The Effect of the Reader's Background on Reading Comprehension Performance

Bushra Saadoon Mohammed Al-Noori

College of Education Ibn-Rushd for Human Sciences English Department, Iraq

E-mail: alnoori.bushra@yahoo.com

Abstract

This study aims at measuring the amount of the effect of the reader's background knowledge on performance in reading comprehension tests through the assessment of information gained in reading comprehension (RC) tests across-four testing techniques, short answer questions ,true-false items , multiple - choice items , and cloze test and re-test. This technique involves the examinees in two types of tasks, i.e. pre-reading and post -reading task. Two hypotheses have been proposed to achieve the aims of this study. They are 1-There are no significant differences between the pre-reading and post-reading performances of examinees on reading comprehension(RC). 2-There are no significant differences in information gain scores across the different techniques of testing reading comprehension (RC) in EFL. To verify the validity of these two hypotheses, a number of statistical procedures have been used such as arithmetical mean, t-test for correlated and independent samples to analyze the performance of third and fourth year College students studying at the Department of English at University of Baghdad /Ibn Rushd College of Education for Humanities on two reading passages taken from TOEFL practice tests (2011). The analysis of the data has shown the following results: 1-The background knowledge has an effect on the performance on reading comprehension (RC). 2-There is a significant difference in students' performance on reading comprehension (RC). 3-The effect of background knowledge is investable on reading comprehension (RC) tests, but it can be identified or neutralized. Based on these conclusions, the researcher presented a number of recommendations.

1. Introduction

The problem of this study stems from the assumption that reading comprehension (RC) tests are both "contaminated" and "based" (Johnston, 1984:79). RC tests are contaminated because they do not test pure (RC) ability. They test background knowledge and linguistic competence as well. They are biased because two readers equal in RC ability but differing in background would show different levels of comprehension of the same text. This difference is likely to show up in the assessment of RC ability.

The notion of information gain IG is used as a means of differentiating between the measurement of BK of the world and /or language competence of a reader before and after reading a passage. Tuinman (1974:113) reports that native speakers of English can score at 65% correct on a RC test without reading the passage on which the comprehension questions are based. Such findings pose a challenge to the construct validity of RC testing, particularly in EFL situations where the chief aim of such tests is to measure the EFL learner's ability to read and understand various types of texts in English.

I t is important to note that the application of the notion of IG to investigate construct validity of EFL RC tests has not been considered in any reliable and empirical type of research in Iraq, therefore, the implementation of the notion of IG in an empirical study on the extent of 'bias' in RC tests may help to fill a gap in our knowledge of constructing EFL RC tests in general, and assessing and validating different techniques of testing RC in particular.

The study aims at (A) Assessing IG in EFL RC tests across four testing techniques, i.e.,

1-short-answer questions,2-true/false items ,3- multiple-choice items, and 4- cloze test, and

(B) measuring the extent of construct validity of these four testing techniques through quantifying the effect of BK tests.

This paper assumes that

1-There is no significant differences between the pre-reading and post-reading subjects performances on RC tests.

2-There are no significant differences in IG scores across the different techniques of testing RC in EFL.

2. Related Literature

Reading specialists such as Goodman (1965,1967), Eskey (1971:78), Cziko(1980:67), Rumelhart (1980:102),Silberstein

(1987:111),and Carrell et al (1988:54) argue that reading is an interactive process between previously acquired knowledge and the content of what is read. The researchers believe that this interaction will enable him to hypothesize about the strategies FL readers try to use when they process the texts they read and investigate the role of BK in comprehension.

Despite the fact that there are many theories which explain the way in which readers extract meaning from what they read, most of these theories can be conveniently classified into three major groups: bottom-up, top-down and interactive views of reading,(Cziko,1980:101).

Concerning the reader's BK ,it is hypothesized that the greater background knowledge a reader has of a text's content area ,the better the reader will comprehend that text(Carrell,1987:147).Such views which stem from the schema theory consider the background a determinant factor in the process of text comprehension and in the evaluation of reading comprehension ability.

According to Grabe (1988:77), no one can deny that the concept of Top-down processing has been considered a revolution which has resulted in exploring the reading process and the possibilities of promoting reading instruction in SL and FL situations and that it has led to a resurgent interest in reading as a whole.

Ludo and Jan,(2008:69) studied specific effects of word decoding , vocabulary and listening comprehension abilities on the development of reading comprehension that were longitudinally examined for a representative sample of 2143 Dutch children throughout the elementary school period. An attempt was made to test two theoretical frameworks for the prediction of the development of reading comprehension: the lexical quality hypothesis in which word decoding and vocabulary are assumed to be critical determinants of reading comprehension, and the simple reading view in which reading comprehension is assumed to be the product of word decoding and listening comprehension. The results showed significant progress across grades on all of the predictor and criterion measures. The stability of the measures was also high across time, which shows the individual differences between students to remain across grades. Word decoding exerted a substantial effect on early reading comprehension and a small effect on later sixth grade reading comprehension. The data provide empirical support for the lexical quality hypothesis as they show knowledge of word forms and word meanings (i.e. Vocabulary) to predict the development of reading comprehension. Support for the simple reading view was also found in that word decoding and listening comprehension significantly predicted reading comprehension as well. As a combined structural model with word decoding, vocabulary and listening comprehension as predictors of reading comprehension at first grade. In subsequent grades, vocabulary is still predicting reading comprehension directly whereas listening comprehension shows a reciprocal relationship with vocabulary.

Arthur(2008,92)presented data, theory and applications of the field of discourse processes .This commentary identifies some of the challenges this field has faced in its attempt to understand the mechanisms of discourse comprehension and production to scale up the research to applications. The prevailing models emphasize the interaction between data-driven and conceptually driven processes and the importance of representational constraints in guiding discourse processing. The field embraces a wide range of methodologies that are illustrated in this special issue, including brain imaging, eye tracking, think aloud protocols ,reading times, recall, summarization , and question answering. The task include text comprehension, search through large repositories of electronic texts and learning from advanced multimedia environments. The major claim in the commentary is that researchers need to conduct deeper analyses of the information in the material (texts and tasks) and word knowledge before they can offer defensible generalizations about cognitive processes and real word applications.

Paul (2008:123) investigated the effect of readers' incorrect knowledge on-line comprehension processes during reading of science texts, with an eye towards examining the conditions that encourage revision of such knowledge. He employed computational (Landscape Model) and empirical (think-aloud and reading times) methods to compare comprehension processes by readers with correct and incorrect background knowledge, respectively. Science texts were presented in either regular or refutation versions; prior research using off line methods suggests that refutation versions promote revision in readers with incorrect knowledge. The results of that study indicate that incorrect knowledge systematically influences both type and content of processing. Moreover, simultaneous activation of correct and incorrect conceptions during reading plays an essential role in knowledge revision. The computational simulation show that refutation texts create optimal circumstances for co-activation of the incorrect and correct conceptions and the empirical data show that such a co-activation is associated with inconsistency detection and revision activities by the readers with incorrect knowledge. These findings provide insights in the effects of misconceptions on the on line texts processing and have important implications for the development of methods for achieving revision during reading.

Gi-Pyo (2008:88) compared L2 listening comprehension with L2 reading comprehension in terms of the roles of linguistic knowledge background knowledge, and question types among 168 university students learning English in Korea. The analyses of the data revealed that L2 listeners processed inferential information more easily than factual information, while the reverse was true for L2 readers. In addition, linguistic and background knowledge exerted significant effects on L2 listening comprehension. In L2 reading comprehension, however, linguistic knowledge played a significant role, while background knowledge played only a moderate role. In terms of the interaction among linguistic knowledge, background knowledge and question types, only linguistic and background knowledge in L2 reading comprehension were significant. Third, linguistic knowledge and background knowledge combined explained a total variance of 14%bin L2 listening comprehension and 20% in L2 reading comprehension. These findings show that

L2 listening comprehension and L2 reading comprehension differ from each other and the comprehension of an oral and written text is a more complex process than the interactive process model holds.

Diana (2005:101) describes L2 vocabulary learning outcomes associated with adult L2 reading comprehension processes, thus connecting L2 learning with the complex cognitive and linguistic processes involved in reading. The study aimed to determine whether BK moderated the relationship between passage comprehension and lexical input processing outcomes such as intake and receptive gain and retention of target -word meanings. The primary theoretical contribution concerns the nature of the relationships obtained between passage comprehension and lexical input processing. Overall, the results suggest that as learners become more efficient in engaging in the various processing activities required during L2 reading, they experience greater memory for linguistic elements encountered during reading such as orthographic forms and semantic aspects of new lexical items. The result support connectionist models of L2 reading and conclusions concerning efficiency in lower level text processing (e.g. See Nassaji, 2002; Koda, 2005). Reading is a complex cognitive activity, involving simultaneous linguistic processing such as pattern recognition, letter identification, lexical access, concept activation, syntactic analysis, propositional encoding ,sentence comprehension, intersentence integration, the activation of prior knowledge, information storage, and comprehension monitoring. The prior knowledge that is accessed is largely determined by the quality of the text base constructed during reading, which is affected by the individual's efficiency in carrying out the various text-processing operations listed above.

The finding that background knowledge did not moderate the relationship between comprehension and receptive retention of meaning provides further support for the robust role of text-processing efficiency in bootstrapping processes such as those involved in mapping new linguistic forms to familiar concepts already stored in memory. The finding with regard to intake of the target words corroborates the main findings, demonstrating that efficient text processing skills are also essential to other aspects of lexical input processing such as establishing and retrieving the connections between new linguistic forms and the specific contexts in which they were encountered. The study also expands upon previous reading and research in the area of methodological innovations applied to incidental learning research paradigms, for instance, by measuring intake and episodic memory and the relationships between text processing and language learning outcomes. In addition, through the concurrent investigation of the impact of several factors on lexical input processing , the study sets the stage for more complex modeling of the processing that contribute to L2 development.

George and Fong (2006:51) investigated the effect of vocabulary glossing on recall and vocabulary learning, as well as learners' performance as to glossing. Eighty-five native speakers of English Spanish at the university level participated in this study. Participants read a Spanish text under one of three treatment conditions: no gloss, English glosses, or Spanish glosses. They then were asked to write what they recalled of the passage, translate a list of the glossed vocabulary, and complete a questionnaire. The translation task was repeated four weeks later.

Results showed that glossing did not significantly affect recall for the participants overall, but that those with higher than average proficiently recalled more if they had read a glossed version of the text. Those who had glosses outperformed their peers on the translation task administered immediately after they had read the text. However, this difference disappeared on the retest. Participants expressed preference for glosses, wished that they be located in the margin and favored Spanish glosses if they were comprehensible.

Rumlhart (1977:47) dealt with an important type of the processing, for him the notion of interactive processing (refers to how readers utilize and integrate their graphic, syntactic and semantic knowledge together with contextual information in extracting and constructing learning from written texts.

Silberstein (1987:31) treated the relation of BK to text processing and information recall. He argued that this relation had been studied under schema theory. This theoretical framework emphasizes the role of pre-existing knowledge structures (schema) in providing the reader with the information that helps him understand what is implicit in a text.

According to Singer and Donlan's (1982:38) studied the role of BK in comprehending texts under the rubric of the schema theory. The basic tenet of this theory is that a text, be it spoken or written ,does not by itself carry meaning .For the sake of fall comprehension, it requires interaction between the reader's BK and the content of the text. This BK (schemata) is seen as providing the framework which enables readers not only to make predictions about what they read, but also to assimilate new information provided by the text, Beck (1981:35).Schema, theorists believe that the process of text comprehension is guided by the principle that every input is mapped against some existing schema and all aspects of that schema must be compatible with the input information.

3. Method

3.1 Sample

The sample of this study includes 70 third and fourth-year students and university teachers of English at the Department of English in the College of Education/Ibn Rushd for Humanities, University of Baghdad. It consists of two groups, namely, the teacher-group sample and the student-group sample.

3.2 The Teacher-Group Sample

The sample consists of 13 university teachers of English who have got MA or Ph.D either in linguistic or in English

language Teaching. Each of them has at least 4 years of teaching experience. This sample has been chosen for only one specific purpose, i.e., to verify the hypothesis which is related to the level of the reader's BK on his/her performance on RC tests.

3.3 The Student-Group Sample

The sample comprises third-year 79 students taken from the same department. Random selection of subjects, in this study, secures different language levels and different BK, which makes the sample a true representative of the real population of advanced learners of English.

4. Test

In the selection of passages for RC tests, test designers are required to consider three criteria, i.e., length, difficulty level and subject matter. Length is associated with the standard of examinees and the time allocated for answering the test. Difficulty level refers to the degree of complexity of structures, vocabulary items and style of the passage. All these aspects should not demand a higher proficiency level than that of the examinees. As for the passage, it should be clear, meaningful and interesting to motive the examinee to do his/her best. These three criteria have been strictly followed when the two test passages were selected. The two passages have been selected from the practice tests of Test of English as a Foreign Language (TOEFL).

Before going into the details of item writing, it is important to describe the structure of the test. The test under this study consists of four sets of subtests of four testing techniques, i.e., short answer questions (SAQs),true false items (TFLs),multiple choice items (MCIs) , and cloze test (CT), to be given on a pre -reading and post-reading tasks. The total number of items is 105 distributed as follows:

1-10 items for SAQs

2-10 items for TFIs

3-10 items for MCIs

4-50 items CT/pre-reading task

5-25 items for CT/post-reading task

5. Pilot Study

To get accurate information regarding the study tool, the investigation should be sharp, valid and reliable. Therefore, the aim behind carrying out the pilot study is to make sure that the investigation tools used in this study possess all the above-mentioned dualities of a good investigation tool. The investigation is limited to three techniques only, i.e., SAQs, TFIs, and MCIs. As for the Ct, it is assumed that any CT is a valid and reliable test (Taylor,1976;Oilier 1979;Anderson,1982) other studies have found high correlation between CT and other techniques for RC tests which range from 0.66-0.81 ,Oiler (1979:26).

To investigate content and face validity, reliability and the suitability of test items, the following procedural measures have been taken.

1-The Jury

Having finished the preparation stage, the researcher have given the final version of the test and the table of specifications to a jury of Ten EFL experts who are well-known for their long experience in the field of teaching EFL. The jury members have been asked to decide on: a) Face and content validity of the test, and b) The suitability of the items to test the points and skills assigned to them .All the jury members have decided that the test has face and content validity, and the test items are suitable for testing the points and skills assigned to them.

2-The Try-Out Sample

A sample of twenty subjects has been chosen to try out the test items for difficulty level, discrimination power and reliability. Ten of these samples are third and fourth college students and ten university teachers of English from other universities.

3-Level of Difficulty and Discrimination Power

As a result of administrating the test to the twenty subjects, it has been found that the three sets (i.e., SAQs, TFIs and MCIs) have:

(a) An acceptable level of difficulty for both passages that ranges from 35-55% for SAQs ; 40-60 % for TFIs ;40-55% for MCIs concerning the first passage ;35-50% for SAQs ;35-60 % for TFIs; and 35-60% for MCIs for the second passage.

4-Evaluation of Distracters

  An additional item analysis has been carried out to check the functioning of distracters for the MC set. Masden (1983:51) points out those weak distracters often cause test questions to have poor discrimination or undesirable level of difficulty. No set percentage of responses has been agreed upon, but examiners usually feel uneasy about a distracter

that is not chosen by at least one or two examinees in a sample of 20 to 30 test papers. Building upon the views of Madsen (1983:52), the result of distracter evaluation has shown that all the distracters are functioning.

Table 1 shows a detailed description of the results of destructors evaluation.

Table 1. The results of destructors evaluation

Item No.	First passage				Second passage			
	A	B	C	D	A	B	C	D
1	3	3	10	4	2	3	3	12
2	2	2	3	13	2	4	12	2
3	3	2	12	3	3	3	10	4
4	2	3	2	13	2	11	5	2
5	3	13	2	2	3	9	3	5
6	13	2	3	2	2	3	12	3
7	2	3	13	2	2	14	2	2
8	2	3	13	2	12	3	3	2
9	13	3	2	2	2	14	2	2
10	2	2	14	2	2	13	2	3

5-Test Reliability

Reliability is a fundamental criterion against which any language test has to be judged, Anastasi (1982:24).It is usually concerned with how far researchers depend on the results that a test produces. According to Weir (1988:34), there are three aspects of reliability that are usually taken into account. The first is concerned with the consistency of scoring different markers, the second is concerned with parallel forms, and the third is concerned with test internal consistency. Of these three aspects, the researcher is concerned with the third aspect for two main reasons.

According to Oller (1979:66), this aspect, test validity .It is also considered to try and ensure that relevant subtest are internally consistent, Weir, 1988:94), and all the subjects presumably measure the same type of skills to be assessed, Oller (1979:67).

The aspects of reliability cannot be measured by test-retest or split -half methods but by working out inter- correlations among those subtest because high correlations between diverse tests must be taken as evidence not only of reliability , but also of substantial test validity .

To calculate test reliability (internal consistency), statisticians usually use Kuder-Richardson formulae , a method is not applicable to the type of reliability required in this study because there are two variables (two subtest) instead of a set of items that belongs to one test. Therefore, the use of Pearson product- moment correlation coefficient is recommended here. The formula for working out this kind of correlation coefficient reads;

$$R = \frac{\sum(X-X)(Y-Y)}{\sqrt{\sum(X-X)^2 \sum(Y-Y)^2}} = \frac{\sum XY}{\sqrt{\sum X^2 \sum Y^2}}$$

Where R= stands for correlation coefficients.

The capital letters X and Y stand for the two variables (two subtests).

The small letters X and Y represent the deviation from the mean X and Y respectively.

The calculation of (r) shows the following results for the two passages:

A-First Passage:

1-SAQs x TFIs =0.83

2-SAQs x MCIs =0.95

3-SAQs x CT   =0.88

4-TFLs x MCIs=0.90

5-TFIsx CT =0.89

6-MCIs x CT =0.94

B-Second Passage :

1-SAQs x TFIs =0.84

2-SAQs xMC =0.85

3-SAQs x CT =0.78

4-TFLs x MCIs =0.87

5-TFIs x CT =0.85

6-MCIs x CT =0.87

It is important to note that the reliability coefficient would be sufficient if it is not less than 0.50 for the standardized tests, Nunnaly (1972:72). However, it is preferable if it reaches 0.68 in unstandardized test ,Hedges (1966:220).

6-Final Version of the Test

As it has become evident, many measures have been taken by the researchers to make sure that the test has acquired the qualities of a valid and reliable test. Such measures are quite important, particularly in quasi-experimental researches where the test is the sole tool for data provision. However, to serve as solid ground for discussing and generalizing the results of this study the qualifications, it is note -worthy to sum up the qualifications secured in the test before the final administration as follows:

1-content validity,

2-face validity,

3-suitable level of difficulty,

4-acceptale power of discrimination,

5-functioning distracters for MCIs ,and

6-high reliability coefficient.

7-Final Administration of the Test

When administrating the two parts of the test to the subjects, the following measures have been taken:

(1) Part one and two have been given on two successive days,

(2)Care has been taken to give clear instructions to avoid any misunderstanding or confusion. For the sake of motivation, the examinees have been told by their teachers that a good score on the test would be taken into consideration in the assessment of the student's class effort.

(3) Suitable examination conditions have been provided for the subjects and complete control has been exercised to ensure that no external factor would interfere as a variable in the administration of the test.

8-Scoring the Test

The scoring of the sets of MCIs and TFIs has been highly objective and straight forward .This is quite usual for fully objective testing techniques. As for scoring the SAQs, any answer which is relevant to the answer provided by the test passage was considered correct and given full marks regardless of any spelling, punctuation or grammatical mistakes .To avoid too much writing, the subjects were told in the rubric to give short answers. The usual scoring procedures for cloze test are (1) the exact word method and (2) scoring for textual appropriateness. In this study, exact and acceptable replacement of words has been considered correct and allotted full marks.

9-Statistical procedures

The researcher used the following statistical procedures to analyze and interpret the test results. Below is a detailed description of these procedures:

1-Arithmetic Mean: It is used to indicate the general averages of the subjects' scores on the pre-reading and post-reading tasks.

2-T-Value: To find out the significances of differences among the average scores on the pre-reading and post-reading tasks for the teachers and students sample, two different t-tests formulae were used.

4. Results

The usual outcome from any RC test generally provides a quantitative assessment of how much a reader comprehends as a results of reading a passage or passages CF a particular RC test ;however , quantitative assessment by itself is not sufficient to give a very clear picture of the subject's performance of RC tests. There should also be room for some sort of qualitative assessment as well. Therefore, a quantitative assessment is made of the performance of the sample under this study on the pre-reading and post-reading tasks.

The quantitative assessment has been carried out through giving numerical indices of the subject's performance and then making multidimensional comparisons for that performance on both tasks.

The analysis of the data shows the following results;

1-The mean of the performance scores on the pre-Rt (30.31) represents the mean score of the amount of Bk (TB) on the whole test.

2-The mean of the performance scores on the post-RT (75.28), represents the mean score of gloss RC of the subjects ability .The difference between the two means(44.97) represents the mean of the IG scores.

3-To calculate the t-value of the difference between the pre-RT and the post-RT performance.

4-The observed t-value is 36.776 while the required t-value for significance is 2.635 at the P 0.01 level and df= 69. Therefore, the difference between the pre-RT performances is statistically significant. The null hypothesis is rejected. The alternative hypothesis should read, "There are significant differences between the pre-RT and post-RT performance of the GGS on RC tests.

A summary of all the results of the performance on the pre-RT and the post-RT for the SGS for the two passages is given in Table 2 below.

Table 2. the Results of the Performance on the Pre-RT and the post-RT for the SGS for the Two Passages

Technique	Pre-RT		Post-RT		IG score or net RC	T-value
	BK(Tb)score		Gross RC Score			
	Mean	SD	Mean	SD		
SAQs	2.885	0.885	8.046	1.410	5.161	29.592
TFIs	3.135	1.035	8.57	1.295	5.435	20.127
MCIs	3.492	1.435	8.20	1.020	4.708	32.134
CT	30.20	9.846	62.121	10.478	32.079	29.891

In addition, the results show that there are no significant differences in IG scores across the different techniques of testing RC in EFL. The IG scores of the four testing techniques are compared with one another to obtain the following findings;

1-5.435 for TFIs

2-5.161 for MCIs

3-4.708 for SAQs

4-32.079 for CT

To calculate the real difference among these mean scores statistically, one should carry out six combinations of comparisons. They are the ;

A-TFL vs. the SAQs,

B-TELs vs. the MCIs,

C-TFIs vs. the CTm,

D-MCIs vs. the SAQs,

E-MCIs vs. the CT, and

F-SAQs vs. the Ct.

The calculations depend on comparing the performance of the SGS on the pre-RT and post-RT of both passages across the four testing techniques.

 A- TFIs vs. SAQs

1-The two means of the performance on the pre-RT for the TFLs and the SAQs are 3.135 and 2.885 respectively. The observed t-value of the difference in this performance on this task is 2.431 in favor of the TFIs. The required t-value for significance of difference is 2.000 a P

<0.05and df 69. Therefore, the difference in performance on the pre-RT for the TFIs and the

SAQs is statistically significant.

2-The two means of the performance on the post-RT for the TFIs and the SAQs are 8.57 and 8.046 respectively. The observed t-value of the difference in this performance is 3.735 in favor of the TFIs. The required t-value for significance of difference is 2.000 at p<0.05 and df 69.Therefore, the difference in performance on the pre-RT for the two testing techniques is statistically significant. Hence, the difference between the two IG scores is statistically significant. The null hypothesis is rejected. The alternative hypothesis should read "There is a significant difference in IG scores

between TFIs and the SAQs.

B-TFIs vs. MCIs

1-The two means of the performance on the pre-RT for the TFIs and the MCIs are 3.135 and 3.492 respectively. The observed t-value of the difference in this performance on this task is 4.731 in favor of the TFIs. The required t-value for significance is 2.000 at P<0.05 and df=69.Therefore, the difference in performance on the pre-RT for the TFIs and the MCIs is statistically significant.

2-The two means of the performance on the post-RT for the TFIs and the MCIs are 8.57 and 8.20 respectively .The observed t-value of the difference between the two IG scores is 0.613.Though this difference is not statistically significant, it is still in favor of the TFIs. Therefore, there is a statistical difference in IG scores between TFIs and the MCIs in favor of the IG score of the TFIS. Hence, the difference in IG score between the two testing techniques is statistically significant. The null hypothesis is rejected .The alternative hypothesis should read" There is a significant difference in IG scores between the TFIs and the MCIs.

C-TFIs vs. CT

It is important to say that standard scores are used whenever there is any kind of comparison with the CT and any other testing teaching in this study.

1-The two means of the performance on the pre-RT for the TFIs and the CT are 3.135 and 30.492 respectively .The observed t-value of the difference in performance on this task is 0.483.The required t-value for significance of difference is 2.000 at P <0.05 and df=69.Though, the difference in performance on the pre-RT for the two techniques is not statistically significant; it is still in favor of the first one.

2-The two means of the performance on the post-RT for the TFIs and the CT are 8.57 and 62.121 respectively .The observed t-value of the difference in this performance is 9.762 in favor the TFIs. The required t. value for significance of difference is 2.000 at P<0.05 and df=69.Therefore, the difference in performance on the post-RT for the two testing techniques is statistically significant. Hence, the difference between the two IG scores is statistically significant. The null hypothesis is rejected .The alternative hypothesis should read 'there is no significant difference in scores between the TFIs and the CT.

D-MCIs vs. SAQs

1-The two means of the performance on the pre-RT for the MCIs and the SAQs are 3.814 and 2.885 respectively. The observed t-value of the difference in this performance on this task is 2.431 in favor of the MCIs .The required t-value for significance is 2.000 at P<0.05 and df= 69.Therefore,the difference is statistically significant.

2-The two means of the performance on the post-RT for the MCIs and the SAQs are 8.2 and 8.046 respectively .The observed t-value of the difference in this performance is 3.048 in favor of the MCIs .The required t-value for significance of difference is 2.000 at P<0.05 and df=69.Therefore, the difference in performance on the post-RT for the two testing techniques is statistically significant. Hence, the difference between, the two testing IG scores is statistically significant. The null hypothesis is rejected. The alternative hypothesis should read "there is a significant difference in IG scores between the MCIs and the SAQs.

E-MCIs vs. CT

1-The two means of the performance on the pre-RT for the MCIs and the CT are 8.2 and 62.121 respectively. The observed t-value of the difference in performance on this task is 5.79 in favor of the former. The required t-value for significance of difference is 2.000 at P<0.05 and df=69.

Therefore, the difference is statistically significant.

2-The two means of performance on the post-RT for the MCIs and the CT are 8.2 and 62.121 respectively. The observed t-value for significance is 2.000 at P<0.05 and df=69.Therefore, the difference in performance on the post-RT for the two testing techniques is statistically significant. Hence, the difference between the two IG scores is statistically significant. The null hypothesis is rejected and the alternative hypothesis should read "There is a significant difference in IG scores between the MCIs and the CT.

F-SAQs vs. CT

1-The two means of the performance on the pre-RT for the SAQs and the CT are 2.885 and 30.657 respectively. The observed t-value of the difference in this performance on this task is 0.555. Though, the t-value is less than the required t-value for significance, it is still in favor of the SAQs.

2-The two means of the performance on the post –RT for the SAQs and the CT are 8:046 and 62.121 respectively. The observed t-value of the difference in this performance is 7.712 in favor of the SAQs. The required t-value for significance is 2.000 at P<0.05 and df=69.Therefore, the difference in performance on the post-RT for the two testing techniques is statistically significant. Hence, the difference between the two IG scores is statistically significant. The null hypothesis is rejected .The alternative hypothesis should read "There is a significant difference in IG scores between the SAQs and the CT.

5. Conclusion

In view of performance of the sample under this study on both the pre-RT and post-RT tasks and across the four testing technique, the following conclusions can be drawn:

1) The overall performance of the whole sample of the tests and across the four testing techniques has been of satisfactory level; buy with varying degree across the four testing techniques.

2) The subjects' performance on the techniques that require recognition of the correct answer (i.e., selection of an option) has been higher than that on the techniques that require producing the correct answer (i.e., writing it). The highest level of performance has been on the TFIs and the lowest on the CT.

3) The performance of the SGS has been comparatively-low on the CT because filling in with the appropriate word in a CT requires mastering the handling of all the types .of contextual constraints.

4) RC tests are biased and contaminated because of the effect of BK on the subjects' performance. This has been exhibited in the significant differences in the performance of the TGS and the SGS.

5) The effect of the reader's BK on performance one RC tests is inevitable , and it should be accepted as such because the role of BK in text processing and comprehension has been established in numerous studies , but this effect does not seriously affect the construct validity of those test this has been exhibited in inter-task comparisons.

6) The amount of IG or net RC ability varies across the four testing techniques. IG score is highly dependent on inter-technique differences rather than any other factors, i.e., recognition-production on item type difference.

7) The information which is compatible with the reader's BK facilitates comprehension and consequently improves performance on RC tests. This has been manifested in the first passage are compatible with what the students have already studied in their course of linguistics. Therefore, their performance on the first passage has been better than that on the second passage whose information is not so common the students as the first passage. Success in RC in EFL in related to the level of proficiency in the language, the higher the level of proficiency, the better the chances that a reader will successfully comprehend the text. This has been evident in the difference in performance of the TGS and SGS on the test in this study.

5.1 Recommendations

In the light of the above-mentioned conclusions, the following views are recommended:

1) When preparing CTs, test designers should specify their aims clearly if they are after assessing general RC ability, they have to select texts that contain information which is compatible with the examinee's BK because compatibility of information facilitates comprehension, but they should select topics that are not quite familiar to the examines if they are after measuring the extent of IG.

2) Student should be trained on how to perform well on different types of items that test RC skills across different types of testing techniques.

3) CTs should receive more attention as an efficient and practicable testing and teaching techniques.

4) When developing or investigating RC skills, teaches and researchers should know that EFL readers should attain an acceptable level of proficiency in the target language before giving them the tasks that require genuine interaction with the texts they read. This is why the researcher has chosen advanced learners of EFL as subjects for this study.

5) EFL readers should be introduced to all the problems that might hinder of wrong schema on comprehension. When preparing exercises for teaching comprehension, authentic texts should be used because the real intentions or ideas of the original writer are marred by the process of simplification or modification.

6) EFL readers should be aware of the fact that effective reading calls for an efficient interaction of all types of text processing.

7) Pre-reading tasks are highly recommended in the teaching of RC because such tasks help readers to activate the right type of schemata which helps them to process and comprehend the text they read.

References

Adams, M., and Collins, A. (1979). *A schema-theoretical view of   freedle*, R. (ed.), New directions of discourse processing, Norwood. Ablex.

Alderson, J., and Urqhart, H. (1988). This test is Unfair: I'm Not An Economist, in Carrell, P.;j. Devine and D. Eskey; (eds), Internet approaches to second language reading, Cambridge CUP.

Alverman, D., Smith, I., and Readence. J. (1985). Prior knowledge activation and the comprehension of compatible and incompatible texts. *Reading research quarterly,20*(4) 420-436.

Anastasi , A. (1982). *Psychological testing.* London , Macmillan.

Arthur, G. (2008). *Advances in text comprehension : commentary and final perspective.* USA. john Wiley & sons , ltd .

Bartlett, F. (1932). *Remembering: A study in experimental and social psychology*. Cambridge, CUP.

Beck, I. (1981). Reading problems and instructional practice. *Reading research quarterly, 2*(1),53-95

Carreil, P.J. Devine and Eskey, D. (1988, eds.). *Interactive approaches to second language reading.* Cambridge , CUP.

Cziko, G. (1980). Language competence and reading strategies. "A comparison of first and second language oral reading errors, language learning. *A Journal of applied linguistics ,30*(1), 101-116.

Devine , J. (1988). A case study of two readers : and reading performance : In Carrell , p., J. Devine and D. Eskey (eds), *Interactive approaches to second language reading* . New York.CUP.

Diana, P. (2005). The relationship between text comprehension and second language incidenl vocabulary acquisition : a matter of topic familiarity? USA, Michigan University press.

Eskey, D. (1790). A new technique for the teaching of reading to advanced student, *TESOL Quarterly , 4*(4),315-210.

Eskey, D., and Grabe, W. (1988). Interactive models of second language reading: perspective on instruction, in: Carrell, p,J Devine and D.Eskey (eds.), *interactive approaches to second language reading*. Cambridge, CUP.

George, J., and Fong, H. (2006). *L1 and L2 vocabulary glosses in reading passage: their effectiveness for increasing comprehension and vocabulary knowledge*. Singapore, Singapore University press.

Gi-pyo, P. (2008). comparison of L2 listening and reading comprehension by university students learning English in Korea, Korea, Soonchunhyang University press.

Good, C. (1959). *A dictionary of education , 3rd education.* New York ,McGraw-Hill Book company.

Goodman, K. (1965). *A linguistic study of cues and miscues in reading elementary English.*

-----------(1988). The reading process in Carrell, P. Deving J. and Eskey , D. (eds.) *Interactive approaches to second language reading.* Cambridge, CUP.

Grabe, W. (1988), Reassessing the term "Interactive" , in Carrell , P. J. Devine , and D. Eskey (eds.) interactive approaches to second language reading Cambridge-,CUP.

Hedges, W. (1966). *Testing and evaluation for the sciences.* California ,wordsworth LTD.

Hudson, T. (1978). *The effect of pre-reading activities on RC And student attitude.* California, UCLA press.

Johnson, P. (1982). Effect on RC of building background knowledge. *TESOL Quarterly, 16*(4), 503-516.

Johnson, P. (1984) Prior knowledge and RC test bias. *Reading research Quarterly, 14*(2), 219-239.

Kintsch, W. (1977). *Memory and cognition.* New York John Wiley and sons Inco.

LaBerge, D. and Samuels, S. (1974). Toward a theory of processes in reading perception.

Ludo, V., and Jan, L. (2008). *Prediction of the development of reading comprehension: A longitudinal study.* The Netherlands., john Wiley & sons ,LTD.

Madson, H. (1983), Technique in testing, Oxford , OUP.

Nunnally, J. (1972). *Education measurement and evaluation*. 2nd Edition , New York , McGraw hill.

Oiler, J. (1972). Assessing competence in ESL reading. *TOSEL Quarterly, 6*, 313-326.

Paul, B. (2008). *Cognitive processes in comprehension of science texts: the role of co-activation in confronting misconception*. Canada, McGill University press.

Rumelhart, D. (1977). Toward an interactive model of reading, in S. Dornic (ed), *Attention and performance.* New York, Academic press.

Samueles, S. and Kamil, M. (1988). Models of the reading process in Carrell, P. Devine and D. Esky (eds.), *Interactive approaches to second language Reading.* Cambridge, CUP.

Silberstein, S. (1987), let's take another look at reading: twenty – five years of reading instruction. *English teaching Forum: 25(*4), 28-35.

Smith, F. (1971). *Understanding reading: A psycholinguistic analysis of reading and learning to read.* New York, Holt Rinehard and Winstoning.

Stanovich, K. (1980). Toward an interactive-compensatory model of individual differences in the development of reading comprehension fluency. *Reading Research Quarterly, 16*(1), 32-71.

Strevens, P. (1977). *New orientation in teaching of English.* London OUP.

Taylor, W. (1953). Cloze procedure: A new tool for measuring readability. *Journalism Quarterly* (30), 415-432.

Thomson, R. (1959). *The psychology of thinking middle sex.* Penguin Books Ltd.