On the Impact of Teacher-Made Vocabulary Tests vs. Standardized Vocabulary Tests on Reading Comprehension Performance of Iraqi Intermediate EFL Learners

Esmaeil Bagheridoust[1] & Zinah Ameen Kadhim[2]

[1] Department of TEFL, College of Foreign Languages and Literature, Islamic Azad University, South Tehran Branch, Iran

[2] Department of TEFL, Islamic Azad University, South Tehran Branch, Iran

Correspondence: Zinah Ameen Kadhim, Department of TEFL, Islamic Azad University, South Tehran Branch, Iran. E-mail: zinaameen569@gmail.com

Abstract

This study Investigates the impact of Teacher-Made Vocabulary Tests (TMVT) vs. Standardized Vocabulary Tests (STVT) on Reading Comprehension Performance of Iraqi Intermediate EFL Learners. After analyzing collected data out of treatment, assessment, and instrumentation, the researcher evaluated and interpreted the score obtained by 66 young Iraqi female students in two language schools in Iraq. The interpretation shows how the participants reacted to Teacher-made vocabulary tests as compared to their reactions to Standardized Vocabulary test. The crux of the matter was the degree to which such vocabulary tests affect the EFL learners' performance in Reading Comprehension tests. The researcher reported that the use of Standardized vocabulary tests DOES play a significantly positive role in promoting the reading comprehension skills of Iraqi EFL college students. She also found that Teacher-made vocabulary tests DO play a significantly positive role in promoting the reading comprehension skills of Iraqi EFL college students. Finally, the researcher found that the t test value of 2.635 indicated a significant difference between the STVT and TMVT. This means that the use of TMVT is more effective than the implementation of STVT in encouraging Iraqi EFL learners to read more successfully.

Keywords: Teacher-Made Vocabulary Tests (TMVT), Standardized Vocabulary tests (STVT), reading comprehension ability, intermediate EFL learners

1. Introduction

1.1 Problem Statement

Indubitably, the nature, function, and type of standardized tests, the internationally qualified exams are determined and designed primarily on the grounds of measuring English language proficiency levels of EFL learners worldwide, owing to the crux of confidentiality policy, the secrets of such standardized tests have always been unidentified and undisclosed to the public, language instructors, language assessors, and language learners. Other way round, the nature of teacher-made tests have always been under scrutiny by other teachers, official testing institutions, families, experts, and students. Researchers has conceived that reading texts followed by questions are better welcomed by the students than most other types of reading passages. But most recently it is suggested that passages followed by teacher-made vocabulary questions or tests are more effective in improving learners' vocabulary development and reading comprehension skills. Recently research on L2 vocabulary learning and reading comprehension has shown that readers need to develop essential vocabulary learning processes and abilities such as rapid word recognition, vocabulary development, text structure awareness and strategic reading. Yet, all researchers recognize that the actual ability to comprehend a text come about through reading, and doing a great deal of it, as the core of reading instruction" (Marianne, C. & Muricia, E. 2011).

Although different strategies have been used to improve vocabulary and reading comprehension, none of them foster a fun learning environment and reduce anxiety towards test-taking. Allowing students to make their own tests while they are reading the text, will give them an opportunity to have more authority over their own learning and feel more autonomous.

If teachers make their own tests from the text, they interact with the text more than usual while teaching the texts. Many EFL instructors feel that their reading comprehension is better than usual when they are able to demonstrate on comprehension tests (Smith, P., 1994). Though researchers have done a lot on other approaches, they have not studied these approaches comparatively.

Accordingly, this study attempted to test the effect of teacher-made tests vs. standardized vocabulary tests on vocabulary development and reading comprehension performance of Iraqi College EFL Students.

1.2 Significance of the Study

This study is expected to concoct grave theoretical and practical significance in foreign language assessment, teaching, and learning. Based on the statements above, the researcher in his research is going to show the effect of Teacher-made test of vocabulary vs. Standardized tests of vocabulary on reading comprehension of Iraqi EFL College students. Based on the questionnaire that Patrick Smith (University of American Puebla), in his article (Learner Self-Assessment in Reading Comprehension: The case for teacher-constructed tests, 1994) mentioned, 76% of teachers felt that the process of writing a test would help them to perform better in teaching vocabulary and reading comprehension passages.

Teacher-made vocabulary Test Questions allow the faculty to assess at least three aspects of student learning. In these questions, teachers see what their students consider the most important or memorable content, what they understand as fair and useful test questions, and how well they can answer the questions they have posed. This information not only provides direction for teaching, but can also alert the teacher when students have inaccurate expectations about upcoming exams. Teacher-made Vocabulary Test Questions help students assess how well they know the material and faculty feedback can refocus their studying.

1.3 Relevant Scholarship (Related Research Results)

Tests are designed to measure the extent to which an individual has "achieved" something, acquired specific information, or mastered a specific skill, usually as a result of planned instruction or training. It is the task of education to bring about deep learning and desirable changes in those who experience the educational process. In addition, for this task to be performed effectively and efficiently, educators must have a useful means of assessing the initial state of a particular individual or group and the changes that have occurred as a result of the instruction. In addition to attempting to qualitatively experience the full extent or direction of change, it is often important for teachers to form a qualitative assessment of the desirability of change (Charles & Richard, 1990). However, all of these assessment methods have found unimpressive student performance on standardized exams administered by public examination boards over the years, and this is a testament to everyone involved in the education sector. Olatoye (2014) found that Nigerian science students consistently performed poorly on internal and external exams. The observed decline in student performance on public standardized tests raises the question of whether high failure rates do not reflect the quality of test use in schools. In other words, tests used by teachers as a means of classroom assessment in the teaching and learning process may explain the observed poor academic performance of students (Charles and Richard, 2010). The standardized test can be used to supplement classroom-her tests created by teachers. Standardized tests, such as tests administered by teachers, help measure student progress and achievement and determine how well students have acquired knowledge and skills. Combining teacher-administered testing and observation along with carefully selected standardized tests provides a comprehensive program of student assessment and program evaluation in education (Okpala, Onocha, and Oyediji, 2012). The process of developing standardized tests provides insight into understanding intended functionality. Test standardization is the process of conducting tests on a representative sample of test takers for the purpose of establishing norms and having clearly defined procedures for administration and evaluation, including normative data. If there is, the test is considered standardized.

Three reading comprehension definitions have been prevalent and known in literacy programs in the United States of America as reported by Foertsch, 1998. Learning to read means learning to pronounce words, as stated by the first definition. Three are a number of other definitions in this area which need to be studied for more understanding of reading concept. Such definitions recognize the importance of teaching skills as part of the reading process (Allington & Cunningham, 1996; International Reading Association and National Early Childhood Education Association, 1998; Maryland State Department of Education, n.d.; Snow, Burns, & Griffin, 1998). It also supports balanced reading instruction for all students (Allington & Cunningham, 1996; Au, 1993; Foertsch, 1998; International Reading Association and National Association for the Education of Young Children, 1998; Snow, Burns, & Griffin, 1998). As we know, comprehension is an important goal for most language learners. As defined by the Partnership for Reading (2005), reading comprehension is the understanding of the text read or the process of creating meaning from the text. Comprehension is a "construction process" since it comprises all the

components of the reading process that work together in reading a text to create an illustration of the passage in the reader's mind.

Comprehension Success in all academic areas is dependent upon the ability to read and comprehend. Reading comprehension is a dynamic process, which requires interaction between the reader and the text as a quintessential part of the activity. Mastery of the other components: phonemic awareness, phonics, vocabulary development, and fluency, facilitates reading comprehension. The act of comprehension is so sophisticated that there is not one instructional approach that can meet the needs of all readers with all texts in all learning situations (Snow, Burns, & Griffin, 1998).

1.4 Research Questions and Formulated Hypotheses

To illustrate the research tenets and assumptions, the researcher poses the following research questions;

R_{Q1}. Does the use of Teacher-made vocabulary tests significantly promote the reading comprehension skills of Iraqi EFL college students?

R_{Q2}. Does the use of Standardized vocabulary tests significantly promote the reading comprehension skills of Iraqi EFL college students?

R_{Q3}. Is there any significant difference between the impact of the use of Teacher-made vocabulary tests vs. Standardized vocabulary in promoting the reading comprehension skills of Iraqi EFL college students?

The corresponding null-hypotheses are:

H_{01}. The use of Teacher-made vocabulary tests (TMVT) does NOT reliably promote the reading comprehension performance of Iraqi EFL college students.

H_{02}. The use of Standardized vocabulary tests (STVT) does NOT reliably promote the reading comprehension performance of Iraqi EFL college students.

H_{03}. There is NO reliable inconsistency between the impact of Teacher-made vocabulary tests (TMVT) vs. Standardized vocabulary tests (STVT) in promoting the reading comprehension performance of Iraqi EFL college students.

2. Method

2.1 Participants

To actualize the purposes of this study, the researcher chose and recruited 60 homogeneous participants from a population of 120 Iraqi EFL students as her study participants. The participants were only female Iraqi EFL learners studying English at an Iraqi College, in Baghdad. The recruited participants were all intermediate EFL learners with almost a homogeneous background in English in Iraq. They study English as an obligatory course in High School and they have no exposure to the English language in any other language school apart from their state school. The participants at this term study the book "Developing Reading Proficiency 2" as a four credit course, provided that they passed the book "Developing Reading Proficiency 1" at first term. The age of participants ranges from 19 to 25. For the researcher to make sure whether the participants were at the same proficiency level, TOEFL Reading Proficiency test was administered. After making sure having 60 homogeneous participants, the researcher at random (by flipping a coin) divided candidates two classes, each having thirty students. One class called the control and the other named experimental.

2.2 Apparatus and Instruments

In this research, these three instrumentations were used:

Teacher-made vobulary tests

Standardized vocabulary tests

TOEFL (a language proficiency test)

Pretest

Post-test

2.2.1 Teacher-Made Vocabulary Tests (TMVT)

Standardized test is constructed by eligible and professional test designers mostly in England and the USA where ESOL and ETS are of huge popularity. IELTS, TOEFL, CELPIP, CAE, FCE, CAEL, and GESE are some of these tests, while a teacher administered test is, as the name suggests, a teacher administered test. Ahman and Glock (1971) argue that a standardized test is one that is almost always carefully designed by groups of people rather than

a single individual (teacher-administered tests). In this study, teacher-administered vocabulary tests are those vocabulary tests developed by the teacher and used experimentally in the classroom in Iraq, where the study takes place, to measure and assess vocabulary knowledge of students in the classroom.

2.2.2 Standardized Vocabulary Tests (STVT)

Regardless of format, vocabulary tests are generally considered interchangeable indicators of vocabulary knowledge, coined by Spearman (1927) as an "indifference indicator." In his extensive analyzes of the factor structure of human abilities, Carroll (1993) concluded, "The exact form in which vocabulary knowledge is measured does not usually change the factor composition of the variables so much that the main characteristic being measured is native vocabulary" (p. 158). The standardized vocabulary tests in this study are those vocabulary tests traditionally used in Iraqi language schools and colleges to measure and evaluate students' vocabulary skills in class.

2.2.3 TOEFL (English Language Proficiency Test)

A TOEFL Reading Proficiency test was administered to find out the homogeneity of the participants. The original test is available at www.iielatinamerica.org. The reliability of this test was 0.93 (93%). The researcher of this study found the reliability of this test through KR-21 rk formula.

2.2.4 Pretest

After dividing homogeneous groups of the participants into two groups as a control and experimental group and piloting the test to a group of 20, the researcher administered a pretest of reading comprehension to the students in both classes. Five reading texts followed by 20multiple-choice and "WH questions" format from the book "Developing Reading Proficiency 2" written by Dr. M. H. Tahririan (2010), published by Payam Noor University, were chosen for pretest.

To construct the pretest, the researcher found the readability of texts through Felsch readability formula. It was done with the Word 2007 software and the mean score was calculated. Therefore, the texts which were chosen for comprehension test, had the readability between 59.77 and 73.81. Then the test was scored.

2.2.5 Post-test

The researcher administered a post-test of reading comprehension to the students in both classes to get the result of treatment. The post-test was five texts that were used in pretest. It means that the texts used in pretest and post-test were the same. The texts followed by 20 questions in the multiple-choice and WH-question format, but these questions were neither the author-made questions that were used for control group nor the student-made questions that students made in treatment group. The post-test test was different test from both the control and treatment group. A pilot teacher-made reading comprehension test was used for post-test, because the aim of post-test was to find out the effect of treatment in this study.

2.3 Study Design

Among several research designs, the one which seems to best fit the purpose of the present study is the quasi-experimental design (The pretest-post-test control group design) (Selinger & Shohamy, 1989). It is illustrated in Table 1 as following:

Table 1.

Groups	Pretest	Treatment	Post-test
Control	T_1		T_2
Experimental	T_1	X	T_2

T1 = Pretest T2 = Post-test X = Treatment

2.4 Procedure

The study was carried out through the comparative-correlative analysis of the scores obtained by the same participants who experience either Teacher-made tests or Standardized tests of vocabulary. Sample size is normally not less than 25. To do this study, the researcher chose 60 homogeneous participants from a population of 120 college students as his study participants by administering a TOEFL. Both genders were recruited whose major was English. They are from Baghdad University. They were studying at second term. The participants at this term study the book "Developing Reading Proficiency 2" as a four credit course provided that they passed the book

"Developing Reading Proficiency 1" at first term. The age of participants ranges from 19 to 25. For the researcher to make sure whether the participants were at the same proficiency level, TOEFL Reading Proficiency test, was administered. The scores were ranged from low to high and 60 participants who scored between the ranges of 60 to 80 were chosen as participants of the study.

After making sure having 60 homogeneous participants, the researcher by dividing the participants into odd and even numbers made two groups of 30 and then at random (by flipping a coin) named these two groups. One group was as control group (having standardized tests of vocabulary in class) and another as experimental group (exposed to teacher-made tests in class). After dividing homogeneous groups of the participants into two classes, the researcher administered a pretest of reading comprehension to the students in both classes. The test for pretest was first piloted to 20 subjects who were at the same level with the participants and then the reliability of the pilot test calculated with split-half method, through Spearman Brown formula.

3. Results

Evidently enough, this study was conducted to investigate the effect of using Teacher-Made Vocabulary Tests vs. Standardized Vocabulary Tests strategy on reading comprehension of Iraqi EFL college students. The researcher has initially assumed that standardized vobulary tests are used conventionally in reading comprehension classes in the college to promote reading comprehension ability. For the same token, the researcher asked the teacher to construct some vocabulary tests to use in the class to promote reading comprehension as the experimental group. The purpose of this research was to see whether using these tests had any significant effect on improving the students' ability to deal with reading comprehension. After testing, scoring, and tabulating the results obtained, the researcher embarked on employing a series of statistical tests and analysis to provide solve research hypotheses.

3.1 Preliminary Data Analysis

3.1.1 Results on TOEFL Test

In order to make sure that the participants the researcher planned to recruit and choose as control and experimental group from a population of 120 to be homogenous, she administered a TOEFL proficiency test. Then the researcher ranged the scores from low to high and chose 60 participants whose scores were ranged between 60 to 80. After making sure having 60 homogeneous participants, the researcher by dividing the participants into odd and even numbers made two groups of 30, and then at random (by flipping a coin) chose one group as control group and another as experimental group.

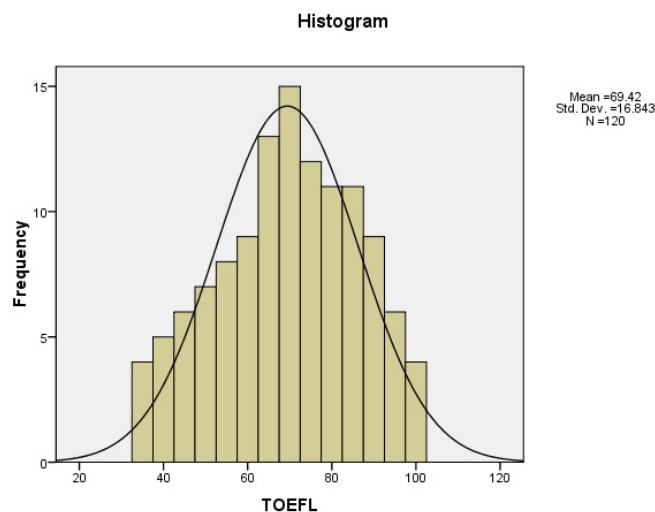The reliability of the TOEFL test was obtained through KR-21 rk formula and that is 0.93 to measure the level of homogeneity of the participants.

Table 2. Descriptive Statistics of TOEFL for homogeneity of participants

	N	Range	Min	Max	Mean	Std. Deviation	Variance
TOEFL	120	65	35	100	69.42	16.84	283.69
Valid N (list wise)	120						

Table 3. Descriptive data analysis of TOEFL for homogenous group

		Frequency	Percent	Valid Percent	Cumulative Percent
	35	4	3.3	3.3	3.3
	40	5	4.2	4.2	7.5
	45	6	5.0	5.0	12.5
	50	7	5.8	5.8	18.3
	55	8	6.7	6.7	25.0
	60	9	7.5	7.5	32.5
	65	13	10.8	10.8	43.3
Valid	70	15	12.5	12.5	55.8
	75	12	10.0	10.0	65.8
	80	11	9.2	9.2	75.0
	85	11	9.2	9.2	84.2
	90	9	7.5	7.5	91.7
	95	6	5.0	5.0	96.7
	100	4	3.3	3.3	100.0
	Total	120	100.0	100.0	

Histogram

Mean =69.42
Std. Dev. =16.843
N =120

The important note is that the obtained reliability index of the TOEFL test based on the standards of ETTS, was reported as to be KR-21 rk = 0.93, which is a high reliability index.

3.1.2 Readability of Texts for the Study

Mousavi (1991) defines readability as "a measure of understandability of written text as given by an analysis of a variety of factors including syntax complexity, vocabulary, thematic expression and continuity of themes." (p.310) In this study the readability indexes of 14 reading texts of the book "Developing Reading Proficiency 2" written by Dr. M. H. Tahririan (2010), published by Payam Noor University were calculated through Felsch Readability Formula. The score in this formula is on a scale of 0 to 100, the lower the score the more difficult the writing is to read. Table 4 shows the Readability of texts.

Table 4. Readability index of study texts

Texts	1	2	3	4	5	6	7	8	9	10	11	12	13	14
	55.2	55.7	57.7	62.6	66.5	66.5	66.6	67.5	68.4	69.5	72.1	72.8	76.7	77.3

As the results show, the obtained mean score of 14 texts is 66.79 and standard deviation is 7.02, so texts which were chosen for study class have the readability between 59.77 and 73.81.

Table 5. Descriptive Statistics of readability of 14 texts of study

	N	Min	Max	Mean	Std. Deviation
Texts	14	55.20	77.30	66.7929	7.02265
Valid N (list wise)	14				

Readability and text selection for pretest and post-test.

Five texts were used for pretest and post-test. Table 6 shows the readability of each text.

Table 6. Descriptive Statistics of texts for pretest and post-test

	N	Min	Max	Mean	Std. Deviation
Texts	5	60.00	73.20	67.5400	5.61765
Valid N (list wise)	5				

3.1.3 Pilot of Pretest

Initially, to determine the reliability of the reading comprehension test, a group of 20 people were given reading comprehension in a pilot study, after which the researcher investigated the reliability of reading comprehension using the Split-half method and Spearman Brown correlation. factor It is a measure of the strength of the relationship between two sets of data. According to this, correlation coefficients between two variables give a value between -1.00 and 1.00, the closer the value is to 1.00, the more reliable the test.

Table 7. Descriptive Statistics of pilot of pretest

	N	Minimum	Maximum	Mean	Std. Deviation	Variance
Whole Score in Pretest	20	7	18	13.10	3.553	12.621
Valid N (list wise)	20					

Table 8. Case Processing Summary

		N	%
	Valid	20	100.0
Cases	Excluded	0	.0
	Total	20	100.0

a. List wise deletion based on all variables in the procedure.

Table 9. Reliability Statistics for Pretest

Cronbach's Alpha				Total N of Items	Correlation Between Forms	Spearman-Brown Coefficient		Guttmann Split-Half Coefficient
Part 1		Part 2				Equal Length	Unequal Length	
Value	N of Items	Value	N of Items					
1.000	1a	1.000	1b	2	.707	.829	.829	.824

a. The items are: Even Scores in Pretest

b. The items are: Odd Scores in Pretest

3.1.4 Pretest of the Study

After dividing homogeneous groups of the participants into two groups as a control and experimental group and piloting the test to a group of 20, the researcher administered a pretest of reading comprehension to the students in both classes. Five reading texts followed by 20 multiple-choice and WH questions from the book "Developing Reading Proficiency 2" written by Dr. M. H. Tahririan (2010), published by Payam Noor University, were chosen for pretest.

Table 10. Descriptive Statistics of pretest in control group

	N	Range	Min	Max	Sum	Mean	Std. Deviation	Variance
Pretest in Control Group	30	50	40	90	2030	67.67	12.981	168.506
Valid N (list wise)	30							

Table 11. Descriptive Statistics of pretest in experimental group

	N	Range	Min	Max	Sum	Mean	Std. Deviation	Variance
Pretest in Experimental Group	30	45	45	90	2095	69.83	12.211	149.109
Valid N (list wise)	30							

We employed Levene's test to compare mean average of both tests to figure out the equality of variances and to indicate that scores are consistent. Then the t test is used to determine the equality of means.

H0: The variances are equal.

H1: The variances are not equal.

Table 12 indicates equal variances because the p value is 0.73 which is higher than $\alpha = 0.05$.

Table 12. Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means					9595% Confidence Interval of the Difference	
		F	Sig.	T	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
Reading Comprehension	Equal variances assumed	.113	.738	-.666	58	*.508	-2.167	3.254	-8.680	4.347
	Equal variances not assumed			-.666	57	.508	-2.167	3.254	-8.680	4.347

*P > .05 =It doesn't show significant difference

The P value which is .50 is more than .05. Also the t-value observed is .66 which is less than the t- critical at the 0.05 level of significance which is 1.67. Therefore, we can safely claim that the mean score of pretest in control and experimental group are not different.

3.1.5 Pilot of Post-test

First, to determine the reliability and validity of the reading comprehension test, a reading comprehension test was administered to a group of 20 as a pilot study, and then the researcher examined the reliability of the reading

comprehension test by Spearman Brown using splint. - half method. correlation coefficient. It is a measure of the strength of the relationship between two sets of data. According to this, correlation coefficients between two variables give a value between -1.00 and 1.00, the closer the value is to 1.00, the more reliable the test.

Table 13. Descriptive Statistics of pilot   of post-test

	N	Range	Minimum	Maximum	Mean	Std. Deviation	Variance
Whole Scores in Post-test	20	11	8	19	13.55	3.252	10.576
Valid N (list wise)	20						

Table 14. Reliability Statistics for Post-test

Cronbach's Alpha				Total N of Items	Correlation Between Forms	Spearman-Brown Coefficient		Guttmann Split-Half Coefficient
Part 1		Part 2				Equal Length	Unequal Length	
Value	N of Items	Value	N of Items					
1.000	1a	1.000	1b	2	.727	.842	.842	.842

a. The items are: Even Scores in Post-test

b. The items are: Odd Scores in Post-test

3.1.6 Post-test of the Study

The researcher administered a post-test of reading comprehension to the students of both classes to obtain the treatment result. There were five texts in the posttest that were used in the pretest. This means that the texts used in the pre- and post-test were the same. The texts were followed by 20 questions in the form of multiple choice and WH questions, but these questions were not the questions asked by the control group, nor the questions asked by the students in the treatment group. The post-test questions were different questions for both the control and treatment groups. In the post-test, a comprehension test text prepared by the pilot teacher was used, because the purpose of the post-test was to find out the effect of the treatment in this study.

Table 15. Descriptive Statistics of post-test in control group

	N	Range	Mini	Max	Sum	Mean	Std. Deviation	Variance
Post-test in Control Group (Exposed to Standardized Vocabulary Tests in class)	30	50	45	95	2160	72.00	12.972	168.276
Valid N (list wise)	30							

Table 16. Summary analysis of post-test in control & experimental group

	Group	N	Mean	Std. Deviation
Reading Comprehension	Control Group (STVT)	30	72.00	12.972
	Experimental Group (TMVT)	30	80.17	10.945

H0: Post-test scores are equal in two groups.

H1: Post-test scores are not equal in two groups.

Table 17 indicates equal variances because the p value is .29 which is higher than α = 0.05.

Table 17. Independent Samples Test

		Levene's Test for Equality of Variances		of t-test for Equality of Means					95% Confidence Interval of the Difference	
		F	Sig.	t	Df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
Reading Comprehension	Equal variances assumed	1.139	.290	-2.635	58	.011	-8.167	3.099	-14.370	-1.964
	Equal variances not assumed			-2.635	56	.011	-8.167	3.099	-14.373	-1.960

* P< .05 =It shows significant difference

Table 17 provides sufficient evidence to reject the null hypothesis of the third null hypothesis of the study namely; (There is no significant difference between the effects of teacher-administered vocabulary tests and standardized vocabulary in promoting Iraqi EFL students' reading skills) because the P-value, which is 0.011, is less than 0.05. A t-value of 2.63 is also observed which is greater than the t-critical value of 0.05 at the level of significance which is 1.67. Therefore, we can safely say that the use of teacher-constructed vocabularies in the classroom is much more effective than the use of standardized vocabulary tests in Iraqi EFL learners' reading comprehension skills in the classroom.

4. Discussion

In this section, the researcher discussed some findings of his research with some findings of other researchers. In recent years, a great deal of research in L1 and L2 fields has been conducted on reading strategy training via vocabulary test exposure. Strategy training comes from the assumption that success in learning mainly depends on appropriate strategy use and that unsuccessful learners can improve their learning by being trained to use effective strategies (Dansereau, 1985; Weinstein & Underwood, 1985). Many studies have shown that vocabulary tests can be taught to students, and when taught, vocabulary help improve student performance on tests of reading comprehension and recall (Carrell, 1985; Brown & Palincsar, 1989; Carrell, Pharis, & Liberto, 1989; Pearson & Fielding, 1991). We can consider comprehension strategies as plans or procedures that readers use and apply when they hear text read aloud, when they read text with a teacher, and when they read independently. According to Iwasaki, I. (TESL Journal, 2008)," using student-made quizzes as a method of learning in the EFL classroom, instead of as a mere method of testing proficiency in the L2, may be one way to foster a fun learning environment and reduce anxiety toward test taking." So using the policy of teacher-made vobulary tests as well as standardized vocabulary tests can be one way to help EFL students comprehend the text better and foster a fun learning environment and reduce anxiety toward test-taking. Many studies have shown that students, who answer the question while studying a text, perform better than students who do not (See Hamaker, 1986, Hamilton, 1985). Teacher-made vocabulary tests is a simple but productive way to support reader engagement with the text. According to Poway Unified District (PUSD, 2005), while reading chunk of text, teachers write down vocabulary questions they have about what they teach and what will happen next in the reading comprehension classroom. As mentioned earlier, this study investigated the effect of teacher-made vocabulary tests vs. Standardized vocabulary tests on reading comprehension of Iraqi college EFL students. Using teacher-made tests vocabulary tests in class in the experimental group led to a better performance in reading comprehension as compared to the standardized vocabulary tests in post-test.

5. Conclusion

Although different techniques and strategies have been used to improve reading comprehension, few of them foster a fun learning environment and reduce anxiety towards test-taking. Allowing teachers to make their own vocabulary tests while using standard vocabulary tests for the ease of work, will give them an opportunity to foster more authority over their own teaching plans. If teachers make their own tests from the texts, they interact with the text more than usual while teaching the tests to the students. Many students of EFL feel that their reading

comprehension is better than usual when they are able to demonstrate on comprehension tests (Patrick Smith, 1994). Though researches have done a lot on other approaches, they have not studied these approaches comparatively. As a result, this study attempted to show the effect of teacher-made vocabulary tests vs, standardized vobulary tests on the ways on how Iraqi EFL students deal with reading comprehension passages. The main concern of this study was to study this assumption whether or not using teacher-made tests strategy can bring about any positive effect on the ability of students in dealing with the reading comprehension passages of Iraqi College EFL Students.

To assure and determine any significant changes in our participants, the results of each group in pretest were compared with post-test applying t test. Statistics show reliable growth in participants of experimental group; students in experimental group benefited highly from the treatment. Moreover, t test indices allowed researchers to reject the null hypotheses and thus, they found proper answers to their questions. This study is by nature a two-fold study, which covers both a macro study on teaching professionalism and micro study of learning development. The inquirer, to approve of her claims, piloted the research questions and tests, followed the situation specific statistical procedures, and came up with the statistical results elaborated comprehensively above. In this section, the researcher also summarizes the investigation procedure along with the findings based on the results and later mentions the pedagogical implications of the study. In the final section, she has provided the readers with some suggestions for further study. Based on what we can read and understand from the t test table reflected earlier and the results obtained, the researcher has come up with the following findings:

The use of Standardized vocabulary tests DOES play a significantly positive role in promoting the reading comprehension skills of Iraqi EFL college students. Based on the difference between the t test results of pretest and post-test data, the mean score of Control Group before treatment in pretest was 67.67, and after treatment of STVT it increased to 72.

The use of Teacher-made vocabulary tests DOES play a significantly positive role in promoting the reading comprehension skills of Iraqi EFL college students. Based on the difference between the t test results of pretest and post-test data, the mean score of Control Group before treatment in pretest was 69*.83, and after treatment of TMVT it increased to 80.17.

By comparing the t test and F-value indices of pre and posttest independent samples t test analysis, we can see that the t value index has increased from .66 to 2.635 and the f-value has increased from .113 to 1.139. The results obtained rejects the first, the second and the third null-hypotheses.

The third null-hypothesis is rejected reliably, since the t test value of 2.635 indicated a significant difference between the STVT and TMVT. This means that the use of TMVT is more effective than the use of STVT in a reading comprehension class.

References

Ahman, J.S., & Glock, M.D. (1971). *Evaluation Pupil Growth, Principle of Tests and Measurements,* 4th Ed. Boston: Allyn and Bacon. Inc. https://doi.org/10.31327/jee.v2i2.289

Allington, R. L., & Cunningham, P. M. (1996). *Schools that work: Where all children read and write*. New York: HarperCollins College Publishers. https://doi.org/10.7202/1036112ar

Carrell, P. (1993). Introduction: Interactive approaches to second language reading. In P. Carrell, J. Devine, & D. Eskey (Eds.), *Interactive approaches to second language reading* (pp. 1-7). Cambridge, England: Cambridge University Press. https://doi.org/10.1017/CBO9781139524513.003

Carrell, Patricia L. (1985). Facilitating ESL reading by teaching test structure. *TESOL Quarterly, 19*(4), 727-45. https://doi.org/10.2307/3586673

Celce-Murcia, M. (1991). L*anguage teaching approaches: An over view in teaching English as second or foreign language*. Edited by M. Celce. Muricia. Boston, MA: Heinle & Heinle. https://doi.org/10.4236/jss.2016.42021

Celce-Murcia, M. (2001). *Teaching English as second or foreign language*. 3rd ed. New York.

Dansereau, D. (1985). Learning strategy research. In J. Segal & S. Chipman (Eds.), *Thinking and learning skills*. London: Lawrence Erlbaum Associates. https://doi.org/10.4236/ce.2022.134090

Foertsch, M. (1998). *A study of reading practices, instruction, and achievement in District 31 schools*. Oak Brook, IL: North Central Regional Education Laboratory. Available online: http://www.ncrel.org/sdrs/areas/31abs.htm. https://doi.org/10.4236/ce.2023.144042

Hamaker, C. (1986). The effects of adjunct questions on prose learning. *Review of Educational Research, 56*(2), 212-242. https://doi.org/10.4236/ojml.2023.132016

Olatoye, R. A., & Aderogba, A. A. (2012). Harnessing the Power of Emotional Intelligence, Scientific Literacy and Problem-Solving Skills for Successful Living. *Pacific Journal of Science and Technology, 13*(1), 403-417. https://doi.org/10.4236/jss.2014.22007

Olatoye, R.A. (2002). *A Causal Model of School Factors as Determinants of Science Achievement in Lagos State Secondary Schools*. An Unpublished PhD Thesis, University of Ibadan, Nigeria.

Pearson, P. D., & Fielding, L. (1991). Comprehension instruction. In R. Barr, M. Kamil, P. Mosenthal, & P. Pearson (Eds.), *Handbook of reading research* (pp. 815-860). White Plains, NY: Longman. https://doi.org/10.4236/am.2016.712119

Pearson, P. D., & Fielding, L. (1991). Comprehension instruction. In R. Barr, M. Kamil, P. Mosenthal, & P. Pearson (Eds.), *Handbook of reading research* (pp. 815-860). White Plains, NY: Longman.

Smith, P. (1994). *Learner self-assessment in reading comprehension: The case for student-constructed tests*. University of American Puebla. https://doi.org/10.4236/ojml.2017.72006

Smith, P. (1994). *Learner self-assessment in reading comprehension: The case for student-constructed tests*. University of American Puebla.

Snow, C. (2002). *Reading for understanding: Toward an R&D program in reading*. https://doi.org/10.4236/ce.2020.118095

Snow, C., Burns, M., & Griffin, P. (1998). *Preventing reading difficulties in young children.* Washington, DC: National Academy Press.

Spearman, C. (1927). *The Abilities of Man. New York: MacMillan*. York: Simon & Schuster Macmillan. https://doi.org/10.4236/crcm.2022.118046

Copyrights