

### III. METHODOLOGY

This chapter describes the design of the research, sample, data collecting technique, research procedure, analyzing the data, instrument. This chapter also describes the criteria of good test, validity and reliability, data analysis, and hypothesis testing.

#### 3.1. Research Design

The design of this research is *one group pretest posttest design* (Hatch and Farhady, 1982: 20) to investigate whether there was an improving in students' reading comprehension achievement through SQ3R technique. In this design, pretest and posttest was administered whether SQ3R technique can be used to improve students' reading comprehension achievement.

The pretest was conducted to measure students' reading comprehension achievement before treatment and the posttest was conducted to find out students' reading comprehension achievement after being taught using SQ3R technique. Then, the means of both pretest and posttest were compared to find out the progress before and after the treatments.

This research used one class as sample of research. This class had both pretest and posttest and three treatments. The design of the research was described as follows:

$$\mathbf{T1 \quad X \quad T2}$$

Where:

- T1** : The pretest
- X** : The treatments
- T2** : The posttest

(Hatch and Farhady in Setiyadi 2006: 131)

### **3.2. Sample**

The population of this research was the second grade students of SMAN 1 Muara Enim. There are six classes of the second grade. Each class consists of 30 students. Based on the teacher's instruction the researcher chosen XI IPS A class as sample of research, and XI IPS B class as try out class.

### **3.3. Data Collecting Technique**

This research used reading test as the instrument in collecting the data. The reading tests included pretest and posttest. The test was multiple choices and some reading texts. The questions had for each (A, B, C and D), one as the correct answer and the rest are the distracters. Therefore, if one participant answers all the items correctly s/he got 100 points. The treatment also used reading text. The texts were taken from English for the second grade students. The length of time in

collecting the data was three weeks. The data was gained from the content and construct validity in which the question represents five of sort reading skills' i.e. determining main idea, finding the detail information, reference, inference and vocabulary (Nuttal, 1985).

1. Pretest

The researcher administered the pretest before the researcher gave the treatment. Pretest was given to know how far the competence of the students in reading comprehension before the treatments were conducted. It required 45 minutes for the test. The test was multiple choices that consist of thirty items with the option A, B, C, D. The materials were report text.

2. Posttest

The posttest gave after the researcher applied the treatments to the subject of the research. The posttest was done after three meetings of the treatments. Posttest was designed to measure the improvement of the students' reading comprehension achievement after the treatment was given. It required 45 minutes for the test. The test was multiple choices that consist of thirty items with the option A, B, C, D. The materials were report text.

The treatments were classroom activities which used and applied SQ3R technique in reading the texts. The texts were taken from students' English book for second grade. The treatment using SQ3R technique was given in three meeting and 90 minutes in each.

The test was administered to find the quality of the test before it was used, whether the test items were good or not in validity, reliability, level of difficulty and the discrimination power. The researcher administered the try out using reading texts with 50 items of multiple choices in 90 minutes. The maximum score was 100; each correct answer had 2 points.

### **3.4. Research Procedure**

The procedures in administering the research were as follows:

1. Identifying the problem

The first step to be made in this research was identifying the core of the problem.

2. Determining the sample of the research

Based on the teacher's instruction the researcher chosen XI IPS A class as the sample of research and XI IPS B class as try out class. There are six classes of the second grade. The researcher taken one class as the sample of the research.

3. Determining the research instrument

The instrument was taken from students' text book. The texts were used for pretest and posttest was report text.

4. Administering the try out test

The researcher administered the try out using reading text and 50 items of multiple choices. The maximal points is 100, each correct answer had 2 points. It was taken 90 minutes. The test was given to find the quality of the test before it was used in order to get the data on the research. It was to

find out whether the test items were good or not validity, reliability, level difficulty and the discrimination power. The researcher used split-half method to measure the reliability in which required her to provide the items into two same groups, first half and second half.

Some items were dropped and revised to be administered in pretest and posttest. The difficulty level of try out considered of 10 easy items (4, 6, 22, 23, 29, 34, 39, 43, 45, 50), 30 satisfactory items (1, 2, 5, 7, 9, 10, 11, 12, 13, 14, 16, 17, 21, 24, 25, 26, 27, 30, 31, 32, 33, 37, 38, 40, 41, 44, 46, 47, 48, 49), 10 difficult items (3, 8, 15, 18, 19, 20, 28, 35, 36, 42). Meanwhile, for discrimination power, 3 bad items (4, 22, 35), 9 poor items (3, 6, 8, 18, 20, 28, 29, 36, 39), 38 good items (1, 2, 5, 7, 9, 10, 11, 12, 13, 14, 15, 16, 17, 19, 21, 23, 24, 25, 26, 27, 30, 31, 32, 33, 34, 37, 38, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50) were satisfactory. Items with negative zero discrimination power were dropped, meanwhile for the items with satisfactory level of difficulty and satisfactory discrimination power were administered. Moreover, the items that were difficult and easy in the difficult level but had satisfactory discrimination power were revised. The items that were satisfactory in level difficult but had discrimination index were revised.

##### 5. Determining final test of the instrument

In this step, the researcher revised the instrument based on the result of try out. The revision was done by changing the ambiguous statements, distracters, double correct answer, etc.

#### 6. Administering pretest

The pretest was aimed to find out the students' basic reading comprehension achievement. The researcher administered pretest before giving the treatments by using reading text and 30 items multiple-choice test. It was administered for about 45 minutes in the subject class of research.

#### 7. Conducting the treatments

After giving the pretest to the students, the researcher conducted the treatment for three meetings, which took 90 minutes in every meeting. The researcher taught reading comprehension of report text by using SQ3R technique to the students in the sample of research.

#### 8. Administering Post Test

The posttest was aimed to find out the students' reading comprehension achievement after giving the treatment which was used as their reading technique. The test using reading text and 30 items multiple-choices test. It was administered for about 45 minutes in the subject class of research.

#### 9. Analyzing the data

Both of pretest and posttest results of the class were analyzed by using Repeated Measures t-test to compare the data of the two means score (Hatch and Farhady, 1982:108). The researcher analyzed the improvement by comparing the scores of pretest and posttest from the sample of research. If the score of posttest is better than pretest, it means that there is a progress of the students' reading comprehension achievement.

#### 10. Concluding the results

After analyzing the results of both pretest and posttest, the conclusion was drawn.

#### 11. Reporting the results

The results of this research were reported in the script including the suggestion from the researcher.

### **3.5. Analyzing the Data**

The researcher analyzed the data by comparing the average score (mean) of the pretest and posttest to know whether there is improvement of students' reading ability through SQ3R technique.

### **3.6. Instrument**

The two reading tests were given to the students to check their reading comprehension achievement. They were pretest and posttest. The researcher used objective test. It was multiple choice (MC) tests consist of four options (A, B, C, D), to make it easy to correct and to give score. The material was about report text. The researcher gave 30 items for pretest and 30 items for posttest. The purpose of pretest was to know the students' reading comprehension achievement before treatments. The purpose of posttest was to know the students' improvement of reading comprehension achievement after treatments.

### **3.7. The Criteria of Good Test**

A test will be said have a good quality if it has good validity, reliability, level difficulty and discrimination power (Heaton, 1991: 5).

#### **1. Validity**

Validity refers to the extent to which the test measures what was intended to be measure. This means that is 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. To measure whether the test has good validity, the researcher used content and construct validity since the other two were considered be less needed. Face validity only concerns with the lay out the test. Criterion-related validity is concerned with measuring the success in the future, as in replacement test (Hatch and Farhady, 1982: 251). The two type uses in this research are:

##### **a. Content Validity**

Content validity is the extent to which the test measures a representative sample of the subject matter content. The focus of the content validity is adequacy of the sample and not simply on the appearance of the test (Hatch and Farhady, 1982: 251). Content validity is intended to know whether the test items are good reflection of what will be covered. The test items are adapted from the materials that have be taught to the students should be constructed as to contain a representative sample of the course. (Heaton, 1975: 160). Therefore,



since the test instrument was conducted to get the data of the students' reading comprehension achievement, the content validity of the test items were conducted by including reading materials which were arranged based on the materials already given and it was suitable with the curriculum. Thus, if the measuring instrument has represented all the ideas that connected with the materials that be measured, that measuring instrument has fulfilled the aspect of content validity.

b. Construct Validity

Construct validity is concerned with whether the test is actually in line with the theory of what it means to know the language (Shohamy, 1985: 74). Regarding the construct validity, it measures the construction had already referred to the theory, meaning that the test construction has already in line with the objective of the learning (Hatch and Farchady, 1982: 251). It means that the items should really test the students whether they have mastered the reading text. Basically, the construct and content validity are overlap. It is a representative of the material from the subject. In line with Nuttal (1985) the relation validity of the instrument refers to construct validity in which the question represents five of sort reading skills, i.e. determining main idea, finding the detail information, reference, inference and vocabulary. Skills of reading in the test are a part of the construct validity and the item numbers are a part of the content validity. The composition of the test items was presented in table 1: table of specification below.

**Table 1. Specification of the Validity Test**

<b>No</b>	<b>Skill of Reading</b>	<b>Item Numbers</b>	<b>Percentage of Items</b>
1	Determining main idea	3, 5, 8, 13, 24, 29, 38, 42, 47	18%
2	Finding specific information	1, 6, 14, 18, 19, 23, 28, 33, 34, 37, 39, 43	24%
3	Inference	4, 9, 17, 22, 27, 32, 36, 41, 46, 50	20%
4	Reference	2, 11, 15, 21, 26, 30, 40, 45, 48	18%
5	Vocabulary	7, 10, 12, 16, 20, 25, 31, 35, 44, 49	20%

In order to measure the content and construct validity, *inter-rater* analysis was used to make the reading test instrument more valid. Thus, four English teachers of SMAN 1 Muara Enim such as Dra. Hj. Konatira, M.Pd., Aries Oktaviany, S.Pd., Rosita Iriani, S.Pd., and Dien Noveta, S.Pd. taken part in measuring the content and construct validity of the test instrument. If the percentage of one item is >50%, it means that the item test would be taken.

Table 2. Inter-rater Analysis of the Try Out

N O	Teacher	Skill Of Reading/ Indicator				
		Determining main idea	Finding specific information	Inference	Reference	Vocabulary
		3, 5, 8, 13, 24, 29, 38, 42, 47	1, 6, 14, 18, 19, 23, 28, 33, 34, 37, 39, 43	4, 9, 17, 22, 27, 32, 36, 41, 46, 50	2, 11, 15, 21, 26, 30, 40, 45, 48	7, 10, 12, 16, 20, 25, 31, 35, 44, 49
1	A	25%	25%	10%	25%	25%
2	B	25%	25%	25%	25%	25%
3	C	15%	25%	15%	25%	25%
4	D	25%	25%	25%	25%	25%
<b>Total Percentage</b>		85%	100%	75%	100%	100%

## 2. Reliability

Reliability refers to whether the test is consistent in its score and gives us an indication of how accurate the test score are (Shohamy, 1985: 70). There are three ways to find out the reliability of a test, i.e. test-retest, parallel tests, and internal consistency methods (Hatch and Farhady, 1982: 24). The test-retest methods is not use since it is not impracticable, certain students would benefit more than others by familiarity with the type and format of the test, moreover, personal factors such as motivation and differential maturation would also account for differences and the performances of certain students. The second method, parallel test, is not be used either. This method also seemed impracticable since two similar versions of particular test must be constructed: such test must be similar versions of a particular test must be constructed: such test must identical in the nature of their sampling, difficulty, length, rubrics, etc. only after a full statistical analysis of the test and all items contain in them can the test

safely be regarded as parallel (Heaton, 1991: 163). Thus the third method, namely split method, was used in this research.

Split half method was used by the researcher to estimate the reliability of the test. This formulation was simple to use since: (1) it avoids troublesome correlations and (2) in addition to the number of items in the test, it involves only the test, mean and standard deviation. Both of which are normally calculated anyhow as a matter of routine, (Heaton, 1991: 164). To measure coefficient of the reliability the first and second half group, the researcher used the following formula:

$$r_1 = \frac{\sum XY}{\sqrt{\sum X^2 \sum Y^2}}$$

Where:

$r_1$  : The coefficient or reliability between first half and second half group

$X$  : The total numbers of first half group

$Y$  : The total numbers of second half group

$X^2$  : The square of  $X$

$Y^2$  : The square of  $Y$

(Lado in Hughes, 1991: 3)

The researcher used “Spearmen Brown’s Prophecy formula” (Hatch and Farhady, 1982: 268) to know the coefficient correlation the whole items.

The formula is as follows:

$$rk = \frac{2rl}{1 + rl}$$

Where:

$r_k$  : The reliability of the test

$r_l$  : The reliability of the half test

The criteria of reliability are:

0.90 – 1.00 : high

0.50 – 0.89 : moderate

0.00 – 0.49 : low

(Hatch and Farhady, 1985:247)

If the reliability the test reach 0.05 the researcher would consider that it has been reliable. The result of the computation by using *Pearson Product Moment formula* showed that the reliability of the half test ( $r_l$ ) was 0.9779023 (See Appendix 11). Then, by using *Spearman Brown's Prophecy formula*, it was found that the reliability of the whole test items ( $r_k$ ) was 0.9888277 (See Appendix 11). According to criteria of the reliability test the reliability of 0.9888277 point belongs to high level, so it indicated that the data collecting instrument in this research was reliable and good. Therefore, the results of the test were believed as the reflection of their reading ability which was accurate and consistent.

### 3. Level of Difficulty

To see the level of difficulty, the researcher used the following formula:

$$LD = \frac{R}{N}$$

Where:

$LD$  : Level of difficulty

$R$  : The number of the students who answer correctly

$N$  : The total number of the student following the test

The criteria are:

$< 0.30$  : difficult

$0.30 - 0.70$  : average

$> 0.70$  : easy

(Shohamy, 1985: 79)

Based on the try out test related to those criteria there was 10 easy items (4, 6, 22, 23, 29, 34, 39, 43, 45, 50), 10 difficult items (3, 8, 15, 18, 19, 20, 28, 35, 36, 42), 10 average items (1, 2, 5, 7, 9, 10, 11, 12, 13, 14, 16, 17, 21, 24, 25, 26, 27, 30, 31, 32, 33, 37, 38, 40, 41, 44, 46, 47, 48, 49).

#### 4. Discrimination Power

To see the discrimination power, the researcher used the following formula:

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

Where:

$DP$  : Discrimination power

$U$  : The proportion of the upper group students

$L$  : The proportion of the lower group students

$N$  : the total number of the students

The criteria:

- a. Of the value is positive discrimination, it means that more high level students than low level students. Of the value is zero, no discrimination.
- b. If the value is negative, means that more low level students than the high level students get the item correct.
- c. In general, the higher the discrimination index, the better. In classroom situation most items should be higher than 0.20 indexes.

(Shohamy, 1985: 81)

According to the try out there were 3 bad items (4, 22, 35), 9 poor items (3, 6, 8, 18, 20, 28, 29, 36, 39), 38 good items (1, 2, 5, 7, 9, 10, 11, 12, 13, 14, 15, 16, 17, 19, 21, 23, 24, 25, 26, 27, 30, 31, 32, 33, 34, 37, 38, 40, 41, 42, 43, 44, 45, 56, 47, 48, 49, 50) were satisfactory.

## 5. Scoring System

In scoring students result of the test, the researcher used Percentage Correct (Lyman, 1971: 95). The percentage correct was used in reporting the result of classroom achievement tests. The researcher calculated the average of the pretest and posttest by using this formula:

$$X_{\%c} = 100 \frac{R}{T}$$

Where:

- $X_{\%c}$  : Percentage of correct score  
 $R$  : Number of right answer  
 $T$  : The total items of items test

(Lyman, 1971: 95)

### 3.8. Data Analysis

Analysis means categorizing, ordering, manipulating and summarizing of data obtain answers to research questions (Kerlinger, 1988:125). The purpose of analysis was to reduce data to be intelligible and interpretable so that the relation of research problem can be studied.

In order to see whether there is a significant improvement of the students' reading comprehension achievement, the researcher examined the students' score using the following steps:

1. Scoring the pre test and post test.
2. Tabulating the result of the test and calculating the score of the pretest and posttest.
3. Drawing conclusion from the tabulate result of the pre test and posttest administer, that is by statistically analyzing the data using statistical computerization i.e. *Repeated Measure T-Test of Statistical Package for Social Science (SPSS) version 17.0 for windows* to test whether the improvement of students' gain is significant or not, in which the significance was determined by  $p < 0.05$ . It is used as the data come from the same sample or know as paired data. (Hatch and Farhady, 1982: 114).



### 3.9. Hypothesis Testing

The hypothesis is used to prove whether the hypothesis proposed in this research is accepted or not. The hypothesis of this research is there is any improvement of students' reading comprehension achievement of report text after being taught through *SQ3R technique*.

The hypothesis is analyzed by using *Repeated measures T-Test* with Statistically Package for Social Science (SPSS) version 17.0. The level of significance is 0.05, and the probability of error in the hypotheses is 5%.