

CHAPTER IV

DATA PRESENTATION AND DISCUSSION

This chapter presents the analysis and the interpretation of the data from the pilot-testing, the computation of the pre-test score, the computation of the post-test score, the computation of the experimental group score, and the computation of the control group score. The analysis of those sources covers a research questions as stated in chapter III, “Is Total Physical Response storytelling effective in improving English vocabulary mastery of fourth graders of an elementary school?”. It discusses a quantitative data analysis.

This chapter is presented based on data collection techniques and a research question in order to provide comprehensive discussion and valid conclusion (Emilia, 2008:204). Data presentation and discussion of this study is presented simultaneously in this chapter. The chapter is divided into five main parts: data and discussion from the pilot-testing, data and discussion from the computation of pre-test score, data and discussion from the computation of post-test score, data and discussion from the computation of the experimental group score, moreover data and discussion from the computation of the control group score. The findings are discussed, interpreted, and compared with theoretical foundations as presented in chapter II. Meanwhile, the discussion section elaborates the findings to answer the research question.

4.1 Data and discussion from the pilot-testing research instrument

4.1.1 The item analysis

Before conducting the pre-test, the test items which were used in the pre-test were tried out to investigate the element of a good test; such as, validity, difficulty index, discrimination index and reliability of the test items (Arikunto, 2003). The try out test was conducted on 26th of April 2010. In trying out the instrument, the vocabulary achievement test which consisted of 20 multiple-choice test items and 20 instruction test items was given to student from another class that was not involve in this study; they were 20 students of fourth graders of SDN Raya Barat 1.

There were some steps in pilot-testing research instrument, described as follow: (1) arranging try out score of the students, (2) calculating the validity and reliability of the instrument, (3) determining the difficulty index of each item, and (4) determining the discrimination index of each item. The computation was done by Microsoft Excel and SPSS 16.0 for Windows Program.

The calculation of the validity of the test item was gained by using *Person product moment* correlation at level of significance 0.05. After r coefficient correlation value was calculated, the t value was calculated. And then compared the t value to the t table, if t value > t table it means that the item is valid and if the t value is < t table, the item is not valid. The result of the computation on the pilot testing can be seen in the following table:

Table 4.1
The Result of Item Analysis

	Item numbers	Score	Interpretation
Validity	1,3,5,6,7,8,9,10,11,12,14,15,20,21,25,26,29,30,31,32,33, 35,36,39,40,	> 0.444	Valid
	2,4,13,16,17,18,19,22,23,24,27,28,34,37,38.	< 0.444	Invalid
Difficulty	4,13,27,28,29,40	0.00-0.30	Difficult
Index	1,3,5,6,8,9,11,12,16,17,18,20,23,24,25,26,30,31,32,33,34,35,36,37,39.	0.30-0.70	Moderate
	2,7,14,15,19,21,22,38.	0.70-1.00	Easy
	31,33.	0.70-1.00	Excellent
Discrimination index	1,3,11,18,20,24,26,35,36,39.	0.40-0.70	Good
	9,10,12,15,19,21,22,23,25,29,30,32,37,38,40.	0.20-0.40	Moderate
	6,7,8,13,14,17,27,28,34,	0.00-0.20	Poor
	2,4,5,16.	<0.00	Bad

Note: the raw score of validity is 0.444. It is obtained from $r_{critical}$ in the table of Pearson Moment Correlation with $N=20$, and the level significance of 0.05.

Table 4.1 displayed the results of the validity, difficulty index, and discrimination index of the test items. There were 25 items (1,3,5,6,7,8,9,10,11,12,14,15,20,21,25,26,29,30,31,32,33, 35,3,39,40.) that was used in the pre-test as the research instrument. Whereas, the rest of the items (2,4,13,16,17,18,19,22,23,24,27,28,34,37,38.) were not appropriate to be used as the research instrument. The details on validity, difficulty index, and discrimination index of the test items can be seen in the appendix.

In addition, the test items mentioned above were not only adapted from *Storytelling with Children* book but also were composed based on the standard in Indonesia national curriculum of teaching English for fourth graders of elementary school. The table which shows the competences and indicators of the test items can be seen in chapter III.

After conducting the validity, difficulty index, and discrimination index of the test items, calculation of reliability of the test items was done. In calculating reliability of the test items, Spearman-Brown formula was used. The reliability value for the instrument of the study was 0.719 (Spearman-Brown Coefficient Equal Length). Meanwhile, $r_{critical}$ (2-tailed) at the level significance of 0.05 with $N=20$ is 0.444. It means that the research instrument has moderate reliability so that it is

appropriate to be used as the research instrument. The result of reliability of the instrument obtained by computing the data which was the result of split half method into SPSS 16.0. The table of the result reliability using SPSS 16.0 is presented in the appendix.

4.2 Data and discussion from the computation of the pre-test score

The pre-test was conducted to control group and experimental group on May 11th, 2010. This test focused on identifying the prior vocabulary knowledge of the students. It consisted of 25 items (13 multiple choice tests and 12 instruction tests) that were arranged after analyzing try-out test. As stated in chapter III, in the pre-test, students have to choose one correct answer of the four options based on picture that was presented in each sentence. And they have to do the right physical response based on teachers' instruction. The pre-test score of the students from the experimental and control group was gained from the students' raw score which was transformed into scale scores 0-100. The scores of both groups can be seen in the appendix.

In analyzing the pre-test score of experimental and control group, independent t-test formula was used to compare the means of both groups. Before t-test was performed, the data from experimental and control group pre-test have to be normal and homogeneous so that the calculation of the normal distribution and homogeneity of variance could be performed.

4.2.1 Normality distribution test

The first step in testing the normality of pretest score was stating the hypothesis as follows:

Ho: the scores of the experimental and the control group are normally distributed.

After stating the hypothesis, the next step was calculating the result of normality test by using Kolmogorov-Smirnov test at level of significance (0.05). Table 4.2 below shows the result of normality test in pretest data score for both experimental group and control group.

Table 4.2

The result of normality distribution test

Groups	Kolmogorov-Smirnov ^a			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
Score Experimental	.104	30	.200*	.975	30	.673
Control	.124	30	.200*	.962	30	.343

a. Lilliefors Significance Correction

*. This is a lower bound of the true significance.

From the table 4.2, it could be seen the significance value of both the experimental group and the control group was 0.20. Since the significance value for both groups were higher than the level of significance (0.05), it could be concluded that the score of experimental and control groups were normally distributed. In other word, the null hypothesis was accepted.

4.2.2 Variance homogeneity test

In testing the homogeneity, firstly, the hypothesis was stated as follows:

Ho: The variances of the experimental and control group are homogenous.

The next step was calculating the result of homogeneity test with the level of significance at 0.05. Table 4.3 below shows the Result of Homogeneity Test.

Table 4.3
The result of homogeneity test

	Levene Statistic	df1	df2	Sig.
Score Based on Mean	.541	1	58	.465
Based on Median	.324	1	58	.571
Based on Median and with adjusted df	.324	1	57.908	.571
Based on trimmed mean	.484	1	58	.489

Table 4.3 described the significance value of the test was 0.465. Since the significance value is higher than the level of significance (0.05), it could be concluded that the students' score on the pretest had homogenous variances. In other word, the null hypothesis was accepted.

4.2.3 t-test Analysis

After analyzing the normality distribution and homogeneity of variances, the data were analyzed by using t-test formula. Independent t- test formula in SPSS 16.0 for windows was used to analyze the significant differences between the pretest means of experimental and control groups. Firstly, the hypothesis was stated as follows:

H_0 : there is no significant difference between means of experimental and control groups.

The table 4.4 below shows the calculation result.

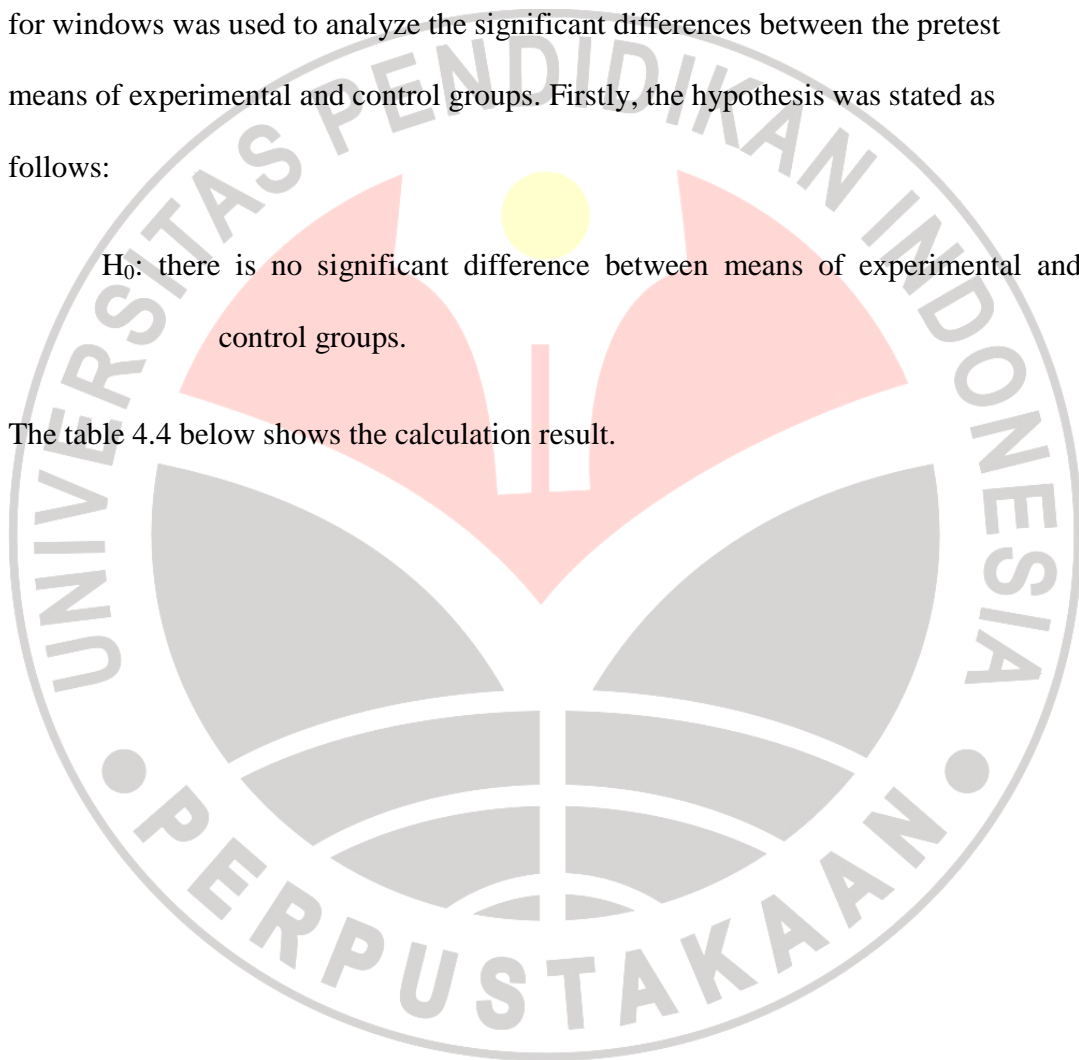
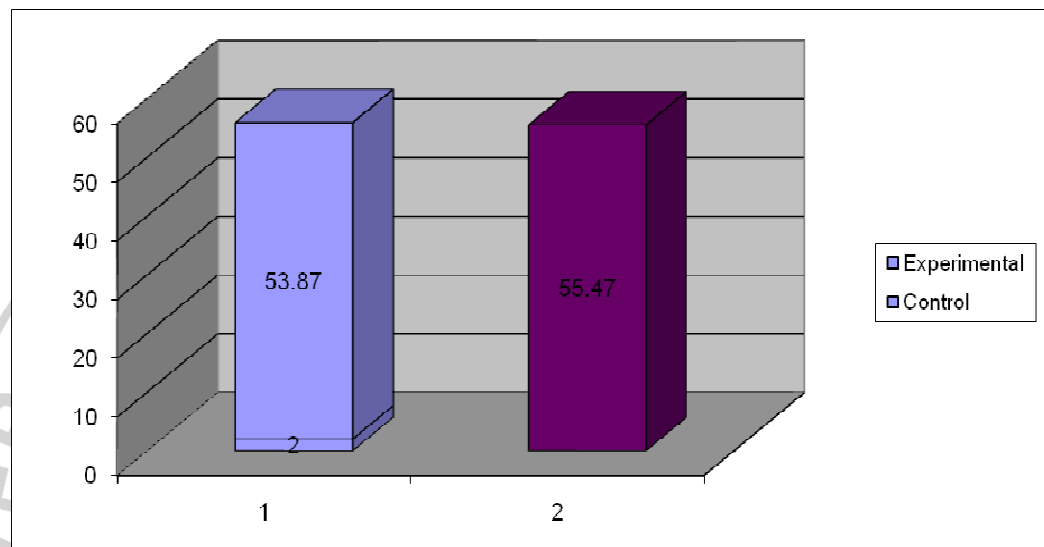


Table 4.4
The result of t-test on the pretest

Independent Samples Test										
		Levene's Test for Equality of Variances		t-test for Equality of Means						
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
									Lower	Upper
Pretest	Equal variances assumed	.556	.459	-.357	58	.722	-1.600	4.483	-10.574	7.374

Chart 4.1

The Graphic Bar of Mean of Experimental and Control Group in Pre-test



4.3 Data and discussion from the computation of the post-test score

The post-test was administered to control group on May 24th, 2010 and to experimental group on May 29th, 2010. The test was carried out to examine the progress of the student's vocabulary mastery after conducting treatments. By comparing the result of the post-test and the pre-test score, it was known the difference exists between experimental groups' score which used TPR storytelling method and control groups' score which used traditional method. The materials of post-test were similar with pre-test. The post-test score of the students from the experimental and control group was gained from the students' raw score which was

transformed into scale scores 0-100. The scores of both groups can be seen in the appendix.

The procedure of posttest data analysis was quite similar with the pretest data analysis. First, the normality distribution of both experimental and control group was determined. After that, the homogeneity variance was determined. finally, t-test formula was conducted to test the hypothesis.

4.3.1 Normality distribution test

The first step in testing the normality of pretest score was stating the hypothesis as follows:

Ho: the scores of the experimental and the control group are normally distributed.

After stating the hypothesis, the next step was calculating the result of normality test by using Kolmogorov-Smirnov test at level of significance (0.05). Table 4.2 below showed the result of normality test in pretest data score for both experimental group and control group.

Table 4.5

The result of normality distribution test

Tests of Normality

Group		Kolmogorov-Smirnov ^a			Shapiro-Wilk		
		Statistic	df	Sig.	Statistic	df	Sig.
Score	Experimental	.154	30	.067	.957	30	.260
	Control	.109	30	.200*	.961	30	.319

a. Lilliefors Significance Correction

*. This is a lower bound of the true significance.

From the table above the significance value of the experimental group was 0.067 and the control group was 0.20. Since the significance value for both groups were greater than the level of significance (0.05), it could be concluded that the score of experimental and control group was normally distributed. In other word, the null hypothesis was accepted.

4.3.2 Variance homogeneity test

The next step after calculating the normality test, was analyzing the variance of homogeneity test. Firstly, the hypothesis is stated as follows:

H_0 : the variance of control and experimental groups are homogenous

Table 4.7 below showed the result of homogeneity of variance for post-test which was analyzed by using Levene test SPSS 16.0 program for windows.

Table 4.6
The Result of Homogeneity of variance

Test of Homogeneity of Variance

	Levene Statistic	df1	df2	Sig.
Score Based on Mean	2.162	1	58	.147
Based on Median	1.754	1	58	.191
Based on Median and with adjusted df	1.754	1	57.937	.191
Based on trimmed mean	2.180	1	58	.145

From the table above showed that the significance value of variance homogeneity test was 0.147. Since the significance value is greater than the level of significance (0.05), it could be concluded that the students' score on the posttest had homogenous variances. In other word, the null hypothesis was accepted.

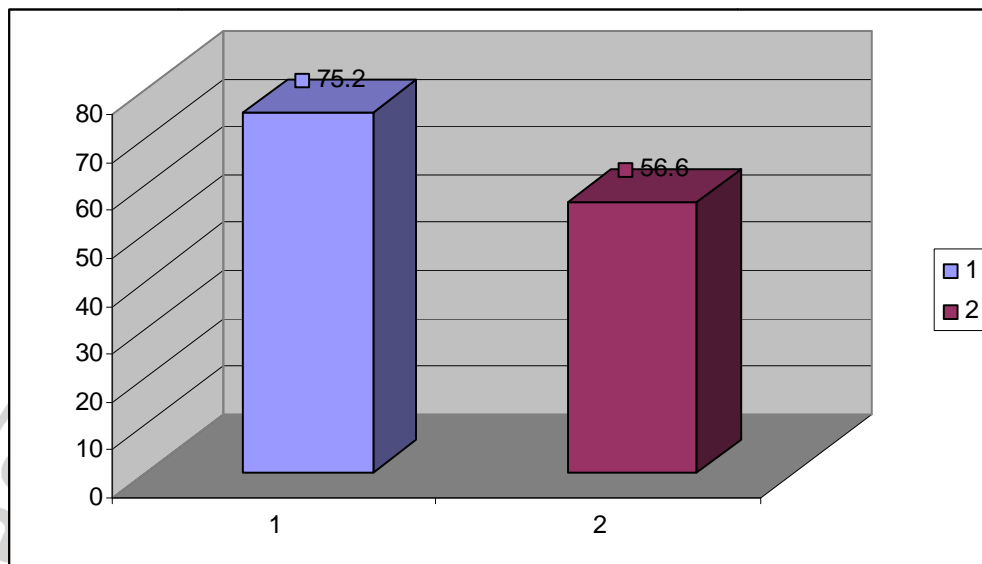
4.3.3 t-test analysis

Post-test	Equal variances assumed	.010	.919	4.612	58	.000	18.600	4.033	10.527	26.673
	Equal variances not assumed			4.612	57.999	.000	18.600	4.033	10.527	26.673

Table 4.7 showed the result of t-test. Based on the t table the $t_{critical}$ for $df = 58$ at the level significance 0.05 is +2.000 and -2.000. the table 4.7 showed that the t_{obtain} was +4.612 and p value was 0.000. since $-t_{obtain} < -t_{critical}$ and $+t_{obtain} > t_{critical}$ ($-4.612 < -2.000$ and $4.621 > 2.000$) and p value was lower than 0.05 ($0.000 < 0.05$). It means there was significance different between the pretest mean value and post test mean value of the class. The difference of means of pretest of experimental and control groups could be seen in chart 4.2 below:

Chart 4.2

The Graphic Bar of Mean of Experimental and Control Group in post-test



4.3.4 Effect Size

The effect size was calculated to find out how strongly the independent variable affected the dependent variable. The result of the computation was presented below:

$$r = \frac{t^2}{t^2 + df}$$

$$r = \frac{(4.612)^2}{(4.612)^2 + 58}$$

$$r = \frac{21.270544}{79.270544}$$

$$r = \sqrt{0.268}$$

$$r = 0.518$$

Thus, from the result above, it could be seen that a very large effect size was observed. In other word, there was a great effect of TPR Storytelling method in improving students' vocabulary mastery.

4.4 Data and discussion from the computation of the experimental group's score

The aim of the research is to find out the effectiveness of Total Physical Response storytelling of an elementary school in teaching vocabulary. As stated in chapter III, in this research, there are two groups, namely experimental and control group. The TPR storytelling was given only to the experimental group. To accomplish the aim of the research, therefore, the research used paired sample t-test formula in SPSS 16.0 for windows to observe whether there is significant difference of experimental group and control group. Before comparing the two groups, the computation of the experimental group scores was done. It was aimed to compare the score of pre-test and post-test in the experimental group. The result of the statistical computation is as follows:

On average, the participants of experimental group got significant effect of the use of TPR method on their vocabulary mastery ($M = 77.20$, $SD = 13.084$, $SE = 2.389$) rather than their score on the pretest, in which they had not yet experienced learning English using TPR Storytelling method.

Table 4.4

The Result of the Experimental Group's Score Computation

	Paired Differences					t	df	Sig. (2- tailed)
	Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference				
				Lower	Upper			
Pair posttest 1 - pretest	-21.333	13.629	2.488	-26.422	-16.244	-8.574	29	.000

Table 4.4 showed the result of paired sample t-test. Based on the t table the $t_{critical}$ for $df = 29$ at the level significance 0.05 is +2.045 and -2.045. the table 4.7 showed that the t_{obtain} was -8.574 and p value was 0.000. Since $-t_{obtain} < -t_{critical}$ and $+t_{obtain} > t_{critical}$ ($-8.547 < -2.045$ and $8.547 > 2.045$) and p value was lower than 0.05 ($0.000 < 0.05$). It means there was significance different between the pretest mean value and post test mean value of the class. It indicated that there was significant

difference between students' pretest and posttest scores in experimental group. Thus, it could be said that the students in experimental group experienced an improvement after the treatment.

The improvement of student's vocabulary mastery was affected by the use of TPR Storytelling method which involves three main aspects. The first aspect is dealing with stories selection, the teacher as the storyteller had to pay attention to certain aspects from both the students' side and the storyteller's side. From the students' side, the teacher should choose appropriate stories for the students. Teacher needs to choose stories which will engage the children. Moreover, the stories should have a rich experience of language to children and lack of long descriptive passage (Wright, 1995). In the classroom, the writer used some stories which have more dialogue than descriptive passage, so that, children did not get bored and they could understand the story easily. From the storyteller side, besides considering the appropriateness of the stories to the students, the storyteller choose stories which she likes and she felt she could tell well (Wright, 1995), so that, the storyteller could involve her emotion in each story.

The second aspect is dealing with story aids, this research tried to meet the students' characteristic as children who are in concrete operational stage of intellectual development that is sensory input (Brown, 2001). Therefore, teacher should bring the visual and auditory modes that are sufficient for a classroom because in this stage they can not generalize their understanding. In the classroom the writer

used some media such as flash card and puppet to help the students comprehend the stories easily. Flashcard and puppet that are involved into visual media seemed to give benefits to the students' learning. Those media could attract and lead the students to pay more attention to the story being told. Besides, the flashcard and puppet were able to illustrate context of what was being told as well. As a result, the students could optimally grasp and remember the words conveyed through flashcard and puppet.

The third aspect is dealing with storyteller. Before delivering stories, storyteller should have a lot of practice in using vocabulary. The writer practiced telling the stories many times before delivering them to the students. She was concerned with pronunciation in order not to transfer wrong meaning of the introduced words to the students. Besides, she practiced by using facial expression, gestures, and other story aids for each story to avoid stiffness when telling the stories in the classroom.

Above all, the experiment worked quite well by following the guidance and tips recommended by some experts in chapter II. It seemed that good atmosphere was created in the classroom. Most of the students actively participated during the treatment without being a shamed and being afraid of making any mistake. They seemed to be happy to learn English by TPR storytelling. Moreover, in teaching process, TPR storytelling make students interested and enthusiastic to follow the teaching and learning process because TPR storytelling made the students free from stressful situation, so that they motivate to be engaged actively in the lesson. Besides,

the students not only enjoyed the learning process but also were able to do the teacher's instruction well.

4.1.5 The computation of the control group's score

The procedure of the computation of the control group scores was quite similar with the computation of the experimental group's score. The computation was done by using paired sample t-test in SPSS 16.0. It was aimed to compare the score of pre-test and post-test of control group's. The result of the statistical computation was presented in the table in appendix.

On average, the post-test score of control group was ($M = 55.47$, $SD = 18.033$, $SE = 3.287$) and pre-test score of experimental group ($M = 56.6$, $SD = 16.700$, $SE = 3.049$). It indicates that the score of the group is constant because they did not experience learning English using TPR Storytelling method.

Table 4.14

The Result of the Experimental Group's Scores Computation:

Paired Samples Test

	Paired Differences					t	df	Sig. (2- taile d)
	Mea n	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference				
				Lower	Upper			
Pair 1 pretest - posttest	- 1.13 3							
		10.683	1.950	-5.122	2.856	-.581	29	.566

Moreover, the table 4.8 showed that the probability was greater than the level of significance ($0.566 > 0.05$) and t_{obt} (0.581) was less than t_{crit} at the 0.05 level significance to the line $df = 29$ (2.045). It indicated that there was no significant difference between students' pretest and posttest scores in control group because they did not experience the treatment.

In summary, quantitative data analysis (in t-test and effect size computation) indicated a significant development of students' vocabulary mastery at the end of the teaching program. The finding lead to the conclusion that there was an improvement on the students' vocabulary test score after the implementation Total Physical Response storytelling were able to help students improve their vocabulary mastery.

In addition the finding supports previous research on the effectiveness of Total Physical Response storytelling to improve students' vocabulary mastery. This is in line with the study that was executed by Michael K. Brune (2004), he found that TPR storytelling was very effective in teaching foreign language to young learners; McKay (2000) In his research, he found that TPR storytelling method was very effective in increasing English vocabulary; Sumiati (2006) who investigate the effectiveness TPR storytelling in teaching vocabulary to second grade students of junior high school.

