CHAPTER III
METHODOLOGY

Methodology is covered in this chapter. It described the instrument, research design, population, research instrument, data collection methods, data analysis techniques, and research methodology used in this study.

3.1 Research Deisgn

The researcher, in this study conducted quantitative research. It was reasonable to use a quasi-experimental design; according to Arry (1985; p. 26), a research study using a quasi-experimental design involves the modification and control of one or more independent variables, followed by observation of the pendents variable for change that is dependent with the impact of the dependent variables (Hidayati et al., 2019b). The researcher's use of this quasi-experimental design is to know the effect of using storytelling techniques through Thai folktales to improve students vocabulary in a Thai EFL elementary school. By using this quasi-experimental design, the researcher's purpose is to find variables involving the control group as well as the experimental group. This can provide insights into the real-life effectiveness of the intervention as it reflects the diverse educational contexts in which students are taught.

A comparison group that is as close to the group receiving treatment as possible is known as a quasi-experimental design. A pre-test was given to both the experimental and controlled groups of students at Ban Elert elementary school in Phukradeung, Loei province, Thailand, to measure students vocabulary level before applying a treatment. Next, the treatment is given to the experimental group by using Thai folktale stories with storytelling techniques. Then, the control class will tell the English story in general using storytelling techniques. Both groups receive the post-test following the completion of the treatment. The following are the procedures used in this quasi-experimental design:

1. Pre-test

    The pre-test was conducted in the first meeting to measure student level in vocabulary before the treatment. The researcher gave a matching item, the clozest test, chose true and false, and created sentences.

2. Tretament

The two group were given a different trearment. It could be explained as follows as:

a. Experimental class:

In this group the teacher given the students some stories text about folktales, which will be given in written text on the paper. The students were taught some important ways to retell the folktales story by using storytelling techniques. The Thai folktales that students will retell are "*Horse-Faced Woman*" and "Frightened *Rabbit.*"

b. Controll Group:

In this group the teacher only shown and explained about the English story in general, namely "*The Bowman and Lion*" and "*Sleeping Beauty*." Then the students are given some exercises, such as reading the narrative text about those English stories in general, and work on some exercises that the teacher has given them.

c. Post- test

The pos-test was conduct in the last meeting to measure students vocabulary masteres after getting the tretament. The researcher gave a test again, similar to the pre-test, the students both in the experimental and control classes were test and asked some questions, but the topic used was the folktales that the teacher had given before.

3.2 Population and Sample

3.2.1 Population

This subject in this research is students at Ban Elert elementary school in Phukradeung, Loei, Thailand. The participants of this research were the students in Ban Elert of 4th grade A & B elementary schools in Thailand. The reseracher took two class 4A, as a experimental group and 4B as a controlled group. Here are the population table of grade 4A & B in Ban Elert Elementary school Thailand.

CLASS	NUMBER OF STUDENTS
4 A	22 Students

4 B	22 Students
Total	*44 Students*

As you can see from the table, there are only 22 students with whom the writer took part in this research. Because of the reality in the Thai school, the class that provides a lot of students is in grades 4A and 4B. Another class, like in grades 5 and 6, is less than 20. It also happens at Juniour High School, which only provides 18 students in each class. So that's why the writer took Grades 4A and 4B as subjects for this research.

3.2.2 Sample

In this research, for the sample, the researcher applied purposive sampling (sampling considerations). In which the researcher takes two class, forst class for the experimental class, which is grade 4A, and second class for the control class, which is grade 4B. In each class, they have 22 students, so the total of those two groups is 44 students for the sample that the researcher used for this research.

3.3 Data Collection

The data were collected using pre- and post-tests. A pre-test was given to the two groups before treatment in order to know and measure the students background vocabulary related to activities, animals, and places (matching items, the clozest test, choosing true and false, and creating sentences). The test will be given to both experimental and controlled classes. Moreover, after the pretest, tretament will be implemented in the experimental class in order to teach vocabulary through retelling the folktale story. The given treatment involved implementing folktale stories using storytelling techniques to teach vocabulary twice. The researcher uses "*Horse-Faced Woman*" and "*Frightened Rabbit*." The students will feel familiar with the story that the researcher chose, the Thai folktales.

During treatments, the controlled class was taught another English story in general, namely "*The Bowman and Lion*" and "*Sleeping Beauty*," using storytelling techniques without any treatment as an experimental class. In other words, a controlled class will be taught without using Thai folktale stories but using English stories in general. After giving

some treatments, a post-test was given in order to measure the ability of students to understand the vocabulary and content of the story that had been given. to give a test (the test is the same as in pre-test questions) after giving different treatments to experimental and controlled classes.

3.4 Research Instrument

An instrument is a tool to gather data for a study. According to Creswell (2011), research instrument is the general term for an instrument for measuring, observing, or documenting quantitative data (Hidayati et al., 2019b). The instruments used in this study were tests that consisted of pretest and post-test forms. Tests are useful tools for measuring knowledge in educational research. As a result, the test plays a crucial part in gathering data.

There are a wide range of methods available for evaluating vocabulary, including resolved times, matching objects, cloze tests, multiple choice, and more. The researcher of this study employed multiple choice, cloze and matching item vocabulary exams. The topic of the pre-test is about describing the world around us (students), and the topic of the post-test is about folktale stories that the teacher has given before. The topic of the pre- and post-test for Ban-Elert students in elementary school in Thailand is about places, activities, and animals. The folktales that the researcher uses here are "*Horse-Faced Woman"* and "*Frightened Rabbit"* and other stories in English in general, namely "*The Bowman and Lion"* and "*Sleeping Beauty"*.

Pre- Test & Post- Test	
Experimental- group	Controll- group
o Pre- test: students give assigment (Matching item, Clozest test, Chose true and false, and creating a sentences) the topic is *Describing Around Us (Place, Activities, and Animals)*. ▪ Give the *Treatments:* The teacher taught the students narrative text about folktales story is will given in written text or listen	o Pre- test: students give assigment (Matching item, Clozest test, Chose true and false, and creating a sentences) ▪ No need to give *Treatment* In control class, the teachers will teach and explain about narrative text, but here the students will give another story of English in general, namely "The Bowman and the Lion"

from the speaker. The the students were taught some important ways to retell the folktales story by using storytelling techniques. The Thai folktales that studnts will retell is horse faced woman story, Frightened Rabbit, The bowman and The lion, and Sleeping Beauty.

- This research in experimental classes will be held twice: in the first meeting, using "Horse Faced Woman," and in the second meeting, using "Frightened Rabbit Stories. This sotries is from Thai folktales stories.

- Each meeting will have 45 minutes.

o Post- Test

The researher give a test again such as in pre- test. The students give a written text, and try to complete the missing word, after that they will retell the story after finish complete the task.

and "Sleeping Beauty." Students are also given some oral exercises, such as reading the narrative test about English stories in general, and they will still retell the stories too.

- This research in the control class will be done twice. In the first meeting, they will learn " The Bowman and The Lion" stories, and in the second meeting, they will learn " Sleeping Beauty". This stories is from English storues in general.

- Each meeting will have 45 minutes.

o Post- Test

The researher give a test again such as in pre- test. The students give a written text, and try to complete the missing word, after that they will retell the story after finish complete the task.

3.4.1 Instrument Validity Test

An instrument is distributed in valid if the instrument can appropriately measure what it wants to measure. The validity test of this study was calculate using SPSS.

There are other events that are simpler and easier to use interpretation of the coefficient of correlation obtained, or the value of r. The interpretation is as follows:

3.1 Interpretation Table

Magnitude Of r Value	Interpretation
Antara 0,800 sampai dengan 1,00	Very High
Antara 0,600 sampai dengan 0,800	High
Antara 0,400 sampai dengan 0,800	Medium
Antara 0,200 sampai dengan 0,400	Low
Antara 0,000 sampai dengan 0,200	Very Low (Does Not Correlate)

(Arikunto, 2013, hlm. 319)

The instrument validity test used is a multiple-choice instrument totaling 40 questions. The questions were tested in grades 4A and 4B according to the number of students in each class a total of 22. The following are the results of the instrument question trial:

3.2 Results of Question Trial Validity

Soal	Pearson Correlation	Sig	Valid	Interpretasi
Soal_1	0,258	0,140	Valid	Low
Soal_2	0,485	0,004	Valid**	Medium
Soal_3	0,428	0,012	Valid*	Low
Soal_4	0,393	0,022	Valid*	Low
Soal_5	0,258	0,140	Valid	Low
Soal_6	0,097	0,584	Invalid	Very Low
Soal_7	0,514	0,002	Valid**	Medium
Soal_8	0,622	≤ 0,001	Valid**	Medium
Soal_9	0,002	0,990	Invalid	Very Low
Soal_10	0,328	0,058	Valid	Low
Soal_11	0,545	≤ 0,001	Valid*	Medium
Soal_12	-0,103	0,561	Invalid	Very Low
Soal_13	-0,050	0,777	Invalid	Very Low
Soal_14	-0,024	0,893	Invalid	Very Low
Soal_15	0,002	0,990	Invalid	Very Low
Soal_16	0,016	0,928	Invalid	Very Low
Soal_17	0,286	0,100	Valid	Low
Soal_18	0,129	0,467	Invalid	Very Low
Soal_19	0,072	0,684	Invalid	Very Low
Soal_20	0,463	0,006	Valid**	Medium
Soal_21	0,445	0,008	Valid**	Medium

Soal_22	0,055	0,757	Invalid	Very Low
Soal_23	0,149	0,400	Invalid	Very Low
Soal_24	0,339	0,050	Valid*	Medium
Soal_25	0,003	0,986	Invalid	Very Low
Soal_26	0,188	0,288	Invalid	Very Low
Soal_27	0,497	0,003	Valid**	Medium
Soal_28	0,467	0,005	Valid**	Medium
Soal_29	0,372	0,030	Valid*	Medium
Soal_30	0,345	0,045	Valid*	Medium
Soal_31	0,219	0,212	Valid	Very Low
Soal_32	0,337	0,051	Valid	Very Low
Soal_33	0,541	≤ 0,001	Valid**	Medium
Soal_34	0,416	0,015	Valid*	Medium
Soal_35	0,375	0,029	Valid*	Medium
Soal_36	0,022	0,903	Invalid	Very Low
Soal_37	0,055	0,757	Invalid	Very Low
Soal_38	0,193	0,275	Invalid	Very Low
Soal_39	0,156	0,378	Invalid	Very Low
Soal_40	0,055	0,757	Invalid	Very Low

The question item can be said if r counts ≥ r table. If the Sig value ≤ 0.05 then the question instrument can be said to be "*Valid*". If the Sig value ≥ 0.05 then the instrument can be said to be "*Invalid*". It can be seen from the table above that the question items are calculated > r table and are said to be valid totaling 16 questions.

3.4.2 Instrument Reliability Test

According to Sugiarto and Situnjuk (2006),), a reliability test determines if the data collecting tool used to gather the information can be relied on to provide accurate results. When a respondent consistently answers the same questions on an instrument all through time, it's considered reliable.

3.3 Table Clasification Reliability

Range	*Category*
0,86 - 100	Very High
0,66 – 0,85	High
0,36 – 0,65	Low
0,20 – 0,35	Very Low
0,00 – 0,19	Unreliable

(Source: Creswell:2012-347)

22

The Effect Using Storytelling Techniques to Enhance Students Vocabulary in Thai EFL Elemantary School-Anisa Dwi Fitria 2023

After being tested for validity by producing 16 valid questions, the question items were tested again with a reliability test. The following are the results of the reliability test of the question items:

3.4 Question Item Reliability Test Results

Reliability Statistics

Cronbach's Alpha	Cronbach's Alpha Based on Standardized Items	N of Items
.800	.803	16

Summary Item Statistics

	Mean	Minimum	Maximum	Range	Maximum / Minimum	Variance	N of Items
Item Means	.789	.618	.912	.294	1.476	.006	16

From the table data above, we can calculate the reliability of the 40 questions that have been identified as valid, namely, 16 questions.

3.4.3 Differentiating Power Test

The distinguishing power of the question is the ability of a problem to distinguish between high-ability students and low-ability students (Sundayana, 2016, p. 76). In calculating differentiating power test the writer using SPSS. Here the classification to make the basic decision about differentiating power test.

3.5 Differentiating Power Classification Table

Range	Category
DP ≤ 0,00	Very Bad
0,00 < DP ≤ 0,20	Bad
0,21 < DP ≤ 0,40	Enough
0,41 < DP ≤ 0,70	Good
0,71 < DP ≤ 1,00	Very Good

(Sundayana, 2016, hlm. 77)

The following are the results of the instrument differentiating power test. The test question items are adjusted to the vulnerability of the differentiating power classification:

3.6 Table of Distinguishing Power Test Results of Question Items

Question Number	R count (Out SPSS)	The Power of Different Question Points
2.	0,485	Good
3.	0,428	Good
4.	0,393	Enough
7.	0,514	Good
8.	0,622	Good
11.	0,545	Good
20.	0,463	Good
21.	0,445	Good
24.	0,339	Enough
27.	0,497	Good
28.	0,467	Good
29.	0,372	Enough
30.	0,345	Enough
33.	0,541	Good
34.	0,416	Good
35.	0,375	Enough

From the differentiating power table above, we can know the value of differentiating power and its interpretation in 16 test questions. We can know that there are 5 points of questions with sufficient distinguishing power interpretation, and there are 11 points of questions with good discriminating power interpretation. So from the 16 questions tested from the differentiating power, we can use 16 questions for test trials with notes that the 5 questions can be used but with a note of improvement in the questions.

3.4.4 Difficulty Test

The level of difficulty is a question item that indicates the percentage of students who correctly answered the question item presented. The difficulty level is calculated by using SPSS. Here the classification table to making basic category in calculate difficulty test.

3.7 Question Item Difficulty Classification Table

Rentang	Ctegory
0,00 – 0,15	Very Difficult

0,16 - 0,30	Difficult
0,31 - 0,70	Enough
0,71 – 0,85	Easy
0,86 – 1,00	Vry Easy

The following are the results of the calculation of the difficulty level of the test instrument test questions that are adjusted to the classification of difficulty levels:

3.8 Test Table Difficulty Test Questions

Question Number	Mean (OutPut Spss)	Difficulty Level
2.	0,79	Easy
3.	0,76	Easy
4.	0,79	Easy
7.	0,88	Very Easy
8.	0,76	Easy
11.	0,62	Medium
20.	0,76	Easy
21.	0,76	Easy
24.	0,76	Easy
27.	0,74	Easy
28.	0,88	Very Easy
29.	0,91	Very Easy
30.	0,91	Very Easy
33.	0,79	Easy
34.	0,71	Easy
35.	0,76	Easy

After seeing the difficulty level of the trial questions, there are 11 questions that fall into the easy question category, 4 questions that fall into the very easy question category, and 1 question that belongs to the very easy question category.

3.5 Data Analysis

The writer have conducted the test, pretest and post-test. The data are compared from the mean of the score from pretest and post-test. After getting the data from the score, then the data were analyzed and processed by using statistic calculation of T-test formula with significance degree and gained scores. T-test in this research used to test the average difference count, was there a significant difference or not between the experimental group and the controlled group. The statistics used to analyze the data

were arithmetic mean score, standard deviation, and t-test. The analyzed results were then presented in a descriptive manner.

While the gained score is the difference between pretest and post-test score of each class of the experimental and the control groups. Gain scores are used to determine the increase or decrease in scores and to determine the effectiveness of the strategy used. However, prior to the tests the hypothesis necessary analysis prerequisite tests first, namely the distribution *normality* test and *homogeneity* test.

3.5.1 Data normality test

The normality test is used to determine whether the amount of data is normally distributed or not. The normality test can be performed using the Chi-squared formula (Sugiyono, 2016, p. 172). The following is the decision-making process for the normality test according to Kolomogrov-Smirnov using SPSS:

1. If the significance value (sig) on the Based on Mean > 0.05 then the research data is *Normal* distributed.
2. If the significance value (sig) on Baesd on Mean < 0.05 then the research data is distributed *Not Normal*.

The following are the results of the normality test on the results of the data in the pre-test and post-test, namely:

3.9 Pre-Test and Post-Test Normality Test Table

Tests of Normality

	Kelas	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
		Statistic	df	Sig.	Statistic	df	Sig.
Hasil Belajar Siswa	PreTest Eksperimen	.143	22	.200[*]	.959	22	.465
	PostTest Eksperimen	.154	22	.187	.922	22	.083
	PretTest Controll	.151	22	.200[*]	.902	22	.033
	PostTest Controll	.128	22	.200[*]	.964	22	.580

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

a. Lilliefors Significance Correction

It can be seen from the table data above that the value (sig) of the pre-test for the experiment class is 0.200, the post-test for the experiment class is 0.187, the pre-test

for the control class is 0.200, and finally the post-test for the control class is 0.200. From all the normality data that we have tested through SPSS that can be interpreted as student learning outcomes in pre-test and post-test in the experimental class and control class, the distribution is "*Normal*".

3.5.2 Test homogeneity of Data

One of the test requirements that must be met in the use of parametic statistics is the homogeneity test. The homogeneity test is carried out to determine whether or not the variance of samples taken from the same population is the same (Arifin, 2014, p. 286). In this study, researchers used SPSS to test homogeneity data from pre- and post-test activities, following the steps of the homogeneity test using SPSS according to Santoso (2017, p. 203).

The following is the basis for decision-making in the homogeneity test, namely:

3. If the significance value (sig) on the Based on Mean > 0.05 then the research data is homogeneously distributed.
4. If the significance value (sig) on Baesd on Mean < 0.05 then the research data is distributed Inhomogeneously.

3.9 Table of Results From Homogeneity Tests on Data Results in Pre-Test and Post-Test

Test of Homogeneity of Variance

		Levene Statistic	df1	df2	Sig.
Hasil Belajar Siswa	Based on Mean	.445	1	42	.509
	Based on Median	.243	1	42	.624
	Based on Median and with adjusted df	.243	1	37.115	.625
	Based on trimmed mean	.469	1	42	.497

It can be seen from the table above that Based on Mean has a value (sig) of 0.509, in Based on Median it has a value (sig) of 0.624, then in Based on Median and with adjusted df it has a value (sig) of 0.625, and finally in the column Based on Trimmed Mean has a value (sig) of 0.497. We can conclude from the homogeneity test table data that the data exceeds > 0.05, so the distribution in this study is "*Homogeneous*".

3.5.3 Idependent Test Sample T-test

If the data is *normal* and *homogeneous*, a parametic test in the form of a t-test is performed. Before calculating the t-test, the data obtained must be normally distributed. Because in this study it is normally distributed and homogeneous, it can use the independent sample T-test using parametics. This research to calculate the t-test using SPSS, according to Santoso (2017, p. 265).

The basis for decision-making for the independent sample t-test is as follows:

1. If the significance value (2-tailed) < 0.05, then Ho is rejected and Ha is Accepted
2. If the significance value (2-tailed) > 0.05, then Ho is accepted and Ha is rejected

3.10 Here's the Idependent Test Table T-test Sample

Group Statistics

	Model Pembelajaran	N	Mean	Std. Deviation	Std. Error Mean
Hail Belajar Siswa	Model Pembelajaran A	22	77.50	10.884	2.320
	Model Pembelajaran B	22	67.95	9.084	1.937

Based on the output table "Group Statistic" above, it is known that the number of learning outcomes for Groups A and B is 22. The average score of student learning outcomes for group A was 77.50, while for group B it was 67.95. Thus, from descriptive statistics, it can be concluded that there is a difference in the average learning outcomes of students between groups A and B. Here are some data results for the independent sample t-test.

3.11 The Table of Idependent Sample T-test

Hasil Belajar Siswa	Nilai sig (2-tailed)
Equal variances assumed	0,003
Equal variances not assumed	0,003

We can conclude from the data table The independent sample t-test refers to the basis of making the test decision that we get the value we get, which is 0.003. With the formula, if nilia sig (2-tailed) < 0.05, then there is a significant difference between learning outcomes in class A and learning outcomes in class B. It can be interpreted as saying that Ha is accepted and Ho is rejected.

So it can be concluded that, with the statement There is a significant effect using storytelling techniques through Thai folktales to enhance students vocabulary for elementary students in Thailand, then the answer is *"yes"*, there is a significant difference in using storytelling techniques through Thai folktales to enhance students vocabulary in Thailand.