The Effect of Content Familiarity and Gender on EFL Learners' Performance on MC Cloze Test and C-Test

Omid Tabatabaei

English Department, Najafabad Branch, Islamic Azad University

Najafabad, Iran

E-mail: tabatabaeiomid@yahoo.com

Shiva Shakerin (Corresponding author)

English Department, Najafabad Branch, Islamic Azad University

Najafabad, Iran

E-mail: Shiva_Shakerin@yahoo.com

Abstract

The present study has investigated the effects of familiar and unfamiliar content on Iranian intermediate EFL learners' performance on Multiple- choice (MC) cloze test and C-test. Also, it has intended to compare the effects of gender on the learners' performance on MC cloze test with familiar and unfamiliar content as well as its effects on their performance on C-test with familiar and unfamiliar content. During the testing period, 107 examinees took an Oxford quick placement test and a homogeneous sample including 30 males and 30 females was selected. They were from Islamic Azad University Najafabad Branch ranged in age from 20 to 35. The first session of the research process was dedicated to MC cloze test with familiar and unfamiliar content and a week later, the second session to C-test with familiar and unfamiliar content. The results showed that the participants had more successful performance on two tests with familiar content. Moreover, the results indicated that gender had no significant effect on the participants' performance on two tests.

1. Introduction

English is now the widespread language of international academic publications. According to Huttner (2008), the importance of publications in English language and the need for language learner to read those texts in various types like research articles, abstracts, stories, recipes and etc. have been confirmed. A very important part of comprehension depends on the prior knowledge which is made up of a person's experience with the world including what that person has read and the person's concepts about how a text is organized. According to Feng (2011), text and background effects may have a crucial role in the development of reading tests. Tests are fundamental parts of a successful curriculum and one of the main partners in language process (Brown, 2004). Based on the importance of testing in language learning, many studies in L2 testing have examined the factors which may influence language test performance. In ESL/ EFL contexts, reading comprehension tests consist of a number of related items which are dependent upon the same reading passage (Lee, 2004).

One of the most common types of reading assessment task is cloze procedure (Brown, 2004). In MC cloze test, students are not required to make a set of words to fill the blanks in the passage. Rather, they are given three, four, or five choices for each blank and they are asked to select the most suitable choice. According to Jonz (1976), MC cloze tests are easy in completing and scoring. The other type of test methods for assessing reading comprehension is C-test. Raatz and Klein-Braley (1981) developed C-test as a replacement for cloze test from both psychometric and theoretical perspectives. A C-test usually consists of four to six short texts in each of which the first sentence is left intact. Then beginning with the second word of the second sentence, the second half of every second word is deleted. Baghaei (2011) described that in addition to the ease of application, scoring and the short time needed for administering; C-tests are valuable proficiency measurements in L2 research.

Another factor considered in this study was the effect of gender on test taker's performance; it is one of the reader variables which may have an effect on reading comprehension. Some studies showed that male and female learners differ in their knowledge, interest, and experiences, and consequently their performance on various reading topics are different (Brantmeier, 2002; 2003; Bugel and Buunk 1996; Schueller, 1999; Young and Oxford, 1997). Paying attention to social studies implies that girls and boys grow and rise differently and these gender differentiations may influence the way they perceive studying foreign languages. Bachman (2000) pointed out that investigation of certain variables' impact on test taking process is necessary. Variables such as gender and content familiarity might influence the performance of individuals differently and can be considered as a threat to the validity of the test. The importance of the external factors brings about the questions of if and to what extent these factors can influence text comprehension and whether these issues can affect testees' achievement on two kinds of tests, namely, Mc cloze and C-test.

2. Literature Review

2.1 Cloze Procedure

According to Oller and Perkins (1978), the philosophy behind the goals of cloze procedure is

the integrative philosophy. The origin of the term 'cloze *procedure*' goes back to more than half a century ago. It was firstly introduced by Wilson Tylor in 1953. Rye (1982) describes that the term '*cloze*' took from the Gestalt psychology concept of 'closure'. 'Gestalt psychology applies to the human tendency to complete a familiar but not quite finished pattern to see a broken circle as a whole one, for example, by mentally closing the gaps... the same principle which is applied to language' (Tylor, 1953, cited in Bastidas, 1984, p. 20). In addition to Gestalt theory, Redundancy Information and Pragmatic Expectancy Grammar theory also tried to describe the cloze test. Oller (1972) claims that those tests in which the reduced redundancy principle is applied begins with the assumption that adult educated native speakers of a language can generally make use of the reduced redundancy of their language to restore damaged messages through their knowledge of the rules, patterns and idioms of their own language and culture, and their competence. Pragmatic expectancy grammar is another theory which supports cloze procedure. This theory was developed by Oller in 1970s to describe 'pragmatic mapping' of a text or discourse onto facts or experiences known in some other manner like the extra linguistic context. Oller (1979) states that there is an inner ability in human beings which causes to predict what will com next in any given sequence of elements.

2.2 Cloze Test

The cloze test was described in some studies (Brown 1994; Cohen, 1980, 1991; Oller, 1979) as a test in which learners are asked to complete the words that are systematically deleted in a passage which deletion may be based on every nth word such as every 6th, 7th, or 8th word. The deleted words are replaced by blanks of standard length and learners are required to complete them with appropriate words; either the original word or a suitable equivalent word.

2.2.1 Types of Cloze Test

There are three types of cloze test: 1) the fixed ratio cloze; 2) the rational cloze; and 3) the MC cloze which are explained in the following part. In fixed ratio cloze deletion of the words are systematically and based on a fixed pattern. The deletion rate is usually from fifth, sixth or seventh every word. This kind of deletion is called random deletion because it deleted every nth word consistently, therefore all types and classes of the words have the same chance of being deleted. According to Sadeghi (2010), this kind of cloze has been widely concentrated on because it retains the original concept of the term cloze itself. The problem which fixed ratio cloze is confronted with is the blanks that can be too easy or too difficult. Therefore, deletion based on selecting the words was suggested and then rational cloze was created.

In rational cloze, selecting the items for deletion is depended on what the test is intended to measure; content words or function words. In this kind of cloze, a special type of word is deleted based on a linguistic principle, such as nouns, verbs, adjectives, etc. (Lu, 2006). Discourse or rational cloze is used to measure specific linguistic abilities in reading assessment, for instance, grammatical features (Lee, 2008). Filling the gaps based on the learners' understanding of the passages is required in both fixed ratio cloze and rational cloze. To remove this problem, another format of cloze test was suggested, namely the MC cloze.

MC cloze test was first suggested by Jonz in 1976. According to Hinofotis and Snow (1980), MC cloze tests are a modified form of cloze test in which the deleted words are substituted with a number of possible choices instead of just using a blank like in traditional completion cloze tests. Deleting words can be based on fixed ratio cloze or rational cloze. In fact this type of cloze is a combination of traditional cloze procedure and MC test. Farhady (1996) mentioned to the advantage of MC cloze tests over the standard cloze tests in a way that the task of comprehension is easier than that of production and students will get higher scores than on an open ended form.

Some studies (e.g., Alderson, 1979; Shanahan, Kamil, & Tobin, 1982) reported that cloze tests did not tap processing abilities beyond the clause level, whereas other studies (e.g., Bachman, 1982; Bachman, 1985; Chihara, 1994; Oller & Taira, 1994) found that cloze tests tapped processing abilities beyond clause boundaries. Some empirical researches have shown that cloze tests can measure both lower and higher level reading abilities depending on the particular word deleted (e.g., Bachman, 1982; 1985; Chavez- Oller, 1994) and on the proficiency level of participants (e.g., Sasaki, 1993, 1996).

2.3 C- Test

C-test is an alternative to cloze test and usually involves four to six short passages in each of which second half of every second word in each text is deleted and the first and the last sentence of the passage are remained intact which is called rule of two (Klein-Braley & Raatz, 1984). Khodadadi and Hashemi (2011) claim that the rule of two or 'C-principle' is the defining feature of C-test. By this rule, test takers are allowed to get familiar somewhat with the topic and style of the text. C-test was proposed by Klein-Braley in 1980s. The rationale behind the C-test is the reduced redundancy principle. It was claimed that C-tests are the best in the family of reduced redundancy tests such as cloze, clozentropy and noise tests. This rule explains that native speakers are able to reconstruct the missing or distorted texts by restoring to different textual information and using natural redundancy in the text (Khodadady & Hashemi, 2011).

Some studies praised C-test due to its ease of construction and administration, scoring and its high reliability and validity (e.g. Connalley, 1997; Dornyei & Katona, 1992; Eckes & Grotjahn, 2006; Grotjahn, 1992, 1994, 1996, 2002, 2006, 2010; Klein-Braley, 1997; Klein-Braley & Raatz, 1984; Sigotte, 2004). However, few studies have been done about its merits such as lack of face validity (Bradshaw, 1990; Jafarpur, 1995; Kontra & Kormos, 2006), poor item discrimination (Cleary, 1988; Kamimoto, 1993; Sigotte & Korbel, 1996), and more importantly, unclear construct validity (Grotjahn, 1986; Jafarpur, 1995; Kamimoto, 1992).

2.4 Reading Comprehension and Schema

For many people especially for language learners, reading is the most important skill among the four skills in second or foreign language for both getting information and pleasure and also it is a means of extending their knowledge of target language. Anderson (2003) argued that being skillful in reading can help ESL and EFL learners achieve success not only in

learning English but also in other content- based classes where reading English proficiently is a necessity. Reading comprehension, as defined by Grabe (1991), is "a combination of identification and interpretation skills" (p. 125). More than just reinforcement of oral communication, fluent reading is done when new information interacts with previous knowledge.

The role that background knowledge plays as one of the external text factors in language comprehension is described in theoretical model called schema theory. The schema theory was developed by cognitive scientists concerning the understanding of how comprehension occurs. This theory in reading explains the process by which readers combine their own background knowledge with the information in a text to comprehend that text. Some researches have been investigated the effects of text familiarity on reading comprehension (Bensoussan, 1998; Kang, 1992; Pulido, 2004; Salmani-Nodoushan, 2003). Their results show that texts with culturally familiar content schema are easier to process.

2.5 Gender

Gender is one of the influential reader variables which makes a sociocultural difference between men and women based upon the characteristics and behavior that are conventionally considered as traits of the two groups of people. There are three main views about gender differences in second language learning. The first one emphasizes the outperforming of female learners in comparison with male learners (Baker & MacIntyre, 2000; Schueller, 2000). The sociolinguistic findings claim that females use a higher frequency of the incoming forms in linguistic changes than males and it seems that it has been used as a first theory for SLA research. The second view holds that male learners outperform female learners in different SLA activities or skills (Brantmeier, 2000; D'Ailly, 2002; Hassan, 2001; Kasanga, 1996). And the last view considers a middle- ground position; it claims that there is no significant difference between the performance of males and females in SLA (e.g. Piske, MacKay & Flege, 2001).

The present study aims to investigate the role of content familiarity and gender as two external factors on Iranian Intermediate EFL learners' performance on MC cloze test and C-test. Therefore, through this study, the authors try to explore the following research null hypotheses:

H_{01}: There is no significant difference between the performance of intermediate EFL learners on MC cloze test with familiar and unfamiliar content.

H_{02}: There is no significant difference between the performance of intermediate EFL learners on C-test with familiar and unfamiliar content.

H_{03}: There is no significant difference between male and female intermediate EFL learners' performance on MC cloze test with familiar and unfamiliar content.

H_{04}: There is no significant difference between male and female intermediate EFL learners' performance on C-test with familiar and unfamiliar content.

3. Methodology

3.1 Participants

Sixty male and female Iranian EFL learners studying at Islamic Azad University, Najafabad branch took part in this study; all of them were native speakers of Persian. Learners were intermediate students whose age ranged from 20 to 35. They were undergraduate and graduate students majoring in English language teaching. They were selected based on their performance on Quick Placement Test (QPT) (2001), version II which was developed by Oxford University Press and Cambridge ESOL. The scores were analyzed and out of 107 examinees, 60 whose scores were between one standard deviation above and one standard deviation below the mean were selected randomly to take part in the study.

3.2 Instruments

For the purpose of data collection, a number of instruments were utilized as presented:

3.2.1 Oxford Quick Placement Test

To determine the general proficiency level of the participants and to screen them, QPT version II, which is a paper and pencil test developed by Oxford University Press and Cambridge ESOL (2001) was administered. It needed to be completed in 30 minutes. It includes 60 questions of vocabulary, grammar and reading cloze test.

3.2.2 Researcher-made Rational Cloze Tests in Multiple-choice Format

Two reading comprehension texts were chosen and changed into cloze tests in MC format each involving 14 items. The two tests were prepared based on two reading passages which were different in content; one with familiar content related to the participants' major, and they studied the similar materials in their previous courses. It was taken from the book; *Reading through Interaction* (Wegmann, Knezevic, & Bernstein, 2007). Another cloze test in MC format with unfamiliar content was also prepared. Its content was not part of the participants' major and they did not study it in their previous courses. It was taken from an English textbook for medical students, *English for the Students of Medicine* (Tahririan & Mehrabi, 1996).

3.2.3 Researcher-made C-tests

Two tests were prepared based on two sets of short reading passages which each set was different in content; one set of four short reading passages with familiar content was related to the participants' major, and they studied the similar materials in their previous courses. This set was taken from the book; *Reading through Interaction* (Wegmann, Knezevic, & Bernstein, 2007). The other C-test was prepared based on unfamiliar content. The material included four short reading passages which were taken from an English textbook for medical students, *English for the Students of Medicine* (Tahririan & Mehrabi, 1996). Therefore, the content was not part of the participants' major and they did not study it in their previous courses. Each of these sets involved 150 deletions separately.

3.2.4 Text Selection

The basis for selecting passages with the proper readability for EFL learners was related to Farhady, Jafarpur and Birjandi's claim in 1994. They pointed out some tips in this relation: 1) Determining the content of the materials studied or is going to be studied and is expected to be tested; 2) Selecting a random sample of passages from the content; 3) Running a certain readability formula for these passages; 4) Considering the average readability index as the criterion for selecting an unseen passage for the group. Therefore, to prepare the familiar cloze test, four passages from the book; *Reading through Interaction* (Wegmann, Knezevic, & Bernstein, 2007) were selected randomly. To determine the readability, the Gunning Fog's index was used and the standard deviation was calculated. Therefore, among four passages, a passage with the difficulty level within the range of 13.405 ± 1.26 was chosen and changed into a cloze test in the form of MC.

Then, to select an unfamiliar passage with the same readability level of the familiar text from the book; *English for the Students of Medicine* (Tahririan & Mehrabi, 1996), the same process of selection of the familiar one was utilized. The average readability and the standard deviation were measured. Therefore, a passage with the readability within the range of 12.76 ± 1.41 was selected and changed into MC cloze test. In each familiar and unfamiliar text, 14 content words were removed based on the rational deletion method. Content words were chosen to remove due to the meaning laden they carry. And the rational deletion method was chosen because based on this method; it is the tester who decides on issues of deletion such as the number of deletions, the selection of word types (content or function words) and also the length of the passage.

Then, according to Klein-Braley (1994), to make C-test, test developer should prepare a number of texts more than what is necessary. Therefore, 22 short passages were totally selected from both familiar and unfamiliar content to decide on 8 final C-tests. Based on Klein-Braley (1997) recommendations, for making C-test, between four and six short and authentic texts should be selected by the test designer. Among those short passages, 16 ones enjoyed the same readability which was determined by the Gunning Fog's index. The readability, mean and standard deviation were measured and then, the texts whose readability were between the range of 13.99 ± 0.79 were selected. These 4 familiar texts were chosen from the book; *Reading through Interaction* (Wegmann, Knezevic, & Bernstein, 2007). Then, to select unfamiliar texts the same process was done but from the book; *English for the Students of Medicine* (Tahririan & Mehrabi, 1996). The average and the standard deviation were calculated. Therefore, the texts with the range between 13.71 ± 0.89 were chosen. After selecting the texts, the second half of every second word in the passage was deleted with the first and the last sentence left intact. The words which consisted of odd numbered letters, the larger half of them was removed and the words which had only one letter were ignored by the researcher. The number of the mutilated words reached 150 in each familiar and unfamiliar group of texts.

3.3 Procedures

3.3.1 Pilot Study

At first, QPT was administered to a population of 40 English language learners studying at Najafabad Azad University. After selecting of the intermediate participants based on their performance on Quick Placement Test (QPT) who had the similar characteristics to the main group, the newly developed MC cloze tests and C-tests which each consisted of two texts with familiar and unfamiliar content were given to them. The KR- 21 reliability formula was used in order to estimate the reliability of four sets of tests, according to which the reliability of familiar MC cloze test was 0.85, unfamiliar MC cloze test 0.82, familiar C-test 0.78 and unfamiliar C-test 0.77. To ensure about the content validity, the newly constructed tests was judged by some EFL professors who had published related articles in some reputable journals.

3.3.2 Main Study

Data collection of the main study was done in three sessions. On the first session, to homogenize the participants, QPT version II was run and a random sample of intermediate learners whose scores were between one standard deviation above and below the mean were selected as the homogeneous group of this study and they were 30 male and 30 female learners. A week after administration of the proficiency test, the group was given the MC cloze test with familiar and unfamiliar content. Based on pilot testing it was made clear that 35 minutes was enough to answer the tests. Each correct answer was given one point and the incorrect one was ignored. A week later in the third session, the group was given the C-test with familiar and unfamiliar content which each set consisted of 150 deletions. According to pilot testing, 40 minutes was enough to complete the tests. To score the C-test, the sum of the correct reconstructed words by the examinees was considered.

4. Results

4.1 Addressing Hypothesis One

To test the first hypothesis, the performance of the participants on MC cloze test with familiar texts had to be compared with their performance with unfamiliar texts. Table 4.1 indicates the descriptive statistics for this comparison, and Figure 4.1 illustrates the means graphically.

Table 4.1. The Descriptive Statistics for Hypothesis One

Groups	No.	Min	Max	Mean	SD
Cloze-familiar	60	50	120	82.83	18.512
Cloze-unfamiliar	60	20	100	63.17	16.621
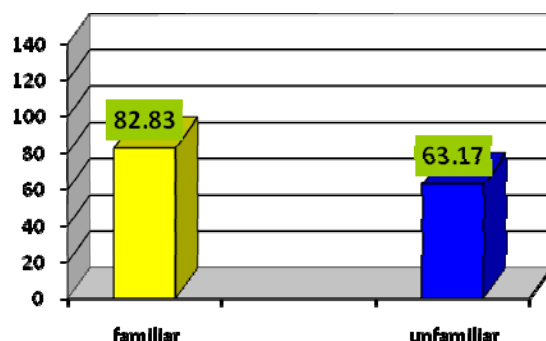
Figure 4.1. Graphical representation of the means for hypothesis one

It can be seen in Table 4.1 and Figure 4.1 that the two means differ from each other. In order to check and see if the difference was significant or not a paired-sample t-test was employed. Table 4.2. presents the results of the t-test.

Table 4.2. The Results of the Paired-Sample t-test for Hypothesis One

Mean Difference	SD	t	df	Sig.
19.67	19.912	7.650	59	.000

By investigating Table 4.2, it can be understood that the amount of t-observed (t-observed= 7.650) is statistically significant because the amount of significance level is p= .000, which is much lower than .05 level. Therefore, the first null hypothesis which states that, "there is no significant difference between the performance of intermediate EFL learners on MC cloze test with familiar and unfamiliar content" is rejected, and, it can be claimed that the opposite is true; in other words, in MC cloze test the participants performed better on familiar texts than unfamiliar ones.

4.2 Addressing Hypothesis Two

For hypothesis two the same procedure which was followed in testing hypothesis one was repeated, but this time with the results of the C-test. Table 4.3 depicts the descriptive statistics for this comparison, and Figure 4.2 shows the graphs for the means of the two sets of scores.

Table 4.3. Descriptive Statistics for Hypothesis Two

Groups	No.	Min	Max	Mean	SD
C-test-familiar	60	46	82	64.47	9.880
C-test-unfamiliar	60	34	86	60.71	12.162
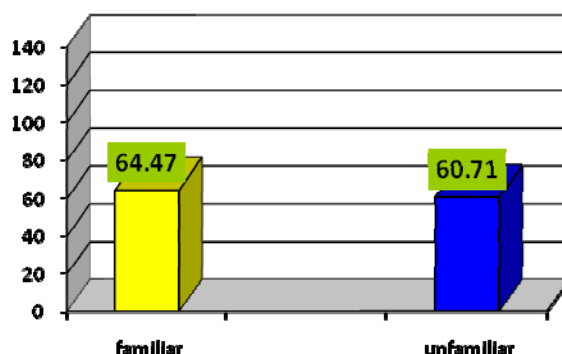
Figure 4.2. Graphical representation of the means for hypothesis two

Table 4.3 and Figure 4.2 confirm that the two means are different from each other. It was decided to check the two sets of scores by employing a paired-sample t-test and see the difference was significant or not. Table 4.4 indicates the results of the *t* test.

Table 4.4. The Results of the Paired-Sample t-test for Hypothesis Two

Mean Difference	SD	t	df	Sig.
3.75	10.635	2.731	59	.008

Table 4.4 shows that the amount of t-observed (t-observed= 3.731) is statistically significant at the probability level of p= .008, which is much lower than .05 level. As a result, the second null hypothesis stating that "there is no significant difference between the performance of intermediate EFL learners on C-test with familiar and unfamiliar content" is also rejected; in other words, like with MC cloze test, on C-test the participants' performance with familiar texts was better than their performance with unfamiliar texts.

4.3 Addressing Hypothesis Three

For this hypothesis two variables were to be considered, gender and (un)familiar texts. Therefore, the performance of the male participants had to be compared with those of the female participants across familiar and unfamiliar texts in MC cloze test. Table 4.5 presents the descriptive statistics for this hypothesis, and Figure 4.3 indicates the means in graphs.

Table 4.5. Descriptive Statistics for Hypothesis Three

Group	Group	N	Min	Max	Mean	SD
Familiar	Male	30	50	110	82.00	18.458
	Female	30	50	120	83.67	18.843
Unfamiliar	Male	30	20	90	63.67	15.643
	Female	30	40	100	63.67	17.905
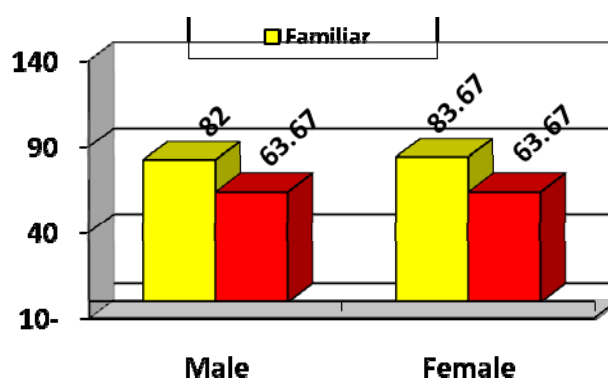
Figure 4.3. Graphical representation of the means for hypothesis three

According to Table 4.5. and Figure 4.3, the means are different from each other. In order to find out if the differences are significant or not, a two-way ANOVA was employed. Table 4.6 shows the results of this ANOVA.

Table 4.6. *The* Results of Two-way ANOVA for Hypothesis Three

Source	SS	df	MS	F	Sig.
Group_F_U	11603.333	1	11603.333	36.917	.000
Group_M_F	53.333	1	53.333	.170	.681
Group_F_U * Group_M_F	3.333	1	3.333	.011	.918
Error	36460.000	116	314.310		
Total	687600.000	119			

Table 4.6 reveals that the amount of F-observed (F-observed= 36.917) for familiar-unfamiliar variable is significant (p= .000), but the amount of F-observed (F-observed= .170) for gender is not significant (p= .681). Moreover, the amount of F-observed (F-observed= .011) for interaction between variables is not significant either (p= .918). Therefore, the third hypothesis stating that "there is no significant difference between male and female intermediate EFL learners' performance on MC cloze test with familiar and unfamiliar content" is retained. In other words, it can be claimed that gender does not play an important role in participants' performance on MC cloze test with familiar and unfamiliar texts.

4.4 Addressing Hypothesis Four

The purpose of this hypothesis was to consider the performance of the male and female participants on the C-test with (un)familiar content. Table 4.7 depicts the descriptive statistics for this hypothesis, and Figure 4.4 illustrates the means graphically.

Table 4.7. Descriptive Statistics for Hypothesis Four

Group	Group	N	Min	Max	Mean	SD
Familiar	Male	30	46	81	64.00	10.342
	Female	30	46	82	64.93	9.549
Unfamiliar	Male	30	34	86	63.37	12.249
	Female	30	36	78	58.07	11.677
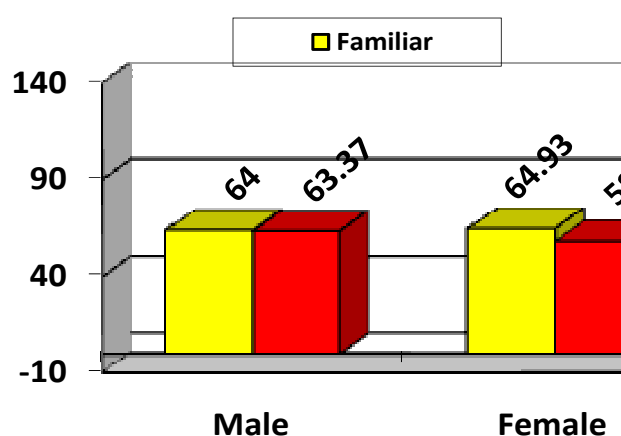
Figure 4.4. Graphical representation of the means for hypothesis four

As it can be seen in Table 4.7, the four means seem to be very close to each other; however, in order to understand if they are significantly different, another two-way ANOVA was applied. Table 4.8 presents the results of this ANOVA.

Table 4.8. The Results of Two-way ANOVA for Hypothesis Four

Source	SS	Df	MS	F	Sig.
Group_F_U	421.875	1	421.875	3.483	.065
Group_M_F	143.008	1	143.008	1.181	.279
Group_F_U * Group_M_F	291.408	1	291.408	2.406	.124
Error	14050.700	116	121.127		
Total	485033.000	119			

By considering Table 4.8, it can be understood that the amount of F-observed is not significant for the two variables, that is, gender and familiarity with the texts (they are 3.483 and 1.181, respectively). Therefore, the fourth hypothesis which states that "there is no significant difference between male and female intermediate EFL learners' performance on C-test with familiar and unfamiliar content" is also retained. In other words, it can be claimed that, like hypothesis three, gender does not play an important role in participants' performance on C-test with familiar and unfamiliar texts.

5. Discussion and conclusion

5.1 Discussion

This study aimed at investigating whether content familiarity and gender could influence Iranian EFL learners' performance on two types of tests, namely, MC cloze test and C-test. Based on the results of data analysis, a number of findings emerged which are discussed in this section.

The first null hypothesis proposed that content familiarity does not have any effect on the performance of EFL learners on MC cloze test. However, the statistical measures show the opposite, therefore this hypothesis is rejected. In other words, in MC cloze test the participants performed better with familiar texts than with unfamiliar ones. These results appeared to support the valuable role of background knowledge and schema theory of reading in test performance. Based on schema theory of reading, familiarity and knowledge about text content can facilitate comprehension during the encoding/ decoding process. This process is done by providing a knowledge structure to which readers can compare and join the pieces of incoming information; therefore, it makes it for the reader possible to understand the text without the need to know all words and phrases in the text (Al-Shumaimeri, 2006). The results also support Paltridge's (1996) claim in which he said learners' knowledge and awareness about the goal and structure of the genre can be helpful in confronting the unfamiliar content and vocabulary in the testing materials. Furthermore, the findings are in line with Aitken's (cited in Atai & Soleimany, 2009) claim that "content can be a source of

bias in scores".

The second null hypothesis predicted no effects of the content familiarity on Iranian EFL learners' performance on C-test. This hypothesis is rejected too, because the results showed that when the content is familiar to the students they have more comprehension of what they read than when the content is unfamiliar. Therefore, like MC cloze test, the performance of participants on C-test with familiar texts was better than their performance with unfamiliar texts. The results show that EFL learners' content familiarity plays an important role in their performance on both MC cloze test and C-test. According to Bachman's (1990, p. 118) model of language ability, "the nature of input the test taker receives" was one of the classifications of method facets which was concentrated on in this study. The findings validate this model and prove Bachman's (1990) theory that if the language of the input in a given test is the feature of a genre that is not familiar to the testee, the tasks that are dependent on the interpretation of that input would be fairly difficult.

According to Alderson and Urquhart (1983, 1985), the testees who are familiar with the content area can answer direct and overview questions with the same ease; and when the test takers lack content familiarity, their ability to answer direct questions is more than that of answering overview questions. Similarly, the findings support Urquhart and Weir's (1998) suggestion that it is the sufficient content familiarity which permits the deployment of suitable skills and strategies to comprehend the text. The results of the present study related to content familiarity also support the following suggestions. Al-Shumaimeri (2006) captured that content familiarity has important effects on the students' comprehension performance and Afghari and Tavakoli (2004) found that when students are familiar with the content, their performance on the test is better than when they are not familiar with content. Tavakoli, Ahmadi and Bahrani (2011) proved that genre familiarity has a great effect on EFL learners' performance. They found that genre familiarity plays a significant role in the performance of learners on two types of tests; cloze test and C-test while the present study considered the effect of content familiarity on the performance of participants in MC cloze test and C-test. However, the findings of this study are in line with those studies.

The third null hypothesis which concerns the effects of male and female intermediate EFL learners' performance on MC cloze test with familiar and unfamiliar content is retained. In other words, it can be suggested that gender does not play an important role in participants' performance on MC cloze test with familiar and unfamiliar texts.

The forth null hypothesis concentrating on the effect of male and female intermediate EFL learners' performance on C-test with familiar and unfamiliar content is also retained. To state the matter differently, the same as hypothesis five, gender does not play a crucial role in participants' performance on C-test with familiar and unfamiliar texts.

The results related to the third and fourth hypothesis of the present study disagree with many reported studies that claim there is a significant difference between the performance of males and females in English language (e.g. Daly, 1999; Gamble & Hunter, 2000; Keshavarz & Ashtarian, 2008; Schueller, 2000; Skarbrevik, 2002). Experimental data about the effect of gender on different types of test of reading comprehension have shown different results. For

instance, Brantmeier (2002, 2004, and 2007) found females outperformed males in their comprehension of given passages in a written recall task. Sotudehnama and Asadian (2011) stated that females and males performed better on topics related to their gender but both males and females performed the same on two types of tests; MC and recall with neutral text. In another study done by Sankarakumar, Chandrakanthi and Malathy (2012) the performance of male and female students on a reading test including different types of questions (identifying topic sentences, MC questions, inserting the sentence, matching the sentence and identifying true/ false statement) has been considered.

The findings revealed that there was no significant difference between the performance of male and female in testing items like MC questions. The results of the study by Habibian (2012) showed that the gender of the subjects has an important role on the overall performance of the subjects on reading comprehension test and cloze test and female students perform better than male students. The findings related to the third and fourth hypotheses of the present study differ from the view of those who suggest that females often outperform males in English tests. It, however, agrees with the view that there is no significant difference in the performance of males and females in English language test (Jegede, 1994; Kincade & Kleine, 1990; Madu & Kasanga, 2005).

5.2 Conclusion

The results of the present study show that familiarity with content has a significant role in the performance of intermediate EFL learners in two types of tests, namely, MC cloze test and C-test. Considering male and female performance on two types of tests; MC cloze test and C-test in this study reveals that gender has no effect on the performance in these two types of tests.

References

Afghari, A., & Tavakoli, M. (2004). Learner variables and ESP test performance. *Iranian Journal of Applied Linguistics* (*IJAL), 7*(1), 1-24.

Alderson, J. C. (1979). The cloze procedure and proficiency in English s a foreign language. *TESOL Quarterly, 13,* 219-223. http://dx.doi.org/10.2307/3586211

Alderson, J. C., & Urquhart, A. H. (1983). The effect of student background discipline on comprehension: a pilot study. In A. Hughes, & D. Porter (Eds.).

Alderson, J. C., & Urquhart, A. H. (1985). The effect of students' academic discipline on their performance on ESP reading tests. *Language Testing, 2*(2), 192-204. http://dx.doi.org/10.1177/026553228500200207

Al- Shumaimeri, Y. A. N. (2006). The effects of content familiarity and language ability on reading comprehension performance of low and high ability Saudi tertiary students studying English as a foreign language. *Educational Science & Islamic Studies, 18*(2), 1-19.

Anderson, N. J. (2003). "Reading". *Practical English Language Teaching*. Ed. D. Nunan. New York: McGraw Hill. pp. 67-86.

Atai, M. R., & Soleimany, M. (2009). On the effect of text authenticity & genre on EFL learners' performance in C-text. *Pazhuesh-e Zabanha-ye Khareji, 49,* 109-123.

Bachman, L. F. (1982). The trait structure of cloze test scores. *TESOL Quarterly, 16*(1), 61-70. http://dx.doi.org/10.2307/3586563

Bachman, L. F. (1985). Performance on cloze tests with fixed-ratio and rational deletions. *TESOL Quarterly, 19*(3), 535-555. http://dx.doi.org/10.2307/3586277

Bachman, L. F. (1990). *Fundamental Consideration in Language Testing.* Oxford: Oxford University Press.

Bachman, L. F. (2000). Modern language testing at the turn of the century: Assuring that what we count counts. *Language Testing, 17*, 1-42.

Baghaei, P. (2011). Optimal Number of Gaps in C-Test Passages. *International Education Studies, 4*(1), 166-171.

Baker, S. C., & MacIntyre, P. D. (2000). The role of gender and Immersion in communication and second language orientations. *Language Learning*, *50*(2), 311-340. http://dx.doi.org/10.1111/0023-8333.00119

Bastidas, J. (1984). The cloze procedure as a teaching technique to improve reading comprehension. *English Teaching Forum, 22*(2), 20-24.

Bensoussan, M. (1998). Schema effects in EFL reading comprehension. *Journal of Research in Reading*, *21*(3), 213-227. http://dx.doi.org/10.1111/1467-9817.00058

Bradshaw, J. (1990). Test-takers´ reactions to a placement test. *Language Testing, 7*(1), 13-30. http://dx.doi.org/10.1177/026553229000700103

Brantmeier, C. A. (2000). The relationship between readers' gender, passage contents, comparison and strategy use in reading Spanish as a second language. Dissertation Abstracts International Section A: Humanities and Social Sciences, 61 (4-A): 1376.

Brantmeier, C. (2002). The effects of passage content on second language reading comprehension by gender across instruction levels. In J. Hammadou Sullivan (Ed.), *Research in second language learning*. Greenwich, CT: Information Age Publishing.

Brantmeier, C. (2003). Does gender make a difference? Passage content and comprehension in second language reading. *Reading in a Foreign Language*, *15*(1), 1-23.

Brantmeier, C. (2004). Building a comprehensive theory of adult foreign language reading: A variety of variables and research methods. *The Southern Journal of Linguistics*, *27*, 1-7.

Brantmeier, C. (2007). Adult second language reading in the USA: The effects of readers' gender and test methods. *Forum on public policy, 14*, 1-34.

Brown, H. D. (1994). *Principles of language learning and teaching* (3rd Ed.). Englewood Cliffs, NJ: Prentice Hall Regents.

Brown, H. D. (2004). *Language assessment: principles and classroom practices*. White Plains, NY: Pearson Education.

Bügel, K., & Buunk, B. P. (1996). Sex differences in foreign language text comprehension: The role of interests and prior knowledge. *Modern Language Journal*, *80*(1), 15-31.

Chihara, T. (1994). When are cloze items sensitive to higher order processing abilities? *Language Learning, 35*(2), 181-206.

Cleary, C. (1988). The C-Test in English: Left-hand deletions. *RELC Journal*, *19*(2), 26-38. http://dx.doi.org/10.1177/003368828801900203

Cohen, A. (1980). *Testing language ability in classroom.* Massachusetts: Newbury House Publishers, Inc.

Cohen, A. (1991). Second language testing. In Celce- Murcia, Editor, *English as a second language,* USA, Newbury House.

Connelly, M. (1997). Using C-tests in English with post- graduate students. *English for specific purposes, 16*, 139-150. http://dx.doi.org/10.1016/S0889-4906(96)00018-X

D'Ailly, H. (2002). A distinctive cultural and gender difference in children's interest and effort in learning: The impact of choice and testing. Paper presented at the Annual Meeting of the American Educational Research Association, New Orleans, LA, April 1-5.

Daly, C. (1999). Reading boys, changing English. *Studies in Reading and Culture, 6*(1), 7-18.

Dornyei, Z., & Katona, L. (1992). Validation of the C-test amongst Hungarian EFL learners. *Language testing, 9*, 187-206. http://dx.doi.org/10.1177/026553229200900206

Eckes, T., & Grotjahn, T. (2006). A closer look at the construct validity of C-Tests. *Language Testing*, *23*, 290–325. http://dx.doi.org/10.1191/0265532206lt330oa

Farhady, H. (1996). Varieties of cloze procedure in EFL education. *Roshd Foreign Language TeachingJournal, 12*(44), 217-229.

Farhady, H., Jafarpur, A., & Birjandi, P. (1994). *Testing Language Skills from Theory to Practice.* Tehran: SAMT Publication.

Feng, L. (2011). A short analysis of the text variables affecting reading and testing reading. *Studies in Literature and Language, 2*(2), 44-49.

Gamblell, J., & Hunter, D. (2000). Surveying gender differences in Canadian schools literacy. *Journal of Curriculum Studies, 32*(5), 689-720. http://dx.doi.org/10.1080/00220270050116941

Grabe, W. (1991). Current developments in second language reading research. *TESOL Quraterly, 25*, 375- 406. http://dx.doi.org/10.2307/3586977

Grotjahn, R. (1986). Test validation and cognitive psychology: Some methodological considerations. *Language Testing, 3*, 159-185.

http://dx.doi.org/10.1177/026553228600300205

Grotjahn, R. (Ed.) (1992). *Der C-Test. Theoretische Grundlagen und praktische Anwendungen, (1*). Bochum: Brockmeyer.

Grotjahn, R. (Ed.) (1994). *Der C-Test. Theoretische Grundlagen und praktische Anwendungen.*, *(2)*. Bochum: Brockmeyer.

Grotjahn, R. (Ed.) (1996). *Der C-Test. Theoretische Grundlagen und praktische Anwendungen, (3).* Bochum: Brockmeyer.

Grotjahn, R. (Ed.) (2002). *Der C-Test. Theoretische Grundlagen und praktische Anwendungen*, *(4).* Bochum: AKS-Verlag.

Grotjahn, R. (Ed.) (2006). *Der C-Test: Theorie, Empirrie, Anwendungen/The C-test: theory, empirical research, applications*. Frankfurt/ M: Peter Lang.

Grotjahn, R. (Ed.) (2010). *Der C-Test: Beiträge aus der aktuellen Forschung/The C-Test: Contributions from Current Research.* Frankfurt am Main: Peter Lang.

Habibian, M. (2012). The investigation of reading comprehension test and cloze test among learners who learn English as a foreign language with respect to their language proficiency and gender. *Journal of Social Sciences and Humanities, 7*(1), 177-186.

Hassan, B. A. (2001). Extraversion/ introversion and gender in relation to the English pronunciation accuracy of Arabic speaking college students. Research Report 34.

Hinofotis, F. B., & Snow, B. G. (1980). An alternative cloze testing procedure: Multiple choice format. In J.W. Oller & K. Perkins (Eds.), *Research in language testing*, 129-133. Rowley, MA: Newbury House.

Huttner, J. (2008). The genres of student writing: developing writing models. *International Journal of Applied Linguistics*, *18*(2), 147- 165.

Jafarpur, A. (1995). Is C-testing superior to cloze? *Language Testing, 12*, 194-216. http://dx.doi.org/10.1177/026553229501200204

Jegede, J. O. (1994). Influences of motivation and gender on secondary school students' academic performances in Nigeria. *Journal of Social Psychology, 134*(5), 695-698. http://dx.doi.org/10.1080/00224545.1994.9923000

Jonz, J. (1976). Improving on the basic egg: The M-C cloze. *Language Learning, 26*, 255-265. http://dx.doi.org/10.1111/j.1467-1770.1976.tb00276.x

Kamimoto, T. (1992). An inquiry into what a C-test measures. *Fukuoka Women's Junior College Studies, 30*(12), 67-79.

Kamimoto, T. (1993). Tailoring the test to fit the students: Improvement of the C-test through classical item analysis. *Fukuoka Women's Junior College Studies, 30*(11), 47-61.

Kang, H. (1992). The effects of culture-specific knowledge upon ESL reading comprehension.

School Education Review, 4, 93-105.

Kasanga, L. A. (1996). Effects of gender on the rate of interaction. Some implications for second language acquisition and classroom practice. I.T.L. *Review of Applied Linguistics,* 111/112: 155-192.

Keshavarz, M. H., & Ashtarian, S. (2008). The relationship between Iranian EFL learners' gender and reading comprehension of three different types of text. *IJAL, 11*(1), 97-113.

Khodadady, E., & Hashemi, M. (2011). Validity and C-test: The role of text authenticity. *Iranian Journal of Language testing, 1*(1), 30-41.

Kincade, K. M., & Kleine, P. F. (1990). Children's' Reading Task by Gender Interactions: implications for research and practice. (Paper presented at the annual meeting of the American Educational Research Association, Boston).

Klein-Braley, C. (1994). *Language testing with the C-Test: A linguistic and statistical investigation into the strategies used by C-Test takers, and the prediction of C-Test difficulty*. Duisburg: Habilitationsschrift.

Klein-Braley, C. (1997). C-tests in the context of reduced redundancy testing: An appraisal. *Language Testing, 14*(1), 47-84. http://dx.doi.org/10.1177/026553229701400104

Klein-Braley, C., & Raatz, U. (1984). A survey of research on the C-test. *In Language Testing, 1*(2), 134-146. http://dx.doi.org/10.1177/026553228400100202

Kontra, E. H., & Kormos, J. (2006). Strategy use and the construct of C-Test. In R. Grotjahn (Ed), Der C-Test: *Theorie, Emprie, Anwendungen/The C-Test: Theory, empirical research, applications* (pp. 121-138). Frankfurt/am: Peter Lang.

Lee, S. H. (2008). Beyond reading and proficiency assessment: The rational cloze procedure as stimulus for integrated reading, writing and teacher-student interaction in ESL. *Elsevier (system), 36*, 642-660.

Lee, Y. (2004). Examining passage-related local items dependence (LID) and measurement construct using Q3 statistics in an EFL reading comprehension test. *Language Testing, 21*(1), 74-100. http://dx.doi.org/10.1191/0265532204lt260oa

Lu, G. (2006). Cloze tests and reading strategies in English language taching in China. Unpublished, MA Thesis. University of Western Cape, China.

Madu, B. N., & Kasanga, L. A. (2005). Sex differences in the acquisition of English as a second language. *Gender and Behaviour, 3*, 442-452. http://dx.doi.org/10.4314/gab.v3i1.23337

Oller, J. W. Jr, & Perkins, K. (1978). *Language in Education: Testing the Tests.* Rowley, Mass.: Newbury House.

Oller, J.W. Jr. (1972). Scoring methods and levels of difficulty levels for cloze tests of proficiency in English as a second language. *The Modern Language Journal. 56*, 151-158.

http://dx.doi.org/10.2307/324037

Oller, J. W. Jr. (1979). *Language test at school: a pragmatic approach.* London: London Group Limited.

Oller, J. W., & Taira, T. (1994). Cloze and episodic organization. In J. W. Jr. Oller, & J. Jonz (Eds.), *Cloze and coherence.* London: Associated University Press, pp. 345-69.

Paltridge, B. (1996). *Genre and the language- learning classroom.* Ann Arbor: University of Michigan Press.

Piske, T. MacKay, I. R. A., & Flege, J. E. (2001). Factors affecting degree of foreign accent in an L2: A review. *Journal of Phonetics, 29*(2), 191-215. http://dx.doi.org/10.1006/jpho.2001.0134

Pulido, D. (2004). The relationship between text comprehension and second language incidental vocabulary acquisition: A matter of topic familiarity? *Language Learning, 54*(3), 469-523. http://dx.doi.org/10.1111/j.0023-8333.2004.00263.x

Raatz, U., & Klein- Braley, C. (1981). The C-test — a modification of the cloze procedure. In T. Culhane, C. Klein- Braley, & D. Stevenson (Eds.), *Practice and problems in language testing* (pp. 113-148). Colchester: University of Essex.

Rye, J. (1982). *Cloze procedure and the teaching of reading.* London: Heinemann.

Sadeghi, K. (2010). Cloze validation against IELTS reading paper: Doubts on correlational validation. *Journal of English Teaching and Learning, 53*, 217.

Salmani-Nodoushan, M. A. (2003). Text familiarity, reading tasks, and ESP test performance: A study on Iranian LEP and Non- LEP university students. *The Reading Matrix, 3*(1).

Sankarakumar, S., Chandrakanthi, S., & Malathy, M. (2012). Assessing the Performance of Male and Female Students in Reading Tests. *European Journal of Scientific Research*, *84*(1), 22-35.

Sasaki, M. (1993). Realtionships among second language proficiency, foreign language aptitude, and intelligence: A protocol analysis. *Language Learning, 43*, 469-505. http://dx.doi.org/10.1111/j.1467-1770.1993.tb00625.x

Sasaki, M. (1996). *Second language proficiency, foreign language aptitude, and intelligence: Quantitative and qualitative analysis*. New York: Peter Lang.

Schueller, J. M. (1999). *The effect of two types of strategy training on foreign language reading comprehension. An analysis by gender and proficiency.* Unpublished doctoral dissertation, The University of Wisconsin, Madison.

Schueller, J. M. (2000). The effects of two types of strategic training on foreign language reading comprehension: An analysis by gender and proficiency. Dissertation Abstracts International Section A: Humanities and Social Sciences, 60(7-A), 2472.

Shanahan, T., Kamil, M. L., & Tobin, A. W. (1982). Cloze as a measure of intestential

comprehension. *Reading Research Quarterly, 17*, 229-255. http://dx.doi.org/10.2307/747485

Sigott, G. (2004). *Towards identifying the C-Test construct*. Frankfurt/M: Peter Lang.

Sigott, G., & Koberl, J. (1996). Deletion pattern and C-test difficulty across languages. In Grotjahn (ed.), *Der C-test: Theoretische Grundlagen und Praktische Anwendungen* (pp. 159-172). Bochum: Brockmeyer.

Skarbevik, K. J. (2002). Gender differences among students eligible for special education. *European Journal of Special Needs Education, 17*(2), 97-108. http://dx.doi.org/10.1080/08856250210129038a

Sotoudenama, E., & Asadian, M. (2011). Effect of gender- oriented content familiarity and test type on reading comprehension. *The Journal of Teaching Language Skills (JTLS), 3*(2), 155- 178.

Tavakoli, M., Ahmadi, A., & Bahrani, M. (2011). Cloze Test and C-test Revisited: The Effect of Genre Familiarity on Second Language Reading Test Performance. *Iranian Journal of Applied Linguistics (IJAL), 14*(2), 173-204.

Taylor, W. L. (1953). Cloze procedure:a new tool for measuring readability. *Journalism Quarterly, 30,* 415-433.

Urquhart, S., & Weir, C. (1998). *Reading in a Second Language: Process, Product and Practice.* London: Longman.

Young, D. J., & Oxford, R. L. (1997). A gender based analysis used to process written output in the native and foreign language. *Applied Language Learning*, *8*(1), 43-73.