The Effect of Topic Interest and Gender on Reading Test Types in a Second Language

Sıla Ay
Ankara University

Özgür Şen Bartan
Ankara University Primary School

ABSTRACT

This study explores how readers' interest, gender, and test types (multiple-choice questions, Yes/No questions, and short-answer formats) affect second language reading comprehension in three different levels and five different categories of topics. A questionnaire was administered to 168 Turkish EFL students to find out the gender-oriented topic interests, and five different topics were identified. For each of the three levels (A2.1, A2.2, and B1 in the CEFR), five passages with the determined topics were selected. Comprehension was evaluated by three different assessment types. The findings of this study present the impact gender-oriented topic interest and assessment types have on reading comprehension.

INTRODUCTION AND LITERATURE REVIEW

Foreign language reading comprehension is believed to be affected by gender and interest. However, there has been little research regarding the correlation of different test types, gender, and topic interest. This study explored the relationship between gender and topic interest in foreign-la nguage reading comprehension at three language levels and with three different test types.

Studies by Christianson (1992), Eysenck (1982), Hidi (1990), and Thayer (1989) suggest that positive emotion, activation, and concentration each play an important role in the learning process. It was assumed that interest is also related, and that these variables mediate significantly in both first and second languages.

Reading Comprehension and Gender

There are some findings linking gender with interest in reading specific types of text. For example, according to Norvell (1958) and Thorndike (1941), males prefer adventure, sports, science, and information, while females prefer mystery and romance; males are less likely to read about female protagonists, whereas females are likely to read about male protagonists.

There have also been studies that document how individual differences affect reading comprehension in a foreign language (Brantmeier, 2002, 2003, 2005; Bügel & Buunk, 1996; Chavez, 2001). The results of some of these studies provide evidence that interest by gender has

a facilitating effect on the reading process (Grabe & Stoller, 2002; Koda, 2005; Pae, 2004). In a study on gender differences in achievement test performance at the college level, Doolittle and Welch (1989) found notable gender differences for items associated with specific texts, reporting that females scored higher than males with humanities-oriented reading texts, but lower than males with science-oriented texts.

Lynn and Mikk (2009) made a contribution to the study of international gender differences in reading ability by examining gender differences in means and variances in the PISA (Program for International Student Assessment) and the PIRLS (Progress in International Reading Literacy Study). One of the PISA studies done with 15-year olds in 40 countries was carried out in 2003, and in all these countries females obtained a higher average score than males. The other PISA study of 15-year-olds in 56 countries was carried out in 2006, and again, in all countries females obtained higher average scores over males. A world-wide measurement of reading ability was conducted, also by the PIRLS, evaluating the reading achievement of fourth-grade students (10-year olds) in 35 countries (Mullis, Martin, Gonzalez, & Kennedy, 2003). In that study, the reading-test scores of the females in all the countries were significantly higher than the males, and the situation in Turkey, where our study took place, as can be seen in Table 1, also lends support to those results.

Table 1. Turkish Readers' Scores in PISA 2003, 2006, and PIRLS 2001

PISA 2003				PISA 2006			PIRLS 2001					
Males' mean	Females' mean	Effect size G-B	Var, ratio B/G	Males' mean	Females' mean	Effect size	Females' mean	Females' SD	Males' mean	Males' SD	Effect size G-B	Var. Ratio B/G
426	459	0.36	1.28	427	471	0.47	459	84	440	87	0.22	1.07

(Lynn & Mikk, 2009, pp. 7-9)

According to Lynn and Mikk (2009), females' superiority in reading comprehension grows with age and practicing the language (p. 10). It is interesting that in average science-score attainment in the PISA, there was no difference between males and females, but in reading females were better ($0.4d$) than males. Lynn and Mikk correlated this difference with the living conditions and activities of males and females. For instance, females were reading more at home and school, while males were busy with computers and DVD players (p. 11).

Reading Comprehension and Interest

Alexander and Jetton (2000) gave a detailed review of definitions and categories of interest (p. 298). They introduced two main dimensions of interest involved in the reading process: situational and individual. Individual interest subsumes the readers' preferences for certain text topics or subject matter, and this interest exists before reading a particular text (Hidi, 1990; Schiefele, 1992). Situational interest refers to interest caused by situational factors, such as the text and test; a particular situation can influence situational interest, and is therefore not enduring (Krapp, Hidi, & Renninger, 1992; Wade, 1992).

Conversely, Hidi and Anderson (1992) defined text-based situational interest as interest that is activated by text through topics or ideas that are of universal appeal. Hidi (1990) noted that, whereas individual interest research tends to focus on individual differences, research that deals with the "interestingness" of a situation focuses on the effect of interest across individuals. Similarly, recognizing that interest can be evoked by sources other than text characteristics, she concluded that the term *situational interest* should be adopted to describe all environmentally-triggered interest of which text-based interest should be considered as one sub-type. In addition to these two categories, topic interest, that is, interest triggered when a specific theme is presented, was yet another topic investigated.

Reading Assessment

Coombe, Folse, and Hubley (2007) summarize the points that need to be remembered about reading assessment. First, testers should test the skills that are taught, and use 4 to 10 items for each test type. Second, questions ought to cover all sections of a text. Third, grammar and vocabulary should be asked in context. Fourth, challenging questions that involve inferencing and critical thinking should be asked (p. 64). Fifth, Bloom's Taxonomy of Educational Objectives (Bloom, Engelhart, Furst, Hill, & Krathwohl, 1956) reflects different abilities and behaviors that are related to thinking processes, which was helpful in writing the items and objectives of this study's reading-assessment questions. Equally, a test should always be constructed with explicit specifications, addressing both cognitive and linguistic abilities (Weir, 2005, p. 14). In our study, all of these considerations were taken into account while writing the items of the reading-assessment tests.

As Coombe and Hubley (2004) stated, it is not possible to observe reading behaviors directly, but in assesment there are major and minor reading skills. Major skills include reading quickly to skim for gist, scanning for specific details, and establishing the overall organization of the passage; reading carefully for main ideas, supporting details, author's argument and purpose; recognizing the relationship of the paragraphs and fact versus opinion, and transfering information from nonlinear texts. Minor skills include understanding at the sentence level; recognizing syntax, vocabulary, and cohesive markers; understanding at the inter-sentence level; understanding references and discourse markers; understanding the components of nonlinear texts and the meaning of graph or chart labels, keys, and the ability to find and interpret intersection points.

According to Bernhardt (1991), in order to understand the complete picture and to examine the validity of reading comprehension tests, a variety of assessment tasks is needed. In Bernhardt's study, three test types—multiple-choice questions, dichotomous items (True/False, Yes/No), and short-answer formats—were selected from the list (see Appendix B) of the reading test-item types (Grabe, 2009). Coombe et al. (2007) mentioned that multiple-choice and True/False question types can be used to test the main idea, supporting details, writer's purpose, and the relationships and facts between paragraphs. Written-recall questions that are focused on details are especially good for checking the gist, intensive understanding of a text, and the who-what-where-when content. Bernhardt recommends written-recall questions to test the information transfer from non-linear texts (pp. 35 and 54).

Despite the fact that multiple-choice questions (MCQs) are widely used, they have limitations as well as strengths. Some of the strengths are that they have retrieval cues, and the answers are predetermined with no ambiguity in the scoring. In addition, learning outcomes, from simple to complex, can be measured; the tasks are clear and structured; scores are less influenced by guessing than True/False items, and a lot of material can be covered efficiently. One of the limitations of MCQs, however, is that they are ineffective for measuring certain types of problem solving and the ability to organize and express ideas. Another problem with this type is that it often focuses on testing factual information and fails to test higher levels of cognitive thinking (Zimmaro, 2004). Coombe et al. (2007) mention that MCQs do not lend themselves to the testing of productive language skills or language as communication, and add that MCQs can also be used to test higher-order thinking skills. Moreover, MCQs have four response options, but with listening, just three options are recommended (Coombe et al., 2007, p. 19). They advise the MCQs writers to take background knowledge into account, provide as much context as possible, keep sensitivity and fairness issues in mind, standardize the number of response options, and one response option should be an unambiguous correct or best answer. Writers of MCQs should make all response options similar in length and level of difficulty, avoid using distracters like *none of the above*, have correct answers appear equally in all positions, make the

stem positive, make sure all questions are independent of one another, move recurring information in response options to the stem, avoid writing absurd or giveaway distracters, and avoid extraneous clues (Coombe et al., 2007, p. 26).

Yes/No questions are considered to be the second-most popular assessment type. They are typically written as statements, and the students' task is to decide whether they are true or false. It is possible with this type to test large amounts of content (Coombe et al., 2007, p. 27). Another advantage of this type of assessment is scoring; it is quick and reliable, and it can be accomplished efficiently and accurately. However, it has some disadvantages, the most important one being that there is a 50% guessing factor that the choice will be correct. To alleviate this handicap, it is recommended that teachers use a third response category, called *Doesn't Say* or *Not Given*, so that the diagnostic value of this type can be 33.3% of the test.

Short-answer items require students to answer questions in a few words, phrases, or sentences, and they are generally focused on details. One advantage of this assessment type is that they encourage students to truly know the answer rather than just recognize or guess. The other advantage is that higher-order thinking skills can be evaluated (Coombe et al., 2007, p. 34). Short-answer items are also used for assessing a quantity of ideas (Brantmeier, 2002, 2003).

Grabe (2009, p. 454) suggests a list of recommendations for reading-assessment practices:

1. Students should be tested on a range of relevant skills.

2. Students should be encouraged to read longer texts (for advanced assessment, 700-1000 words, assuming 120-150 wpm).

3. Background knowledge influences all comprehension, and needs to be accounted for in a positive way (multiple topics, multiple tasks, general topics limited interdependence of items within some subsets of tasks).

4. Group tasks might be used to engage discussions of reader interpretations of texts.

5. Extensive reading should not be discouraged by assessment procedures.

6. The importance of identification and fluency skills needs to be explored (reading word lists, oral reading for one minute, silent reading on computer, timed reading, assessment of rereading).

7. Tests could explore ways to assess synthesis skills, evaluation skills, strategies, meta-cognitive knowledge, and skills monitoring (text monitoring while reading).

8. Reading should be tested within a context-focused battery (but item interdependence has to be a content).

9. Tests should consider item types that take advantage of computer interfaces (e.g., allow a text to disappear after reading, use a few hypertext links in a test passage, or combine information from multiple texts to complete a task).

Many skills might be measured usefully through informal assessment options in classroom contexts. What one loses in reliability and objective controls could be countered by the many formal and informal assessments that can be made in the classroom.

Reading Assessment and Topic Interest

While Hidi and McLaren (1991) have considered topic interest to be a form of situational interest, others have viewed topic interest as a form of individual interest (Schiefele, 1996;

Schiefele & Krapp, 1996). The latter research indicated that both situational and individual factors contributed significantly to topic interest, and topic interest influenced affect, which in turn influenced persistence, and persistence was significantly related to learning. Topic interest may have a critical role in reading, as students usually encounter the topic at the very beginning of their reading tasks (Hidi, 2001, pp. 194-195).

However, the impact of interest on reading test types of comprehension has not been widely examined. This issue was discussed in the study of Bray and Barron (2004, p. 121) who investigated the relationship between students' interest in reading texts and their performance on comprehension test items. They indicated that there was a small but significant relationship between interest and test performance: it was stronger for girls and for students of higher ability levels (p. 107).

Studies by Alexander, Jetton, and Kulikowich (1995) and Anderson, Shirey, Wilson, and Fielding (1987) have found a positive relationship between interest and recall. However, there is some evidence that interest may be even more strongly related to deeper levels of discourse processing (Schiefele, 1996). Moreover, Bray and Barron (2004) used MCQs to assess not only the literal understanding of the texts but also higher-order comprehension, such as the ability to make inferences and generalizations.

THE STUDY

This study was conducted to examine the relationship between readers' gender, topic interest, and foreign-language reading comprehension in three different language levels and with three different test types. The primary research question was: What is the effect of topic interest and gender on reading test types in a second language?

The Pilot Study

To determine the gender-oriented topic interests, a pilot study was conducted in which an interest scale presenting 23 topics (see Appendix A) was administered to 168 Turkish EFL students in another private primary school. The students of the pilot study were at the same level (but different school) as the study participants. The topics of the scale were selected from the topics in the recent textbooks on the market. Five different texts with determined topics were given, and comprehension was evaluated by the three different assessment types. According to the results of this scale, five topics of the highest interest were identified; the male topic was technology, the female topic was fashion, the topic that both genders found interesting was humor, the least interesting topic for both genders was fairy tale, and the neutral topic was art.

Reading Texts

Fifteen reading texts (five texts in the mentioned topics for each of the three levels) were chosen from textbooks and educational websites. The guidelines for the selection of materials were linguistic complexity, reference to the Cambridge English Lexicon and structure, and the number of words used. According to their levels, texts contained a certain number of words (approximately 150 words for A2.1, 300 words for A2.2, and 450 words for B1; see Appendix B). Reading texts can be categorized as narrative and expository, but the genres of the texts were not taken into consideration as the main focus was the interestingness of the topics.

Comprehension Assessment Tasks

To check the clarity and comprehensibility of the assessment tasks, the pilot group read the texts and answered the comprehension questions. The assessment types were considered and decided according to the participants' levels. While the Cambridge ESOL exam of A2.1 (Flyers) incorporates a sentence-completion task to assess reading comprehension, the exam for A2.2 (KET) includes a written recall task to check reading comprehension. Likewise, although A2.2 (KET) covers *Yes/No* and *Doesn't Say* options, neither A2.1 (Flyers) nor B1 (PET) embraces the *Doesn't Say* option. The assessment types were adapted from these guidelines (see Table 3).

Table 2. Assessment Tasks and Scoring

Level	Assessment Tasks			Total
A2.1	Sentence completion (10 pts)	Yes/No (10 pts)	Multiple choice (10 pts)	30 pts
A2.2	Short answer formats (10 pts)	Yes/No/Doesn't Say (10 pts)	Multiple choice (10 pts)	30 pts
B1	Short answer formats (10 pts)	Yes/No (10 pts)	Multiple choice (10 pts)	30 pts

Table 2 shows the aims of the questions that were prepared for each assessment type (MCQs, Yes/No questions, and short-answer formats). For instance, the MCQ questions accompanying the reading texts aimed at assessing the vocabulary (in context) and sentence/ paragraph comprehension, understanding the main message of the text, and comprehending the inferences in context. Therefore, the goal was to test the same specifications in all three different levels (for an example of the assessment tasks, see Appendix C).

For each type of reading assessment, different specific aims were tested. However, for all three levels, the reading skills tested were almost the same (e.g., Table 3 shows the MCQs specifications). Although five different texts and three different levels were tested, all MCQs were tested to find the same specifications which are shown in Table 3. All three question types have both advantages and disadvantages, but they all need to be prepared with some specification in order to be valid and reliable.

Table 3. Assessment Tasks and Scoring

Table of specifications: MCQs **Total points**: 10 points **Time:** 10 minutes **Text type:**	Narrative
Level:	A2.1, 183 words
Topic:	Technology (the male topic): **The Car of the Future**
Reading skills task:	Main idea, supporting details, inferences, vocabulary in context, text coherence were tested with five MCQs with three choices; the questions were in order with the text, and there were one or two distracters.
The focused skill for each item:	The questions were written to comply with the following rules: 1. to find the real information in the sentence. 2. to find the sentence which is similar to the other. 3. to find the general meaning of the text or paragraph. 4. to find the inference of the marked word. 5. to find the vocabulary in context.
Scoring:	Each question was one point, and testers circled the right answer.

The results of the pilot study informed a revision of the test items as well as the rubric of the sentence-completion and short-answer tasks so that the answers to the open-ended questions were more predictable.

Participants

The participants of this study were 159 students (80 males and 79 females) between the ages of 12 and 14 from Ankara University Primary School, Turkey. There were three different groups, who were assumed to have achieved the levels of A2.1, A2.2, and B1 of CEFR (Common European Framework of Reference for Languages).

Table 4. Participants According to Their Levels and Gender

|	**B1**	**A2.2**	**A2.1**
Males	14	15	51
Females	26	20	33
Total	40	35	84

Study Procedure

The readings were then given to the study participants by their English teachers in separate sessions and on different days, and each participant read one text per week for five weeks. MCQ and True/False questions were scored by the two researchers, while written-recall questions were scored by the two researchers as well as their regular teachers to ensure that the scores awarded were accurate. During the assessment of the comprehension tasks, sentence-completion items were scored according to the possible answers observed during the pilot study. Data management and analysis was performed using SPSS 11.5.

FINDINGS

Gender

Females generally attained higher scores overall (in all three levels, five topics, and three test types), but the difference between male and female scores was especially significant in the short-answer formats (see Figure 1).

Figure 1. Test Type and Gender Interaction
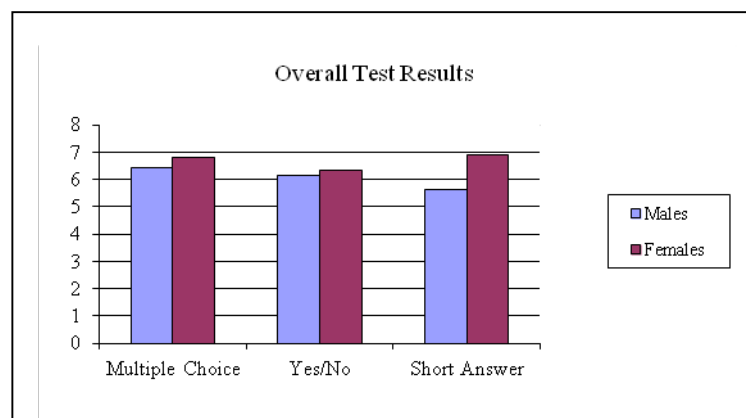
When the scores were analyzed, females had mostly higher grades, except for the topic of technology where males had a slightly higher mean grade (see Figure 2).

Figure 2. Gender and Text Topic Scores
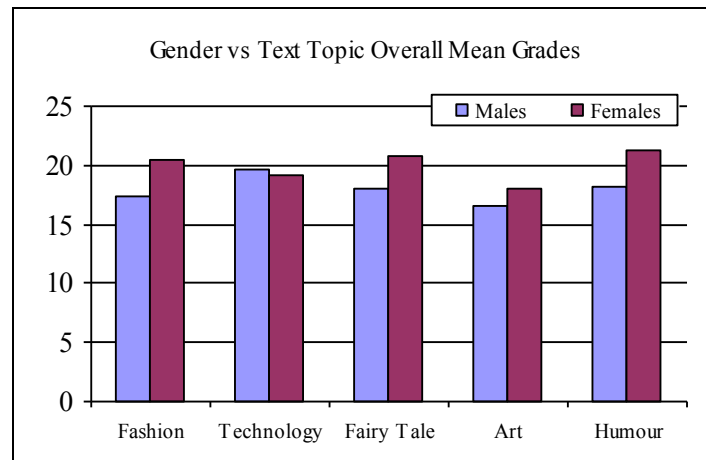
Topic

To find out the difference between topics in overall grading, the paired-samples *t* test was used. The pairs with a *p* value smaller than .05 show a significant difference in mean scores. The scores taken from the art texts were significantly lower than all other scores. There is no other significant difference between the text topics in overall grading (see Table 5).

Table 5. Paired-Samples Table

|		Paired Differences							
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference		t	df	Sig. (2-tailed)
Pair 1	Fashion/ Technology	-2,523	29,04996	2,3333	-7,133	2,085	-1,082	154	,281
Pair 2	Fashion /Fairy Tale	-2,425	26,73748	2,1476	-6,667	1,817	-1,129	154	,261
Pair 3	Fashion /Art	5,642	24,03765	1,9307	1,828	9,456	2,922	154	,004
Pair 4	Fashion/ Humor	-3,251	25,91880	2,0818	-7,364	,861	-1,562	154	,120
Pair 5	Technology/ Fairy Tale	,098	35,59704	2,8592	-5,550	5,746	,034	154	,973
Pair 6	Technology/ Art	8,166	32,90019	2,6426	2,945	13,38	3,090	154	,002
Pair 7	Technology/ Humor	-,727	31,04239	2,4933	-5,653	4,197	-,292	154	,771
Pair 8	Fairy Tale /Art	8,067	24,99529	2,0076	4,101	12,033	4,018	154	,000
Pair 9	Fairy Tale/ Humor	-,826	29,79106	2,3928	-5,553	3,900	-,345	154	,730
Pair 10	Art /Humor	-8,894	27,76334	2,2300	-13,29	-4,488	-3,988	154	,000

Level

Pairwise comparison results show that there is a significant difference between levels. A2.1 students had significantly lower scores than A2.2 and B1 students, but no significant difference was observed between A2.2 and B1 students' scores.

Test Type

MCQ scores of both genders were higher than the scores of the other two test types. The lowest MCQ scores were taken from the fairy tale topic, which was the least interesting topic for both genders.

**Figure 3.** Topic and MCQ Mean Grades
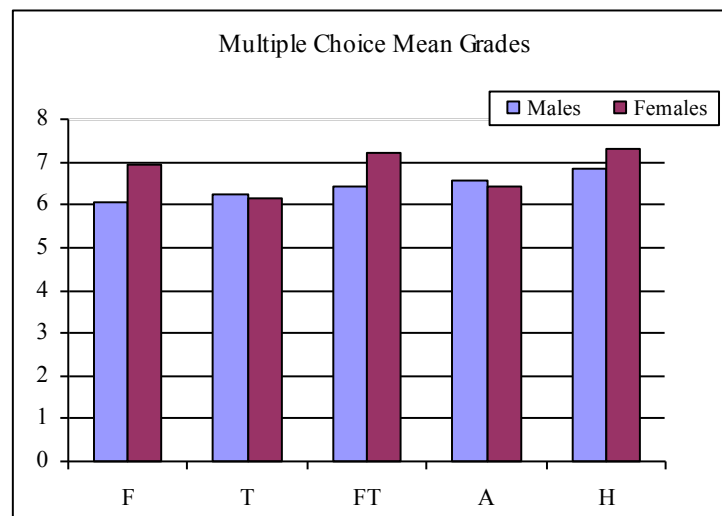
In the Yes/No test type, males received mostly lower scores compared to females, showing that the difference between genders is more significant than the two other test types. This finding is difficult to explain because in this type of assessment there is a 50% guessing factor that the choice will be correct. Nevertheless, males got the highest score in this test type from their highest-interest topic. Therefore, it seems that this finding is more likely to be the effect of topic interest, not the test type.

Figure 4. Topic and Yes/No Mean Grades
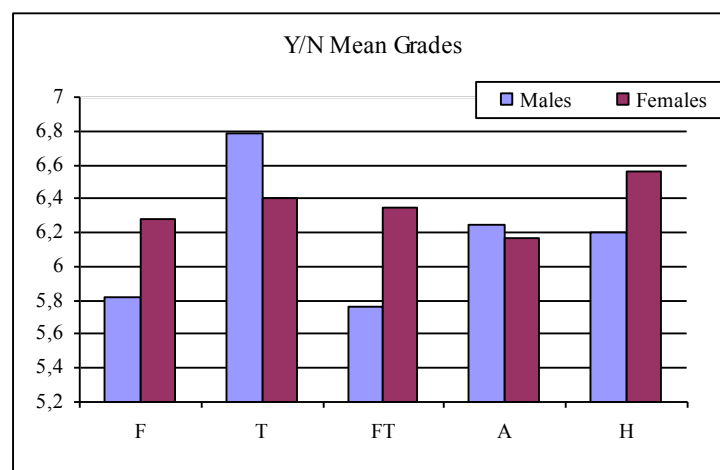
The difference between genders was more significant when short-answer questions were examined (Figure 5). Both genders obtained higher Yes/No scores from the topics in which they had the greatest interest (i.e., males received higher scores from technology, and females received higher scores from the topic of fashion).

Figure 5. Topic and Short-Answer Mean Grades
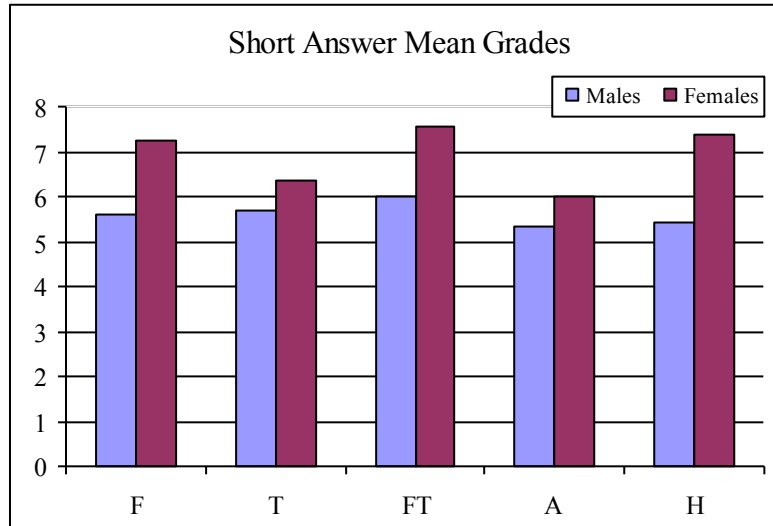
When gender is taken into consideration, males received lower scores on short-answer format questions, whereas females recxeived higher scores from this test type. Unlike the other two test types, males did not obtain higher scores in any topic, even from their highest-interest topic (see Figure 4).

Gender, Level, and Test-Type Interactions

For the female students of the A2.1 level, being assessed by three different test types was not effective for the topic of humor, as shown in Figure 6. For the male students of the same level, however, it seems that being assessed by three different test types was effective for all five topics. From the art and humor topics, similar scores were taken from the MCQ and Yes/No test types.

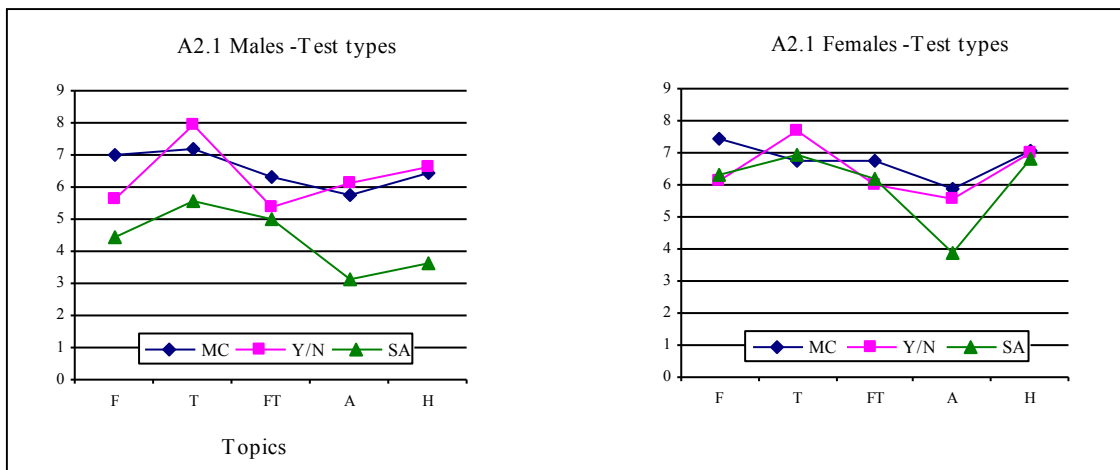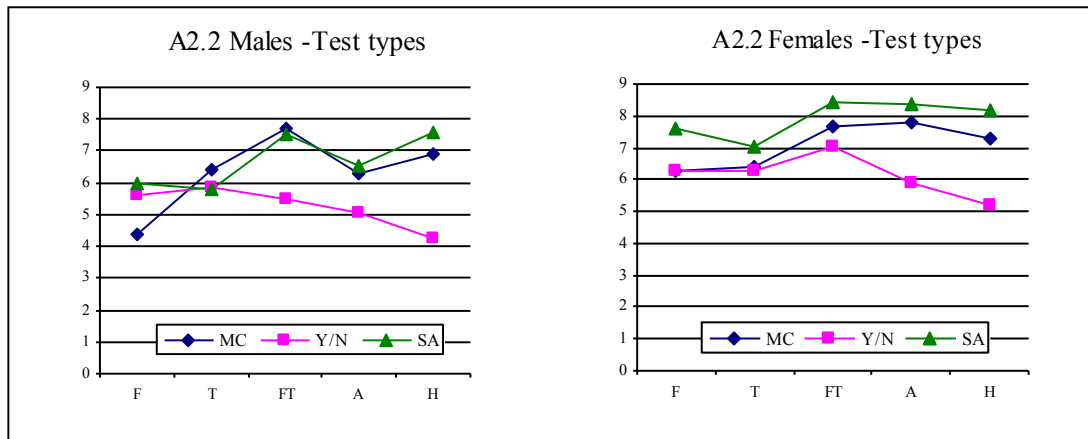**Figure 6.** Males and Females — A2.1 Test Types

Figure 7. Males and Females — A2.2 Test Types

As shown in Figure 7, for the female students of the A2.2 level, being assessed by three different test types was effective. The highest scores were achieved on the short-answer questions, and the lowest scores were obtained on the Yes/No questions. For the male students of the same level, it can be clearly seen that being assessed by three different test types was not effective. However, from the humor, art and fairy tale topics, Yes/No question scores were lower.

Figure 8. Males and Females — B1 Test Types
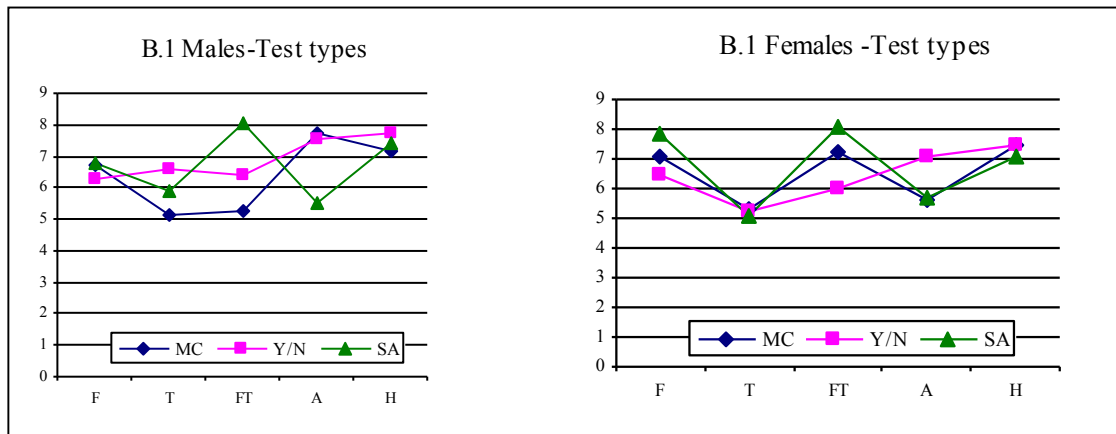
For the female students of the B1 level, being assessed by three different test types was effective. Similar scores were obtained on the three different test types only for the technology and humor topics. For the males the fairy-tale topic scores show the effect of being assessed by different test types. Again, for the topic of humor and fashion, the three test types' scores were similar (see Figure 8).

DISCUSSION AND CONCLUSION

Regarding the effect of the test types, the lowest scores were taken from the short-answers, and what is more, among the short-answer format questions, the lowest scores were from the text on humor, which was the most-liked topic of both genders. It is hard to explain this gap between the scores on humor and fashion, or humor and technology. The highest scores were

taken from the MCQs regardless of the variables (gender, level, or topic). This could be a result of Turkish students' being accustomed to MCQ types in all subjects throughout their education.

Females were found to be more successful with all of the text topics in this study, but both genders got the highest scores from the text of their highest-interest. The present findings seem to be consistent with other research which shows that the performance of males and females tends to be similar in their topics of interest (Anderson et al., 1987; Asher & Markell, 1974; Hidi & Berndorff, 1998; Hidi, Berndorff, & Ainley, 2002). Females got higher scores than males in every topic when assessed with short-answer questions (which require higher-order thinking and verbalizing skills), which confirms the results of previous studies (Doolittle & Welch, 1989; Ehrman & Oxford, 1998; Locke, Ginsborg, & Peers, 2002) reporting that females are better in linguistic tasks. Males got lower scores from the short-answer questions on all of the topics. The lowest scores were from the texts on art; the second-lowest scores were taken from fashion. From the text on technology, however, males achieved slightly higher scores on this test type. Therefore, this finding also demonstrates the importance of topic interest over test type for males.

Test takers had a clue as to what they were going to read with the help of the titles for the topics of fashion and technology. However, no clues were given about the humor texts, except for the A2.2 level text. In one of their studies, Rinehart, Gerlach, Wisell, and Welker (1998) found that the adolescent readers were able to draw important clues from titles, summaries, and cover illustrations. With this information, they accurately predicted whether or not they would like to read the text. According to Ainley, Hillman, and Hidi (2002), "When a student is presented with the title of a literary text and some basic information about its content, this triggers an immediate reaction. The reaction is an expectancy or anticipatory response and at that point represents the students' triggered interest in the text topic" (p. 412).

In this study, the titles for the technology texts were, *The Car of the Future* (A2.1), *Where Real Reality Ends* (A2.2), and *Virtual Communities* (B1). The titles for the fashion texts were, *The Clothes Make the Man* (A2.1), and *Fashion* (for both A2.2 and B1). The humor texts' titles were, *Doughnuts* (A2.1), *A Funny Thing Happened to Me* (A2.2), and *Job at the Zoo* (B1). It can be seen from the fashion and technology titles that they reflected the content clearly, but among the humor texts, only the title of the A2.2-level text reflected the content for the readers. It is important to note here that both males and females obtained higher scores on this text (A2.2), especially on the short-answer format, which supports the idea that a title that hints at the content is likely to trigger interest in a person who is fond of that subject.

The study results reveal that different assessment types do not effectively assess language skill, as similar patterns were observed in all three levels' of reading-comprehension scores. In other words, no correlation between the language levels and the types of reading assessment used was found. Overall, the results of this study show that it is not the test type, but rather the topic interest that affects comprehension most significantly.

The results of this study have enhanced our understanding of the effect of topic interest on pre-adolescent learners, but some limitations should be borne in mind. With a small-size sample and a certain age group, caution must be applied, as the findings might not be transferable to the other test types, such as cloze, gap filling, information transfer, matching, summary, portfolios, editing, and other informal assessment formats. Another issue that could be a limitation of this study, and that merits further investigation, is the effect of language proficiency. The study was limited to three ability groups which had similar reading competences. Future studies may reveal more about this aspect if done with two other levels, such as A1 and C1 of CEFR, in which students have significant proficiency differences. Lastly, as the current study was not specifically designed to evaluate factors related to the text title, further research regarding its role on different test types in reading-comprehension performance is strongly recommended.

Sıla Ay *is an assistant professor whose Ph.D. was in foreign language reading skills. She has been a lecturer at Ankara University Faculty of Letters, Applied Linguistics Department, Turkey since 1995. Her research interests focus on individual differences in foreign language learning, language teaching methodology, reading skills, and dyslexia.*

Email: silaay@gmail.com

Özgür Şen Bartan *has been teaching English in the Ankara University Primary School. Her teaching career began nine years ago upon graduating from the English Linguistics Department, and then pursuing her M.A. in Sociology, both at Hacettepe University. She is currently pursuing her Ph.D. at Ankara University in the Foreign Language Teaching Division.*

Email: ozgursen1@yahoo.com

REFERENCES

Ainley, M., Hillman, K., & Hidi, S. (2002). Gender and interest processes in response to literary texts: Situational and individual interest. *Learning and Instruction, 12,* 411-428.

Alexander, P. A., & Jetton, T. L. (2000). Learning from a text: A multidimensional and developmental perspective. In M. L. Kamil, P. B. Mosenthal, P. D. Pearson, & R. Barr (Eds.), *Handbook of reading research: Vol. 3.* Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Alexander, P. A., Jetton, T. L., & Kulikowich, J. M. (1995). Interrelationship of knowledge, interest, and recall: Assessing a model of domain learning. *Journal of Educational Psychology, 87,* 559-575.

Anderson, R. C., Shirey, L. L., Wilson, P. T., & Fielding, L. G. (1987). Interestingness of children's reading material. In R. E. Snow & M. J. Farr (Eds.), *Aptitude, learning, and instruction: Vol. 3. Cognitive and affective process analyses*. Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Asher, S. R., & Markell, R. A. (1974). Sex differences in comprehension of high- and low-interest reading material. *Journal of Educational Psychology, 66,* 680-687.

Bernhardt, E. B. (1991). *Reading development in a second language*. Norwood, NJ: Ablex.

Bloom, B. S., Engelhart, M. D., Furst, E. J., Hill, W. H., & Krathwohl, D. R. (1956). *Taxonomy of educational objectives: The classification of educational goals, by a committee of college and university examiners. Handbook 1: Cognitive domain.* NY: Pearson Longman Publishing.

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*. 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*, 1-27.

Brantmeier, C. (2005). Effects of reader's knowledge, text type, and test type on L1 and L2 reading comprehension in Spanish. *The Modern Language Journal*, *89*, 37-53.

Bray, G. B., & Barron, S. (2004). Assessing reading comprehension: The effects of textbased interest, gender, and ability. *Educational Assessment, 9*(3/4), 107-128.

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

Chavez, M. (2001). *Gender in the language classroom.* NY: McGraw Hill.

Christianson, S. A. (Ed.). (1992). *The handbook of emotion and memory: Research and theory*. Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Coombe, C., Folse, K., & Hubley, N. (2007). *Assessing reading: A practical guide to assessing English language learners*. Ann Arbor, MI: The University of Michigan Press.

Coombe, C., & Hubley, N. (Eds.). (2004). The corner stone of language testing: A practical guide for teachers in the gulf. In D. Lloyd, P. Davidson & C. Coombe (Eds.), *Fundamentals of language assessment* (pp. 3-11). Dubai, United Arab Emirates: TESOL Arabia.

Doolittle, A., & Welch, C. (1989). *Gender differences in performance on a college-level achievement test* (ACT Research Rep., series 89-90). Iowa City, IA: American CollegeTesting Program.

Ehrman, M., & Oxford, R. (1998). Effects of sex differences, career choice, and psychological type on adult language learning strategies. *The Modern Language Journal, 72*(3), 253-265.

Eysenck, M. W. (1982). *Attention and arousal*. Berlin: Springer.

Grabe, W. (2009). Teaching and testing reading. In M. H. Long & C. J. Doughty (Eds.), *The handbook of language teaching,* Chichester, UK: Blackwell.

Grabe, W., & Stoller, F. L. (2002). *Teaching and researching reading*. London: Pearson Education/Longman.

Hidi, S. (1990). Interest and its contributions as a mental resource for learning. *Review of Educational Research*, *60*, 549-572.

Hidi, S. (2001). Interest, reading and learning: Theoretical and practical considerations. *Educational Psychology Review, 13*(3), 194-195.

Hidi, S., & Anderson, V. (1992). Situational interest and its impact on reading and expository writing. In A. Renninger, S. Hidi, & A. Krapp (Eds.), *The role of interest in learning and development*. Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Hidi, S., & Berndorff, D. (1998). Situational interest and learning. In L. Hoffmann, A. Krapp, K. A. Renninger, & J. Baumert (Eds.), *Interest and learning: Proceedings of the Seeon conference on interest and gender*. Kiel, Germany: Institute for Science Education at the University of Kiel.

Hidi, S., Berndorff, D., & Ainley, M. (2002). Children's argument writing, interest and self-efficacy: An intervention study. *Learning and Instruction*, *12*(4), August 2002, 429-446. doi:10.1016/S0959-4752(01)00009-3

Hidi, S., & McLaren, J. (1991). Motivational factors and writing: The role of topic interestingness. *European Journal of Psychology of Education, 6*(2), 187-197.

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

Krapp, A., Hidi, S., & Renninger, K. A. (1992). Interest, learning and development. In A. Renninger, S. Hidi, & A. Krapp (Eds.), *The role of interest in learning and development*. Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Locke, A., Ginsborg, J., & Peers, I. (2002). Development and disadvantage: Implications for the early years and beyond, *International Journal of Language & Communication Disorders, 37*(1), 3-15.

Lynn, R., & Mikk, J. (2009). Sex differences in reading achievement. *Trames*, *13*(63/58), 3-13.

Mullis, I. V. S., Martin, O. M., Gonzalez, E. J., & Kennedy, M. A. (2003). *PIRLS 2001 international report: IEA's study of reading literacy achievement in primary schools*. Chestnut Hill, MA: Boston College. Retrieved May 26, 2008, from http://pirls.bc.edu/pirls2001i/PIRLS2001_Pubs_IR.html

Norvell, G. W. (1958). *What boys and girls like to read?* Morristown, NJ: Silver Burdett Company.

Pae, T. (2004). Gender effect on reading comprehension with Korean EFL learners. *System*, *32*(3), 265-281.

Rinehart, S. D., Gerlach, J. M., Wisell, D. L., & Welker, W. A. (1998). Would I like to read this book?: Eighth graders' use of book cover clues to help choose recreational reading. *Literacy Research and Instruction, 37*(4), 263-280.

Schiefele, U. (1992). Topic interest and levels of text comprehension. In A. Renninger, S. Hidi, & A. Krapp (Eds.), *The role of interest in learning and development*. Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Schiefele, U. (1996). Topic interest, text representation, and quality of experience. *Contemporary Educational Psychology, 12*, 3-18.

Schiefele, U., & Krapp, A. (1996). Topic interest and free recall of expository text. *Learning and Individual Differences, 8*, 141-160.

Thayer, R. E. (1989). *The biopsychology of mood and arousal.* New York, NY: Oxford University Press.

Thorndike, R. L. (1941). *Children's reading interests.* New York, NY: Columbia University Press.

Wade, S. E. (1992). How interest affects learning from a text. In A. Renninger, S. Hidi, & A. Krapp (Eds.), *The role of interest in learning and development*. Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Weir, C. J. (2005). *Language testing and validation: An evidence-based approach.* Basingstoke, Hants, UK: Palgrave MacMillan.

Zimmaro, D. M. (2004). Writing good multiple-choice exams. Retrieved April 16, 2012, from http://ctl.utexas.edu/assets/Evaluation--Assessment/Writing-Good-Multiple-Choice-Exams-04-28-10.pdf

APPENDIX A

Topics Included in the Topic Interest Scale

1. Biography/Autobiography	6. Fantasy	11. History	16. Legends	21. Horror
2. Detective Stories	7. Folk Tales	12. Sports	17. Technology	22. Food
3. Romantic Stories	8. Humour	13. Culture	18. Health	23. Art
4. Science Fiction	9. Social Studies	14. Fairy Tales	19. Science	
5. Daily Routines	10. Fashion	15. Nature	20. Hobbies	

APPENDIX B

Fashion Text for B1 Level

There was a time, not long ago, when urban fashion did not exist. Fashion meant expensive designer outfits that only the very rich could afford. Think Paris catwalks, Vogue magazine and very uncomfortable shoes. And it was all about women. Even though the outfits were normally designed by men, there was very little interest in menswear. The outfits were very feminine, very expensive, and impractical for anything other than walking in a straight line.

All of this has changed. Young people no longer dress like their parents. They wear comfortable, practical, and affordable clothes which have probably been designed by somebody young: Urban fashion has taken over the world. **It** began, perhaps, with the hippies of the 1960s, who rejected mainstream society's values and dress codes with their long hair and scruffy image. Next, came the punks of the 1970s, who went one step further with shocking, scary outfits that deliberately offended ordinary people. Even today, if you go to London's Camden Town on the weekend, there will be plenty of people with lots of tattoos, piercings and chains hanging out of their trousers. This punk-influenced style is all about looking different from everybody else—but exactly the same as your friends. Certainly, by the 90s, high street shops like *Gap* quickly followed the trends from the streets, and designers like Stella McCartney began to sell street-influenced clothes, such as ripped T-shirts, for huge amounts of money.

Music has always played a part in urban fashion. Nu-metal fans, for example, in addition to wearing Marilyn Manson t-shirts, tend to look different from people who like chart music, and many of today's pop stars get their ideas from people on the street rather than the other way around. In California, surfers needed comfortable clothing to wear on the beach, and the surf look spread throughout America via companies such as *Quicksilver*, *Stüssy*, *O'Neil,l* and similar companies that originally produced clothes for sports. But the boom in extreme sports such as skateboarding and snowboarding has influenced urban fashion. Baggy jeans are better if you are wearing inline skates, for example, and **hoodies** are influenced by practical snowboarding outfits. Of course there are plenty of people who never do any sport at all and are still happy to wear the latest sporty urban fashion, just as most people who wear trainers rarely go near a sports centre. (Kane, 2003)

A. Choose the best answer. Circle a, b, c or d.

1. What played an important role in the development of urban fashion?

    a) The large amount of people dealing with sports in general.
    b) The increasing popularity of extreme sports.
    c) The huge number of people doing outdoor sports.
    d) The snowboarding outfits.

2. What did you learn about the style of the teenagers today?

    a) They wear very different clothes than their parents did.
    b) They dress the same as their parents.
    c) Some of them like to dress like their parents.
    d) They wear expensive clothes.

3. Which of the following is the key word for urban fashion?

    a) Elegance
    b) Discomfort
    c) Comfort
    d) Cheap price

4. In the sentence, "**It** began, perhaps, with the hippies of the 1960s,…" "**It**" refers to…

    a) The phase of change in urban fashion.
    b) The change of the colours.
    c) Fashion
    d) The explosion in the popularity of urban fashion.

5. In the sentence, "and **hoodies** are influenced by practical snowboarding outfits." "**Hoodies**" means…

    a) Tight tops that keep warm
    b) Sunglasses
    c) Large comfortable sweatshirts
    d) A pair of mittens

B. Circle true (T) or false (F).

    1. T F "Urban fashion" is a recent concept.
    2. T F In the past, fashion was associated with ordinary people like you and me**.**
    3. T F Most fashion designers were women.
    4. T F Only women's wear attracted people's attention at the time.
    5. T F Urban fashion cannot be dissociated from music.

C. Answer the questions.

    1. How could hippies affect the fashion?
    …………………………………………………………………………………………………
    2. Which generations played an important role in this new vision of fashion, and how?
    …………………………………………………………………………………………………
    3. What does Gap sell and how are its prices?
    …………………………………………………………………………………………………
    4. What do Nu-metal fans wear? What don't they wear?
    …………………………………………………………………………………………………
    5. What is the writer's main purpose in writing this text?
    …………………………………………………………………………………………………

APPENDIX C

Reading-Test Item Types (Grabe, 2009)

1. Cloze formats
2. Gap-filling formats
3. MCQformats
4. Sentence completions
5. Matching techniques
6. Classification into groups
7. Text-segment ordering
8. Dichotomous items (True/False, Yes/No)
9. Editing formats
10. C-tests
11. Cloze-elide formats
12. Text-gap formats
13. Short-answer formats
14. Free-recall formats
15. Summary formats
16. Information-transfer formats
17. Choosing from a "heading bank" for identified paragraphs
18. Portfolios
19. Project performance
20. Informal assessment methods