Scientific Research

The Effect of Task Characteristics on IELTS Reading Performance

Behnam Behfrouz, Elham Nahvi

Faculty of Persian Literature and Foreign Languages, Islamic Azad University, Tehran, Iran
Email: Behnambehfrouz@hotmail.com, E.nahvi@yahoo.com

One of the skills that involve thinking processes is reading. Reading is a skill which allows students to become familiar with other people's ideas; compare and contrast different ideas; examine and evaluate arguments, interpretations, beliefs, or theories; make inferences, predictions, or interpretations; and explore implications and consequences. In essence, reading is a skill that compels students to think critically. Thus, the present study aims to investigate the effects of task characteristics on the learners' performance in the reading section of IELTS. To accomplish the purpose of the research, a sample of 50 male and female students participated in this study. In order to ensure the homogeneity of the participants and determine the participants' reading proficiency level, a reading part of the IELTS test was conducted at the outset of the study. Based upon the results, 30 participants were selected to be included within this study. Next, five IELTS tests were given to the participants. The subsequent data was then analyzed using a one-way ANOVA and the Sheffe test. Data revealed that participants performed differently on the IELTS test due to different task types given as a treatment. The results suggest that task characteristics have a significant effect on IELTS reading test performance.

Keywords: Task Characteristics; Performance; Reading Section; IELTS

Introduction

Reading materials have been considered to be one of the primary sources of acquiring comprehensible input and therefore reading has always been a significant aspect of language learning (Chastain, 1988). Krashen (1988) believes readers recreate the message while reading. According to Rumelhart (1977; cited in Chastain, 1988), the reader, the text, and the interaction between the reader and the text are involved in reading. In this interactive process of reading, meaningfulness plays a very important role in better comprehension of the texts. From the 1960s most cognitive psychologists have focused on the role of meaningful learning and the organization of background knowledge.

Reading comprehension skill has been the home of choice for language testers and has attracted a lot of attention as a prime source of input for the ELT researchers and teachers (Amiri & Maftoon, 2010). Reading is the most available and foremost source of information and necessary input for the EFL students (Bachman & Palmer, 2010) and mostly second or foreign language learners find it difficult to comprehend the texts by reading them just once or without looking up the words in a dictionary. IELTS examinees' performance in the reading section could be affected by variety of factors, one of which is "task characteristics". The characteristics of the used items and the nature of the test in general, affect learners' answers. The present study is an attempt to focusing on the specific assessment tasks employed in the reading section of IELTS and their possible impact on the examinees' (learners') performance.

Theoretical Background

Reading

One of the skills that involve thinking processes is reading. Reading is a skill which makes students familiar with other people's ideas; compare and contrast different ideas; examine and evaluate arguments, interpretations, beliefs, or theories; make inferences, predictions, or interpretations; explore implications and consequences, and in short it is a skill which makes students to think critically. According to Hudelson (1994) cited in Edigar (2001: p. 154), in reading "an individual constructs meaning through a transaction with written text that has been created by symbols that represent language. The transaction involves the reader's acting on or interpreting the text and the interpretation is influenced by reader's past experiences, language background, and cultural framework, as well as the reader's purpose for reading".

There are different definitions for reading comprehension. However, the nature of the process of reading is not known exactly. Some of the common definitions are as follow; in very simple terms reading is "perceiving a text to understand its contents" (Richards & Schmidt, 2002: p. 443).

According to Chastain (1988), "reading is a process involving the activation of relevant knowledge and related language skills to accomplish an exchange of information from one person to another. Reading requires that the reader focus attention on the reading materials and integrate previously acquired knowledge and skills to comprehend what someone else has written" (p. 216).

Also Bowen et al. (1985) claimed that reading is a problem

solving activity that involves the reader in the process of deriving the meaning from the text. According to Carrell (1991) reading comprehension is a complex behavior, which involves conscious and unconscious use of various strategies in order to get the intention of the writer. He claimed that they decode meaning, respond to what they read and select for particular aspect for consideration.

Reading is the activation of related knowledge and language skills to perform an exchange of information from one person to another. Reading is a receptive skill; it means that the reader is receiving a message from a writer. Reading is a complementary skill for language learning. Second language learners are required to learn to read for communication and to read greater number of authentic materials (Chastain, 1988). Nuttall (1996) considered reading as convey of meaning from mind to mind; the movement of message from writer to reader.

Reading is a language process. Students should be assisted to answer the visual symbols which depict the same auditory signals to which they had answered before (Finocchiaro & Bonomo, 1973).

Barnett, cited in Hadley (2003: p. 163) listed reasons for including reading skill development as an important part of second language curriculum: 1) reading is still essential in the teaching of literature, which remains an important goal in many programs; 2) it is a skill that can be maintained after students complete formal language study; 3) it fosters the development and refinement of literary skills.

There are various theories to explain what is involved when we read. Current research shows that reading is an interactive, socio-cognitive process which involves a text, a reader and a social context within which the activity of reading takes place (Bernhardt, 1991, cited in Edigar, 2001).

Widdowson (1979) cited in Grabe (1988) stated that reading is the process in which textual information is combined with the information a reader brings to a text. In this view the reading process is not simply a matter of extracting information from the text. Rather, it is the process in which a range of knowledge in the reader's mind is activated and this knowledge may be influenced by the new knowledge. Grabe (1991) cited in Edigar (2001: p. 154) listed six general component skills and knowledge areas, within the complex process of reading:

1) Automatic recognition skills—a virtually unconscious ability, ideally requiring little mental processing to recognize text, especially for word identification;

2) Vocabulary and structural knowledge—a sound understanding of language structure and a large recognition vocabulary;

3) Formal discourse structure knowledge—an understanding of how texts are organized and how information is put together into various genres of text (e.g., a report, a letter, a narrative);

4) Content/world background knowledge—prior knowledge of text-related information and a shared understanding of the cultural information involved in text synthesis and evaluation skills/strategies—the ability to read and compare information from multiple sources, to think critically about what one;

5) Read, and to decide what information is relevant or useful for one's purpose;

6) Meta-cognitive knowledge and skills monitoring—an awareness of one's mental processes and the ability to reflect on what one is doing and the strategies one is employing while reading.

He said that when a fluent reader reads, he brings together all

of these components into a complex process.

Reading Purposes

Rivers and Temperly cited in Nunan (1999: p. 251) suggested seven main purposes for reading:

1) To obtain information about some topic;

2) To obtain instructions about doing some task;

3) To act in a play, play a game, do a puzzle;

4) To keep in touch with friends by correspondence or to understand business letters;

5) To know when or where something will happen or what is available;

6) To know what is happening or has happened;

7) For enjoyment or excitement.

Chastain (1988) stated that the ultimate reading goal must include other factors that convert a laborious problem solving task into a viable skill. Language learners have to learn in order to dealing with linguistic materials over which they have no control. They must learn to interact with the reading as a productive way to determine meaning even when some of the word endings and forms are not meaningful.

Task Characterization and Reading Comprehension

A lot of researchers agreed that reading is rapid, purposeful, and interactive (Grabe, 1999; Alderson, 2000). At first level, the different components and the levels of reading passage might interact with each other. Second, reading is interactive in that the reader's background knowledge and other attributes interact with the content of the text.

With regard to the fact that reading involves interaction between readers and texts, it can be concluded that characteristics of both, the reader and the text, can affect the reading process (Alderson, 2000).

According to Bachman and Palmer (1996, 2010) task characteristics can be synonymous with test method facets. Therefore, the characteristics evaluation of the reading comprehension tests may be the optimal means for the assessment of the learners' task performance.

Of greater relevance here is the point that text characteristics are describable in a systematic manner. Bachman and Palmer (1996) referred to tasks as the activities which can be done in proper settings for gaining specific objectives. With reference to their definition of the task, they concluded that tasks are associated with specific situations and the active participation of the language users is goal-oriented. In other words, the participants perform a specific task in a particular setting in order to achieve a specific goal. This definition of language use tasks entails both the specific activity and the setting in which it takes place. Bachman (1990), and later modified in Bachman and Palmer (1996), presented a model of language ability which draws attention to a range of factors which can affect test performance. Bachman and Palmer (1996) also posited the importance of method facets, which they now term "task characteristics" as follows:

Language use involves complex and multiple interactions among the various individual characteristics of language users, on the one hand, and between these characteristics and the characteristics of the language use or testing situation, on the other. Because of the complexity of these interactions, we believe that language ability must be con-

sidered within an interactional framework of language use (p. 62).

Bachman and Palmer (1996) test task characteristics model provides a means for systematically describing various characteristics of tests and testing procedures. Bachman and Palmer (2010) described the characteristics of the setting for tasks and discussed that "the setting comprises the circumstances under which either language use or testing takes place" (p. 68).

The characteristics of the setting based on what Bachman and Palmer (2010) presented, included the physical characteristics, the participants, and the time of task.

Bachman (1990) classified test method facets into five categories: 1) testing environment; 2) test rubrics; 3) the nature of the input; 4) the nature of the expected response; and 5) the interaction between the input and the response. According to Bachman (1990), these factors can affect test performance; it is important for testers to be aware of their influences and, if possible, to minimize them.

Based on the concept of task characteristics, Carr (2006) used a new methodological approach to describe variation in test task characteristics and explored how differences in these characteristics might relate to examinees' performance. He constructed an expanded test task characteristics instrument and added a large number of syntactic features to this instrument (Celce-Murcia & Larsen-Freeman, 1999). Ratings and numerical counts were gathered for three forms of the Test of English as a Foreign Language (TOEFL) Reading Comprehension Section. The items were the objects of the measurement and based on them, the results were used in a series of exploratory and confirmatory factor analyses, together with IRT (item response theory) parameter estimates for the items in question.

The most important finding of this study was that key sentence variables were the most important determining factors in the measures of difficulty of reading items. Key sentence variables accounted for 15 of the 20 variables in the final task characteristics model. Carr (2006) asserted that the findings are particularly significant because other studies have generally ignored key sentences as a unit of analysis.

Bachman and Palmer's model (1996) is initially intended as a tool for designing and constructing language tests, but because of the detailed descriptions of tasks, the model also provides a useful research model for empirical studies and other research. A number of language testing studies have made use of this model (Clapham, 1996) or other systems of describing test task characteristics (Freedle & Kostin, 1993).

Task Characteristics and Learners' Performance

The characteristics of tasks could be taken into account as highly influential factors affecting the learners' performance on the test items. Task characteristics, reading comprehension nature, and task difficulty are among the concepts which require clarification before beginning to conduct a research on the effect of task characteristics on the performance of learners or examinees on the tests. IELTS as the high-stake proficiency test taken into consideration in the present study also requires some description.

Bachman (1990) provides a framework of task characteristics. This framework includes a set of features that describe five aspects of tasks: setting, test rubrics, input, expected response, and relationship between input and response.

Characteristics of the setting refer to all physical conditions under which testing takes place. This includes the physical setting, participants, and the time of task. The second aspect to mention here is called the characteristics of the test rubric. Test rubric refers to those features that show how the test takers should proceed during the test to accomplish the tasks. The characteristic of rubric include: the organization (structure) of the test, instructions, the duration of the test as a whole and of the individual parts, and how the language that is used is evaluated and scored. In a test task the instructions should be as explicit and clear as possible because students should know exactly what to do and just in this way the teachers can make safe decisions based on their performance on the test. Structure of the test task shows how the different part of the test are put together and presented to the test takers. This aspect deals with the number of parts or tasks, the salience of the parts/tasks which shows how the different parts are clearly distinguished from each other, the order or sequence of parts/tasks, relative importance of tasks, and the number of items in each part. Time allotment is the amount of time for each part of test and the entire test. Scoring method specifies how numbers will be assigned to test takers' performance.

The third aspect of test task characteristics which is utilized in this study is characteristic of the input. Input consists of the material contained in a given test task which the test takers are going to process in some way and to which they are expected to respond. Input can be studied from different aspects namely format and language. By format Palmer means the way in which the input is presented. Format includes channel, form, language, length, type, degree of speediness, and vehicle. Channel determines the way of presenting the input which can be aural, visual, or both. Also input can take the form of language, non-language, or both. If the form is language, it can be the testees' native language or target language. In terms of length, input can be just a single word, a phrase, a sentence, a paragraph, or an extended discourse. Type of input may be either an item or a prompt. The purpose of an item is to elicit either a selected or a limited response. An example of a test item is the familiar multiple-choice question. The purpose of a prompt is to elicit an extended production response. A test task as summary writing contains input of this type. Degree of speediness refers to the rate at which the test taker has to process the information. Vehicle is the means by which input is delivered which can be live or reproduced.

Characteristics of the expected response are another aspect of test task features. Here again two factors are important: format and type of response. The explanation of format is the same as what was said about input. Genesee and Upshur (1996) state that when selecting language tests, the response characteristics or demands of the test task should be taken into consideration. They state that based on the response characteristics, test methods can be described as closed-ended, limited, and open ended. For describing the type of response other measurement specialists use three categories: selected, limited production, and extended production. In a selected response no production is needed. The test taker just chooses or selects one response from among two or more that are provided. A typical example of selected response is a multiple-choice test task. A limited production response includes a single word or phrase, and may be as long as a single sentence or utterance. In short completion items and cloze test this type of response is required. If the test taker has to provide a response that is longer than a single

sentence and its range is from two sentences to a whole composition, then the response is named an extended production one. The summary writing test task used in this study requires an extended response and the test takers are to provide a summary of about 100 words. Other features relating to response are the degree of speediness and the language of expected response.

Task characteristics model presented by Bachman and Palmer (1996) and modified as test task characteristics (Bachman and Palmer, 2010), was developed based on "test method facets model" presented by Bachman (1990). This notion provides a means for systematically describing various characteristics of tests and testing procedures. It is primarily intended as a tool for designing and constructing language tests, but because of the detailed descriptions of tasks, the model also provides a useful research paradigm for empirical validation studies and other research. A number of language testing studies have made use of this model (e.g. Bachman et al., 1996; Clapham, 1996) or other systems of describing test task characteristics (e.g. Brown, 1989; Freedle & Kostin, 1993; Perkins et al., 1995).

Most researchers agree that reading is rapid, purposeful, and interactive (Alderson, 2000). Specifically, this interactivity occurs at two levels. First, regardless of the exact components or levels that they posit, models of the reading process depict these components or levels as interacting with each other. Second, reading is interactive in that the reader's background knowledge and other attributes interact with the content of the text. Given that reading involves interaction between readers and texts, it logically follows that characteristics of both the reader and the text will affect the reading process (Alderson, 2000). Reader characteristics have been examined in previous research, in terms of both readers' background characteristics and their abilities (Bachman, 1990). Of greater relevance here is the point that text characteristics are describable in a systematic manner, using frameworks such as the Bachman and Palmer (1996) model. This has been demonstrated in a number of previous studies, some of which have also examined interactions between these characteristics and examinee performance as indicated by item difficulty and/or discrimination (e.g. Bachman et al., 1996; Clapham, 1996).

Method

Participants

To accomplish the purpose of the research, a sample of 50 male and female students participated in this study. The subjects were between 19 and 28 years old. These participants were studying English as a foreign language. After giving a reading part of the IELTS test, 30 students whose scores were between one standard deviation above and below the mean were chosen as the subjects of this study. They were undergraduate students from Azad universities of Tehran, the major field of study was English translation and English language teaching.

Instruments

First of all, a proficiency test on the basis of IELTS reading section was held in order to homogenize the participants' levels of reading proficiency.

Five original tests of IELTS were utilized to achieve the purpose of this study. These tests were administered in 2009 and 2011 (Cambridge IELTS 8, 2011; Cambridge IELTS 8,

2007). IELTS is a standard test which is provided to measure the students English proficiency level. It consisted of four sections: 1) listening (40 items); 2) reading comprehension (40 items); 3) writing; and 4) speaking. The total time for this test was 90 minutes. Since listening, writing and speaking sections, not being the focus of the present study, were deliberately omitted for practical reasons.

Procedures

To collect appropriate data for this study several steps were taken. In order to ensure the homogeneity of the participants and determine the participants' reading proficiency level a reading part of IELTS test was conducted at the outset of the study. Those participants whose scores were one standard deviation above and below the mean were selected as a homogenized group.

After homogenizing the students 30 of them were selected to be included in this study. During five sessions, five IELTS tests were taken from the participants. The purpose of this study was to compare the performance of the participants on different tasks of the IELTS.

Three kinds of tasks out of twelve different tasks were selected and included in the study: 1) Matching heading; 2) True/false/not given; 3) Multiple-choice. These tasks were separated and all the participants were scored based on their tasks. The statistics of the distribution of different tasks in the tests are as follows:

Task	Number of Questions
Matching heading	32
True/false/not given	27
Multiple-choice	23

Data Analysis

In order to test the Research Hypothesis, the following statistical techniques were utilized:

1) Descriptive statistics and Standard deviation were utilized to homogenize the participants based on their reading proficiency.

2) A one-way ANOVA was run to compare the mean scores of the groups, to see the effect of the task characteristics on the IELTS tests.

3) Multiple-comparisons were conducted using Sheffe test to distinguish the exact difference of the groups on the IELTS tests.

The purpose of the present study was to determine the effect of task characteristics on the learners' performance in the reading section of IELTS. Considering the purpose and based on the problem specified above, the following research question was raised:

Do task characteristics affect the learners' performance in the reading section of IELTS?

Considering this question the following null hypothesis was formulated:

H_0: Task characteristics DO NOT affect the learners' performance in the reading section of IELTS?

To test the above hypothesis, the researcher conducted a series of statistical procedures which are elaborated in the following parts of this section.

In order to validate the IELTS test with Iranian students, the test was administered to 30 subjects the same as our sample. The reliability coefficient of this test is in **Table 1**.

To homogenize the group, they were pre-tested through the reading section of IELTS. The descriptive statistics for the pretest is presented in **Table 2**. The mean score is 25.44, and the SD is 3.27.

Figure 1 presents the bar graph of the scores with normal curve in pretest.

After pretest the group took five different IELTS reading tests. The intended task types were separated and scored differently. The results were calculated based on the different task types. The statistics of the raw scores of task types are as follows.

The descriptive statistics for the raw scores of matching tasks is presented in **Table 3** the mean score for the matching task is 24.16, and the SD is 2.64.

The bar graph of **Figure 2** represents the scores with normal curve in matching task.

The descriptive statistics for the raw scores of T/F/NG tasks is presented in **Table 4** the mean score for the matching task is 21, and the SD is 2.62.

The following bar graph, i.e. **Figure 3** deals with the raw scores of participants in T/F/NG task.

The descriptive statistics for the raw scores of multiple-choice task is presented in **Table 5** the mean score for the multiple-choice task is 15.83, and the SD is 3.16.

In **Figure 4**, some information about the raw scores of M/C task has been stated.

After selecting task types they were scored. Since the number of items for each task is not the same, all the scores were converted to percentage.

The descriptive statistics for the percentage of the scores of matching task is presented in **Table 6** the mean score for the matching task is 75.52, and the SD is 8.25.

The following **Figure 5** is about the percentage of scores with normal curve in matching task.

The descriptive statistics for the percentage of scores of T/F/NG task is presented in **Table 7** the mean score for the T/F/NG task is 77.77, and the SD is 9.72.

Figure 6 is the bar graph of the percentage of the scores with normal curve of T/F/NG task.

The descriptive statistics for the percentage of the scores of multiple-choice task is presented in **Table 8** the mean score for the multiple-choice task is 68.84, and the SD is 13.75.

Figure 7 also stated the percentage of the scores in M/C task as follow:

A one-way ANOVA was run to compare the mean scores of the percentage of the scores. The descriptive statistics for one-way ANOVA is presented in **Table 9**. The mean scores for the matching, T/F/NG and multiple-choice tasks are 75.52, 77.77, and 68.84 respectively. As it can be figured out the participants did better in T/F/NG then matching and multiple-choice.

As displayed in **Table 10**, the F-observed value is 5.52. This amount of F-value at 2 and 87 degrees of freedom is higher than the critical value of F, i.e. 3.11. As it is shown the probability level is lower than .05 i.e. .006.

Therefore it can be concluded that task type has a significant effect on the participants' reading skill. So our null hypothesis is rejected. We can conclude that task type has a significant effect on the Iranian EFL learners' reading skill.

Table 1.
Reliability coefficient of the pretest.

	N	Minimum	Maximum	Mean	Std. Deviation	Reliability Coefficient
IELTS	30	25	35	30.46	4.77	.83

Table 2.
Descriptive statistics of the pretest.

		Valid	50
N		Missing	0
Mean			**25.44**
Std. Error of Mean			.46
Median			25
Mode			25
Std. Deviation			**3.27**
Variance			10.7
Skewness			.39
Std. Error of Skewness			.33
Kurtosis			−.05
Std. Error of Kurtosis			.66
Range			13
Minimum			20
Maximum			33
Sum			1272
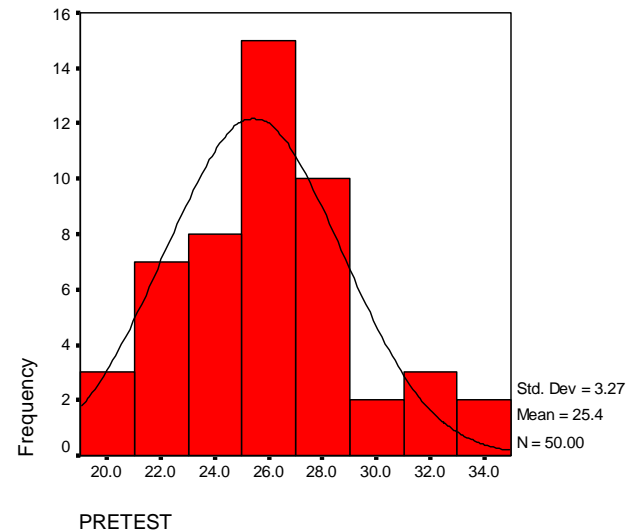
Figure 1.
Bar graph of the pretest.

In order to distinguish the difference among the performance of the group on different task types, multiple comparisons are conducted using Sheffe test. As it is represented on **Table 11** the main difference is between multiple-choice group and the other groups.

Discussion

As mentioned previously, the present study aimed at providing answers to the addressed research question:

"Do task characteristics affect the learners' performance in

Table 3.
Descriptive statistics of raw scores of matching.

N	Valid	30
	Missing	20
Mean		24.16
Std. Error of Mean		.48
Median		24
Mode		25
Std. Deviation		2.64
Variance		6.97
Skewness		.1
Std. Error of Skewness		.42
Kurtosis		−.13
Std. Error of Kurtosis		.83
Range		11
Minimum		19
Maximum		30
Sum		725

Table 4.
Descriptive statistics of raw scores of T/F/NG.

N	Valid	30
	Missing	62
Mean		21
Std. Error of Mean		.47
Median		21
Mode		21
Std. Deviation		2.62
Variance		6.89
Skewness		−.18
Std. Error of Skewness		.42
Kurtosis		−.94
Std. Error of Kurtosis		.83
Range		9
Minimum		16
Maximum		25
Sum		630
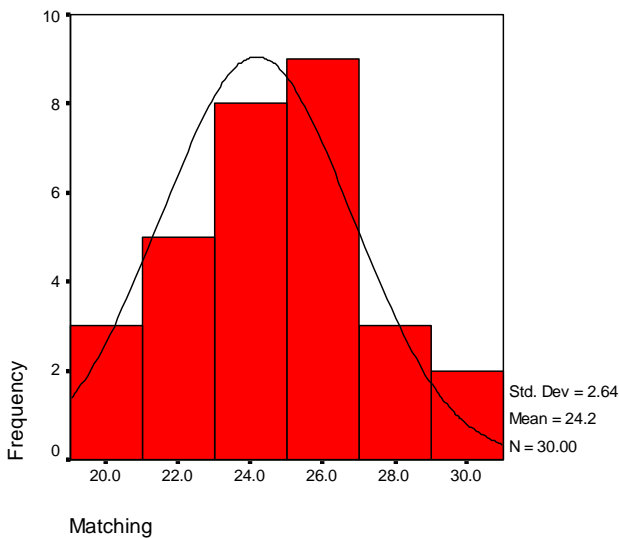
Figure 2.
Bar graph of the raw scores of matching.
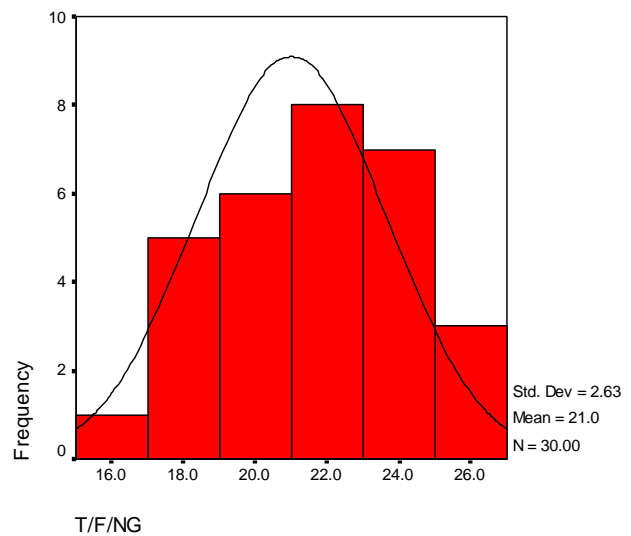
Figure 3.
Bar graph of the raw scores of T/F/NG.

the reading section of IELTS?"

The comparison made between matching and T/F/NG had revealed that there is not a significant difference between the performances of the participants on the test in these two groups, the probability level is .723. And the comparison made between matching and multiple-choice showed that the performance of the participants on the test is not significant too, the probability level is .063. The difference between the performance of the participants of the test, in T/F/NG and multiple-choice is significant i.e. .008.

As the results showed, task characteristics have a significant effect on the learners' performance in the IELTS reading tests. All teachers and test developers ought to pay attention to test characteristics especially those which were utilized in this study, i.e. the degree of being communicative, and the difference in

expected response. Teachers should be aware of task characteristics and try to use the test which best suit the students' needs.

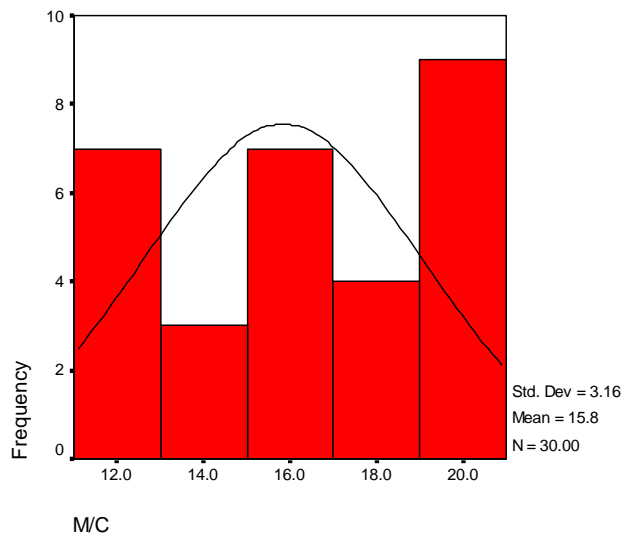The participants did better in T/F/NG then matching and multiple-choice.

The participants on the T/F/NG and matching did better than the multiple-choice group. It can be inferred that multiple-choice task is a kind of traditional test that students need more practice in mastering the strategies to answer.

Conclusion

Based on the results, it can be concluded that the participants performed differently on the IELTS test due to the different task types utilized as the treatment; therefore, the null hypothesis of the study is rejected.

Table 5.
Descriptive statistics of raw scores of M/C.

N	Valid		30
	Missing		62
Mean			15.83
Std. Error of Mean			.57
Median			16
Mode			20
Std. Deviation			3.16
Variance			10
Skewness			−.12
Std. Error of Skewness			.42
Kurtosis			−1.38
Std. Error of Kurtosis			.83
Range			9
Minimum			11
Maximum			20
Sum			475

Table 6.
Descriptive statistics of the percentage of the scores of matching.

N	Valid		30
	Missing		62
Mean			75.52
Std. Error of Mean			1.5
Median			75
Mode			78.13
Std. Deviation			8.25
Variance			68.07
Skewness			.1
Std. Error of Skewness			.42
Kurtosis			−.13
Std. Error of Kurtosis			.83
Range			34.38
Minimum			59.38
Maximum			93.75
Sum			2265.63

Figure 4.
Bar graph of the raw scores of M/C.

Figure 5.
Bar graph of the percentage of scores of matching.

"*Task characteristics affect the learners' performance in the reading section of IELTS*?"
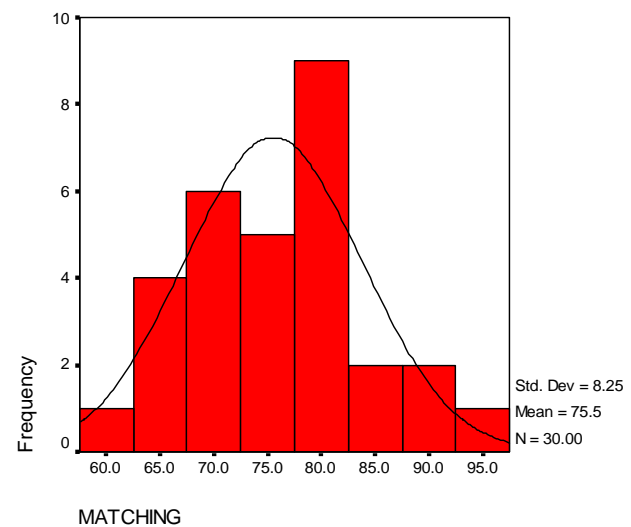
Implications and Applications

The findings confirm the previous research on the effect of task characteristics on the reading ability.

These findings have some implications and applications:

Testing characteristics have a significant effect on the performance of the learners on the test. Task types should be in accordance to the class procedures and the specified goals. In a well-organized curriculum all the teaching materials and procedures should be predefined and based on the goals of the course the appropriate task type should be selected. In choosing a suitable test, some of the characteristics of tasks which are setting, input and expected response should be considered.

The results of the current study can be of use in all educational centers. It has direct or indirect implications and applications in teaching, learning, test development, syllabus design, and material development. Considering the important decisions which are made based upon the tests and their possible influence on students' fates, there is no doubt that some task types which are formative, motivating, and anxiety-reducing methods, should be incorporated into the syllabuses and lesson plans of the educational centers. Since in recent years, the focus has shifted from the products of instruction to the processes, the type of task used in the classes should be considered. Hence, considering the instructional as well as evaluation value of task and their positive effects, course designers should plan for their proper use in instruction.

In our own experience in test development, we have found that task characteristics are useful for describing the character-

Table 7.
Descriptive statistics of the percentage of scores of T/F/NG.

N	Valid	30
	Missing	62
Mean		77.77
Std. Error of Mean		1.77
Median		77.77
Mode		77.78
Std. Deviation		9.72
Variance		94.6
Skewness		−.18
Std. Error of Skewness		.42
Kurtosis		−.94
Std. Error of Kurtosis		.83
Range		33.33
Minimum		59.26
Maximum		92.59
Sum		2333.33

Table 8.
Descriptive statistics of the percentage of the scores of M/C.

N	Valid	30
	Missing	62
Mean		68.84
Std. Error of Mean		2.51
Median		69.56
Mode		86.96
Std. Deviation		13.75
Variance		189.14
Skewness		−.12
Std. Error of Skewness		.42
Kurtosis		−1.38
Std. Error of Kurtosis		.83
Range		39.13
Minimum		47.83
Maximum		86.96
Sum		2065.22

Std. Dev = 9.73
Mean = 77.8
N = 30.00
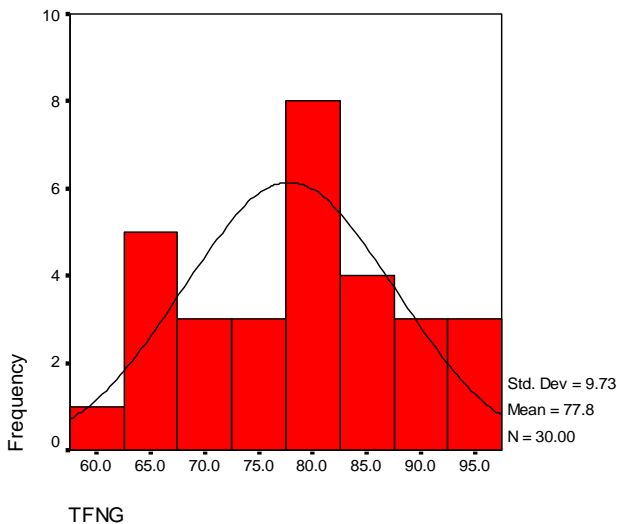
Figure 6.
Bar graph of the percentage of the scores of T/F/NG.

Std. Dev = 13.75
Mean = 68.8
N = 30.00
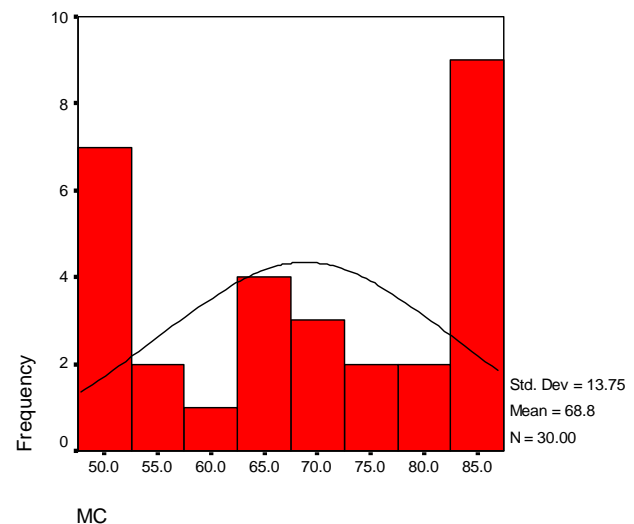
Figure 7.
Bar graph of the percentage of the scores of M/C.

istics of task types. Language teachers intuitively realize that the types of tasks that are included in language tests are important. Frequently one of the first questions asked in our testing classes is about our opinions of the best way to test a particular area of language ability. These teachers may not yet have refined their thinking as to what specific characteristics might make one test task more appropriate for a given purpose than another. However, they are clearly aware that the way they test language ability affects how their students perform on language tests and hence the quality of the information obtained from their tests.

There is also considerable research in language testing that demonstrates the effects of task type on test performance. This research and language teachers' intuitions both lead to the same conclusion: the characteristics of the tasks used are always likely to affect test scores to some degree, so that there is virtu-

ally no test that yields only information about the ability we want to measure. The implication of this conclusion for the design, development, and use of language tests is equally clear: since we cannot totally eliminate the effects of task characteristics, we must learn to understand them and to control them so as to insure that the tests we use will have the qualities we desire and are appropriate for the uses for which they are intended.

When we think about the different types of tasks that are commonly used for language tests, we realize that they are not single wholes, but rather collections of characteristics. Consider, for example, the multiple-choice item, which has often been used as a kind of test task: multiple-choice test items vary in a number of ways, such as in their length, syntactic complexity, level of vocabulary, topical content, and type of response required, to name but a few. Similarly, the "composition" task

Table 9.
Descriptive statistics of one-way ANOVA.

	N	Mean	Std. Deviation	Std. Error	Minimum	Maximum
Matching	30	**75.52**	**8.25**	1.5	59.38	93.75
T/F/NG	30	**77.77**	**9.72**	1.77	59.26	92.59
Multiple-choice	30	**68.84**	**13.75**	2.51	47.83	86.96
Total	90	**74.04**	**11.36**	1.19	47.83	93.75

Table 10.
One-way ANOVA.

	Sum of Squares	Df	Mean Square	F Critical	F Observed	Sig.
Between Groups	1295.93	2	647.96	**3.11**	**5.52**	**.006**
Within Groups	10202.96	87	117.27			
Total	11498.89	89				

Table 11.
Sheffe test.

(I) GROUPS	(J) GROUPS	Mean Difference (I-J)	Std. Error	Sig.
Matching	T/F/NG	−2.25	2.79	**.723**
Matching	Multiple-choice	6.68	2.79	**.063**
T/F/NG	Multiple-choice	8.93(*)	2.79	**.008**

Note: *The mean difference is significant at the .05 level.

type encompasses a wide variety of prompts that can differ in characteristics such as the intended audience, purpose, and specific organizational pattern requested. It is thus clear that we cannot characterize test tasks precisely if we think of them only as holistic types. Therefore we need a descriptive framework of task characteristics.

This framework also enables test developers to better understand which specific characteristics can be varied, and to suggest ways in which these can be varied, thus providing a valuable tool for tailoring tests appropriately for specific groups of test takers, and so enabling them to perform at their best.

The researcher therefore believes that attempting to control the test task characteristics by design provides the most useful and practical means for maximizing the usefulness of our tests for their intended purposes.

Suggestions for Further Research

The results of the study can help teachers, administrators, test task writers, curriculum designers and EFL textbook/material writers in decision making in the classroom teaching, test development, curriculum design, and textbook writing. The study is intended to serve as a guide to orient teachers to task selection.

The researcher hopes that the results are actively and carefully employed by teachers, and teachers and test task writers benefit from the results when they choose activities, and it also gives them insight to what goes on in the learning process, thus shedding much needed light on some of the dark areas created in the classroom. It is also very practical in the sense that by showing the different phases of the process, it gives teachers and test task writers clues to precisely what the difficulties are, and, indirectly, what they can do to help students overcome them and it also may assist test task writers and curriculum designers when they are trying to gauge the level and impact of

a task and to determine just what difficult or easy means. Examining task types helps educators get a better idea of how students approach and complete tasks and what they learn by doing so.

In conducting any research project, there exist a number of limitations which can be overcome in future studies. It is suggested that the present study can be replicated with various methods and techniques of testing, larger samples, different levels of language proficiency and other language skills. The followings are some further suggestions:

1) Is there any relationship between consciousness-raising and task type?

2) Is there any relationship between task characteristics and students' personality character?

3) Do task types show the exact weaknesses of the learners' reading comprehension skill?

4) Is there any relationship between task type and learners' learning styles?

REFERENCES

Alderson, J. C. (2000). *Assessing reading*. Cambridge: Cambridge University Press. doi:10.1017/CBO9780511732935

Amiri, M., & Maftoon, P. (2010). Awareness of reading strategies among Iranian high school students. In *EDULEARN10 Proceedings CD* (pp. 6782-6791). Valencia: IATED.

Bachman, L. F. (1990). *Fundamental considerations in language testing*. Oxford: Oxford University Press.

Bachman, L. F., & Palmer, A. S. (1996). *Languag testing in practice*. Oxford: Oxford University Press.

Bachman, L. F., & Palmer, A. S. (2010). *Language assessment in practice*. Oxford: Oxford University Press.

Bowen, J. D., Maadsen, H., & Hilferty, A. (1985). *TESOL techniques and procedures*. Rowley: Newbury House.

Carr, N. T. (2006). The factor structure of test task characteristics and examinee performance. *Language Testing, 23,* 269-289.

doi:10.1191/0265532206lt328oa

Carrell, L. P. (1991). Second language reading: Reading ability or language proficiency? *Applied Linguistics, 9,* 233-242.

Celce-Murcia, M., & Larsen-Freeman, D. (1999). *The grammar book: An ESL/EFL teacher's course*. Boston, MA: Heinleand Heinle.

Chastain, K. (1988). *Developing second-language skills*: *Theory and practice* (3rd ed.). New York: Harcourt Brace Jovanovich.

Clapham, C. (1996). *The development of IELTS*: *A study of the effect of background knowledge on reading comprehension*. Cambridge: University of Cambridge Local Examinations Syndicate.

Edigar, A. (2001). Teaching children literacy skills in a second language. In M. Celce-Murcia (Ed.), *Teaching English as a second or foreign language* (pp. 153-169). Boston: Heinle & Heinle.

Finocchiaro, M., & Bonomo, M. (1973). *The foreign language learner*: *A guide for teachers*. NewYork: Regents.

Freedle, R., & Kostin, I. (1993). The prediction of TOEFL reading item difficulty: Implications for constructvalidity. *Language Testing, 10,* 133-170.

Grabe, W. (1988). What every EFL teacher should know about the reading in English. *TESOL Journal, 31,* 130-148.

Grabe, W. (1999). Developments in reading research and their implications for computer-adaptive reading assessment. In M. Chalhoub-Deville (Ed.), *Issues in computer-adaptive testing of reading proficiency* (pp. 11-48). Cambridge: University of Cambridge Local Examinations Syndicate.

Hadley, A. O. (2003). Teaching language in context (3rd ed.). Boston: Heinle & Heinle.

Nunan, D. (1999). *Second language teaching & learning*. Boston, MA: Heinle & Heinle.

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

Perkins, K., Gupta, L., & Tammana, R. (1995). Predicting item difficulty in a reading comprehension test with an artificial neural network. *Language Testing, 12,* 34-53.
doi:10.1177/026553229501200103

Richards, C. J., & Schmidt, R. (2002). *Longman dictionary of language teaching & applied linguistics*. London: Pearson Education.

Krashen, D. S. (1988). *Second language acquisition and second language learning*. Englewood Cliffs, NJ: Prentice-Hall International.

Genesee, F., & Upshur, J. A. (1996). *Classroom-based evaluation in second language education*. Cambridge: Cambridge University Press.