

Scrambled Cloze Procedure: Does It Influence EFL Learners' Reading Comprehension and Writing Performance?

Jahanbakhsh Nikoopour

Department of English Language Teaching, North Tehran Branch, Islamic Azad University, Tehran, Iran

Mino Bargnil

Department of English Language Teaching, North Tehran Branch, Islamic Azad University, Tehran, Iran

Abstract—The present study investigated the effect of using scrambled cloze procedure on reading and writing among intermediate EFL learners. To fulfill the purpose of the study, 60 participants were tested on a PET test, from whom 48 homogeneous participants were selected and assigned randomly to two groups of 25 and 23, as experimental and control group. Then, both groups were given a pre-test of writing and reading comprehension. During 14 sessions, the experimental group were provided with scrambled cloze samples and tasks while practicing paragraph writing and reading, whereas the control group did not, while they were practicing paragraph writing and reading. At the end, the two groups were tested in reading comprehension and writing, based on a posttest. Two raters scored the writing samples collected from two groups in the pre and posttest. Pearson correlation was used to compute the inter-rater reliability. The results showed the experimental group outperformed the control group in paragraph writing and reading comprehension. However, it was found that gender did not make a significant difference in reading comprehension and writing performance.

Index Terms—scrambled cloze procedure, reading comprehension, paragraph writing

I. INTRODUCTION

Based on the previous studies, a cloze test can measure three types of knowledge: linguistic knowledge, textual knowledge, and knowledge of the world. A standard cloze can be used to measure almost all these types of knowledge (Cohen, 1980), that is, the testees use their linguistic knowledge to deal with the text such as comprehending the text. Textual knowledge is used by the language users to detect the logical intersentential relationships and understand the references within the text. These two types of knowledge are mapped to the content of the text. These different types of knowledge can be tested through using different types of cloze procedure. For instance, in a cloze test in which certain functional words such as articles or prepositions are deleted can measure the linguistic knowledge, or a cloze test in which cohesive devices such as conjunctions, or connectors are deleted can test the textual knowledge.

An important implicit assumption is that cloze tasks produce changes in comprehension over time. For example, some scholars investigated the impact of using cloze tests on comprehension during a period of time (Bloomer, 1962; Heitzman & Bloomer, 1967; Kennedy & Weener, 1973; Schneyer, 1965). However, cloze tasks may have temporary effect on the reading comprehension of the language users, and may not have permanent change in the development of their reading comprehension in the long term. Bloomer (1966) attempted to investigate the learners' reading comprehension after providing them with a single exposure to a cloze task. Unlike past studies, more regular investigations of using cloze tests in reading comprehension and writing classes are required if cloze procedures are considered as a teaching technique for reading comprehension and writing. As Bloomer (1996) claimed, previous studies used questions to assess reading comprehension, which may not provide enough detailed information to assess changes in comprehension processes.

The reputation of cloze tests was in computing readability of texts and language assessment. The validity of cloze tests is under discussion as a measure of comprehension (e.g., Brown, 2013; Chen, 2004; Gellert & Elbro, 2013; Greene, 2001; Kobayashi, 2004; Oller & Jonz, 1994a; O'Toole & King, 2010, 2011; Trace, Brown, Janssen, & Kozhevnikova, 2017). Some other researchers (Alderson, 1979a; Kintsch & Yarbrough, 1982; Klein-Braley & Raatz, 1984; Pearson & Hamm, 2005; Shanahan, Kamil, & Webb Tobin, 1982) claimed that only grammatical and linguistic knowledge can be gauged whereas cloze cannot measure intersentential relationships in a text. Cloze is not a valid measure of text comprehension because it does not measure discourse level knowledge. On the other hand, some researchers (Baker, 2011; Eckes & Grotjahan, 2006) asserted that cloze tests are integrative alternatives to measure reading comprehension. Oller (1973, 1976, 1979) claimed that cloze procedure can measure reading comprehension as well as other language abilities integratively. Brown (2004) stated that in a cloze test, a language user is expected to make correct guesses, which depends on his/her ability to listen, speak, and write. He added that cloze procedure has regained its importance

and become a valid assessment procedure for measuring language skills.

In L1 and L2 setting, while being a controversy among researchers, cloze procedure has widely been used as an integrative measure of reading comprehension. Some researchers claimed that cloze can be an objective, and reliable test of language learners' general comprehension (Bachman, 1982; Clarke, 1979; Cziko, 1978; Oller, 1973). Others questioned the validity of cloze procedure to tackle the intersentential relationships, and they hypothesized that in a cloze test, language users process tasks mainly at the sentence or sub sentence-level and fill in the blanks but they do not reach the intersentential comprehension (Alderson, 1979; MacGinitie, 1961; Miller & Coleman, 1967; Shanahan, et al., 1982). Considering the cloze test as a measure of overall reading comprehension, some researchers (e.g., Klein-Braley & Raatz, 1984; Pearson & Hamm, 2005; Shanahan, et al., 1982) argued that cloze procedure fails to test it. Of course, most of these investigations were done on standard cloze tests in which every nth word is deleted.

However, the proponents of the cloze procedure argued that for performing on a cloze test, testees require to process information at the supra-sentential level (Brown, 2002; Chihara, Oller, Weaver, & Chávez-Oller, 1994; Cziko, 1983; Gellert & Elbro, 2013; Jonz, 1994; McKenna & Layton, 1990). Brown (2004) detected that in standard cloze tests, 56-70% of cloze items are cohesive devices. Jonz (1994) explored that 32% of the cloze items require information beyond sentence levels. Finally, the analyses of item difficulty show that passage-level variables influence the difficulty of individual cloze gaps (Chávez-Oller, Chihara, Weaver, & Oller, 1994; Kobayashi, 2002a; Oller & Chen, 2007; Trace et al., 2017). These critics asserted that standard cloze tests measure para-sentential knowledge.

Reading Comprehension and Cloze Procedure

Many language learners at various levels get into trouble while reading English texts. Several studies demonstrated that most language learners are not able to read English with perfect comprehension (e.g., Lijuan, 2007; Wahyuni, 2000; Wallace, 2001). Such learners are to be helped in comprehending the texts. Assessing reading ability is not limited to only measuring comprehension. There are some more important factors which should be considered while assessing the learners' full understanding in formative classroom assessments. Brown (2004) claimed that if readers fail to comprehend the text, they need to develop some reading strategies to enhance their reading comprehension. For instance, a student may comprehend an academic technical report at the sentence level, but may have a misunderstanding at the discourse level because the learner has not exercised certain strategies. It can be claimed that scrambled cloze tasks involve the learners' mind beyond the separate sentence level, and they need to process the text as a whole, which would be a good practice for them to read and understand the text as a whole or to rewrite or unscramble the sentences as a text.

During the past years, various definitions have been offered for reading comprehension. Reading was defined as "the process of understanding and interpreting information from the text" (Shanahan, 2005; p. 175). Readers are expected to process the integration of several variables in order to understand this complicated language skill. Different researchers considered these factors from separate perspectives which may influence reading comprehension: the familiarity with text, structure and topic, the use of reading strategies, and word recognition (Pang, 2008), lower-level and higher-level processes (Ajideh & Sattarpour, 2014; Grabe & Stoller, 2002), micro skills and macro skills (Brown & Abeywickrama, 2010), and task types (Grabe, 2009).

An important goal in ELT programs will be improving reading comprehension. Thus, many activities in learning to read are organized to develop reading comprehension. Cloze procedure, as a technique used to improve comprehension, requires learners to read and understand a text while filling in the blanks within the text at regular intervals (Bormuth, 1966), so that the completed text could make sense. Cloze procedure was originally designed to gauge readability (Taylor, 1953); later was used to help language learners use contextual clues (Kennedy & Weener, 1973); to increase learners' motivation to read (Bloomer, 1966; Heitzman & Bloomer, 1967); and even to help learners improve their reading ability in many studies (Rye, 1982). Many researchers and teachers have used it successfully in improving learners' reading ability.

Heaton (1991) believes that the goal of using cloze procedure is to measure readability and reading comprehension. It indicates that performance on the cloze procedure reflects the learner's ability to use the language context (Rye, 1982), and reflects learners' success in reading (Steinman, 2002). Hence, if the reader can supply the missing words in a cloze, it shows that he/she could understand the text (Daines, 1982). In other words, cloze passage can be used to assess reading comprehension (Yamashita, 2003). Also, cloze procedure is used to assess the learning or reading strategies used by language learners. As Steinman (2002) asserted while practicing cloze procedure, readers get more awareness of meaning, and other reading skills such as scanning and searching, which are not tackled in normal reading. Ahluwalia (1993) claimed that in practicing cloze procedure, readers get involved in more active reading, and are more conscious of their reading strategies during this process. In other words, teaching cloze strategies is, in fact, teaching reading strategies.

Since reading is a very complex and important skill, some measures have been introduced to best measure it. These methods are divided into two main types: traditional and modern test methods (Brown & Abeywickrama, 2010). As Alderson (2000) asserted, in traditional test methods, learners' reading abilities are measured separately and lower-order reading skills such as detecting dates, linking the referring expressions with the antecedents, and finding the grammatical relations are being emphasized. According to Fulcher (2010), based on the traditional test methods, learners' understanding of the text will be generally indicated; however, the cognitive and metacognitive processes

required to comprehend a text will not be revealed. However, more recent approaches are more holistic and integrative by nature, requiring readers to employ a vast set of resources to grasp the text meaning and understand the intersentential relationships (Bachman & Damböck, 2018).

Research in cloze procedure concentrated on reviewing the cloze tests carefully and refining its procedure. Alderson (1979) detected that the deletion rate could affect learners' comprehension. Brown (2002) found different patterns of word deletion could result in different results. Other researchers found that shorter texts could make the cloze test less reliable (Woods, Grafton, Watson, Sicotte & Mazziotta, 1998); and cloze tests would frustrate some testees compared with other measures (Brown, 2002). Reacting to these critical reviews, researchers focused on modifying the original format of cloze tests, while introducing other alternatives for cloze tests. Several studies attempted to refine the cloze procedure, but very little is known if using cloze tests can improve students' reading comprehension. Apsari (2016) found that learners instructed by cloze passages significantly were better than those taught by full passages in terms of reading comprehension. Thus, cloze passages need to be provided in teaching reading.

Writing Skill and Cloze Procedure

Writing proficiency develops gradually during a period of time. Writing, as a language skill, needs to be instructed. Writing involves conveying the ideas, which is hard for a writer to do. In fact, writing will be a personal device, at the most advanced stage, for expressing one's own knowledge and experience (Bereiter, 1980). The more writers get more proficient in writing, the more they move from knowledge-telling to knowledge-transformation (Bereiter & Scardamalia, 1987). The former is the signal of less proficient writers, and the latter is more complex. As knowledge transformer, writers express ideas and reasoning as a vehicle for the development of personal awareness, philosophical ideas, and knowledge.

An autonomous use of lower-level skills such as spelling, grammar, handwriting, punctuation, keyboarding, and proper diction paves the way for the easy utilization of the higher-level writing macro-strategies used for planning, generating, and revising a text (Bereiter & Scardamalia, 1987). If writers are not skilled in the lower-level skills, they will find it difficult to use macro-strategies to produce a coherent summary or a persuasive essay. Command in both these lower- and higher-level skills is vital for writers to be able to write. According to Li (2000), writing is a complex process particularly for elementary and intermediate language learners, which requires learners to focus on only one aspect of syntactic complexity or grammatical accuracy, etc. at the expense of the other one. In fact, language learners find it difficult to attend to form and meaning at the same time.

Creating a successful written text is a complex process which requires "simultaneous control over a number of language systems as well as an ability to take into consideration the ways the discourse must be shaped for a particular purpose" (Celce-Murcia, 2001; p.48). On the other hand, cloze procedure has been extensively investigated from both theoretical and methodological perspectives. Heaten (1990) stated, "performance on cloze test correlates highly with the reading, writing and speaking abilities" (p.17). Hughes (2003) declared, "in doing cloze tests, learners make use of the abilities that underlie all their language performance" (p.139). Nunan and Carter (2001) found strong relationship between testers' performance in integrative tests (such as cloze test) and in the subcomponents of various other test batteries, testing other language skills, such as writing and speaking.

The purpose of this study was to provide information for effective use of a certain type of cloze; namely, the scrambled cloze procedure, in increasing comprehension and writing ability. The study then focused on these questions; what is the impact of using scrambled cloze procedure on Iranian intermediate EFL learners' paragraph writing/reading comprehension? Is there any significant difference between writing/ reading of male and female Iranian EFL learners who practice scrambled cloze procedure and that of those who do not?

II. METHOD

Research Design

The study was based on an experimental design, in which two intact groups participated in the study. The participants were not *selected* randomly; hence, the design was by nature a quasi-experimental one. However, the participants were randomly *assigned* to an experimental and a control group. The independent variable was the mode of instruction, which influenced the two dependent variables under the study. In one group, the researchers presented scrambled cloze tasks to the participants, whereas in the other, they used the current prevalent techniques for teaching reading and paragraph writing. These two dependent variables were measured before and after the treatment to make intra- and inter-group comparisons possible.

Participants

To carry out the study, the researchers selected a sample of 48 participants based on non-random sampling, including both male and female Iranian EFL learners who enrolled in two English language institutes in Tehran. They were intermediate adult learners within the age range of 18-30, who had been learning English for some years. They were divided into two groups: experimental (N=23) and control group (N=25). Both groups took a pretest of reading comprehension and paragraph writing. The writing pretest consisted of a topic, about which the participants were supposed to write a paragraph. The reading pretest consisted of four short reading passages followed by 20 multiple-choice items. The participants in the experimental group received the treatment, which was practicing scrambled cloze tasks during 14 sessions. The control group received no treatment; however, they practiced the traditional way of

learning reading and writing. At the end, both groups sat for the posttest.

Instruments

Language Proficiency Test: A standardized general proficiency test (PET) was used to help the researchers select the homogeneous participants in terms of their language proficiency. From among the participants, a proper sample was needed to be selected for the study. Accordingly, the test was administered to a group of 60 EFL learners. Then, those learners whose scores fell within one SD below and above the mean ($N=48$) were selected. Afterward, the researchers randomly assigned them to two groups. The test consisted of 60 items in different sub-sections, which were administered to the participants in the first session.

Reading Comprehension Tests: Two teacher-made identical tests of reading comprehension were used to measure the participants' reading comprehension before and after the treatment. The difficulty level of the reading passages of the tests were computed based on Fog Index of readability, which was adapted to be at the level of their textbook. Each of the two identical reading comprehension tests consisted of four short reading passages followed by 20 multiple-choice items. The rationale for selecting such tests was to minimize the practice effect in the study. The reliability coefficients for the two reading tests were computed through using Cronbach alpha.

Paragraph Writing Tests: Two teacher-made writing tests were used to measure the participants' performance in paragraph writing before and after the treatment. Since students in both groups were taught how to write paragraphs during the treatment, the researchers designed two identical tests of writing in which they were required to write a paragraph. Their writing samples in the two tests (pre- and posttest) were scored analytically based on Baily and Brown's (1984) scoring scale, in which five main components of writing were given credits. These five components were organization (introduction, body and conclusion), logical development of ideas (content), structure, mechanics (punctuation, spelling, capitalization ...), variety and quality of expression. Each writing task was scored by two raters to ensure the inter-rater reliability.

Procedure

At first, the research instruments were prepared: a test for measuring learners' language proficiency, a pre- and posttest for reading comprehension, and for paragraph writing. The language proficiency test (PET) was administered to the two groups to help the researchers select the homogeneous learners for the two groups.

The pre-tests of reading comprehension and paragraph writing were administered to both groups to enable the researchers compare the two groups before presenting the treatment. Having compared the mean scores of the two groups, the researchers made sure about the homogeneity of the two groups in terms of reading comprehension and paragraph writing.

The treatment was presenting the scrambled cloze passages and tasks to the participants in the experimental group. The treatment lasted for 14 sessions. In the first session, the researchers administered the reading and writing pretest to them. Then, from the second session on, the researchers presented scrambled cloze procedure to the participants in the experimental group. In 14 sessions, fourteen passages were presented to them in the form of scrambled cloze tasks, in which the sentences were scrambled. The participants were required to fill in the blanks first, and then put the scrambled sentences back into the correct order to form a unified and coherent paragraph.

Having accomplished the treatment in 14 sessions, the researchers administered a posttest of reading comprehension and paragraph writing to the participants. This enabled the researchers to collect data from the two groups in the posttest. Thus, the scores in the pretest and posttest were the basis for intra- and inter-group comparisons.

III. RESULTS

The Analysis of the Language Proficiency Test (PET)

The researchers had to check the homogeneity of the two groups in terms of their language proficiency. As shown in Table 1, the descriptive statistics showed that the mean scores of the two groups ($M= 51.21$, $M= 50.32$) and the standard deviations of the two groups ($SD= 3.75$, $SD=2.86$) were close to each other.

TABLE 1
DESCRIPTIVE STATISTICS FOR THE LANGUAGE PROFICIENCY TEST (PET)

Groups	N	Mean	Std. Deviation	Std. Error Mean
Experimental group	23	51.21	3.75	0.782
PET Control group	25	50.32	2.86	0.573

In order to compare the mean scores of the two intact groups, the researchers used an independent samples t-test. As shown in Table 2, the difference between the two groups was not statistically significant, $t(46) = 2.458$, $p= 0.118$. The results showed that there was no significant difference between the two groups on the PET. Accordingly, the researchers could safely carry out the study with the two homogeneous groups.

TABLE 2
INDEPENDENT SAMPLES T-TEST RESULTS FOR LANGUAGE PROFICIENCY TEST (PET)

		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
PET	Equal v assumed	.015	.902	2.458	46	.118	2.65	1.08	0.481 4.83
	Equal v not assumed			2.457	45.6	.118	2.65	1.08	0.480 4.83

The Analysis of the Writing and Reading Pretests

Since the main focus was on the EFL learners' reading comprehension and paragraph writing, the researchers had to check the participants' homogeneity in these two dependent variables. Hence, a writing and a reading pretest were administered to the participants. The descriptive statistics as shown in Table 3 indicated that the mean and SD of the experimental group (M= 13.89, SD= 2.20) and control group (M= 12.90, SD= 1.89) were close to one another in writing. In reading pretest, the results showed that these measures in the experimental group (M= 14.45, SD= 2.06) and control group (M= 15.20, SD= 1.94) were also close to one another.

TABLE 3
DESCRIPTIVE STATISTICS FOR THE TWO PRETESTS

Groups	N	Mean	Std. Deviation	Std. Error Mean
PW pretest	Exp group	2313.89	2.20	.459
	Control group	2512.90	1.89	.378
RC pretest	Exp group	2314.45	2.06	.430
	Control group	2515.20	1.94	.388

In order to make sure that there was no statistically significant difference between the two groups in these two skills, an independent samples t-test was utilized. The result of the Levene test indicated that the condition of the homogeneity of the variances was assumed, $p > 0.05$. As shown in Table 4, there was no statistically significant difference between the two groups with regard to their paragraph writing, $t(46) = 1.67$, $p = 0.101$. The difference between the two groups was not statistically significant. Also, the difference between the two groups in reading comprehension was not statistically significant, $t(46) = 1.28$, $p = 0.205$. Thus, the researchers could safely start to present the treatment to the candidates. The treatment was using scrambled cloze samples and tasks in the experimental group, while using common prevalent traditional techniques for teaching paragraph writing and reading comprehension in the control group.

TABLE 4
INDEPENDENT SAMPLES T-TEST RESULTS FOR THE WRITING AND READING PRETEST SCORES

		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
PW pretest	Equal v assumed	.657	.422	1.67	46	.101	.991	.59	-.199 2.182
	Equal v not assumed			1.66	43.5	.103	.991	.59	-.209 2.192
RC pretest	Equal v assumed	.001	.975	-	46	.205	-.743	.57	-1.908 .421
	Equal v not assumed			-	45.0	.207	-.743	.58	-1.911 .424

The Analysis of the Writing Posttest (Testing the First Null Hypothesis)

An independent samples t-test was utilized to compare the mean scores of the two groups in order to test the first null hypothesis. The descriptive statistics as illustrated in Table 5 showed the means and SDs of the experimental group (M= 16.00, SD= 1.38) and of the control group (M= 14.30, SD= 1.37). The means are somehow different, whereas the SDs are close to one another.

TABLE 5
DESCRIPTIVE STATISTICS FOR THE PARAGRAPH WRITING POSTTEST

Groups	N	Mean	Std. Deviation	Std. Error Mean
PW posttest	Exp group	2316.00	1.38	.28
	Control group	2514.30	1.37	.27

In order to check whether there was a statistically significant difference between the two groups in their paragraph writing after the treatment, an independent samples t-test was run. The result of the Levene test indicated that the condition of the homogeneity of the variances was assumed, $p > 0.05$. As shown in Table 6, the difference between the

two groups in their paragraph writing, $t(46) = 4.26$, $p = 0.000$, was statistically significant. That is, the experimental group outperformed the control group in paragraph writing; thus, the first null hypothesis restated as “Using scrambled cloze procedure makes no significant difference in Iranian intermediate EFL learners' paragraph writing” is rejected.

TABLE 6
INDEPENDENT SAMPLES T-TEST RESULTS FOR THE PARAGRAPH WRITING POSTTEST SCORES

		Levene's Test for Equality of Variances		t-test for Equality of Means					
		F	Sig.	t	df	Sig. (2-tailed)	(2-Mean Difference)	Std. Difference	Error95% Confidence Interval of the Difference
PW post	Equal v assumed	.182	.671	4.2646	.000	1.70	.39	.89	2.50
	Equal v not assumed			4.2645	6.000	1.70	.39	.89	2.50

The Analysis of the Reading Posttest (Testing the Second Null Hypothesis)

In order to test the second null hypothesis, an independent samples t-test was used to compare the mean scores of the two groups in reading comprehension. The descriptive statistics included in Table 7 shows the mean scores and SDs of the experimental group ($M = 16.15$, $SD = 1.14$) and of the control group ($M = 14.04$, $SD = 1.39$).

TABLE 7
DESCRIPTIVE STATISTICS FOR THE READING COMPREHENSION POSTTEST

Groups	N	Mean	Std. Deviation	Std. Error Mean
RC posttest				
Exp group	23	16.15	1.14	.23
Control group	25	14.04	1.39	.27

These measures in the two groups are different; however, the inferential statistics were used to compare the two groups statistically. In order to check whether there was a statistically significant difference between the two groups in reading comprehension, an independent samples t-test was run. The results of the Levene test indicated that the condition of the homogeneity of the variances was assumed, $p > 0.05$. As shown in Table 8, there was a statistically significant difference between the two groups in reading comprehension, $t(46) = 5.70$, $p = 0.000$. The difference between the two groups was statistically significant. That is, the experimental group outperformed the control group in reading comprehension; thus, the second null hypothesis restated as “Using scrambled cloze tasks makes no significant difference in Iranian EFL learners' reading comprehension” is rejected.

TABLE 8
INDEPENDENT SAMPLES T-TEST RESULTS FOR THE READING COMPREHENSION POSTTEST SCORES

		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
RC post	Equal v assumed	.090	.765	5.7046	.000	2.11	.37	1.36	2.85
	Equal v not assumed			5.7445	3	.000	2.11	.36	1.37

Gender Analysis (Testing the Third and Fourth Null Hypothesis)

In order to test the third and fourth research null hypothesis, an independent samples t-test was utilized to compare the mean scores of the male and female participants in paragraph writing and reading comprehension.

The descriptive statistics as demonstrated in Table 9 show the mean scores and SDs of the female group ($M = 15.11$, $SD = 1.72$) and of the male group ($M = 15.12$, $SD = 1.44$) in paragraph writing posttest. Furthermore, the mean scores and SDs of the female group ($M = 15.25$, $SD = 1.50$) and of the male group ($M = 14.67$, $SD = 1.89$) in reading comprehension posttest are also displayed.

TABLE 9
DESCRIPTIVE STATISTICS FOR GENDER DIFFERENCES IN POSTTESTS

	Gender	N	Mean	Std. Deviation	Std. Error Mean
PW posttest	Female	31	15.11	1.72	.30
	Male	17	15.12	1.44	.34
RC posttest	Female	31	15.25	1.50	.27
	Male	17	14.67	1.89	.45

As shown in Table 10, there was not a statistically significant difference between the female and male participants in paragraph writing, $t(46) = .010$, $p = 0.992$. The difference between the two groups was not statistically significant. Thus, the third null hypothesis restated, as “There is no significant difference between the paragraph writing of male and female Iranian EFL learners who practice scrambled cloze procedure and that of those who do not” is not rejected. That is, male and female participants performed similarly in paragraph writing.

TABLE 10
INDEPENDENT SAMPLES T-TEST RESULTS FOR THE GENDER DIFFERENCES IN POSTTEST SCORES

		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
PW post	Equal v assumed	.665	.419	-	46	.992	-.004	.49	-.99 .98
	Equal v not assumed			.010	38.2	.992	-.004	.46	-.94 .93
RC post	Equal v assumed	.889	.351	1.16	46	.249	.581	.49	-.42 1.58
	Equal v not assumed			1.09	27.2	.285	.581	.53	-.51 1.67

Based on Table 10, the difference between female and male participants in their reading comprehension was not statistically significant, $t(46) = 1.16, p = 0.249$. Accordingly, the fourth null hypothesis restated as “There is no significant difference between the reading comprehension of male and female Iranian EFL learners who practice scrambled cloze procedure and that of those who do not” is not rejected. Female and male participants show homogeneity not only in their writing performance but also in their reading comprehension after practicing scrambled cloze samples and tasks during the study.

IV. DISCUSSION

The findings provided a strong support for the claim that scrambled cloze techniques in language learning classes does improve the participants' reading comprehension and paragraph writing. However, male and female learners performed similarly not only in reading comprehension but also in paragraph writing.

The results of the study were in agreement with those of many studies in the past (Alderson, 1979; Brown, 2002; Woods, et al., 1998). Based on some critical reviews done in cloze-related studies, some modifications in the cloze procedure have been made and some other varieties have been proposed, which have been used in language learning programs. Despite the fact that many investigations were done on the cloze procedure, it is still miscellaneous whether using cloze tests or tasks in language classrooms can help improve reading comprehension (Chávez-Oller, et. al., 1994; Kobayashi, 2002a; Oller & Chen, 2007; Trace, et. al., 2017).

Like scrambled cloze activities, other effective strategies were used by many researchers to improve reading comprehension such as, Concept-Oriented Reading Instruction (Guthrie, et al., 1996, 1999, 2000), Students' Achieving Independent Learning (Pressley, Harris & Marks, 1992), Summarizing-Contextualizing-Infering-Monitoring-Corroborating (Hicks, Doolittle & Ewing, 2004), Survey-Question-Read-Recite-Review-Reflect (Richardson & Morgan, 1997), and the Mind Map technique (Buzan, 2000). In most of these studies, the learners' reading comprehension was enhanced. The findings of the present study confirmed the results of the previous studies, which indicated that cloze tasks produce superior comprehension (Bloomer, 1962, 1966; Kennedy & Weener, 1973). Based on the available evidence from the previous studies done on cloze procedure, it has been widely used as an effective technique for developing reading comprehension and writing performance in language classes (Baker, 2011; Eckes & Grotjahan, 2006).

On the other hand, there have been some disagreement in using cloze tests or tasks in language classes to help improve reading comprehension, which may be because of some deficiencies in research methods in this regard (e.g., Brown, 2013; Chen, 2004; Gellert & Elbro, 2013; Greene, 2001; Jongsmá, 1971; Kobayashi, 2004; Oller & Jonz, 1994a; O'Toole & King, 2010, 2011; Trace, et. al., 2017). Concerning the effectiveness of cloze procedure, cloze tests and tasks will produce changes in comprehension, which may lead to an improvement in comprehension. In such studies, however, the researchers did not make clear that cloze tests or tasks can change which reading strategies or processes. The results of the present study contradicted the findings of some other studies which claimed that other reading activities make more effective comprehension than cloze procedure (Heitzman & Bloomer, 1967; Schneyer, 1965).

The change of cloze tasks in comprehension occurs during time. That is, when cloze tasks were used in language classes during several weeks, researchers came up with some change in comprehension (Bloomer, 1962; Heitzman & Bloomer, 1967; Kennedy & Weener, 1973; Schneyer, 1965). Based on these critical views, cloze tasks may have temporary effect on the reading comprehension of the language users, and may not have permanent change in the development of their reading comprehension in the long term. Cloze tests are not sensitive to intersentential constraints and they measure only lower-order skills; that is, measuring grammatical and linguistic knowledge (e.g., Alderson, 1979a; Kintsch & Yarbrough, 1982; Klein-Bráley & Raatz, 1984; Pearson & Hamm, 2005; Shanahan et al., 1982), rather than gauging intersentential and discoursal knowledge.

V. CONCLUSION

In the beginning of the study, the two groups were homogeneous in terms of their reading comprehension. After the treatment, the inter-group comparisons showed that the reading comprehension of the participants in the treatment group differed significantly from those in the control group. Also, the findings of this study showed that the approach in

enhancing writing performance through scrambled cloze activities among learners produced better outcomes. The use of scrambled cloze activities as a technique in language learning classes might be considered as the merits of study, which increased their skill in writing. The finding obtained through the research procedure led to the conclusion that this study supports the use of scrambled cloze procedure.

The results of the study indicated that when samples of scrambled cloze tasks are carefully constructed and the items are supplied with proper words, and when the scrambled sentences are put together to form a unified and coherent paragraph, they can help increase both reading comprehension and writing potential. This teaching technique is quite helpful for poor readers and/or writers. Some more studies are required to investigate the impact of using scrambled cloze procedure on reading and writing in the long term.

The present study aimed to investigate the gender differences while the intervention was being done during the 14 sessions. There was a uniform procedure for measuring the reading comprehension and writing skill of the two groups. Having all the learning conditions and exam administration environments the same, the researchers found that the two groups of learners; that is, the male and female, showed homogeneity in their reading and writing posttests.

The findings of this study are applicable for devising policies and decision making in EFL settings. Teachers and supervisors can include some varied activities in the EFL learners' language learning programs not only to give variety to their materials, but also to motivate them to practice language learning beyond the sentence level; that is, practice language items at the text or discourse level. The findings showed that authorities, material developers and policy makers in charge of EFL instruction should devise new ways to change the attitudes and perceptions of learners from mere traditional ways of practicing reading or writing. They should try to instill the same mentality in learners' minds that exists in their minds. Another implication is the necessity of familiarizing learners with new techniques, strategies and activities for practicing language skills and components, and giving them a realistic picture of how EFL instruction operates. Through teacher education programs, on-the-job training workshops, teachers' and supervisors' meetings, and journal papers used to update teachers' teaching practice and professional development, teachers, supervisors and mentors could share their views, attitudes, and perceptions to updating classroom techniques, which could lead to discussions, constructive arguments, and finally reaching agreements on points of difference.

Interested researchers are recommended to investigate the impact of other types of cloze test procedure such as catered technique, summary cloze tasks, summarizing tests, cloze elide tasks, editing tasks, etc. on language skills and components. Further research could select the sample purposefully from learners of different proficiency levels, and compare them together while including the language level as a moderator variable. Doing a research while using a single research instrument may not be convincing for the researchers who advocate mixed method designs. It is recommended to do some similar studies while using triangulation either in tools, in data or in methods. Some variables have not been controlled in the study, which may be the source of variability in the results compared with some previous studies. It is suggested to replicate the study while controlling for such variables as intelligence, learning style, personality type, attitude, etc.

REFERENCES

- [1] Ahluwalia, N. (1993). Major issues in cloze procedure. *Indian Journal of Applied Linguistics*, 19(1), 81-96.
- [2] Ajideh, P., & Sattarpour, S. (2014). Investigating the relatedness of cloze-elide test, multiple-choice cloze test, and C-test as measures of reading comprehension. *Journal of English Language Teaching and Learning*, 6(13), 89-112.
- [3] Alderson, J. C. (1979). The cloze procedure and proficiency in English as a foreign language *TESOL Quarterly*, 13(2), 219-227.
- [4] Alderson, J. C. (2000). *Assessing reading*. Cambridge: Cambridge University Press.
- [5] Apsari, Y. (2016). Cloze passage in improving students' reading comprehension. *Eltin Journal, Journal of English Language Teaching in Indonesia*, 4(2), 53-62.
- [6] Bachman, L., & Damböck, B. (2018). *Language assessment for classroom teachers: Assessment for teachers*. Oxford: Oxford University Press.
- [7] Bachman, L. (1982). The trait structure of cloze test scores. *TESOL Quarterly*, 16(1), 61- 70.
- [8] Baker, B. A. (2011). Use of the cloze-elide task in high-stakes language proficiency. *Spain Fellow Working Papers in Second or Foreign Language Assessment*, 9, 1-16.
- [9] Bereiter, C., & Scardamalia, M. (1987). An attainable version of high literacy: Approaches to teaching higher-order skills in reading and writing. *Curriculum Inquiry*, 17(1), 9-30.
- [10] Bereiter, C. (1980). Development in writing. In Gregg, LQ & Steinberg, ER (Eds.). *Cognitive processes in writing* (pp. 73-93). Routledge.
- [11] Bloomer, R. H. (1962). The cloze procedure as a remedial reading exercise. *Journal of Developmental Reading*, 5(3), 173-181.
- [12] Bloomer, H. A. (1966). A critical evaluation of diuresis in the treatment of barbiturate intoxication. *The Journal of Laboratory and Clinical Medicine*, 67(6), 898-905.
- [13] Brown, H. D., & Abeywickrama, P. (2010). *Language assessment principles and classroom practice* (2nd ed.). White Plains, NY: Pearson Education.
- [14] Brown, J. D. (2002). Do cloze tests work? Or, is it just an illusion? *Second Language Studies*, 21(1), 79-125.
- [15] Brown, H. D. (2004). *Language assessment: Principles and classroom practice*. White Plains: NY: Longman/ Pearson Education.
- [16] Brown, J. D. (2013). My twenty-five years of cloze testing research: So what. *International Journal of Language Studies*, 7(1), 1-32.
- [17] Buzan, T. (2000). *The Mind Map Book: Unlock Your Creativity, Boost Your Memory, Change Your Life*. Pearson Education

Ltd.

- [18] Celce-Murcia, M. (2001). Language teaching approaches: An overview. *Teaching English as a second or foreign language*, 2(1), 3-10.
- [19] Ch ávez-Oller, M. A., Chihara, T., & Weaver, K. A. (1994). When are cloze items sensitive to constraints across sentences? In Oller, J.W. Jr. & Jonz, J., editors, *Cloze and coherence*. London: Associated University Press, 229-245.
- [20] Chen, L. (2004). On text structure, language proficiency, and reading comprehension test format interactions: a reply to Kobayashi, 2002. *Language Testing*, 21(2), 235-244.
- [21] Chihara, T., Oller, J.W. Jr., Weaver, K. & Ch ávez-Oller, M.A. (1994). Are cloze items sensitive to constraints across sentences? *Language Learning*, 27(1), 63-73.
- [22] Clarke, M. A. (1979). Reading in Spanish and English: Evidence from adult ESL students. *Language Learning*, 29(1), 121-150.
- [23] Cohen, A. D. (1980). *Testing Language Ability in the Classroom*. Rowley, MA: Newbury House.
- [24] Cziko, G. A. (1978). Differences in First-and Second-Language Reading: The Use of Syntactic, Semantic and Discourse Constraints. *Canadian Modern Language Review*, 34(3), 473-89.
- [25] Cziko, G. A. (1983). Commentary: Another response to Shanahan, Kamil, and Tobin: Further reasons to keep the cloze case open. *Reading Research Quarterly*, 18(3), 361-365.
- [26] Daines, D. (1982). *Reading in the content areas: Strategies for teachers*. Scott, Foresman.
- [27] Eckes, T., & Grotjahn, R. (2006). A closer look at the construct validity of C-tests. *Language Testing*, 23(3), 290-325.
- [28] Fulcher, G. (2010). *Practical Language Testing*. London: Hodder Education/Routledge.
- [29] Gellert, A. S., & Elbro, C. (2013). Cloze tests may be quick, but are they dirty? Development and preliminary validation of a cloze test of reading comprehension. *Journal of Psychoeducational Assessment*, 31(1), 16-28.
- [30] Grabe, W., & Stoller, F.L. (2002). *Teaching and researching reading*. Harlow, Essex, UK: Pearson Education.
- [31] Grabe, W. (2009). *Reading in a second language: Moving from theory to practice*. Cambridge: Cambridge University Press.
- [32] Greene, B. (2001). Testing reading comprehension of theoretical discourse with cloze. *Journal of Research in Reading*, 24(1), 82-98.
- [33] Guthrie, J. T., Meter, P., McCann, A. D., Wigfield, A., Bennett, L., Poundstone, C. C. & Mitchell, A. M. (1996). Growth of literacy engagement: Changes in motivations and strategies during concept - oriented reading instruction. *Reading Research Quarterly*, 31(3), 306-332.
- [34] Guthrie, J. T., Wigfield, A., Metsala, J. L., & Cox, K. E. (1999). Motivational and cognitive predictors of text comprehension and reading amount. *Scientific Studies of Reading*, 3(3), 231-256.
- [35] Guthrie, J. T., Wigfield, A., & VonSecker, C. (2000). Effects of integrated instruction on motivation and strategy use in reading. *Journal of Educational Psychology*, 92(2), 331-348.
- [36] Heaton, J. B. (1990). *Classroom testing*. Longman Group Ltd. UK.
- [37] Heaton, J. B. (1991). *Writing English language Tests*. New York: Longman Group Ltd.
- [38] Heitzman, A. J., & Bloomer, R. H. (1967). The effect of non-overt reinforced cloze procedure upon reading comprehension. *Journal of Reading*, 11(3), 213-223.
- [39] Hicks, D., Doolittle, P.E., & Ewing, T. (2004). The SCIM-C strategy: expert historians, historical inquiry, and multimedia. *Social Education*, 68 (3), 221-225.
- [40] Hughes, A. (2003). *Testing for Language Teachers*. Cambridge: Cambridge University Press.
- [41] Jongsma, E. (1971). *The Cloze Procedure as a Teaching Technique*. Newark, Del.: International Reading association. ERIC ED 050 893.
- [42] Jonz, J. (1994). Textual sequence in linguistic semiosis. In J. W. Oller, Jr. and J. Jonz (Eds.), *Cloze and coherence* (pp. 287-315). London: Associated University Press.
- [43] Heitzman, A. J., & Bloomer, R. H. (1967). The effect of non-overt reinforced cloze procedure upon reading comprehension. *Journal of Reading*, 11(3), 213-223.
- [44] Kennedy, D. K., & Weener, P. (1973). Visual and auditory training with the cloze procedure to improve reading and listening comprehension. *Reading Research Quarterly*, 8(4), 524-541. <https://doi.org/10.2307/747173>.
- [45] Kintsch, W., & Yarbrough, J. C. (1982). Role of rhetorical structure in text comprehension. *Journal of Educational Psychology*, 74(6), 828-834. <https://doi.org/10.1037/0022-0663.74.6.828>.
- [46] Klein-Braley, C. & Raatz, U. (1984). A survey of research on the c-test. *Language Testing*, 1(2), 134-146.
- [47] Kobayashi, M. (2002a). Cloze tests revisited: Exploring item characteristics with special attention to scoring methods. *The Modern Language Journal*, 86(4), 571-586.
- [48] Kobayashi, M. (2004). Investigation of test method effects: text organization and response format: a response to Chen, 2004. *Language Testing*, 21(2), 235-244.
- [49] Li, Y. (2000). Linguistic characteristics of ESL writing in task-based e-mail activities. *System*, 28(2), 229-245.
- [50] Lijuan, S. (2007). An introduction to the experiences of popularization activities of reading in Hong Kong and Macao. *Library Journal (China)*, 25(5). http://en.cnki.com.cn/Article_en/CJFDTotal-TNGZ200705018.htm.
- [51] MacGinitie, W. H. (1961). Contextual constraint in English prose paragraphs. *The Journal of Psychology*, 51(1), 121-130.
- [52] McKenna, M. C., & Layton, K. (1990). Concurrent validity of cloze as a measure of intersentential comprehension. *Journal of Educational Psychology*, 82(2), 372-377. <https://doi.org/10.1037/0022-0663.82.2.372>.
- [53] Miller, G. R., & Coleman, E. B. (1967). A set of thirty-six prose passages calibrated for complexity. *Journal of Verbal Learning and Verbal Behavior*, 6(6), 851-854.
- [54] Nunan, D., & Carter, R. (Eds.). (2001). *The Cambridge guide to teaching English to speakers of other languages*. Cambridge: Cambridge University Press.
- [55] OToole, J. M., & King, R. A. (2010). A matter of significance: can sampling error invalidate cloze estimates of text readability? *Language Assessment Quarterly*, 7(4), 303-316.
- [56] OToole, J. M., & King, R. A. R. (2011). The deceptive mean: Conceptual scoring of cloze entries differentially advantages more able readers. *Language Testing*, 28(1), 127-144.

- [57] Oller Jr, J. W., & Chen, L. (2007). Episodic organization in discourse and valid measurement in the sciences. *Journal of Quantitative Linguistics*, 14(2-3), 127-144.
- [58] Oller, J. W., & Jonz, J. (1994). Cloze and coherence (pp. 1-20). Cranbury, New Jersey: Bucknell University Press.
- [59] Oller, J. W. (1973). Cloze tests of second language proficiency and what they measure. *Language Learning*, 23(1), 105-118.
- [60] Oller, J. W. (1976). Evidence for a general language proficiency factor: An expectancy grammar. *Die Neueren Sprachen*, 75(2), 165-174.
- [61] Oller, J. W. (1979). *Language tests at school*. London: Longman Group Ltd.
- [62] Pang, J. (2008). Research on good and poor reader characteristics: implications for L2 reading research in China. *Reading in a Foreign Language*, 20(1), 1-18.
- [63] Pearson, P. D., & Hamm, D. N. (2005). The assessment of reading comprehension: A review of practices-past, present, and future. In *Children's reading comprehension and assessment* (pp. 31-88). Routledge.
- [64] Pressley, M., Harris, K. R., & Marks, M. B. (1992). But good strategy instructors are constructivists!. *Educational Psychology Review*, 4(1), 3-31.
- [65] Richardson, J. & Morgan, R. (1997). *Reading to learn in the content areas*. Belmont, CA: Wadsworth Publishing Company.
- [66] Rye, J. (1982). *The cloze Procedure and the Teaching of Reading*. London: Oxford: Heinemann Educational Books Ltd.
- [67] Schneyer, J. W. (1965). Use of the cloze procedure for improving reading comprehension. *The Reading Teacher*, 19(3), 174-179.
- [68] Shanahan, T. (2005). The National Reading Panel Report. Practical Advice for Teachers. *Learning Point Associates/North Central Regional Educational Laboratory (NCREL)*.
- [69] Shanahan, T., Kamil, M. L., & Tobin, A. W. (1982). Cloze as a measure of intersentential comprehension. *Reading Research Quarterly*, 17(2), 229-255. <https://doi.org/10.2307/747485>.
- [70] Steinman, L. (2002). Considering the cloze. *Canadian Modern Language Review*, 59(2), 291-301.
- [71] Taylor, W.L.(1953). Cloze procedure: A new tool for measuring readability. *Journalism Quarterly*, 30(1), 414-433.
- [72] Trace, J., Brown, J. D., Janssen, G., & Kozhevnikova, L. (2017). Determining cloze item difficulty from item and passage characteristics across different learner backgrounds. *Language Testing*, 34(2), 151-174.
- [73] Wahyuni, P. (2000). *Teacher's Style in Presenting Expository Texts: A Qualitative Study on Teaching Learning Process in Reading at the First Grade Students of High School at PP Darul Arqam Muhammadiyah Garut*. Unpublished Thesis.
- [74] Wallace, B. (Ed.). (2001). *Teaching thinking skills across the primary curriculum: A practical approach for all abilities*. London: David Fulton Publishers.
- [75] Woods, R. P., Grafton, S. T., Watson, J. D., Sicotte, N. L., & Mazziotta, J. C. (1998). Automated image registration: II. Intersubject validation of linear and nonlinear models. *Journal of Computer Assisted Tomography*, 22(1), 153-165.
- [76] Yamashita, J. (2003). Processes of taking a gap-filling test: comparison of skilled and less skilled EFL readers. *Language testing*, 20(3), 267-293.



Jahanbakhsh Nikoopour is an assistant professor in Applied Linguistics at Islamic Azad University, North Tehran Branch. He was born in Tehran in 1966. He finished his BA studies in TESOL at University for Teacher Education in Tehran in 1990, then accomplished his MA studies at Tarbiat Modarres University in Tehran in 1994. He got his Ph.D. from Islamic Azad University, Science and Research Campus, Tehran, Iran in 2005. His dissertation was “The Wash back Effect of the University Entrance Examinations on EFL Education in Iran.”

Dr. Nikoopour is a faculty member in TEFL Department at Islamic Azad University, North Tehran Branch. He has published several papers in domestic and international journals so far. He is also a member of the Editorial Board for some national and international journals related to TESOL. His research interests include language assessment, teacher education, language learning strategies, and learner variables.



Mino Bargnil is an MA holder in TESOL, graduated from Islamic Azad University, North Tehran Branch. She was born in 1988 in Tehran. She finished her BA studies in English Translation Studies in 2014, and accomplished her MA studies in TEFL in 2018. She has been teaching English at various levels for about ten years. She has shown great interest and creativity in her teaching career. Her research interests include language testing and assessment, teacher education and factors influencing language learning.