Effect of Gender – Oriented Content Familiarity and Test Type on Reading Comprehension

E. Sotoudehnama [*]
Assistant Professor
Alzahra University, Tehran
email: esotoude@alzahra.ac.ir

M. Asadian
M. A., TEFL
Alzahra University, Tehran
email: foggy_dew1283@yahoo.com

Abstract

One important issue that any material or test designer should bear in mind is to be careful about developing teaching and testing reading materials which would not be to the benefit of a group of readers or examinees and to the detriment of some others. One of the factors which may cause difference between the performances of English learners or test takers is the selection of reading tests with which some learners have familiarity and others not. This study sets forth to analyze whether gender and topic-familiarity can be determining factors in the differences among the performances of foreign language learners on reading comprehension tests. For this purpose, 64 (34 females and 30 males) intermediate students of the Bayane Salees Institute in Tehran completed a topic-familiarity questionnaire and took three tests of reading, composed of one male-oriented text (Car Maintenance), one female-oriented text (Women's Shoes) and one neutral text (Beliefs about Numbers). Two different measures were used to assess comprehension: multiple-choice and free recall. Findings revealed that males had better performance on the male-oriented text and females did better on the female-

oriented text. The performances of males and females on the neutral text were seen to be the same as expected. It was also evidenced that males did better on the male-oriented and females did better on the female-oriented test on multiple choice tasks; on the neutral test, both genders performed the same. However, interestingly, in the free recall tasks, the males outperformed females in all three tests. This result brings about the assumption that males might be better at free recall kind of reading measurement, though this conclusion requires more investigation.

Keywords: gender- oriented texts, content familiarity and reading comprehension, test type and reading comprehension

1. Introduction

Reading comprehension is affected by different variables among which, passage content or topic familiarity, gender, and test types can be named as a few. One of these factors, i.e., previously acquired knowledge (Carrell & Eisterhold, 1983 in Alderson, 2000) or as called schema theory (Ruddell, 2005) defined as organized and related units in which knowledge is stored so as to facilitate recall and play as a basis for the incoming information to fit into (Ajideh, 2006; Nuttall, 1996; Chastain, 1988) Is considered to be culture-specific as well (Carrell 1981; Carrell and Eisternhold 1983).

While Carrell (1983a) believes in formal and content schemata, Alderson (2000) divides content schemata into two subcategories of background knowledge and subject-matter knowledge. Brantmeier (2003c, p. 2) defines content schema as the "knowledge relative to the content domain of the reading passage that the reader brings to a text". Also, Carrell (1983b) speaks of content schema as a subcategory of schema which relates to the knowledge of the topic of the text. Thus the concepts of topic familiarity and passage content are closely interrelated. Many scholars such as Brantmeier (2003c) along with others (Steffenson, et al, 1979; Hudson, 1982; Johnson, 1982; Carrel, 1981, 1983a; Carrell & Eisterhold, 1983; Pritchard, 1990) have emphasized the influential role of topic familiarity in L2 comprehension.

Two factors related to background knowledge which affect the comprehension of L2 readers are mentioned as a kind of reading test and gender. The research conducted shows that test type affects the result of reading comprehension ((Johnson 1981, 1982; Karakas 2002; Mohammad & Swales 1984; Alderson & Urquhart 1984, cited in Alderson, 1984; Floyd & Carrell 1987).

Each of the two kinds of tests used in this study, i.e., multiple choice and free recall, has its own merits and demerits. Multiple-choice questions are a very common device for testing reading comprehension. Some scholars recently have criticized this technique for many reasons. One shortcoming of this test, according to Koda (2005), is the necessity for several task-specific operations, which normally do not occur in real-life reading. For example, matching which requires students to match multiple-choice items with relevant text elements, in both their surface forms and their meanings. Another reason for which it is criticized is that the "well-constructed distracters can alter stored information extracted from the text" (Koda, 2005, p. 239) being too traditional and being unable to "capture the complex processes that take place between learner and text" (Bernhardt, 1991; cited in Heinz, 2004; p. 1). The critics of this method prefer free recall as an integrative and a more authentic technique.

In free recall (sometimes called immediate-recall test) "test takers are asked to describe everything they remember from the text read. Little time is required to prepare the task, but scoring involves an extensive analysis of recall protocols"(Koda 2005, pp. 236-237). As any other kind of testing instrument, it has its own advantages and disadvantages but as Brantemeier, (2004) mentions, many researchers (such as Barnett, 1986; Carrell, 1983b; Lee, 1986a; 1986b; 2002; Brantemeier, 2002, 2003b; among others) utilize this procedure because when students are allowed to write freely they are not confined by the pre-determined and created assessment tasks. So, in this study free recall–as a kind of direct and integrative assessment task- is used as a complementary procedure

besides multiple-choice technique which is regarded as a traditional testing method.

Considering the other factor, gender, a variety of studies have shown differences between the performance of males and females on topics which are toward either gender. As an example, Bugel and Buunk (1996; cited in Macaro, 2003) found that a topic of a text is an important factor in explaining gender-based differences in L2 reading. In another study conducted by Doolittle and Welsch (1989), who examined differences of performances on specific passages, they found out that females scored higher than males on humanities-oriented readings whereas males did better than females on passages with science-oriented topics.

Brantemeier (2002, 2003a, 2003b, 2003c, 2003d) utilized a passage on boxing and another passage on midwives. Although at higher levels of instruction, passage content made no significant effect on L2 reading comprehension by gender, it showed to be an effective factor among readers at intermediate level. Other studies in this regard showed gender differences in reading comprehension (Chiu & McBride-Chang, 2006; Pae, 2004; Chavez, 2001, cited in Brantmeier, 2003a; Maccoby, 1998).

Since content schema is part of the individual cultural orientation (Al-Issa, 2006), it can be concluded that the readers' performances on gender-oriented topics cannot be the same in different countries. The fact that this area has not been investigated in Iran has made the impetus for the present research. Of course, a series of studies conducted in Iran (Fazilatfar 1989; Davanlou 1995; Gholampour 1993; Madanian 1995; Nafar 2002; Tahmasbi 1999; Afzali Shahri 1995; Ahmadi 2001) deal somehow with background knowledge but none of these studies concern the area of passage content in relation to gender and reading comprehension test.

2. Objective

The present study aimed at examining the effects of reader's gender and passage content on L2 readers' comprehension among Iranian

intermediate learners of English as a foreign language. The following research questions guided the study:

1. Does gender-oriented passage content (male, female and neutral) reading text affect learners' comprehension (multiple-choice and free recall)?

2. Is there any significant difference between EFL males and females in free recall and multiple-choice reading comprehension performances scores in three different gender-oriented texts?

The null hypotheses related to the questions investigated are as follows:

1. Gender-oriented passage content (male, female and neutral) reading text does not affect learners' comprehension (multiple-choice and free recall)?

2. There is no significant difference between EFL males and females in free recall and multiple-choice reading comprehension performances scores in three different gender-oriented texts?

3. Method

3.1 Participants

The participants of this study were 64 (34 females and 30 males) EFL intermediate level students ranging between 16 to 35 years old studying in an institute (Bayane Salees) in Tehran. Selection of the intermediate level participants was based on the scores between one standard deviation above and below the mean gained in a reading subtest of a TOEFL (2004) test.

3.2 Instrumentation

3.2.1 Topic familiarity questionnaire

The topic familiarity questionnaire was to help the researchers determine the topics as being gender-oriented or neutral, so that they could be administered as the reading comprehension passages. So, twenty topics were included and the participants determined their level of familiarity with different topics based on a five-point Likert scale ranging from 1(absolutely unfamiliar) to 5 (completely familiar). A group of 50 adult

English learners (25 females and 25 males) who were nearly at the same proficiency level of English (about FCE level) were assigned to answer this questionnaire.

3.2.2 Comprehension assessment tasks

Results of some studies have concluded that the method of assessing reading comprehension influences how readers perform on a test of reading comprehension (Shohamy, 1984; Wolf, 1993). This being the case, readers' performance across two different reading comprehension assessment tasks (i.e. multiple-choice questions and free recall) was checked through three realy passages:

Three passages including "Women's Shoes" as the female-oriented one (Flesch Readability: 73.30), "Car Maintenance" as the male-oriented topic (Flesch Readability: 76.66), and "Beliefs about Numbers" as the neutral topic (Flesch Readability: 72.61). These passages were chosen according to the topic familiarity questionnaire.

3.2.3 Multiple-choice test

For each of the 3 passages a set of multiple-choice questions with three options were designed. Ten questions for the male-oriented passage, 10 for the neutral and 8 for the female-oriented text were developed. These sets of questions were piloted by administering them to a sample group of participants having similar characteristics as the main group.

3.2.4 Recall essay-type test

In the free recall test the focus is on the quantity of the correct information recalled. The recall essay-type test required the learners to write down in their own language (Persian) whatever they could remember from the text -- main ideas as well as details-- without taking a second look at it. They were also told that the more the number of the ideas recalled, the higher the scores would be. The free recall task was administered before the multiple choice questions to ensure that the information within the multiple-choice test would not affect the free recall test.

From among the alternative methods for scoring, counting the 'idea units' (Alderson, 2000; p. 230) was selected. In this study the scoring was

based on the analysis of the propositions in the pausal units or breath groups as provided by Alderson (2000; p. 231). So the idea units of every text were listed. The number of idea units in each text was 36. For a more accurate analysis and listing of the propositions, the passages were given to a native-like speaker to take these measures. The participants' scores were calculated by checking the presence or absence of such units. A proposition unit was counted as correct if the gist of the proposition in the text was expressed and word-for-word recall was not expected. Since this method is subjective, they were judged by two raters and an inter-rater reliability was estimated. The calculation of the inter-rater reliability was done through Pearson Correlation for all the three texts. The inter-rater reliability for the female-oriented text was .987, for the male-oriented text was .964 and for the gender-neutral text was .990.

3.2.5 TOEFL test

To homogenize the participants, a reading subtest of TOEFL (2004) test was used and the participants who scored between the range of one standard deviation below and above the mean were selected as the intermediate ones to participate in this study.

3.3 Procedure

The first 4 units of the Gold FCE (First Certificate in English) book which are covered during their first 2 terms of FCE courses in the institute incorporate five reading comprehension texts. The readability level of these five passages was calculated using Flesch readability formula (Burmeister, 1978; p. 34) and the resulted readability was used in order to determine the readability level at which the passages were to be chosen.

The passages which were selected as the reading comprehension test were considered to have a readability level within the range of 0.5 standard deviation below and above the readability mean of the five texts within their book (Readability Mean=76.17; and 0.5 standard deviation above and below the mean makes the range of 71.58 to 80.76). Based on the readability range which was determined through the stated procedure,

the female and male-oriented as well as the neutral topics which were going to be used in the topic familiarity questionnaire were selected. The topic familiarity questionnaire was made up of about 14 topics including 5 female-oriented, 6 male-oriented, and 3 neutral topics.

Then the questionnaire was administered to two groups of 25 males and 25 female participants with the same language proficiency level (FCE level) similar to the main participants who then were to be selected for the tests.

The questions in the topic familiarity questionnaire included 5 alternatives which ranged from 1 (totally unfamiliar) to 5 (totally familiar). The topic items which were scored above 75 out of 125 (mostly ranked 4 or 5 on the Likert scale) by only male participants was identified as being male-oriented and the ones which were scored above 75 out of 125 by only females were considered as being female-oriented. Neutral topic, on the other hand, was the one which was scored lower than 75 by both males and females. Then female/male-oriented as well as the neutral topics turned out to be respectively as follows: "Women's Shoes" (readability: 73.30), "Car Maintenance" (readability: 76.66) "Beliefs about Numbers" (readability: 72.61).

To make a comprehension test out of these three passages, a total of 28 multiple-choice questions were written (8 questions for "Women's Shoes"; 10 questions for "Beliefs about Numbers" and 10 questions for "Car Maintenance" which were reduced to 9 due to the revisions made after the factor analysis and item analysis). These questions were first evaluated regarding validity and reliability. As for the content validity of the tests, two experts who had Ph.D. degrees in English Teaching verified the content validity of the questions.

Then the Multiple-choice test was piloted in order to be analyzed as for the values of central tendency, dispersion, reliability and validity. The validity of the tests was estimated through a factor analysis, of course, despite the insufficient number of the participants. The pilot test was conducted on 45 females and 30 males who shared the characteristics of the main participants of the study, i.e. they were intermediate students of

the same institute. Finally, based on the results of the pilot test, the multiple-choice questions were revised.

To have a homogenized group, a test of TOEFL (2004) was administered to the participants who were learners of the Bayane Salees Institute in Tehran. The test was given to 53 females and 51 males. Out of this number, the answer sheets of 47 females and 44 males were returned. Again out of this number, the scores of 34 female and 30 male participants were within our desirable limits i.e.,one standard deviation below and above the mean. These participants were the main participants of the study.

4. Data Analysis

Firstly, the content validity as well as the reliability estimates of the three texts was examined. The reliability and validity of the tests were calculated and the content validity of them was confirmed by 2 English Ph.D. graduates. The first research question was dealt with using an independent sample t-test for each of the three passages. So there was one dependent variable-the average score of multiple-choice and free recall–and one independent variable-gender. To explore the second research question and to see the effect of each measurement task on either gender, a series of independent samples t-tests were conducted. The independent variable was gender and there were two dependent variables: free recall score and multiple-choice score.

A one-sample KS test was calculated for each test to find out if the examinees' scores were normally distributed. The scores of the participants in the pilot test on all the passages, i.e., the female-oriented test [$Z=1.291$; $p>.05$], the male-oriented test [$Z=1.073$; $p>.05$], and the neutral test [$Z=1.343$; $p>.05$ showed to have a normal distribution.

The reliability measures of the multiple-choice tests on each of the three topics were calculated through Cronbach's alpha. The multiple-choice questions showed a reliability of 0.702 on the female-oriented, 0.608 on the male-oriented, and 0.727 on the neutral topic.The content validity of the test was confirmed by 2 English Ph.D graduates.

After the administration of the pilot test, the measures of item facility (IF) and item discrimination (ID) were calculated for each question. To analyze the efficiency of the items revised based on the item facility and item discrimination values, this study referred to the criteria presented by Brown (1996). Item 4 of the "Car Maintenance" text had high item facility and item 7 had low item discrimination. Also, item 2 of the neutral text and item 9 of "Car Maintenance" had low discriminative power. Therefore, these items underwent some revisions for the main test except item 9 of "Car Maintenance" which was deleted from the main test due to its very low discriminating power.

4.1 Analysis of the main test

Before the performance of the main test, a TOEFL (2004) test was administered in order to make a homogenized group. A number of 47 female and 44 male participants took this test. The scores were checked against a normal distribution curve. The one-sample Kolmogorov-Smirnov test was checked and the distribution of the scores was shown to be normal [female: Z= 0.908; p= 0.382; male: Z= 1.295; p=0.07]. Since the study was supposed to include intermediate level of English proficiency, thirty-four female and thirty male participants who scored between one standard deviation below and above the mean score (female: M= 29.744; SD= 7.811; Male: M=30.772; SD= 6.205) were the main participants of the study.

To find out whether the mean scores of the participants on the multiple-choice and free recall tasks were distributed normally, a One-Sample Kolmogorov-Smirnov test was calculated. The scores of all three sets of questions were normally distributed [shoes (female text): Z=.531; p=.941], [numbers (neutral text): Z=.817; p=.516], [cars(male text): Z=.572; p=.899].

4.2 Analysis of null hypothesis 1

To answer the first research question a series of independent samples t-test were conducted. Table 1 shows the descriptive statistics for the males

and females' scores in both kinds of tests- multiple choices and free recall- in the three passages including female and male oriented and the neutral one.

Table1. Descriptive statistics of males and females' scores in multiple choice and free recall tests in three types of texts

	gender	N	Mean	Std. Deviation	Std. Error Mean
Mean of MC & FR (Shoes)	female	34	10.3456	2.90768	.49866
	male	30	8.8917	2.78286	.50808
Mean of MC & FR (Cars)	female	34	9.6397	2.90845	.49879
	male	30	13.0417	4.10744	.74991
Mean of MC & FR (Numbers)	female	34	11.4632	4.07883	.69951
	male	30	11.1917	3.52680	.64390

MC stands for Multiple Choice and FR stands for Free Recall

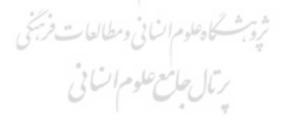The summary of the independent samples t-test with the mean score of both test types are provided in table 2 .

Table 2. Independent samples t-test for mean scores of multiple choice and free recall

	Leven's test for equality of vaiatuion		t-test for equality of means					95% Confidence Interval of the difference	
	F	Sig	t	df	Sig. (2- tailed)	Mean difference	St.Error difference	lower	higher
Mean of MC & FR (Shoes) Equal variances assumed / Equal variances not assumed	.001	.977	2.037 / 2.042	62 / 61.572	.046 / .045	1.45392 / 1.45392	.71389 / .71191	.02688 / .03065	2.88097 / 2.87720
Mean of MC & FR (Cars) Equal variances assumed / Equal variances not assumed	4.033	.049	-3.858 / -3.777	62 / 51.461	.000 / .000	-3.40196 / -3.40196	.88184 / .90065	-5.16474 / -5.20967	-1.63919 / -1.5942
Mean of MC & FR (Numbers) Equal variances assumed / Equal variances not assumed	.034	.854	.283 / .286	62 / 61.980	.778 / .776	.27157 / .27157	.95951 / .95075	1.64646 / -1.62897	2.18960 / 2.17211

MC stands for Multiple Choice and FR stands for Free Recall

It was shown that for the female-oriented text, there is a significant difference [t (62) =2.037; p=.046] between the performance of females and males; female participants (10.3456) outscored males (8.8917). Considering the male-oriented text, too, there is a significant difference [t (51.481) = -3.777; p <.05] and males (13.0417) outperformed females

(9.6397). For the neutral text, there is no significant difference [t (62) = .283; p=.778] between males and females. This similarity of the mean values showed that males and females have performed rather equally on the neutral reading text.

4.3 Analysis of the free recall tasks
To score the free recall task, the idea units in each text was listed based on the definition of "idea units" by Alderson (2000, p.61). Each participant's paper was scored by two raters and to check for subjectivity in scoring, the inter-rater reliability between the scores of the two raters on each text was calculated through a Pearson correlation (Table 3).

Table3. Inter- rater reliability of free recall scores on three texts

		Rater 1	Rater 2
Shoes	Pearson Correlation	1	.987*
	Sig. (2-tailed)		.000
Car Manintenance	Pearson Correlation	1	.964*
	Sig. (2-tailed)		.000
Bielifs about numbers	Pearson Correlation	1	.990*
	Sig. (2-tailed)		.000
N		64	64

*Correlation is significant at 0.01 level

Table 3 shows a significantly high correlation between the scores of the two raters on the topic of Women's Shoes [Pearson correlation= .987], on the topic of "Car Maintenance" [Pearson correlation= .964] and on the neutral topic [Pearson Correlation= .990]. As there was high correlation between the sets of scores on all tests, in order to report a single score for each test, the mean between the scores of the raters on each test was

used. To analyze whether the mean scores on the free recall task for each text makes a normal distribution curve, a One-Sample Kolmogorov-Smirnov test was calculated and for the female-oriented topic [Z= .534; p=.938], for the male-oriented topic [Z= .657; p=.781], and for the neutral topic [Z= .647; p=.797] the results show the distribution curve is also normal.

4.4 Analysis of null hypothesis 2

To find the difference between males and females in the three kinds of texts considering two different kinds of tests, a series of independent samples t-test were conducted. Table 4 provides the descriptive statistics and table 5 indicated the degree of difference.

Table 4. Descriptive statistics on free recall and multiple choice of three texts

Gender		N	Mean	Std. Deviation	Std. Error Mean
FR (Shoes)	female	34	14.2059	5.02738	.86219
	male	30	14.6667	5.70592	1.04175
FR (Cars)	female	34	12.3824	5..21661	.89464
	male	30	19.0667	7..34229	1.34051
FR (Numbers)	female	34	15.9853	7.46252	1.27981
	male	30	16.3833	6.87158	1.25457
MC (Shoes)	female	34	5.47	1.331	.228
	male	30	4.13	1.502	.274
MC (cars)	female	34	6.03	1.000	.171
	male	30	7..97	.809	.148
MC (Numbers)	female	34	6.09	1.990	.341
	male	30	6..97	.277	.277

Table5. Independent samples t-test between males and females' scores on free recall and multiple choice tests

| | Leven's test for equality of vaiatuion | | t-test for equality of Means | | | | | %95 Confidence Interval of the difference | |
	F	Sig	t	df	Sig.(2-tailed)	Mean difference	St.Error difference	lower	higher
FR (Shoes) Equal variances assumed	.011	.918	-.343	62	.732	-.46078	1.34149	-3.14238	2.22082
Equal variances not assumed			-.341	58.298	.735	-.46078	1.35226	-3.16734	2.24577
FR (Cars) Equal variances assumed	4.093	.047	-4.235	62	.000	-6.68431	1.57828	-9.83926	-3.52937
Equal variances not assumed			-4.148	51.592	.000	-6.68431	1.61163	-9.91890	-3.44973
FR (Numbers) Equal variances assumed	.150	.699	-.221	62	.826	-.39804	1.80156	-3.99931	3.20323
Equal variances not assumed			-.222	61.876	.825	-.39804	1.79217	-3.98068	3.18460

| | Leven's test for equality of vaiatuion | | t-test for equality of Means | | | | | | |
	F	Sig	t	df	Sig.(2-tailed)	Mean difference	St.Error difference	%95 Confidence Interval of the difference lower	higher
MC (Shoes) Equal variances assumed / Equal variances not	.063	.803	3.776 / 3.747	62 / 58.443	.000 / .000	1..337 / 1..337	.354 / .357	.629 / .623	2.045 / 2.052
FMC (Cars) Equal variances	.421	.519	-8.450 / -8.563	62 / 61.567	.000 / .000	-1.937 / -1.937	.229 / .226	-2.396 / -2.390	-1.479 / -1.485
MC (Numbers) Equal	2.915	.093	-1.964 / -1.997	62 / 60.806	.054 / .050	-.878 / -.878	.447 / .440	-1.773 / -1.758	.016 / .001

For the free recall test of the female-oriented text no significant difference [t (62) = -0.343; p= 0.732] was observed between the performance of females and males though male participants' mean score (14.6667) preceded the female participants' mean score (14.2059). But for the multiple-choice of the female-oriented text, significant difference [t (62) = 3.776; p< .oo1] was seen. According to the mean scores of males (4.13) and females (5.47), females have outperformed males.

On the free recall test of the male-oriented text, there was significant difference [t (51592) = -4.184; p< .001] between the performances of

females and males; males (19.0667) have highly outperformed females (12.3824) while on the multiple-choice test of the male-oriented text, there was a significant difference [t (62) = -8.450; p<.001] between mean scores of males (7.97) and that of females (6.03).

For the free recall test of the neutral text, the t-test shows that there is no significant difference [t (62) =-.221; p= .826] between males and females but male participants' mean score (916.3833) preceded the female participants' mean score (15.9853). For the multiple choice test of the same neutral text, no significant difference test [t (62) = -1.964; p= .054] was observed. However the means indicate that males (6.97) have performed better than females (6.09).

5. Discussion

What the first research question of this study was aimed to examine was whether gender-oriented passage content of the second language reading text affect learners' reading comprehension. So three texts were provided (One female-oriented, one male-oriented, and one neutral text) and testing reading comprehension was approached via two comprehension measures: Multiple- choice and Free Recall.

Considering the mean scores of females and males on both methods of measurements, the statistical analysis reveals that females outperformed males on the female-oriented topic. On the other hand, males gained higher marks on the male-oriented passage. These results are in accordance with the self reports of males and females on the topic-familiarity questionnaire. Therefore, these results can be attributed to the claims made by them having more familiarity with and interest in these subjects. Also, the mean of the scores males and females gained on the two measurement methods (multiple choice and free recall) did not differ significantly from each other on the neutral text (Beliefs about Numbers). With regard to the reports of males and females on the topic-familiarity questionnaire having similar amounts of topic-familiarity with this subject (and this was the reason for the selection of this text as the neutral text for the study), this was expected to be the outcome of the study.

The results of this study are in line with the studies which claim that readers are expected to have better interpretation of the texts having background knowledge (Shohamy, 1984; Chihala et al., 1989; cited in Alderson, 2000; Johnston, 1984; Fazilatfar, 1989; Brantmeier, 2003a) and interest in the topic (Omaggio, 1993; cited in Brantmeier, 2006).

On the other hand, the result of this study contradicts Young and Oxford's (1997). No significant difference by gender with recall scores on the selected topics was observed. This result contradicts the results of the present study in that, here, males showed to be better at free recall scores on all sorts of topics (gender-oriented and neutral). This contradiction can be traced to the fact that, in the study by Young and Oxford (1997), the amount of familiarity of both genders with the selected topics did not differ significantly. In fact, the passages used in their study were not inclined toward either gender. The topics were mostly neutral. But their study can be a support from the perspective that the reading performance of females and males do not differ significantly on neutral texts.

Contrary to the present study in which multiple-choice and free recall were used for the measurement, the study by Bugel and Buunk (1996; cited in Macaro, 2003) made use of only multiple-choice task. Hence, it can be concluded that although the effect of topic familiarity has influenced similarly on the reading comprehension of females and males, the difference in the results of these two studies has been due to utilizing a different measurement task- i.e. free recall. On the whole, it can be put forward that in addition to topic-familiarity, the type of measurement can also play a role in the reading performance of females and males. Findings of the series of works by Brantmeier (2002; 2003a, 2003b, 2003c, 2003d; 2004a) suggest that in addition to linguistic factors, other variables such as gender, passage content, and topic familiarity may increase the second language reading burden. Those studies also are in harmony with what the present study was aimed to arrive at.

Based on the results of the second hypothesis, in the female-oriented topic, females outperformed males on multiple-choice type of test; in the

male-oriented topic, males had better performance than females on the multiple-choice test; in the neutral text, there was no significant difference between the performances of males and females on the part of the multiple-choice test.

The results may be interpreted from two different aspects: from the gender aspect and the test type. Considering the gender aspect, the results were in accordance with the expected presuppositions; however, regarding the results from the test type it should be noted that in the free recall tasks, in the male-oriented topic, males were better performers; in the neutral topic, there was no significant difference between males and females (although males had slightly higher mean scores than females) and in the female-oriented topic, unexpectedly, males outscored females. This can indicate that males are better at the free recall type of reading comprehension tests in general.

To look at the same results from another perspective, it can also be supposed that males have performed better on all the tests except for the multiple-choice questions of the female-oriented text. As Bernhardt (1991) suggests in multiple-choice kind of measurement the readers are limited to the concepts pointed out in the questions posed to them and this becomes problematic when the readers have comprehended many points in the text most of which are not applied in the multiple-choice questions and on the other hand the questions are about some points which the learners have not understood well. Therefore it can be supposed that males have had a better comprehension of all the three texts but have not managed to show their comprehension through multiple-choice questions on the female-oriented passage.

6. Conclusion

It can be concluded that passage content is for sure of paramount importance in reading comprehension and is truly related to the background knowledge of the readers (as it was seen in this study, females and males performed better on the topics related to their own gender). On the other hand, based on the results of this study, test type is

also related to gender and may play an important role. To ensure that this variable is related to gender, further research is required to show that free recall tests are more appropriate for males than females while of course keeping the content of the text, gender-oriented.

References

Afzali Shahri, M. (1995). *On the relevance of schema theory to the reading modules of the IELTS test.* Unpublished master's thesis. Tarbiat Modarres University, Tehran, Iran.

Ahmadi, H. (2001). *On the effects of content schemata and linguistic simplification on EFL reading comprehension and recall with regard to language proficiency.* Unpublished master's thesis. University of Teacher Education, Tehran, Iran.

Ajideh, P. (2006). Schema-theory based considerations on pre-reading activities in ESP textbooks. *The Asian EFL Journal Teaching Articles*, *16*, 1–19.

Alderson, J. C. (1984). Reading in a foreign language: A reading problem or a language problem. In J. C. Alderson & A. H. Urquhart. (Eds.), *Reading in a foreign language.* London: Longman.

Alderson, C. (2000). *Assessing reading.* Cambridge: CUP.

Al-Issa, A. (2006). Schema theory and L2 reading comprehension: Implications for teaching. *Journal of College Teaching and Learning.* 3 (7): 41-48.

Bernhardt, E. B. (1991). *Reading development in a second language: Theoretical, research, and classroom perspectives.* Norwood, N.J.: Ablex.

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: Literacy and the second language learner* (pp. 149-176). Greenwich, CT: Information Age Publishing.

Brantmeier, C. (2003a). Does gender make a difference? Passage content and comprehension in second language reading. *Reading in a Foreign Language* (*15*)1:1-24.

Brantmeier, C. (2003b). The role of gender and strategy use in processing authentic written input at the intermediate level. *Hispania, 86*(4):844-856.

Brantmeier, C. (2003c). Language skills or passage content? A comparison of native and non-native male and female readers of Spanish. *Applied Language Learning 13*(1):183-205.

Brantmeier, C. (2003d). Beyond linguistic knowledge: Individual differences in second language reading. *Foreign Language Annals*, 36 (1), 33–43.

Brantmeier, C. (2004). Gender, violence-oriented passage content, and reading in a second language. *The Reading Matrix 4*(2):1-19.

Brantmeier, C. (2006). Toward a multicomponent model of interest and L2 reading: Sources of interest, perceived situational interest, and comprehension. *Reading in a Foreign Language*, *18*(2), 89-115.

Burmeister, L. E. (1978). *Reading strategies for middle and secondary school teachers*. Reading Mass: Addison Welsley.

Carrell, P. L. (1981). Culture-specific schemata in L2 comprehension. In R. Orem & J. F. Haskell (Eds.), *Selected papers from the ninth Illinois TESOL/BE* (pp. 123-132). Chicago, Illinois: Illinois TESOL/BE Annual Convension, First Midwest TESOL Conference (pp.123-32). Chicago: Illinois.

Carrell, P. L. (1983a). Some issues in studying the role of schemata, or background knowledge, in second language comprehension. *Reading in a Foreign Language*, 1, 81-92.

Carrell, P. L. (1983b). Three components of background knowledge in reading comprehension. *Language Learning, 33*(2), 183-203.

Carrell, P. L., & Eisternhold, J. (1983). Schema theory and ESL reading pedagogy. *TESOL Quarterly*, *17*(4), 553-73.

Chastain, K. (1988). *Developing second language skill: Theory and Practice.* New York. Harcort Brace Jovanovich.

Chiu, M. M., & McBride-Chang, C. (2006). Gender, context, and reading: A comparison of students in 43 countries. *Scientific Studies of Reading*, *10*(4), 331–362.

Davanlou, M. (1995). *Background knowledge, language proficiency and EFL reading comprehension*. Unpublished master's thesis. University of Teacher Education, Tehran, Iran.

Doolittle, A., & Welch, C. (1989). *Gender differences in performance on a college-level achievement test*. (ACT Research Report Series 89-9). Iowa City: ACT.

Fazilatfar, A. M. (1989). *World knowledge needed for reading comprehension of the target language*. Unpublished master's thesis. Tarbiat Modarres University, Tehran, Iran.

Floyd, P., & Carrell, P. L. (1987). Effects on ESL reading of teaching cultural content schemata. *Language Learning*, 37, 89-108.

Gholampour, S. (1993). *The impact of topics on understanding EFL texts*. Unpublished master's thesis. University of Teacher Education, Tehran, Iran.

Heinz, P. J. (2004). Towards enhanced second language reading comprehension assessment: computerized versus manual scoring of written recall protocols. *Reading in a Foreign Language.16*(2). http://nflrc.hawaii.edu/rfl/october2004/heinz/heinz.html

Hudson, T. (1982). The effects of induced schemata on the 'short circuit' in L2 reading: non-decoding factors in L2 reading performance. *Language Learning, 32*(1), 1-31.

Johnson, P. (1981). Effects on reading comprehension of language complexity and cultural background of a text. *TESOL Quarterly, 15*, 169-181.

Johnson, P. (1982). Effects on reading comprehension of building background knowledge. *TESOL Quarterly*, *17*, 367-88.

Johnston, P. (1984). Prior knowledge and reading comprehension test bias. *Reading Research Quarterly,* 19(2), 219-239.

Karakas, M. (2002). The effects of pre-reading activities on ELT trainee teachers' comprehension of short stories. http://eku.comu.edu.tr/makaleler/3.html

Koda, K. (2005). *Insights into second language reading*: *A cross-linguistic approach*. NY: Cambridge University Press.

Maccoby, E. (1998)*. The two sexes: Growing up apart, coming together*. Cambridge, MA: Belknap Press of Harvard University Press.

Madanian, E. (1995). *The role of topic in EFL reading comprehension*. Unpublished master's thesis. Tehran University, Tehran, Iran.

Mohammad, M. A. H., Swales, J. M. (1984). Factors affecting the successful reading of technical instructions. *Reading in a Foreign Language*, 2:206-17.

Nafar, A. (2002). *The effect of textual annotation and contents schema on ELT and ESP reading comprehension*. Unpublished master's thesis. University of Teacher Education, Tehran, Iran.

Nuttall, C. (1996). *Teaching reading skills in a foreign language*. Oxford: Macmillan Heinemann.

Pae, T., (2004). Gender effect on reading comprehension with Korean EFL learners.

Pritchard, R. (1990). The effects on cultural schemata on reading processing strategies. *Reading Research Quarterly, 25*, 273-295.

Ruddell, M. R. (2005). *Teaching content reading and writing*. NJ: John Wiley& Sons, Inc.

Shohamy, E. (1984). Does the testing method make a difference? The case of reading comprehension. *Reading Research Quarterly*, 17, 229-255.

Tahmasbi, J. (1999). *The effect of background knowledge on text content and structure on the reading comprehension of EFL learners.* Unpublished masters' thesis. University of Teacher Education, Tehran, Iran.

Wolf, D. (1993). A comparison of assessment tasks used to measure FL reading comprehension. *Modern Language Journal*, 77, 473-489.

Young, D. J., & Oxford, R. (1997). A gender-related analysis of strategies used to process input in the native language and a foreign language. *Applied Language Learning, 8*, 43-73.