as discussion features contributing to development of critical thinking (Rezaie et al., 2011).

The results also can be attributed to the potential of collaborative assessment in involving learners in a kind of problem solving activity. Buskist and Irons (2008) articulate, in order to foster critical thinking, it is incumbent upon teachers to engage learners in problem solving both in and out of the classroom. It was observed that collaborative assessment pushed learners to detect the problems, elaborate on by relying on robust evidence, defend them against other opponent views and finally try to settle them down. It can be concluded that since collaborative assessment has a problem posing and problem solving nature, it can act as a stimulus to critical thinking.

The results of the study, further, lend some support to Smith's (1990) claim that involving or observing a situation demanding critical thinking and reasoning in an atmosphere on the basis of respect rather than power can assist critical thinking development. Collaborative assessment pave the ground for doing activities requiring critical reasoning and thinking in a friendly and respectful environment.

4.2 Results and discussion for the third question

This question sought to compare the effect of two types of collaborative assessment on gains in critical thinking. To meet the objective of this question, an independent samples-t-test was conducted on the results for the critical thinking post-test. Descriptive statistics in Table 3 unveiled that there was a mean difference for gains in critical thinking as a result of collaborative assessment type.

Table 3. Descriptive statistics for the effect of two types of collaborative assessment on CT

	collaborative assessment type	N	Mean	Std. Deviation	Std. Error Mean
	CT growth	S-S	18	41.50	6.27
	T-S	18	33.50	6.46	1.52

The independent samples t-test, as portrayed in Table 4, showed that there was a significant difference for S-S collaborative assessment ($M = 41.5$, $SD = 6.27$) and T-S collaborative assessment ($M = 33.5$, $SD = 6.46$); $t(34) = 3.69$, $p < .05$ (two-tailed). The magnitude of the differences in the means was very large ($\eta^2 = .28$). That is, the S-S collaborative assessment group significantly outperformed the T-S group in terms of development in critical thinking.

Table 4. Comparing the effect of two types of collaborative assessment on CT

		Levene's Test for Equality of Variances		t-test for Equality of Means						
		F	Sig.	t	df	Sig. (2- tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
									Lower	Upper
CT growth	Equal variances assumed	.28	.59	3.76	34	.001	8.00	2.12	3.68	12.31
	Equal variances not assumed			3.76	33.96	.001	8.00	2.12	3.68	12.31

This question intended to compare and contrast the effect of two kinds of collaborative assessment, namely, student-student and teacher-student collaborative assessment. The results of the study highlighted that S-S collaborative assessment was more conducive to fostering critical thinking skills. This finding is justifiable on some possible grounds. First, it was noticed that initially learners were more relaxed to challenge their peers' evaluative judgments than those of their teacher, because they have been grown up in an education system that teachers have had the final say in evaluation. In Iranian education system, this is the teacher who is in charge of evaluation and learners are hardly ever allowed to take part in assessment and challenge teachers' evaluative authority. Thus, as voiced by participants in T-S collaborative assessment, they thought that it was useless to disagree with teacher's scores since he most probably would not change his minds; additionally, it was pointed out that the teacher is very skillful at scoring, so he rarely makes mistakes. One of the repetitive current themes in semi-structure interview was that teachers are skillful at evaluating and scoring students' works. One of the students voiced that

At the beginning of this method, it was really difficult for me to argue against teachers' evaluative decisions because I have never did so previously and I thought that it was impolite; further, I wondered whether the teacher would change his mind if I express my disagreement about his scoring. It took some time that I noticed the teacher is ready to listen to our complaints about the scoring and change the score in cases that we were right and argue logically and persuasively. When I talked to my classmates after the class, they pointed out that they had experienced the same concerns at the outset of the class. However, when we realized that the teacher pay respect and credit to those arguing for their own deserved score, we were prompted to evaluate our writing on the basis of the scale more attentively and discuss the identified points of difference with the teacher.

Though it took some time for T-S group to discuss the points of the difference, S-S group was more relaxed to challenge their peers' evaluative judgments from the beginning of the class. The results of the interview indicated that learners in S-S group considered themselves equal to their own peers, so they expected their peers to make mistakes in evaluation process. Additionally, S-S group recurrently articulate that they believed that they could change their peers' decisions, if their arguments rooted in logic and solid evidence. One of them, for instance, highlighted that,

I thought that we could affect our peers with convincing reasons. Whenever I found a distinct difference between my peers' scoring and what I deserved to get, I collected required evidence by perusing my own writing and scoring it according to the checklist and easily discussed it in the next session.

Another explanation behind the superiority of S-S collaborative assessment over T-S collaborative assessment probably resides in the fact that S-S assessment involved learners in two evaluation phases i.e. self- and other-evaluation whereas T-S collaborative assessment deprives the learners of other evaluation processes. That is to say, S-S group initially had the opportunity to evaluate their peers' writing and answer their questions, if any question was raised; later on, they were offered the opportunity to get engaged in their own evaluation and discuss the points of difference.

To summarize, engaging in S-S collaborative assessment was more fruitful than T-S collaborative assessment in honing critical thinking skills owing to the fact that learners gain more opportunities to critically analyze their own work and received scores.

4.3 Results for the fourth and fifth Questions

These two questions were raised to probe whether two types of collaborative assessment had any effect on gains in second language writing. Two paired samples t-tests were run on the results of pre- and post-test writing tests to fulfill the objectives of these two questions.

Table 5. Descriptive statistics for the effect of collaborative assessment types on writing

		Mean	N	Std. Deviation	Std. Error Mean
S-S collaborative assessment	Writing Pre-t of S-S	41.33	18	5.90	1.39
	Writing Post-t of S-S	64.55	18	8.65	2.04
S-T collaborative assessment	Writing Pre-t of T-S	41.27	18	6.14	1.44
	Writing Post-t of T-S	55.16	18	7.51	1.77

The first sample t-test was carried out to evaluate the effect of S-S collaborative assessment on writing development. There was a statistically significant increase in writing proficiency from Time 1 ($M = 41.33$, $SD = 5.9$) to Time 2 ($M = 64$, $SD = 8.65$), $t(17) = -21.68$, $p < .05$ (two-tailed). The eta squared statistic (.96) indicated a large effect size. The second paired-samples t-test indicated that T-S collaborative assessment gave rise to significant increase in gains in writing from Time 1 ($M = 41.33$, $SD = 6.14$) to Time 2 ($M = 55.16$, $SD = 7.51$), $t(17) = -14.27$, $p < .05$ (two-tailed). Additionally, the eta squared statistic (.92) showed a very large effect size. That is to say, the changes in dependent variable are highly attractable independent variable, collaborative assessment.

Table 6. The effect of collaborative assessment types on gains in writing

		Paired Differences					t	df	Sig. (2-tailed)
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference				
					Lower	Upper			
Pair 1	Writing Pre-t of S-S - Writing Post-t of S-S	-23.22	4.55	1.07	-25.48	-20.95	-21.61	17	.00
Pair 2	Writing Pre-t of T-S - Writing Post-t of T-S	-13.88	4.12	.97	-15.94	-11.83	-14.27	17	.00

The results for these questions provided some evidence for Saito's claim that "peer involvement in assessment holds tremendous potential for learning" (Saito, 2008, p. 554). It seems possible that these gains were accrued due to the point, as mention by Saito, that involvement in collaborative assessment prompted reflective learning via observing self and others' performances and getting more awareness about performance and evaluative criteria. Learners were inspired to evaluate and re-evaluate their own or peers' writings reflectively on the basis of the scoring scale. Additionally, in line with Chau, these results can be ascribed to the fact that the learners managed to develop a critical awareness toward evaluative criteria. For example, one of them voiced that

This study helped me to find out on which criteria my writing have been measured so far. Frankly speaking, I was unaware of such criteria prior to this class but involvement in collaborative assessment allowed me to compare teacher's scoring with mine to get a better understanding of evaluating my writing. Previously I only received my paper with a score and some underlined parts or corrections. When knowing about the scale, I made my best to meet the criteria to improve my work and gain a high score. I am not merely concerned with grammatical accuracy while writing; I pay heed to vocabulary, content, cohesion and coherence, and mechanics. Further, from now on, I run my eyes over my writing to check to what extent I have observed the criteria before handing in my writings.

Moreover, the findings are justifiable on the plausible grounds that involving in collaborative assessment equipped learner with a sense of shared responsibility. As pointed out about peer evaluation, cooperative assessment "encourages learners and teachers to regard assessment as a shared responsibility" (Azarnoosh, 2013, p. 8). Learners experienced that they had a say in evaluation process which is listened to and respected. This point was verified by a number of participants; as a good case in point, one of the learners put forward that

I was really delighted to see that I have a right to participate in evaluating my own work and I had the opportunity to change my teacher's score. It is very enjoyable to argue for what you deserve and see the fruits of your logical strong defense. When I convinced my teacher to change my score, I felt proud of myself.

In summary, the results can be explained by drawing on Azarnoosh's argument that "focusing on peers' strengths and weaknesses can enhance students' learning, raise their level of critical thinking, and lead them to autonomy" (2013, p. 8).

4.4 Results and discussion for the sixth question

This question aimed to compare the effect of S-S and T-S collaborative assessment on developing writing proficiency. As can be viewed in Table 7, descriptive statistics indicated that there is a difference between the means for post-test scores. Thus, the inferential statistics were consulted to see whether the difference was significant.

Table 7. Descriptive statistics for the effect of two types of collaborative assessment on writing

	collaborative assessment type	N	Mean	Std. Deviation	Std. Error Mean
WT growth	1	18	64.55	8.65	2.04
	2	18	55.16	7.51	1.77

The results of independent-samples t-test demonstrated that there was a significant difference in scores for S-S collaborative assessment group ($M = 64$; $SD = 8.65$) and T-S collaborative assessment group, $M = 55.16$, $SD = 7.51$; $t(34) = 3.47$, $p < .05$ (two-tailed). The magnitude of differences in the means (mean differences = 9.38, 95% CI: 3.89 to 14.88) was large ($\eta^2 = .41$).

Table 8. Comparing the effect of two types of collaborative assessment on writing

		Levene's Test for Equality of Variances		t-test for Equality of Means						
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
									Lower	Upper
WT growth	Equal variances assumed	.163	.68	3.47	34	.001	9.38	2.70	3.89	14.88
	Equal variances not assumed			3.47	33.34	.001	9.38	2.70	3.89	14.88

The results for the superiority of S-S collaborative assessment over T-S collaborative assessment can be explained on the following reasons. The superiority of S-S collaborative assessment might be due to the fact that it opened the doors for the learners to detect strong and weak points of their peers' writings and learn important lessons from them. This point is well manifested in one of the students' words having the possibility to evaluate a good learner's writing.

I was lucky to read Saman's writings two times. His works, in fact, showed the way of coherent writing to me. He taught me how to make a link between my sentences and paragraphs. Whenever I put the pen to paper the other times, I visualized Saman's writing. How he started, supported and finished his ideas emerged in my mind.

Another possible reason for outperformance of S-S group can be the time spent on collaborative assessment. Learners in S-S had about ten minutes to discuss the controversies over their own writing and their peer writings that they had scored while the learners in T-S had only three minutes to collaborate with the teacher to settle the conflicts. Additionally, as pointed out previously, the learners in the S-S collaborative assessment group were more inclined to collaborate with their peers rather than the learners in T-S collaborative assessment group to interact with their teacher. Therefore, having stronger inclination and more time offered the S-S group to perform better than T-S group.

5. Conclusions

This study investigated the effect of two types of collaborative assessment on fostering critical thinking and writing proficiency. It came to light that S-S and T-S collaborative assessment were able to foster critical thinking and writing proficiency. Additionally, the results of the study revealed that S-S collaborative assessment was more effective than T-S collaborative assessment in developing critical thinking and writing proficiency.

In light of the findings, language teachers are encouraged to instruct learners on how to carry out collaborative assessment and give more credence to S-S collaborative assessment which entails self- and peer-assessment as well as reflection. Further, materials developers are suggested to assign some parts to teaching the principles of collaborative assessment.

However, a number of caveats need to be noted regarding the current study. First, this study lacked a control group; thus, it cannot be claimed that collaborative assessment is more effective than other types of assessment. Further, the sample included in the study was not selected randomly; therefore, the generalizability of the results should be done with caution.

Finally, some fertile avenues for further research are worth mentioning. Due to dearth of empirical studies on the potential of collaborative assessment for fostering different dimensions and components of language learning processes, further research are required to explore this issue. Since this study was primarily quantitative in nature, qualitative investigation of dialogues during collaborative assessment can be a fruitful ground for prospective studies. Additionally, the effect of some individual factors such as gender and learners' proficiency learners on collaborative assessment results is an interesting point for further investigation.

Acknowledgments

We take the opportunity to express our heartfelt thanks to the learners who wholeheartedly followed the procedures all throughout the study.

References

- Arend, B. (2009). Encouraging critical thinking in online threaded discussions. *Journal of Educators Online*, 6(1), 1-23.
- Azarnoosh, M. (2013). Peer assessment in an EFL context: attitudes and friendship bias. *Language Testing in Asia*, 3 (11), 1-10.
- Berg, E. C. (1999). The effects of trained peer response on ESL students' revision types and writing quality. *Journal of Second Language Writing*, 8, 225-241.
- Brown, H. D., & Abeywickrama, P. (2010). *Language assessment. Principles and classroom practices* (2 Ed.). White Plains, NY: Pearson Education.
- Buskist, W. & Irons G.J. (2008). Simple strategies for teaching your students to think critically. In D. S. Dunn, J. S. Halonen, and R. A. Smith (Eds.), *Teaching critical thinking in psychology: a handbook of best practices* (pp.49 -57). UK: Blackwell Publishing Ltd.
- Cha, J. (2005). Effects of collaborative assessment on language development and learning. *Language Learning Journal*, 32, 27 - 37.
- Chance, P. (1986). *Thinking in the Classroom: A Survey of Programs*. New York: Teachers College, Columbia University.
- Charge, N., & Taylor, L.B. (1997). Recent developments in IELTS. *ELT Journal*, 51, (4), 374-380.
- Cheng, W., & Warren, M. (2005). Peer assessment of oral proficiency. *Language Testing*, 22, 93-121.
- Elder, L., & Paul, R. (2004). Critical thinking and the art of close reading (part IV). *Journal of Developmental Education*, 28(2), 36-37.
- Ennis, R. H. (1993). *Critical thinking assessment. Theory into Practice*, 32 (3), 179-186.
- Fahim, M., & Pezeshki, M. (2012). *Manipulating Critical Thinking Skills in Test Taking*. *International Journal of Education*, 1 (4), 153-160.
- Falchikov, N. (1986). Product comparisons and process benefits of collaborative peer group and self assessments. *Assessment and Evaluation in Higher Education*, 11, 146-165.
- Falchikov, N., & Goldfinch, J. (2000). Student peer assessment in higher education: A meta-analysis comparing peer and teacher marks. *Review of Educational Research*, 70, 287-322.
- Floyd, C. B. (2011). Critical thinking in a second language. *Higher Education Research and Development*, 30, 289-302.
- Gadzella, B. M., & Baloglu, M. (2003). Psychometric properties of Watson-Glaser Critical Thinking Appraisal for a sample of education majors. *Psychol Rep*, 92 (3), 1249-1254.
- Gardner, H. (1999). 'Assessment in context'. In Murphy, P. (Eds.), *Learners, learning and assessment* (pp. 90-117). London: Paul Chapman Publishing.
- Gieve, S. (1998). *Comments on Dwight Atkinson's "A critical approach to critical thinking in TESOL"*. *TESOL Quarterly*, 32(1), 123-129.
- Gokhale, A. A. (1995). Collaborative learning enhances critical thinking. *Journal of Technology Education*, 7 (1). Retrieved on Jan 14, 2008 from <http://scholar.lib.vt.edu/ejournals/JTE/v7n1/pdf/gokhale.pdf>.

- Hashemi, M. R., & Zabihi, R. (2012). Does critical thinking enhance EFL learners' receptive skills? *Journal of Language Teaching and Research*, 3(1), 172-179.
- Hopper, C. (2003). Practicing college learning strategies. *Boston: Houghton Mifflin Co.*
- Hu, G. (2005). Using peer review with Chinese ESL student writers. *Language Teaching Research*, 9, 321-342.
- Kabilan, K.M. (2000) Creative and critical thinking in language classroom. *Internet TESL Journal*, 6/6. <http://iteslj.org/Techniques/Kabilan-CriticalThinking.html>
- Liaw, M. (2007). Content-based reading and writing for critical thinking skills in an EFL context. *English Teaching and Learning*, 31(2), 45-87.
- Mahyuddin, R., Lope Pihie, Z. A., Elias, H., & Konting, M. M. (2004). The incorporation of thinking skills in the school curriculum. *Kajian Malaysia, Jld, 22(2)*, 23-33.
- Massa, J., (1997). Alternative assessment of second-language writing: a developmental model. In S. Tchudi (Eds.), *Alternatives to grading student writing* (pp. 77-89). Urbana: NCTE.
- Min, H T. (2006). *The effects of trained peer review on EFL students' revision types and writing quality. Journal of Second Language Writing*, 15, 118-141.
- Paul, R., & Elder, L. (2006). *The miniature guide to critical thinking: Concepts and tools*. Dillon Beach CA: The Foundation for Critical Thinking.
- Pithers, R, T., & Soden, R. (2000). Critical thinking in education: a review. *Educational Research*, 42, 237-249.
- Reid, J. M. (1993). *Teaching ESL writing*. NJ: Prentice Hall Regents.
- Rezaei, S., Derakhshan, A., & Bagherkazemi, M. (2011). Critical thinking in language education. *Journal of Language Teaching Research*, 2(4), 769-777.
- Shirkhani, S., & Fahim, M. (2011). Enhancing critical thinking in foreign language learners. *Procedia - Social and Behavioral Sciences*, 29, 111-115.
- Siegel, H. (1990). The generalizability of critical thinking. *Educational Philosophy and Theory*, 23, 18-30.
- Smith, F. (1990). *To think*. New York: Teachers College Press.
- Sofo, F. (2004). *Open your mind: 7 keys to thinking critically*. Sydney: Allen and Unwin.
- Saito, H. (2008). EFL classroom peer assessment: training effects on rating and commenting. *Language Testing*, 25, 553-581.
- Stanley, J. (1992). *Coaching student writers to be effective peer evaluators. Journal of Second Language Writing*, 1, 217-233.
- Weigle, S.C. (2002). *Assessing writing*. Cambridge University Press.
- Wood, R. (2002). *Critical thinking*. Retrieved July 20, 2012, from: <http://www.robinwood.com/Democracy/GeneralEssays/CriticalThinking.pdf>
- Yuretich, F. R. (2004). Encouraging critical thinking: Measuring skills in large introductory science classes. *Journal of College Science Teaching*, 33(3), 40-46.
- Zainuddin, H., & Moore, R. A. (2003). Enhancing critical thinking with structured controversial dialogues. *The Internet TESL Journal*, 9(6). Retrieved July 10, 2012, from <http://iteslj.org/Technique/ZainuddinControversial.html>
- Zhang, J. (2008). A comprehensive review of studies on second language writing. *HKBU Papers in Applied Language Studies*, 12, 89-123. Retrieved July 20, 2011 from: lc.hkbu.edu.hk/book/pdf/v12_05.pdf.