Investigating the Impact of Collaborative and Static Assessment on the Iranian EFL Students' Reading Comprehension, Critical Thinking, and Metacognitive Strategies of Reading

Masoomeh Estaji
Assistant Professor, Allameh Tabataba'i University
mestaji74@gmail.com

Farzaneh Khosravi
M.A., Khatam University
farzanehkhosravi32@yahoo.com

Abstract

This study was aimed at investigating the effect of collaborative assessment, as compared to that of static assessment, on the EFL learners' reading comprehension, metacognitive strategies, and critical thinking. In this mixed-methods research design study, 62 intermediate EFL learners out of a population of 79 English language learners, were randomly selected and divided into two groups of experimental and control with 31 participants in each group. They were both female and male, who were administered a meta-cognitive strategy and critical thinking questionnaire and a reading comprehension test at the pretest and posttest stage of the study. Furthermore, semi-structured interviews were conducted with 15 participants and six sessions of the course were picked for observation. A multivariate ANOVA (MANOVA) test was run and the obtained results revealed that the students in the experimental group showed statistically significant gains on the reading comprehension and metacognitive strategy as compared with the participants in the control group but did not show any significant difference on their critical thinking. Similarly, the analysis of the interviews and classroom observations provided important themes which revealed that the students found collaborative assessment a very fruitful and practical way of promoting their reading skills and strategies mainly metacognitive strategies.

Keywords: Collaborative Assessment, Static Assessment, Reading Comprehension, Critical Thinking, Metacognitive Strategy

1. Introduction

The notion of collaborative learning, the grouping of the students for the aim of attaining an academic goal, solving a problem, completing a task, or creating a product, has been widely researched and adhered to in the professional literature. One of the most important elements of collaborative learning is Collaborative Assessment (CA) or team testing (Giraund & Enders, 2000; Helmericks, 1993; Muir & Tracy, 1999), which is applied to an assessor and assessee working together to form a mutual understanding of the student's knowledge (Pain et al., 1996). Collaborative assessment is a form of alternative assessment that has been derived from sociocultural theory of Vygotsky's zone of proximal development (ZPD).

The goal of self-, peer-, and collaborative-assessment is integrating learning and assessment and increasing the learners' active engagement in the assessment procedure. In addition, collaborative testing has been shown to enhance content retention, higher order thinking (Stearns, 1996; Yuretich, Khan, Leckie, & Clement, 2001), and the general pleasure of the course (Stearns, 1996). On the other hand, Hollingsworth et al. (2007) and Alhaidari (2006) mention that collaborative learning as a way of teaching turns out to be a worthy means to assist the learners with comprehension strategies while stimulating positive interactions with the peers.

The main problem noticed in this study was that the traditional system of assessment fights against the notion of developing the students' ability to assess their own work in ways which are suitable for their future educational and professional work in any specific discipline. One probable way of changing the existing state is to make the assessment procedure an integral part of the learning procedure since life-long learning requires that the persons be able not only to work independently but also to evaluate their own performance

and progress. Accordingly, new methods are required to encounter the challenges of assessment and evaluation in the field of English Language Teaching (ELT) and education. This study addresses this gap by examining the impact of CA, compared to static assessment, on the EFL learners' reading comprehension performance, metacognitive strategies of reading, and critical thinking.

2. Review of the Related Literature

Collaborative assessment is based on the works of Piaget and Vygotsky. Somervell (1993) looks at collaborative assessment as a teaching and learning procedure in which the students and instructor meet to clarify the goals and standards. In collaborative-assessment (or co-assessment), "the learner(s) and instructor(s) collaborate to clarify the objectives, negotiate the details of the assessment procedure, discuss any misunderstandings that exist, and provide a mutually agreed assessment of the work or the performance of the learner(s)" (Gouli et al., 2008, p. 616).

Although CA is novel to Applied Linguistics, there have been some studies that have discovered the use of this method in L2 educational contexts. One example is the study of McConnell (2002) who examined the different ways in which the learners talk about their perceptions and experience of the collaborative assessment as it happens in e-learning situations. The outcomes of this study showed that a positive social climate is essential for fostering collaborative assessment and this form of assessment assists the learners in decreasing their dependence on the lecturers. In addition, his study revealed that the learners foster language skill and view themselves as competent learners in evaluating their own and each other's work, which are certainly good lifelong learning skills.

Moreover, the outcomes of another study by Chau (2005), regarding the effect of collaborative assessment on language development and learning, revealed a shift in the students' roles from a "passive learner" to an "active participant" and change in the use of learning skills. This study views CA as an experiential journey and a process of development for the learners as participants, without decreasing the evaluative characteristics of assessment.

Similarly, one another study was carried out by Sluijsmans, Dochy, and Moerkerke (1999) on creating a learning condition by employing self-, peer-, and co-assessment. The findings indicated that the administration of such forms of assessments speeds up the development of a curriculum based on competencies (knowledge as a tool) rather than knowledge (as a goal) and leads toward the integration of instruction and assessment in higher education.

As to collaborative assessment techniques, Joshi (2013) carried out a study to assess the English speaking skills of ESL learners. The outcomes of the study evinced that "incorporating collaborative assessment techniques leads to the beneficial and meaningful learning of English language which is the pre-requisite for a successful study, career, and life altogether in the 21st century" (p. 116).

In a different study by Cortright et al. (2003), there was a detailed investigation of collaborative-group testing and the student's retention of the course content. The results of the study revealed that the student's retention of the course content decreased when the students completed the original examinations individually. In sharp contrast, the student's retention improved when the students completed the original examinations in collaboration and within the groups. The findings suggested that "collaborative testing is an effective strategy to promote the student's learning and retention of the course content" (p. 102).

Overall, this literature review shows the benefits obtained by employing collaboration in the learning and assessment procedure. Collaborative assessment has been utilized successfully across the different teaching and learning situations. This background argues that learning is specifically invaluable if the students involve collaboratively in the assessment procedure and reflect critically on their learning in the various contexts.

Despite the contribution of the abovementioned studies, there are still very few studies investigating the effectiveness of CA in promoting the EFL students' reading performance, critical thinking, and metacognitive strategy. Moreover, a detailed account of the learners' strengths and weaknesses can be gained by utilizing CA. Therefore, this study intends to deal with the following research questions.

1. What are the Iranian intermediate EFL students' perceptions of collaborative assessment, mainly the benefits they have experienced through the use of collaborative assessment?

2. Is there any statistically significant difference between collaborative assessment and static assessment in terms of their relationship with the reading performance of Iranian intermediate EFL students?

3. Is there any statistically significant difference between collaborative assessment and static assessment in terms of their relationship with metacognitive strategies of reading performance of Iranian intermediate EFL students?

4. Is there any statistically significant difference between collaborative assessment and static assessment in terms of their relationship with critical thinking of Iranian intermediate EFL students?

3. Method

3.1. Participants and Research Setting

In order to carry out the study, 62 intermediate students (31 male and 31 female), out of a population of 79 English language learners, with the age range of 15 to 25 were randomly selected. To control for and minimize the variations associated with nationality and first language, all the contributors were Iranians and L1 speakers of Persian in the present study. They were studying at the intermediate level in the institute. A PET proficiency test (Cambridge ESOL examination, 2003) was applied to ensure the homogeneity of the subjects. Based on the results, 62 qualified students were picked and randomly divided into experimental and control groups with thirty one students in each group. The experimental group underwent collaborative method of assessment. On the other hand, the control group underwent the traditional method of assessment. The study was conducted in one term including 10 sessions (2 sessions a week) that each session lasted 45 minutes. In each session, one unit of the Select Reading book for intermediate students (Lee & Gundersen, 2011) was focused. In addition, 15 participants were picked through convenience sampling to be interviewed.

3.2. Instruments

This study was conducted in two phases: Quantitative and qualitative. The instruments that were employed in the quantitative phase of the study included a Preliminary English Test (PET, Cambridge ESOL Examination, 2003), reading comprehension test (selected from among the readings of PET, 2003), metacognitive strategies of reading questionnaire (entailing 30 items on a five-point likert type scale which was designed and validated by Mokhtari and

Richard, 2002), and critical thinking questionnaire (comprising 30 items on a five-point likert type scale which was developed by Honey, 2004). Furthermore, the instruments that were employed in the qualitative phase of the study included an observation scheme (developed by the researcher) and a semi-structured interview (entailing questions which were developed by the researcher). It should be noted that the reliability and validity of all research instruments were examined at the pilot-testing stage of the study. The way in which the instruments of the study were developed for gathering the data as to the variables of the study are explained in the following part.

3.3. Data Collection Procedures

In order to conduct the present study, eight steps were taken. Before conducting the main study, all the research instruments were piloted using 30 EFL intermediate students in order to ensure their reliability and 100 EFL intermediate students in order to ensure their validity. Likewise, for ensuring the homogeneity of the participants with regard to their English proficiency level in each group, a PET proficiency test was administered to 79 subjects. Based on their scores, those participants, whose scores were within the range of one standard deviation above and one below the mean, were selected for this study. The qualified subjects (n=62) were randomly divided into the experimental and control group with 31 participants in each group.

The reading comprehension test was administered to the students of both groups to check their levels of reading comprehension before the treatment stage. Furthermore, the metacognitive strategies of reading and critical thinking questionnaires were administered to both experimental and control groups as their pretests. Subsequently, the treatment of the study was operationalized for the members of the experimental group. It is important to

note that the control group was the static test group of the study. The methodology used in this study was based on the principles of collaborative assessment. It included 45 minutes of collaboration in assessment in the classroom which went on for ten sessions. It was conducted the session after the pre-test, based on the self-, peer- and teacher assessment and two collaborative methods of assessment namely learner-learner collaboration and learner-teacher collaboration. The goal was to make the learners more aware of the collaborative assessment strategies required for better comprehension. In contrast, no treatment was given to the participants of the control group, focusing on just the traditional method of assessment.

The treatment of the study was based on a practical framework (Heron, 1981) which was employed in the study by Haughton and Dickinson (1988). Figure 1 displays an operational model of collaborative assessment which was adapted from Haughton and Dickinson (1988).
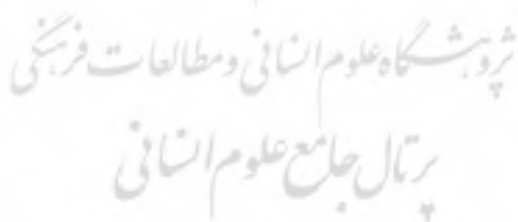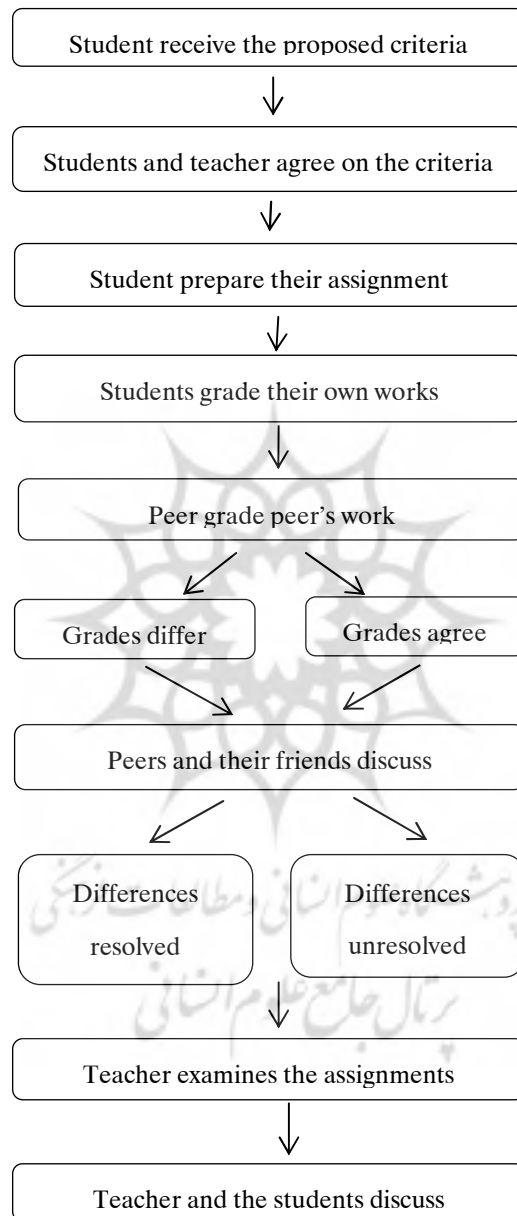
Figure 1. An Operational Model of CA Adapted from Haughton and Dickinson (1988)

Furthermore, six sessions of the course were selected for observation. In each session, an observation scheme was filled by the researcher. During the observations, the researcher took field notes about the learners' behavior and performance. The observations enabled the researcher to verify or differentiate the data collected through the questionnaire and interviews.

At the end, the reading test and the questionnaires were read ministered to the students to compare their meta-cognitive strategies of reading, critical thinking, and reading comprehension achievement. Further, a semi-structured interview was conducted with 15 subjects, to gather more in-depth information from the participants at the end of the course. The semi-structured interview entailed nine questions encouraging the participants to extend their ideas and perceptions of collaborative assessment. In particular the participants were asked to discuss the importance of collaborative assessment in language teaching, and the positive and negative aspects involved in this type of assessment. Each interview lasted 15-20 minutes and the participants' voice was recorded with their permission. At the end of the interview, the recordings were transcribed, coded, and analyzed.

3.4. Data Analysis

This research was carried out within the framework of mixed methods research design. In order to answer the research questions formulated in this study, the gathered data were analyzed through both descriptive and inferential statistics. First, to measure the construct validity and reliability of the questionnaires, the explanatory factor analysis and Cronbach's Alpha coefficient were run. Then, the descriptive statistics of the PET scores and the groups' pre-test and post-test scores, and a multivariate ANOVA (MANOVA), for comparing the

experimental and control groups' gain scores on the reading comprehension test, metacognitive strategy and critical thinking questionnaire, were run. Finally, the content analysis of the data gathered through interviews and observations, using frequency count and descriptive statistic, was carried out.

4. Results

4.1. Quantitative Data Analysis Results

A PET reading comprehension test was administered to the participants of experimental and control groups in the pretest and posttest. Table 1 depicts the descriptive statistics of the two groups' pretest and posttest scores on the PET reading comprehension test, critical thinking, and metacognitive strategies utilized in this study.

Table 1. *Descriptive Statistics of the Groups' Pretest and Posttest Scores*

Dependent Variable	Group		N	Minimum	Maximum	Mean	Std. Deviation
Reading Comp	Experimental	Pretest	31	14.0	28.0	19.87	3.09
		Posttest	31	15.0	30.0	23.67	4.36
	Control	Pretest	31	13.0	31.0	19.32	3.96
		Posttest	31	12.0	30.0	20.29	4.18
CT	Experimental	Pretest	31	46	121	87.57	18.560
		Posttest	31	80	130	107.81	14.418
	Control	Pretest	31	53	146	95.82	23.139
		Posttest	31	49	150	104.09	22.758
Metacog	Experimental	Pretest	31	66	133	103.82	15.02
		Posttest	31	84	150	116.91	16.74
	Control	Pretest	31	63	150	107.23	19.32
		Posttest	31	47	149	107.42	21.66

The mean scores of both the experimental and control groups in reading comprehension test at their pretest were 19.87 and 19.32 respectively. Furthermore, the outcomes of such analysis showed that at the post-test, the mean score of the experimental group changed to 23.67. However, the control group's mean score changed to 20.29 at the post-test. As for the critical thinking pretest and posttest scores, the experimental and control group's mean scores at the pre-test that were 87.57 and 95.82 respectively changed to 107.81 and 104.09 at the post-test. Likewise, the mean scores of both the experimental and control groups in metacognitive strategies questionnaire at the pretest were 103.82 and 107.23 that changed to 116.91 and 107.42 at the post-test. Figures 1 displays the experimental and control groups' gained mean scores on reading comprehension, critical thinking and metacognitive strategies of reading.
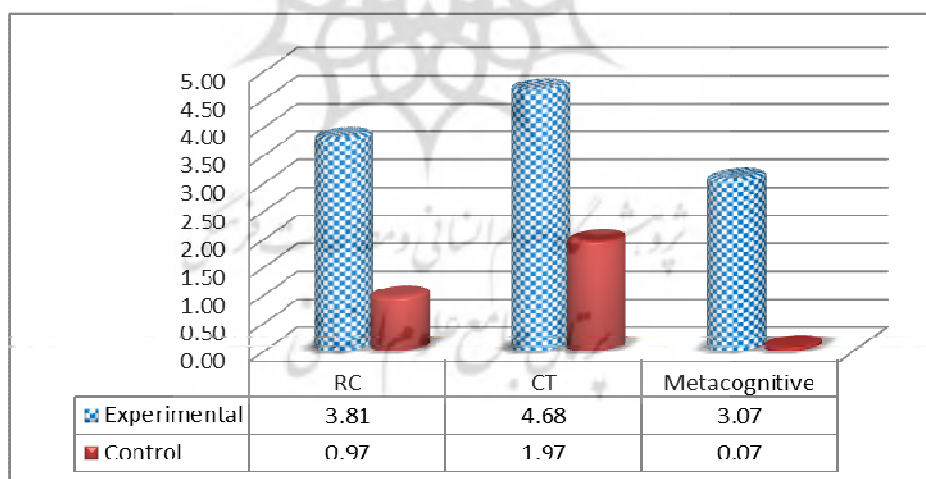

	RC	CT	Metacognitive
Experimental	3.81	4.68	3.07
Control	0.97	1.97	0.07

Figure 2. Gained Mean Scores by Groups

To compare the results, the gained mean scores of the groups were considered for the three tests and analyzed through the multivariate ANOVA (MANOVA) test in order to probe the research questions. In fact, a

28

multivariate ANOVA (MANOVA) was run to compare the experimental and control groups' gained mean scores on the reading comprehension, metacognitive strategies, and critical thinking tests.

Based on the results displayed in table 2, it can be concluded that there were significant differences between the gained mean scores of the experimental and control groups on the reading comprehension, critical thinking, and metacognitive strategies (F $(3, 58)$= 4.15, P< .05, Partial $\eta2$=.17 which represents a large effect size). These significant results justify using separate tests to compare the two groups on the three tests.

Table 2. *Multivariate Tests*

Effect		Value	F	Hypothesis df	Error df	Sig.	Partial Eta Squared
Intercept	Pillai's Trace	.39	12.85	3	58	.00	.39
	Wilks' Lambda	.60	12.85	3	58	.00	.39
	Hotelling's Trace	.66	12.85	3	58	.00	.39
	Roy's Largest Root	.66	12.85	3	58	.00	.39
Group	Pillai's Trace	.17	4.15	3	58	.01	.17
	Wilks' Lambda	.82	4.15	3	58	.01	.17
	Hotelling's Trace	.21	4.15	3	58	.01	.17
	Roy's Largest Root	.21	4.15	3	58	.01	.17

Similarly, the results presented in table 3 confirm to a great extent the results of table 2.

Table 3. *Tests of Between-Subjects Effects*

Source	Dependent Variable	Type III Sum of Squares	Df	Mean Square	F	Sig.	Partial Eta Squared
	Gain-RC	124.90	1	124.90	7.03	.01	.10
Group	Gain-CT	113.80	1	113.80	3.57	.06	.05
	Gain-Meta	139.50	1	139.50	5.10	.02	.07
	Gain-RC	1065.80	60	17.76			
Error	Gain-CT	1911.74	60	31.86			
	Gain-Meta	1639.74	60	27.32			
	Gain-RC	1544.00	62				
Total	Gain-CT	2710.00	62				
	Gain-Meta	1931.00	62				

Based on the results displayed in table 3, it can be concluded that there was a significant difference ($F_{(1, 60)}=7.03$, $P<.05$, Partial $\eta2=.10$ representing a moderate to large effect size) between the two groups' gained means on the reading comprehension. Thus, the first null-hypothesis was rejected. However, based on the results, it can be concluded that there was not any significant difference ($F_{(1, 60)}=3.57$, $P>.05$, Partial $\eta2=.056$ representing an almost moderate effect size) between the two groups' gained means on critical thinking. Accordingly, the second null-hypothesis was supported. As to the third research question, on the other hand, it can be concluded that there was a significant difference ($F_{(1, 60)}=5.10$, $P<.05$, Partial $\eta2=.078$ representing a moderate effect size) between the two groups' gained means on metacognitive strategies. Hence, the third null-hypothesis was rejected.

4.2. Qualitative Data Analysis Results

4.2.1. Interview Results

The interview questions, developed in this study, consisted of three parts. In the first part of the interview, the participants were asked few questions about "whether or not they ever talk about the assessment criteria in the class" and "whether or not it is important for them how the instructor is assessing them." Based on the results, %75 of students mentioned that they have talked about the assessment criteria or the scoring scale in the class in advance whereas %25 of the students claimed that they have not yet talked about the assessment criteria in the class. However, they unanimously believed that it is important for them how the teacher is assessing them.

In the second part of the interview, the students were asked to express their ideas as to being assessed by the partner, the teacher, even by themselves. Further, they were to express their views toward collaborative assessment in general and how collaborative assessment would be different from previous assessment types. The results presented 10 themes for the students' ideas on being assessed by the partner, teacher, even by themselves. In this regard, 5% of the students stated that by CA (Collaborative Assessment) they know what the teacher thinks of them, in other words, what goes in the teachers' mind. However, 15% of the students mentioned that by CA, they can share their information, suggestions, and ideas. Because 35% of the students referred to CA as a great method whereas 5 % of the students mentioned that by CA they became aware of their knowledge and themselves, they can assess each other in a parallel way, and since they are known by their friends more than by the teacher so they are more influenced by them. Besides they stated that scoring in the class is the best way of comparing all the students' performance. Regarding the advantages, 25% of the students referred to improving their

31

knowledge and skills of English by CA and knowing their problems which was mentioned by 10% of the students.

As to understanding of collaborative assessment, 15% of the students defined CA as working in groups and 20% of them defined it as self, peer, and teacher assessment. Further, they differentiated CA from other assessment types. For instance, 20% of the students stated that traditional assessment has been done just by the teachers, but by CA we can listen to my friends' ideas and learn from their experience. 10% of the students asserted that by this method, we can better understand the concepts and meanings, increase our effectiveness regarding CA, 10% of the students saw CA as a professional and modern method. Hereupon, one of them stated that

"Traditional assessment was conducted at the end of the term with strict questions in final exam."

The other student stated that: "By CA students learn and become familiar with the method of teaching at the basic level."

In the next part of the interview, the students were asked few questions about the degree of effectiveness of CA in developing their language proficiency level and skills and whether or not they consider this type of assessment as an efficient way to improve their knowledge of reading comprehension and strategies. Based on the results, 95% of the students considered CA as an efficient way to improve their knowledge of reading comprehension and strategies. Besides, they all found CA a fruitful approach influencing their language proficiency and skills mastery.

In the final part of the interview, 28 themes were explored under three broad categories of "advantages of CA", "disadvantages of CA" and "changes in the assessment approach that the students would like to see." In their responses, nearly 40% of the learners stated that through CA they can

understand what the teacher wants them, they can recognize their problems and solve them, their stress level is decreased, they can also learn many things from their friends, and CA can help them think about things which are expected so that they can become critical thinkers. On the other hand, 15% of the students referred to improving their knowledge of English by CA, encouraging students to work in the class, identifying what they should do to become better, feeling confident, becoming familiar with the assessment criteria and improving their speaking skill. A small number of the students (around 5%) mentioned some other advantages such as becoming smarter than before, learning techniques and strategies, being responsible for their peer's scores, and becoming further familiar with different parts of the passage.

Regarding the disadvantages of CA, the majority of the participants (35%) claimed that CA has no specific disadvantage; however, they referred to time limit as a problem they experienced in the class. A small proportion of the students (nearly 5%), on the other hand, mentioned some of the difficulties of CA as students' preparation for peer assessment, lack of familiarity with the method, its negative impact on the learners' relationship in the group, more competition and high tension among the students, and teacher's lack of control and management in the class.

Finally, as for the changes the students expected in the assessment approach, they pointed on raising the social communicative skills of the learner, assessing more practices done in the class, dedicating more time for complementing the tasks and assessments, and creating a kindly environment (as stated by 5% of the students).

4.2.2. Results of the Classroom Observation

During the study, the teacher played the role of the participant observer and a 5-point Likert-type observation scheme was filled out by the observer every other session. The observation scheme ranges from 1 (not at all) to 5 (well done) that was developed by researchers based on Bloom's Taxonomy of cognitive levels.

In order to analyze the data gathered from the observations, the participants' behavior was grouped under certain themes such as (1) Evaluation: Correct gaps in comprehension, (2) Synthesis: Predict (3) Analysis: Picture, (4) Comprehension: Monitor ongoing comprehension, and (5) Knowledge: Annotate.

The results of the classroom observation showed that, in the early sessions of the study, knowledge and comprehension skills were emphasized and practiced by 70 % of the students since collaborative assessments helped them remember information they had forgotten and better recall or retrieve what had learned before. According to Munzenmaier and Rubin (2013, p. 5) "Knowledge and comprehension are often referred to as lower-order thinking skills." Through recall and comprehension, students may realize what they have learned but cannot recognize when to use their knowledge (Munzenmaier & Rubin, 2013). Moreover in the early sessions of the study, analysis and application skills were employed by the students since collaborative assessments help them interpret information, demonstrate mastery of a concept, apply a skill learned, and recognize the relationships among parts. However in this stage, reading comprehension skills such as evaluation and synthesis were just employed by a small percentage of the students (20%) since they were not thoroughly familiar with collaborative assessment.

Interestingly, by developing the students' self and peer assessment skills at the end of the study, evaluation and synthesis skills were noticed to be utilized by 50% of the students since collaborative assessment helped them measure, choose, revise or solve a problem. Furthermore, observations revealed that by repetition and development of collaborative assessment, the use of analysis and application skills increased from 40% to 70% and comprehension skills developed by 20% at the end of the term. Munzenmaier and Rubin (2013) stated that evaluation and synthesis skills "need higher-order or critical thinking skills. Synthesis calls for a creative behavior because learners produce newly constructed and, many times, unique products. On the other hand, evaluation involves making judgments about a value."

More significantly, despite primary anxieties of getting feedback from the teacher and peers, observations showed more enthusiasm and motivation among the students for collaborative works and assessments. In addition, observations depicted the level of progress in the students' answers to the questions, the ability to better organize the ideas and how to adjust assessment strategies for different purposes, for example, to find the main ideas or the details in the text. In particular, the students were able to better rate their own performance and the performance of their peers using the cooperative assessment rubric owing to their practices and discussions over the rubric. However, observations revealed that some of the learners, whose proficiency level was not optimal, failed to express themselves and the ideas in negotiations, so they required more support and assistance from the instructor to work successfully in the groups.

Overall, the results of the classroom Observation showed that both more proficient and less proficient students benefited from doing the cooperative

assessment tasks in that they had more opportunities to discuss and utilize the language and the students were engaged and more active in their learning.

5. Discussion of Findings

The current study was carried out to investigate the effect of collaborative assessment on enhancing reading comprehension, metacognitive strategies and critical thinking of EFL intermediate students. As depicted earlier, the outcomes of data analysis rejected the first and second null hypotheses of the study while the third one was supported. In other words, the students' performance on the tests demonstrated that collaborative assessment has a meaningful significant effect on improving reading comprehension and metacognitive strategies of the students, whereas it has no significant effect on the students' critical thinking.

The findings confirm some of the findings of earlier studies; however they are in conflict with some others mentioned in the literature review. As for the first research question, the outcome of this study is in line with Wiggins (1990, p. 5) who mentioned that "what you assess is what you get; if you don't assess it, you won't get it. In addition, the result of this study supports the findings of the study by El-Koumy (2009) who suggested that "classroom assessment is less effective in improving the secondary school EFL students' basic reading skills, but more effective in developing their inferential reading skills than traditional assessment" (p. 2). Moreover, the result of the study concurs with what Vangah (2013) and Moheidat and Baniabdelrahman (2011) revealed in their study that self-assessment has a significant effect on the students' reading comprehension and with what Van Zundert et al. (2010) found in their study that the practice of peer assessment ameliorates the learners' performances and positively influences their attitudes toward its practice.

On the other hand, as shown earlier in the literature, the assumptions of collaborative assessment have been derived from collaborative learning. Therefore, the results of the studies by Jalilifar (2010) and Marzban and Akbarnejad (2013), who revealed the significant effect of cooperative learning on improving the EFL students' reading comprehension achievement, can further support the results of this study

Considering the effects of collaborative assessment compared to static assessment on the metacognitive strategy use of the EFL intermediate students, the results of the study represent what Kletzien and Bendar (1990) concluded in that particular strategies can assist the students in solving their reading comprehension problems. Particularly, the strategy analysis in the students' performance of this study indicated that collaborative assessment assisted the students in activating their evaluation, synthesis, analysis, comprehension, and knowledge skills that led to their better understanding of the text. Moreover, the outcome of this study is in line with the idea of Clarke and Silberstein (1977), who inferred that second language students require some cognitive strategies such as estimating from the passage, describing expectations, making interpretation from the passage, and skimming in order to read more effectively. Indeed, what can be interpreted from the research findings is that, collaborative assessment enhances the learners' use of the metacognitive strategies of reading and performance.

In the present study, the collaborative assessment techniques gave the learners the opportunity for analyzing, synthesizing, and evaluating the opinions cooperatively. This group discussion assisted the learners in learning from each other's knowledge, skills, and experiences. The students provided reasons for their judgments and thought about the criteria applied prior to giving them judgments.

Despite the aforementioned activities, the outcome of the study doesn't show the impact of collaborative strategy on critical thinking. These findings are consistent with Hosseini (2009), Shabani (1999) and Wiggs (2011) and are in conflict with the findings of Johnson, Archibald, and Tenenbaum (2010) and with Gokhale (1995) who found that the learners who are involved in collaborative tasks accomplish significantly better on the critical thinking test than those who study individually.

All in all, the general outcome of this study reveals that collaborative assessment has a significant effect on reading comprehension and metacognitive strategies but not on their critical thinking. In addition, the qualitative analysis of the study indicates that collaborative assessment leads to good discussions, helps the students realize what is being asked more clearly than before, which response is correct and why, assists them in having an active role in the class and increases their use of metacognitive strategies of reading.

Particularly, the learners can benefit from collaborative assessment since their reading skills can be more accurately evaluated, which can in turn lead to a higher level of reading comprehension. Therefore, the outcome of this research study can confirm that the collaborative assessment is an effective approach in the area of EFL reading comprehension and metacognitive strategies of reading. More importantly, the findings of this study backup the hypothesis that collaborative assessment is a proper substitution for static assessment.

6. Conclusions and Implications

As depicted earlier, the outcomes of data analysis rejected the first and second null hypotheses of the study while the third one was supported. In other words,

the findings of this study showed that CA had significant effects on promoting the students' reading comprehension and metacognitive strategies but not on their critical thinking. In other words, it was revealed that the students who benefited from collaborative assessment had a better performance in reading comprehension compared with the students who experienced static assessment.

Moreover, the results of this study demonstrated that collaboration in assessment led to both improved performance and enhanced understanding of the processes underlying that performance.

More importantly, this method has had additional benefits compared with statistic assessment methods like an environment for negotiation, social interactions, and opportunities for criticism which all enhanced the learners' disposition toward the use of metacognitive strategies. Therefore, the collaborative assessment technique made the class more active because the students were engaged in the assessment process checking their own and peers' progress and finding their own and peers' strengths and weaknesses. In fact, collaborative assessment became a part of the learners' learning experiences. Throughout the process, the learners learned to foster a high level of responsibility for their learning. Collaborative assessment also provided the learners an environment where they could better realize the nature of assessment. It also assisted them in learning about learning, therefore, developed the metacognitive understanding of their own learning process. It should be noted that since the assessment was integrated with instruction and learning, it assisted the students in realizing their learning potentials. According to McDowell (1995, as cited in Sluijsmans, Dochy, & Moerkerke, 1998), the main strengths of CA are "(1) there is a development of evaluative and critical abilities, (2) there are opportunities for skills development, (3)

knowledge is more integrated, and (4) students highly collaborate and are motivated and satisfied" (p. 27).

Bases on the findings of this study, it can further be concluded that by CA, readers use metacognitive strategies of reading comprehension frequently and adequately; therefore, they can become more successful readers. Therefore, the significant role of CA should be taken into account in the improvement of the students' reading comprehension and use of reading skills.

The educational value of collaborative assessment lies in the fact that the findings of this research encourage the teachers to use collaborative assessment as a practical technique in their classroom and enable them to become facilitators and collaborators. It also encourages the teachers, who still believe in teacher-centeredness in language teaching, to change their viewpoints. The results also help the students to participate effectively in the classroom and improve their involvement, use of metacognitive strategies, and reading comprehension. CA can be used along with other achievement tests to assist instructor to come up with more accurate measurement of the learners' works. Syllabus designers in designing syllabuses should also investigate and consider the students' right for their own suggestions, decisions, and criticism. Additionally, there are a few readily available materials about the use of fruitful assessment methods in the foreign language classroom. For instance, in the teacher's guide, a section can be devoted to explicitly reveal what assessment strategies can be used for each task and how it can help the teachers and students to improve their teaching and learning.

Although this study tightly backs up the positive role of collaborative assessment in L2 reading comprehension development, there is a need for further studies of CA to be conducted not only in the domain of L2 reading comprehension, but also in all other language skills and sub-skills.

Furthermore, CA can be applied to courses other than English. Future research studies can examine the role of individual differences like motivational factors, self-efficacy, or resilience in the learners' acceptance of collaborative assessment.

References

Alhaidari, M. S. (2006). *The effectiveness of using cooperative learning to promote reading comprehension, vocabulary, and fluency achievement scores of male fourth-and fifth-grade students in a Saudi Arabian school.* Doctoral dissertation, The Pennsylvania State University.

Cambridge ESOL Examinations (2003).*Cambridge Preliminary English Test 2 Student's Book: Examination Papers from the University of Cambridge ESOL Examinations (PET Practice Tests).* Cambridge: Cambridge University Press.

Chau, J. (2005). Effects of collaborative assessment on language development and learning. *Language Learning Journal*, *32*(1), 27-37. http://dx.doi.org/10.1080/09571730585200161

Clarke, M. A., & Silberstein, S. (1977). Toward a realization of psycholinguistic principles in the ESL reading class. *Language learning*, *27*(1), 135-154.

Cortright, R. N., Collins, H. L., Rodenbaugh, D. W., & DiCarlo, S. E. (2003). Student retention of course content is improved by collaborative-group testing. *Advances in Physiology Education*, *27*(3), 102-108. http://dx.doi.org/10.1152/advan.00041.2002

El-Koumy, A. S. A. K. (2009). The effect of classroom performance assessment on EFL students' basic and inferential reading skills. Retrieved March 15, 2014, from: http://www.eric.ed.gov/ERICWebPortal/contentdelivery/servlet/ERICServlet?accno=ED514530

Giraund, G., & Enders, C. (2000, April). *The effects of repeated cooperative testing in an introductory statistics course.* Paper presented at the Annual Meeting of the American Educational Research Association. San Diego, CA.

Gokhale, A. A. (1995). Collaborative learning enhances critical thinking. *Journal of Technology Education, 7*(1), 22-30.

Gouli, E., Gogoulou, A., & Grigoriadou, M. (2008). Supporting self-, peer-, and collaborative- assessment in e-learning: The case of the peer and collaborative assessment Environment (PECASSE). *Journal of Interactive Learning Research, 19*(4), 615-647.

Haughton, G., & Dickinson, L. (1988). Collaborative assessment by masters' candidates in a tutor based system. *Language Testing, 5*(2), 233-246. http://dx.doi.org/10.1177/026553228800500208

Helmericks, S. (1993). Collaborative testing in social statistics: Toward gemeinstat. *Teaching Sociology, 21*(3), 287-297. http://dx.doi.org/10.2307/1319027

Heron, J. (1981). Self and peer assessment. In T. Boydell & M. Pedler (Eds.), *Management self-development, concepts, and practices* (pp. 111-128). Farnborough: Gower.

Hollingsworth, A., Sherman, J., & Zaugra, C. (2007). *Increasing reading comprehension in first and second graders through cooperative learning* (Unpublished master's thesis). Saint Xavier University, Chicago, Illinois.

Hosseini, Z. (2009). Collaborative Learning and critical thinking. *Journal of Iranian Psychologists, 5*(19), 199-208.

Honey, P. (2004). Critical thinking questionnaire. Retrieved February 10, 2014, from Peter Honey Learning Website: http://www.PeterHoney.com

Jalilifar, A. (2010). The effect of cooperative learning techniques on college students' reading comprehension. *System, 38*(1), 96-108.

Johnson, T. E., Archibald, T. N., & Tenenbaum, G. (2010). Individual and team annotation effects on students' reading comprehension, critical thinking, and meta-cognitive skills. *Computers in Human Behavior, 26*(6), 1496-1507.

Joshi, N. S. (2013). Collaborative assessment techniques to assess English speaking skill of ESL learners. *International Electronic Journal of the Teachers of English, 3*(2), 103-117.

Kletzien, S. B., & Bendar, M. R. (1990). Dynamic assessment for at-risk readers. *Journal of Reading, 33*(7), 528-533.

Lee, L., & Gunderson, E. (2011). *Select Reading: Intermediate.* (2nd Edition). Oxford: Oxford University Press.

Marzban, A., & Akbarnejad, A. A. (2013). The effect of cooperative reading strategies on improving reading comprehension of Iranian university students. *Procedia-Social and Behavioral Sciences, 70,* 936-942. http://dx.doi.org/10.1016/j.sbspro.2013.01.141

McConnell, D. (2002). The experience of collaborative assessment in e-learning. *Studies in Continuing Education, 24*(1), 73-92. http://dx.doi.org/10.1080 /01580370220130459

Moheidat, A. S., & Baniabdelrahman, A. A. (2011). The impact of Omani twelfth-grade students' self-assessment on their performance in reading in English. *Asian EFL Journal, 13*(1), 48-84.

Mokhtari, K., & Richard, C. A. (2002). Assessing the students' metacognitive awareness of reading strategies. *Journal of Educational Psychology, 94*(4), 249-259. http://dx.doi.org/10.1037//0022-0663.94.2.249

Muir, S. P., & Tracy, D. M. (1999). Collaborative essay testing: Just try it! *College Teaching, 47*(1), 33-35. http://dx.doi.org/10.1080/87567559909596077

Munzenmaier, C., & Rubin, N. (2013). *Bloom's taxonomy: What's old is new again.* Santa Rosa, CA: The Learning Guild Research.

Pain, H., Bull, S., & Brna, P. (1996). A student model 'for its own sake'. In P. Brna, A. Paiva, & J. A. Self (Eds.), *Proceedings of the European Conference on Artificial Intelligence in Education.* Edicöes Colibri, Lisbon.

Shabani, H. (1999). *The effect of group problem-solving on critical thinking and learning achievement in 4-grade primary school students in Tehran* (Unpublished doctoral dissertation). Tarbiyat Modarres University.

Sluijsmans, D., Dochy, F., & Moerkerke, G. (1998). Creating a learning environment by using self-, peer- and co-assessment. *Learning Environments Research, 1*(3), 293-319.

Sluijsman, D., Dochy, F., & Moerkerke, G. (1999). *The use of self, peer-and co-assessment in higher education.* Paper presented at the EARLI-conference, Göteborg.

Somervell, H. (1993). Issues in assessment, enterprise and higher education: the case for self-, peer-, and collaborative assessment. *Assessment and Evaluation in Higher Education, 18*(3), 221-233. http://dx.doi.org/10.1080/0260293930180306

Stearns, S. A. (1996). Collaborative exams as learning tools. *College Teaching, 44*(3), 111-112. http://dx.doi.org/10.1080/87567555.1996.9925564

Van Zundert, M., Sluijsmans, D., & Van Merriënboer, J. (2010). Effective peer assessment processes: Research findings and future directions. *Learning and Instruction, 20*(4), 270-279. http://dx.doi.org/10.1016/j.learninstruc.2009.08.004

Vangah, F. P. (2013). Effect of self-assessment on Iranian EFL learners' reading skill and vocabulary knowledge. *International Research Journal of Applied and Basic Sciences, 4*(3), 676-680.

Wiggs, C. M. (2011). Collaborative testing: Assessing teamwork and critical thinking behaviors in baccalaureate nursing students. *Nurse education today, 31*(3), 279-282. http://dx.doi.org/10.1016/j.nedt.2010.10.027

Wiggins, G. (1990). The case for authentic assessment. *Practical Assessment, Research & Evaluation, 2*(2), 239-244.

Yuretich, R. F., Khan, S. A., Leckie, R. M., & Clement, J. J. (2001). Active-learning methods to improve student performance and scientific interest in a large introductory oceanography course. *Journal of Geoscience Education, 49*(2), 111-119.