The Impact of Collaborative Testing on Iranian EFL Learners’ Course Achievement

Aram Reza Sadeghi Beniss  
*Semnan University, Iran*  
aramsadeghy@semnan.ac.ir

Fatemeh Tajalli  
*University of Semnan, Iran*  
fatimahtajalli@gmail.com

How to cite:  
https://doi.org/10.36892/ijlts.v3i4.256

**ARTICLE HISTORY**  
Received: 22/08/2022  
Accepted: 25/11/2022

**Abstract**  
Students recognize that in order to achieve their academic goals, the other students in the group must also achieve theirs. Active engagement allows students to practice interpersonal skills while accomplishing academic goals. This study aimed to assess the effectiveness of collaborative and individual testing on the course achievement of Iranian EFL learners. To accomplish this purpose, a group of 60 intermediate EFL learners from two distinct language schools were chosen based on convenience sampling. They were randomly divided into two equal groups, namely, collaborative and individual testing. During the research, each group’s members took eight tests. Collaborative group members took the tests in groups, whereas individual group members took the identical tests independently. The findings of the independent sample t-test revealed that there is a significant difference between the collaborative and individual groups’ performance in course achievement, in such a way that the collaborative group outperformed the individual group.

**KEYWORDS**  
Collaborative testing, individual testing, course achievement

1. **INTRODUCTION**

Individual testing is the predominant testing method in Iran. This testing method’s only objective is to assess student achievement and knowledge. In other words, quizzes are viewed as a basis for grades. However, these means can play the role of both a tool for grading and a teaching technique (Murray, 1990). Their time-consuming process makes them potentially used more effectively by turning them into a learning experience. To fulfil teaching expectations adopting a new testing method seems essential.

Collaborative testing, which is also called by other terms like cooperative testing, paired testing, double testing, group examination and dyad testing, is generally defined as an extension of collaborative learning where students are required to work together to complete a test (Durrant Pierson & Allen, 1985; Lusk & Conklin, 2003). Even the given definition sends the message that all theoretical perspectives underlying cooperative learning also come together to
support collaborative testing. The two theories mentioned below, individually or together, provide support for cooperative learning and, consequently, collaborative testing, where students work together to learn.

The cognitive-developmental theory grounded in Piaget and Vygotsky's (Johnson & Johnson, 1999) provides a rationale for the significance and success of collaborative testing. These two linguists ascertained that social interaction is essential for human development. Vygotsky (1978) theorized that when individuals are involved with communication, their cognition develops through discussing and processing the topic verbally. Similarly, Piaget defined cooperation as attaining goals while trying to coordinate feelings and viewpoints with others (Johnson & Johnson, 1999). It was similarly confirmed by Smith and MacGregor (1992) that learning is the outcome of social interaction.

Put another way, an individual’s problem-solving skill and ability to give reasons for responses lead to effective learning, by the same token subjective dimensions. When different learners bring a great deal of knowledge from their diverse backgrounds to their collaborative experiences, they can build new knowledge based on previous knowledge and experiences in this active learning process. When working with others, students listen to others’ perspectives and realize that group work can bring about learning.

Through exploring the value of cooperative group work, cognitive elaboration is another notion that seems relevant to the cognitive perspective. Cognitive elaboration, which is abundantly available while doing collaborative tests, is another ground for this new testing system. This perspective states that working in groups which includes explaining the material to someone else, benefits both the listener and the recaller (O'Donnell, 2000; O'Donnell, 1996). This viewpoint emphasizes the efficacy of explication and elaboration in facilitating the learning and thinking process and the opportunity for practice and reform that elaboration can bring along to improve learning tasks (Slavin, 2011). This opinion was supported by Singhanayok and Hooper (1998) when they considered elaboration as a method of constructing new information on the previous knowledge being restructured during this process. All this can result in a more profound processing of the material. According to Zakaria, Chin and Daud (2010), explaining and clarifying ideas to students will give them a chance to have more successful learning. This can be why recallers benefit more from cooperative group work (O'Donnell, 2000). Research in cognitive psychology has set forth the significance of some form of cognitive reform or elaboration of the content material for the learners to enable the memory to retain information and relate it to existing one (Wittrock, 1986). Therefore, using group work and cooperation may require an efficient elaboration technique in almost all cooperative learning methods (Slavin, 2011).

Currently, language educators challenged themselves to revise curricula and pedagogy through which they successfully employed cooperative language learning techniques. Across disciplines and populations, collaborative learning has been supported by a large body of research (Baumberger-Henry, 2005; Beeken, 1991; Duncan & Dick, 2000; Gokhale, 1995; Rao & DiCarlo, 2000; Burgess & Medina-Smuck, 2017; Caboral-Stevens & Fox, 2020). Collaborative learning has proved to be effective in terms of remarkably higher individual achievement compared to competitive and individualistic approaches (Martin, Friesen & De Pau, 2014). Along with the changes in language teaching methodology, a change in the language testing approach seems necessary.

Although relatively new to TEFL, collaborative testing as an extension of cooperative learning is not new to other disciplines (Lusk & Conklin, 2003; Mitchel & Melton, 2003). Since
The Impact of Collaborative Testing on Iranian EFL Learners’ Course Achievement

1985, in one form or another, collaborative testing has been employed in higher education in different programs and disciplines such as biology, chiropractic medicine, dental hygiene, physiology, psychology, medicine, sociology and physical therapist education (Cortright, Collins, Rodenbaugh, & DiCarlo, 2003; Cortright, Collins, & DiCarlo, 2005; Durrant, Pierson & Allen, 1985; Giuliodori, Lujan & DiCarlo, 2008; Giuliodori, Lujan & DiCarlo 2009; Johnson & Johnson, 1999; Leight, Saunders, Calkins & Withers, 2012; Meseke, Bovee & Gran, 2009; Rao, Collins, & DiCarlo, 2002; Slusser & Erickson, 2006; Zipp, 2007; Rivaz, Momennasab and Shokrollahi, 2015; Caboral-Stevens and Fox, 2020; Merlo, Ediger & Sasaki, 2022) but rarely was it adopted in EFL. Different investigators adopted different forms of collaboration to make the best of collaborative testing. Some of these methods are mentioned below.

Different researchers used the different-number-assignment method. Breedlove (2004) randomly paired partners of the same gender for each of the two exams. Each student was allowed to submit their answer sheet; therefore, reaching an agreement was not a vital factor in his study. Cortright (2003) also considers collaborative testing as pair work. The students, however, were given traditional tests first, followed by several questions allowed to be answered collaboratively. The retention of students was tested four weeks later by an individual retest on the previous subset.

Some researchers integrated individual and collaborative testing to keep the tests valuable as a reliable evaluation method. Lusk and Conklin (2003) employed collaborative testing only to test students’ learning of units; by contrast, traditional testing was used for the final exam. After a 20-minute collaboration, the tests were preceded by an extra 40-min of individual testing. Since there was no need to match answers, each student handed in their own answers. Similarly, Hickey (2006) used a mixed method of individual and collaborative testing except for consensus, which was required on each question. Pray Muir and Tracy (1999) chose pairs of students randomly for all the tests which students took. Their attempt to supervise and control the physical environment included desks and chairs that could be moved when necessary. Russo and Warren (1984) chose to have half of the midterm examination taken collaboratively.

Zimbardo (2003) utilized a different approach in which the students were given three non-comprehensive exams. The first test was taken individually, while on the second and third tests, the students were allowed to choose if they would take the tests solo or in collaborative groups of two. If the students chose collaborative testing, the same score on the answer sheet would be given to both group members. If the students’ choice was individual tests, they would be ushered to a separate room to take their exams. Meseke, Bovee & Gran (2009) took collaborative tests on a weekly basis. They randomly selected two cohort groups of students within one basic science course. 73 students took their quizzes in groups of three which were randomly assigned. All major exams were, conversely, taken individually. Quiz scores of two groups were significantly different while no differences were found in the exams taken from units or the final exam scores.

LoGiudice, Heiz and Kim (2021) conducted a study in which 79 psychology students consented to take two tests in two groups of individual and collaborative testing. Groups of 3 to 4 were randomly assigned to collaborative groups and took a quiz they had been informed about in advance. But, the second test was a pop-up post-test with questions similar to the first quiz for half of the students while the rest took a newly-designed test. The collaborative group students appeared to have better performance in the post-test.
Another study done in 2021 revealed students' perception of their taking tests collaboratively. This study employed collaborative testing for four out of eight tests. However, the rest of the tests were taken traditionally. Physical education students believed that their autonomy and critical thinking were enhanced during the treatment. As an additional benefit, they mentioned a raise in their competence owing to the discussion they took part in while taking collaborative tests (Patiwael, Douma, Bezakova, Kusurkar & Daelmans, 2021).

Rarely has collaborative testing been implemented in the Iranian context. There are some traces of it in nursing. Rivaz, Momennasab and Shokrollahi (2015), for instance, implemented a collaborative testing method in nursing classes at Shiraz University. The results confirmed better achievement and students' positive perception towards this testing method. No literature regarding implementing the collaborative testing method in other disciplines, like EFL in Iran, is available.

Besides better achievement, collaborative testing has also proven to have other benefits. One of the most favourable positive impacts is the retention of course material. Some researchers associated better and deeper learning with collaborative testing (Bloom, 2009; Cortright, Collins, Rodenbaugh & DiCarlo, 2003; Jones & Lishman, 2011), although other researchers reported that improved performance did not continue to exist in final exams (Leight, Saunders, Calkins & Withers, 2012; Meseke, Bovee & Gran, 2009; Molsbee, 2013; Sandahl, 2010; Slusser & Erickson 2006; Wiggs, 2011; Woody, Woody & Bromley, 2008). Other outcomes associated with collaborative testing included improved critical thinking (Gallagher, 2009) and less anxiety (Haberyan & Barnett, 2010; Kapitanoff, 2009; Keselyak, Saylor, Simmer-Beck & Krust Bray, 2009; Mitchel & Melton, 2003; Pandey & Kapitanoff, 2011) improved relationship between peers (Heglund & Wink, 2011; Kapitanoff, 2009; Pandey & Kapitanoff, 2011; Sandahl, 2010) and a positive effect on students' motivation to learn (Slusser & Erickson, 2006).

Although a large body of study has supported collaborative testing across disciplines in recent years, the experience of collaborative testing in EFL is sparse. This study intended to adopt collaborative testing as a new method that breaks all conventional testing rules. Priority of competition over cooperation and considering student discussion of exam content as misconduct would no longer be a test regulation. In addition, it will afford EFL teachers a learning technique to desirably enhance student course achievement. Accordingly, the researcher decided to examine the impacts of this kind of testing on EFL learners. Hence the question for the study was:

- Is there a significant difference between the achievement of students in the individual testing group and the collaborative testing group?

2. METHOD

Collaborative testing was used to investigate changes in student achievement courses while taking tests collaboratively. This study utilised collaborative testing as an ongoing formative assessment process, composing 40% of the final score. 60 intermediate EFL learners from two different language institutes were selected for this study. Students' scores from the previous semester were used as the pretest scores to homogenise the group members. Random assignment was used to form two equal groups of collaborative and individual testing. Students
in both groups were given 8 tests designed by the researcher based on the content lesson of the previous session during the treatment.

The learners of the collaborative group completed the tests in their groups, whereas individual group members took the same tests individually. For collaborative testing, teachers assign students randomly in groups of three. To avoid social loafing, which was the concern of some of the researchers (Desrochers, Fink, Thomas, Kimmerling & Tung, 2007; Lusk & Conklin, 2003), the groups were not chosen to be long-term in the time frame. Using any resource materials was not allowed, but the collaborative group students were required to discuss the questions during the test. Each group was supposed to submit one answer sheet, so the members needed to reach a consensus. But the final exam was taken collaboratively for neither group.

In this study, Oxford placement test was used as the pretest to determine students' English proficiency and select the sample for the study. To evaluate the probable changes in student achievement, the mean score of students' final exam who experienced collaborative testing was compared to the mean final exam scores of another group of students using the standard language school testing procedure, solo testing.

Two additional issues arose because talking and discussing the questions were allowed for collaborative testing groups. The first one was that this type of test-taking was more time-consuming. Therefore, to accomplish the exams, twice as much time as the traditional exams was allotted. Hence, the total amount of time added up to 40 minutes. The second problem was the noise level which dramatically increased. The teacher's role became noise volume control rather than a proctor.

3. RESULTS

The descriptive statistics related to the pretest scores are shown in Table 1.

Table 1. Descriptive statistics of the control and experimental group on pretest

<table>
<thead>
<tr>
<th></th>
<th>N</th>
<th>Minimum</th>
<th>Maximum</th>
<th>Mean</th>
<th>Std. Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Control (Pretest)</td>
<td>30</td>
<td>14.00</td>
<td>21.00</td>
<td>18.000</td>
<td>2.55963</td>
</tr>
<tr>
<td>Experimental (Pretest)</td>
<td>30</td>
<td>13.00</td>
<td>21.00</td>
<td>17.133</td>
<td>2.66178</td>
</tr>
</tbody>
</table>

The results of the learners' performance on the post-test are shown in Table 4.2.

Table 2. Descriptive statistics of the control and experimental group on the post-test

<table>
<thead>
<tr>
<th></th>
<th>N</th>
<th>Minimum</th>
<th>Maximum</th>
<th>Mean</th>
<th>Std. Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Control (Posttest)</td>
<td>30</td>
<td>26.00</td>
<td>37.00</td>
<td>30.741</td>
<td>2.73584</td>
</tr>
<tr>
<td>Experimental (Posttest)</td>
<td>30</td>
<td>27.50</td>
<td>39.00</td>
<td>32.6167</td>
<td>3.79810</td>
</tr>
</tbody>
</table>
To verify the research question of the study in finding whether collaborative testing has any significant impact on Iranian EFL learners’ language achievement, an independent sample t-test between the post-test scores of the control and experimental group. The results are shown in Table 3.

Table 3.  
**Independent sample t-test between the post-test scores of the control and experimental group**

<table>
<thead>
<tr>
<th>Group(s) Equal variances assumed</th>
<th>t-test for Equality of Means</th>
<th>Levene's Test for Equality of Variances</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>F</td>
<td>Sig.</td>
</tr>
</tbody>
</table>
|                                  | 7.28 | .00  | 9  | 2.19 | 4  | .032             | 1.87500        | .85460               | .1643 \( \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \}\n
The results indicated a statistically significant difference between the post-test scores of EFL learners in the control and experimental group \((t = 2.19, p > .05)\). Thus, the first research question of the study is verified.

4. DISCUSSION AND CONCLUSION

The findings of this study, as the results of other studies on collaborative testing, revealed that students benefit more from the collaborative testing experiment. Results from this study confirm previous reports, which clarified that students' course achievement increased when they were engaged in collaborative testing compared to individual examinations. Pray, Muir and Tracy (1999) agreed with the findings of this study while students were paired to take in-class essay exams collaboratively. This study also agreed with the findings of Lusk and Conklin (2003), which followed the same method of grouping students and time frame. This study showed an increase in students’ grades as (Lusk, & Conklin, 2003; Rivaz, Momennasab & Shokrollahi, 2015; LoGiudice, Heiz & Kim, 2021) did. Based on (Wittrock, 1986), this outcome can be the result of discussions that students get involved with in which they need to elaborate and restructure what they had studied to help their group mates and reach an agreement to answer questions which can bring about better retention and consequently better grades. The results of this study were also in line with the study done by (Patiwael, Douma, Bezakova, Kusurkar & Daelmans, 2021), where figures did not measure students' achievement. However, they reported a better sense of competence through answering a questionnaire.
Therefore collaborative testing not only provides teachers with a more compatible testing approach to teaching methods but affords them an effective learning strategy with the capability of boosting student achievement. Overall, collaborative testing can play an efficient role as a learning strategy in enhancing student course achievement.

This study suggests an alternative to individual testing for teachers who keep concerns about collaborative testing to provide a true reflection of students’ capabilities. These teachers do not require carrying over collaborative testing to the final examination. They can keep individual examinations as their final comprehensive test and make the best of collaborative testing as ongoing in-class formative assessments.

The researcher observed a more relaxed test atmosphere, and test anxiety seemed irrelevant in this type of testing. It could be the result of the elimination of the role of the proctor. Since cheating was not a concern during the process of this type of testing, the teacher was not considered a proctor anymore. Or this reduction in test anxiety might originate from students’ assurance of having other students aside in case of need.

Regarding collaborative tests in which the students were given a single copy of the quizzes to complete collaboratively and required students’ agreement, more in-depth discussions among group members were promoted. It can favourably lead to more cooperative people who can contribute their proportion to teamwork.

REFERENCES


The Impact of Collaborative Testing on Iranian EFL Learners’ Course Achievement


