III. RESEARCH METHODS

This chapter discusses the research design, population and sample, data collection and instrument, and the hypothesis testing.

A. Research Design

This research is a quantitative study in which one-group pretest-posttest design is applied. The research design is outlined as follows:

$$T1 \times T2$$

whereT1 refers to the pretest, T2 the posttest, and ×signifies the treatment (i.e. the application of TPS technique for teaching reading comprehension skill of narrative text) (Setiyadi, 2006:131). The pretest is used to find out the students' preliminary ability and the posttest to see whether there is an increase of students' reading achievement after the treatment ×. The pretest, treatment and posttest were conducted in the experimental class.

B. Population and Sample

The population of this research was the second-grade students of MTsN 1 TanjungKarang, Bandar Lampung. Out of the population, two classes were taken, one as the tryout class (Class VIII B), and one as the sample or the experimental

class (Class VIII A), with each class consisting of 30 students. The tryout class received no treatment and was used for instrument tryout only. The experimental or sample class was given the treatment, the pretest and the posttest. Both classes were selected through simple random probability sampling using lottery.

C. Data Collecting Technique

The instrument for data collection (i.e. the pretest and the posttest) wasa series of reading comprehension tests, each test consisting of a number of narrative texts and 30 multiple-choice items, and each item with four options (A, B, C, and D) that comprised one correct response and three distractors. The pretest and the posttest were similar, only the order of the texts and their corresponding items were rearranged for the posttest. The instrument for these two tests was previously qualified through the tryout test.

In collecting the data, the research adhered to the following steps:

1) Trying-Out the Instrument

The tryout test was administered toClass VIII B that consisted of 30 students. This test was given to identify the quality of the test before itwas used to obtain the data for the research. Thetest included 40multiple-choice items, each with four options (A, B, C, and D): one correct response and three distractors.

The items in the tryout test consisted of the following specifications:

Table 1. Table of Specificationof theTryoutTest

No.	Reading Specification	Item Number	Percentage
1	Determining main idea	2, 13, 27, 28, 29, 30, 32	17.5%
2	Inference	6, 12, 17, 19, 21, 22, 24, 36	20%
3	Reference	3, 9, 14, 15, 25, 34, 38, 40	20%
4	Finding detailed information	1, 4, 7, 8, 10, 11, 18, 23, 37, 39	25%
5	Vocabulary	5, 16, 19, 26, 31, 33, 35	17.5%
Total		**40**	**100%**

The larger number of items in the tryout test—i.e. 40 items—than that in the instrument for data collection (the pretest and posttest) was to guard against the presence of poor items, as well as to achieve higher reliability; thus, when such items exist and need dropping, we still have enough number of items to use for reliable data collection. Ten poor items in this tryout test were dropped after item analysis, leaving 30 items for the pretest and the posttest.

2) Conducting the Pretest

For the pretest, the researchtookClass VIII A that consisted of 30 students. The pretest was conducted before the treatment, and was intended for identifying to what extent the students had understood about main idea, detailed information, reference, inference and vocabulary—i.e. the aspects of reading comprehension. The pretest included 30 items that were left after item reduction through the tryout. The table of specification for the pretest is presented on the next page.

Table 2. Table of Specification of the Pretest

No.	Reading Specification	Items	Percentage
1	Determiningmain idea	10,23,24	10%
2	Inference	9, 18	6.67%
3	Reference	2, 7, 11, 12, 21, 25, 27, 29	26.67%
4	Finding detailed information	1, 3, 5, 6, 8, 14, 15, 17, 20, 26, 28, 30	40%
5	Vocabulary	4, 13, 16, 19, 22	16.67%
Total		**30**	**100%**

3) Conducting the Treatment

For the treatment, narrative texts were used as the reading materials. The treatment was conducted to the experimental class—Class VIII A—through the application of Think-Pair-Share technique. The treatment was conducted in three meetings as described below:

a) *The first meeting*, TPS technique was used to help the students in reading comprehension activity. The teacher was explaining about how to determine the main idea of a paragraph or of the whole story. The narrative story used as the material for the first meeting is *The Hermit*.

b) *The second meeting*, TPS technique was used to help the students in reading comprehension activity. The teacher explained about finding reference and identifying inference in reading comprehension. The narrative story used as the material for the second meeting is *The Wolf, the Kid, and the Goat*.

c) *The third treatment*, TPS technique was used to help students in reading comprehension activity. The teacher explained about how to find detailed information and how to find or guess the meaning of words from context (vocabulary) in reading comprehension. For the third meeting, *The Tiger who would be a King* was used as the material of narrative text.

4) Conducting the Posttest

The posttest was conducted after the treatment. It was to identify to what extent the students had understood about the aspects of reading comprehension and how well they could comprehend a narrative text after the application of TPS technique in teaching-learning of reading comprehension.

The items for the posttest were similar to those of the pretest, only the order of the texts and their corresponding items were rearranged. The items were based on the following table of specification:

Table 3. Table of Specification of the Posttest

No.	Reading Specification	Item Number	Percentage
1	Determining main ideas	4, 24, 29	10%
2	Inferences	2, 30	6.67%
3	References	3, 7, 12, 13, 16, 18, 22, 28	26.67%
4	Finding detailed information	1, 5, 6, 9, 11, 14, 15, 17, 19, 20, 23, 27	40%
5	Vocabularies	8, 10, 21, 25, 26	16.67%
Total		**30**	**100%**

D. Instrument Used for Collecting the Data

The instrument for data collection—the pretest and the posttest—was previously qualified through the tryout test. The instrument was tried out to achieve good quality. For such instrument with multiple-choice items, good quality means having satisfactory degree of validity, reliability, level of difficulty and discrimination power (Shohami, 1985).These terms are described on the next page.

1) Validity

Validity refers to the extent to which the test measures what it is intended to measure; in other words, it relates directly to the purpose of the test (Shohamy, 1985:74). There are four types of validity: face validity, content validity, construct validity, and empirical or criterion-related validity. *Face validity* concerns only the surface layout of the test, while *criterion-related validity* with the measurement of future success, such as in replacement test (Hatch and Farhady, 1982:251). In case of the instrument for this research, the other two types of validity—content and construct validities—are given greater consideration as they are stronger and more relevant.

Hatch and Farhady (1982:251) define *content validity* as the extent to which the test measures a representative sample of the subject matter content. The focus of content validity is on the adequacy of the sample and not simply on the appearance of the test. *Construct validity*, on the other hand, is concerned with whether the test is actually consistent with the theory of what it means to know the language (Shohamy, 1985:74). One of the efforts to achieve validity is to construct the instrument based on a table of specification that outlines the content and the aspects of reading comprehension skill to measure. The instrument was constructed based on the initial table of specification presented below:

Table 4. Initial Table of Specificationfor Data Collection Instrument

No.	Reading Specification	Item Number	Percentage
1	Determining main idea	2, 13, 27, 28, 29, 30, 32	17.5%
2	Inference	6, 12, 17, 19, 21, 22, 24, 36	20%
3	Reference	3, 9, 14, 15, 25, 34, 38, 40	20%

No.	Reading Specification	Item Number	Percentage
4	Finding detailed information	1, 4, 7, 8, 10, 11, 18, 23, 37, 39	25%
5	Vocabulary	5, 16, 19, 26, 31, 33, 35	17.5%
Total		**40**	**100%**

2) Reliability

Reliability refers to whether the test is consistent in its scoring and gives us an indication of how accurate the test scores are (Shohamy, 1985:70). To measure the reliability of the instrument, split-half method is calculated using Spearman-Brown split-half coefficient:

$$r_{sb} = \frac{2r_{xy}}{(1 + r_{xy})}$$

In which r_{sb} is the split-half reliability coefficient and r_{xy} represents the correlation between the two halves of the scores on the items (i.e. between the sets of scores on the odd- and even-numbered items). The r_{sb} value is categorized as *high* if falls between 0.90 to 1.00, *moderate* between 0.50 to 0.89, and *low* from 0.49 to zero (Hatch and Farhady, 1985:247). The value of r_{xy} is, in turn, obtained using Pearson correlation coefficient formula:

$$r_{xy} = \frac{SP}{\sqrt{SS_X \times SS_Y}}$$

where:

$$SP = \sum (X - \bar{X})(Y - \bar{Y}) = \sum XY - \frac{(\sum X)(\sum Y)}{N}$$

$$SS_X = \sum (X - \bar{X})^2 = \sum X^2 - \frac{(\sum X)^2}{N}$$

$$SS_Y = \sum (Y - \bar{Y})^2 = \sum Y^2 - \frac{(\sum Y)^2}{N}$$

andwhere *SP*refers to the sums of products, SS_X to the sums of squares of the scores of variable *X*, SS_Y to the sums of squares of the scores of variable *Y*, and *N*to the number of cases.

3) Level of Difficulty

Difficulty level is concerned with how easy or difficult an item is from the point of view of the students who take the test. Difficulty level is important since test items that are too easy can tell us nothing about the differences within the test population. To identify the level difficulty, the following formula is used:

$$FV = \frac{Correct\ U + Correct\ L}{N}$$

where *FV* denotes facility value or index of difficulty, *Correct U*the proportion of upper group students who respond correctly, *Correct L*the proportion of lower group students who respond correctly, and *N*the number of test participants. An item, according to Heaton (1991:178), is categorized as *difficult* if it has a difficulty level value less than 0.30, *average* between 0.30 and 0.70, and *easy* if more than 0.70.

4) Discrimination Power

Discrimination power is used to indicate how accurate the test items can really differentiate between the successful (upper) and the failed (lower) students who take the test. This value is determined using the following formula:

$$DP = \frac{Correct\ U - Correct\ L}{\frac{1}{2}N}$$

Where *DP* means discrimination power, *Correct U* the proportion of upper group students who respond correctly, *Correct L* the proportion of lower group students who respond correctly, and *N* the number of test participants. The following criteria are applied to discrimination power (Heaton, 1991: 80):

1) If the value is positive, a larger number or more knowledgeable students than less-knowledgeable ones get the item correct. If the value is zero, there is no discrimination.

2) If the value is negative, more less-knowledgeable students than knowledgeable ones get the item correct.

3) In general, higher discrimination index is better. In classroom situation, most items should have discrimination indexes higher than 0.20.

Students' responses in the tryout, the pretest, and the posttest were scored using the following formula:

$$S = \frac{R}{N} \times 100$$

In which *S* is the score, *R* the total correct responses, and *N* the number of items.

E. Hypothesis Testing

The hypothesis is tested using paired-samples t-test as the samples (i.e. the means) come from the same subjects. The manual formula of paired-samples t-test is presented below:

$$t = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{\sum d^2 - \frac{(\sum d)^2}{n}}{n(n-1)}}}$$

Where $\bar{X}_1$ and $\bar{X}_2$ represent the means of the samples (pretest and posttest), *d* the difference scores, *Σd* the sum of the difference scores, *(Σd)²* the square of the sum of the difference scores, *d²* the squared difference scores, *Σd²* the sum of the squared difference scores, and *n* the number of participants in each sample. The hypothesis is one-tailed as it predicts a significant increase after the treatment, with the significance level set at *p* = 0.05. If the calculated *t* value is higher than the *t*-table value on specific degree of freedom *df*, the null hypothesis is rejected, and vice versa.

F. Data Analysis

Data analysis was following the steps below:

1) Scoring the pretest and the posttest.

2) Tabulating the results of the pretest and the posttest and calculating the means of both.

3) Testing the hypothesis using paired-samples t-test via SPSS® for Windows®to see whether the improvement gained by the students after the application of TPS technique was significant or not.

4) Interpreting the results of statistical calculations and drawing the conclusion.