

III. RESEARCH METHODS

This chapter describes the following major points: the design of the research, population and sample, data collecting technique, research procedure, scoring system of reading test, try out of research instrument, data analysis, and hypothesis testing.

3.1. Research Design

This research used *one group pre-test post-test design* (Hatch and Farhady, 1982:20). Hatch and Farhady state that this design is an improvement over the one-shot case study because it has measured the gains that the subjects have made rather than just looking at how well everyone did at the end. This design was chosen to find out whether if there is the improvement of students' reading comprehension achievement after being taught through Jigsaw by comparing the average score (mean) of the pre-test with the average score (mean) of the post-test.

The research design could be represented as follow:

T₁ X T₂

T₁ : Pre-test
X : Treatment using Jigsaw
T₂ : Post-test

(Hatch and Farhady, 1982: 22)

Firstly, the pretest was administered to identify the main idea, to determine references, to make inference, to identify supporting detail, and to understand difficult vocabulary before applying the technique. Then, the students were given four treatments by using *Jigsaw Cooperative Learning*. Eventually, a posttest was administered to identify improving aspects of reading comprehension.

3.2. Setting of the Research

The settings in this research consisted of place and time. SMAN 2 Metro was chosen as the research because from the teacher's information in that school, it was reported that most of students still had problems in comprehending the text. Many students had score under the KKM and they had difficulties in comprehending the English text. There were 6 meetings that consisted of 2 sessions for pretest and posttest, 4 sessions for treatment.

3.3. Population and Sample

The population of this research was the second grade students of SMAN 2 Kota Metro in the 2013/2014 academic year. There were six classes in second grade of SMAN 2 Kota Metro and 30 students for each class. The sample was one class as experimental class, which was selected by using Simple Random Sampling. Simple Random Sampling was used if there were not stratified class. It was applied based on the consideration that every student in the population have the same chance to be chosen and in order to avoid the subjectivity in the research (Setiyadi, 2006: 39).

3.4. Data Collecting Technique

In collecting the data, the following technique was employed:

Reading Test

The kind of reading test used was objective test. The reading test was given to identify learners' reading achievement which consisted of pretest and posttest. The pretest was given before the treatment was conducted, firstly, the pretest was administered to find out the students' reading comprehension achievement before treatment. The posttest was administered at the end of treatments in order to find out the results of students' reading comprehension achievement after the 4-time treatments. The test of reading comprehension in this research was used to answer those two research questions.

In selecting reading text, this research considered the text based on themes stated in curriculum for second years of SMA (KTSP 2006). The texts were used taken from any text books and articles on the internet. The composition of the test items was presented in Table 3.1. below.

Table 3.1. Specification of Reading Test

No	Sub-Skills of Reading	Item Number	Percentage
1	Determining main ideas	3,12,17,22, ,28,30,35,38	20%
2	Inferences	1,7 ,14,18, 21, 24,25,31,36	22,5%
3	References	2,6,11, 15, 20, 27,32,37	20%
4	Finding detail information	4, 8, 10, 23, 29,34,40	17,5%
5	Vocabularies	5, 9, 13,16,19 ,26,33,39	20 %
Total		40	100 %

3.5. Try Out

Before conducting the research, the data collecting instrument was tried out in order to make sure whether the instruments were valid and reliable to collect the data. The try out contained 40 items, each item consisted of four options. The type of the instruments was multiple choice test and time allocation was 90 minutes.

The try-out test was administrated first in XI IPA 3 class to analyze the reliability, level of difficulty, and discrimination power to achieve good test instrument criteria. Then the pre-test and post-test were administrated in XI IPA 3 class to analyze the improvement of the students' reading comprehension through Jigsaw technique.

3.5.1. Result of the Try-out Test

Before administrating the pre-test, the try-out test was conducted on May 2nd, 2014 in class XI IPA 3 of SMAN 2 Kota Metro which was chosen randomly to analyze the reliability, level of difficulty, and discrimination power to achieve good test instrument criteria. There were 40 items administrated based on eight different narrative texts. Those items were in the form of multiple choices, which contained four options of answer for each (A, B, C, and D). The time allocated was 90 minutes.

Based on the table in Appendix 9, there were 40 items in the try-out test. After analyzing the criteria of good test by using level of difficulty and discrimination power, it could be seen that 10 items were dropped, such as item number 1, 5, 8,

9, 11, 13, 16, 21, 34 and 40. The try-out test consisted of 5 difficulty items (25, 31, 35, 36, and 38); 24 average items (2, 3, 4, 6, 7, 10, 14, 15, 17, 18, 19, 20, 22, 23, 24, 26, 27, 28, 29, 30, 32, 33, 37, and 39); and 11 easy items (1, 5, 8, 9, 11, 12, 13, 16, 21, 34, and 40). Some of poor and difficult items were revised, 8 revised items (12, 25, 26, 29, 31, 35, 36, and 38) while the average and satisfactory items were administrated in the pre-test and post-test.

In analyzing discrimination power, there were 13 poor items (1, 4, 8, 9, 11, 13, 16, 21, 25, 26, 29, 34, and 40); 6 good items (3, 10, 22, 23, 27, and 28) and 21 satisfactory items (2, 5, 6, 7, 12, 14, 15, 17, 18, 19, 20, 24, 30, 31, 32, 33, 35, 36, 37, 38, and 39). The items that had criteria level of difficulty < 0.30 and $> 0.70 - 1.00$ but had satisfactory and good discrimination were revised, meanwhile the items which had average level of difficulty and good and satisfactory discrimination indexes were administrated for the pre-test and post-test.

After analyzing the level of difficulty and discrimination power, it was found that 30 items were good and administrated for the pre-test and post-test. On the other hand, 10 items were bad and dropped because they did not fulfill the criteria of level of difficulty and discrimination power.

To analyze the reliability of the test, Split-half technique was used to estimate the reliability of the test and to measure the coefficient of the reliability between odd

and even group, *Pearson Product Moment* formula was used. The computation showed that the reliability coefficient of the test was 0.99 (Appendix 12). It could be stated that the test had a high reliability since the range of high criteria in the criteria of reliability was 0.80 – 1.00 (Hatch and Farhady, 1985:247).

3.6. Procedures of Collecting Data

In collecting the data, the following steps are employed.

1. *Determining research instrument:* for both reading tests (pretest and posttest), the materials were taken from students' textbook. In the test, kind of narrative texts was used and in each text used consisted of 3 until 6 paragraphs. It was aimed at making an equal proportion and level of difficulty of both pretest and posttest. The number of the items were arranged in such a way so that the reliability of the tests could be seen through split-half method.
2. *Determining the population and sample of the research:* the sample of the research was determined through simple random probability sampling. It means that the sample was selected randomly by using lottery class.
3. *Administering Try-Out Test:* the try-out test was administered in grade XI. Students were given reading test with 40 items of multiple choices. The test was done in order to measure the level of difficulty (LD) and discrimination power (DP) as well as find out the realibility and validity of the test.
4. *Administering the pretest:* pretest was conducted before the treatments. It was aimed to check students' reading ability in determining mind idea,

determining references, making inference, finding detail information, and vocabulary in texts. Pretest administered for about 60 minutes on first week of the research.

5. *Giving treatment*: three treatments by using jigsaw were given in two weeks. The treatment conducted in three meetings and 90 minutes for each. The treatments were classroom activity, which uses an apply Jigsaw Cooperative Learning in reading.
6. *Conducting posttest*: posttest were conducted after the treatment. Posttest was conducted to find out whether there is a significant students' reading ability after the treatments. It was administered for 60 minutes in experimental group.
7. *Analyzing the result of the Test*: all the data were gathered by the average score (mean) of reading test and questionnaire was analyzed to draw the conclusion.

3.7. Criteria of Good Test of Reading

In this research, to prove whether the test of reading had good quality, it must be tried out first. As a good quality if it has good validity, reliability, and level of difficulty and discrimination power.

3.7.1 Validity

A test can be considered valid if the test measure the object to be measured and suitable with the criteria (Hatch and Farhady, 1982: 250). The discussion of the validity of reading test are provided below.

1. Validity of the Reading Test

a. Content validity

Content validity was concerned whether the test was sufficiently representative and comprehensive for the test. In the content validity, the material was given suitable with the curriculum. Content validity was the extent to which a test measures a representative sample of the subject matter content, the focus of content validity was adequacy of the sample and simply on the appearance of the test (Hatch and Farhady, 1982; 251).

The topic chosen was narrative text. The topics are the representative of reading materials of School Based Curriculum or KTSP as a matter of tailoring the lesson to students' need. To know whether the test has a good content validity, the items of the test discussed with the expert (advisors), the researchers' colleague, and the English teacher of Senior High School.

b. Construct Validity

It was concerning to whether the test is actually based on the theory of which it means to know the language that was being measured. In this research the researcher focused on reading comprehension in the form of narrative texts. Nuttal (1985) states that 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 detail information, reference, inference and vocabulary mastery. Skills of reading in the test were a part of the construct validity and the item numbers are a part of the content validity. The test compared to the table of specification to know whether the test had a good reflection of what has been taught. A table of specification was an instrument that helps the test constructor plans the test.

3.7.2 Reliability

Reliability of the test can be defined as the extent to which a test produces consistent result when administrated under similar conditions (Hatch and Farhady, 1982:243). Split-half technique was used to estimate the reliability of the test and to measure the coefficient of the reliability between odd and even group, *Pearson Product Moment formulais* used as follows:

$$r_l = \frac{\sum xy}{\sqrt{[\sum x^2][\sum y^2]}}$$

r_l : Coefficient of reliability between odd and even numbers items.

x : Odd number.

y : Even number.

x^2 : Total score of odd number items.

y^2 : Total score of even number items.

xy : Total number of odd and even numbers.

(Lado, 1961 in Hughes, 1991:32).

The criteria of reliability are:

- 0.80 – 1.00: high.
- 0.50 – 0.79: moderate.
- 0.00 – 0.49: low.

(Hatch and Farhady, 1985:247).

To know the coefficient correlation of whole items, *Spearman Brown's prophecy formula* was used. The formula is as follows:

$$rk = \frac{2r1}{1 + r1}$$

rk: The reliability of the test.

r1: The reliability of the half test.

(Hatch and Farhady, 1982:246).

3.7.3. Level of Difficulty

Level of difficulty relates to “how easy or difficult the item was in the form of the point of view of the students who took the test. It was important since test items which were too easy (that all students get right) can tell us nothing about differences within the test population (Shohamy, 1985: 79).

Level of difficulty is calculated by using the following formula:

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

LD = level difficulty

R = number of students who answers it right

N = total number of students

The criteria are:

LD < 0.30 = difficult

LD = 0.31- 0.70 = satisfied

LD > 0.71- 1.00 = easy

(Arikunto, 1997:214)

3.7.4. Discrimination power of the Test

Discrimination power refers to “the extent to which the item differentiates between high and low level students on that test. A good item which is according to this criterion, is one in which good students did well, and bad students failed (Shohamy, 1985:81).

The formula is:

$$DP = \frac{\text{Upper} - \text{lower}}{\frac{1}{2}(N)}$$

DP = discrimination power

Upper = proportion of “high group” students getting the item correct

Lower = proportion of “low group” students getting the item correct

N = total number of students

The criteria are follows:

LD = 0.00-0.20 = poor

LD = 0.21-0.40 = satisfactory

LD = 0.41-0.70 = good

LD = 0.71-1.00 = excellent

(Arikunto, 1997:223)

3.7.5. Scoring System

Arikunto’s formula was used in scoring the students’ result of the test. The higher score will be 100

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

Where:

S is the score of the test

R is the right answer

N is the total of the items

3.8. Hypothesis Testing

The hypothesis testing was used to prove whether the hypothesis proposed in this research was accepted or not. The hypothesis of this research was there is any significant improvement of students' reading comprehension achievement after being taught by Jigsaw cooperative learning.

The hypothesis was also statically tested by using statistical computerization (SPSS 16), in which the significance is determined by $p < 0.05$. Therefore, the hypothesis which can be stated is as follows:

- H_0 : There is no significant improvement of the students' reading Comprehension achievement through Jigsaw Cooperative Learning.
- H_1 : There is a significant improvement of the students' reading Comprehension achievement through Jigsaw Cooperative Learning.

3.9. Schedule of the Research

This research was conducted based on sequenced schedule which was appropriate at schedule of English subject in the class. Try out test was administered on Friday, May 2nd, 2014 in XI IPA 3 as try out class to determine the content and construct validity of the text, also the level difficulty and the discrimination power of its. On Wednesday, May 7th, 2014 the pre test was carried out in XI IPA 1 in order to know the students' achievement of reading comprehension before giving treatments. For all treatment, XI IPA 1 class was taken as the experimental class.

The first meeting was on Thursday, Monday 12nd 2014; the second meeting was on Wednesday, May 14th 2014, and the third meeting was on Saturday, May 20th 2014. After the treatments had been administered, the post test was given in that class on Monday, May 21th 2014 in order to know the gain of the students' reading comprehension achievement after being taught using Jigsaw Cooperative Learning technique. The schedule of the research can be seen in the following table:

Table 3.2. Research Schedule in Conducting Research at SMAN 2 Metro.

No	Date	Activities
1	Friday, May 2 nd 2014	Try out test in XI IPA 3
2	Wednesday, May 7 th 2014	Pretest in XI IPA 1
3	Monday, May 12 th 2014	First Meeting in XI IPA 1
4	Wednesday, May 14 th 2014	Second Meeting in XI IPA 1
5	Saturday, May 17 th 2014	Third Meeting in XI IPA 1
6	Monday, May 19 th 2014	Fourth Meeting in XI IPA 1
7	Saturday, May 21 th 2014	Posttest XI IPA 1