

III. METHODS

This chapter discusses certain points: methods deals with design, variable, and source of data, instrument, procedure, data analysis, and hypothesis testing.

3.1. Research Design

This research was a quantitative research, in which data tend to use statistic as measurement in deciding the conclusion (Hatch and Farhady; 1982). In conducting this research, the researcher used Static Group Comparison Design that deals with two groups, the first one (E_1) as experimental group one that was given treatment using non-authentic material and another one (E_2) as experimental group two that was given treatment using authentic material.

The pretest was done first before the treatment. It was done to know the basic of students' reading ability in comprehending texts. Experimental class 1 (E_1) was needed for comparison purposes because it lets the writer interpret her findings more confidently.

Based on Hatch and Farhady (1982: 22), the researcher used the following design:

E1 (Random)	:	T1	X1	T2
E2 (Random)	:	T1	X2	T2

Notes:

E_1 = experimental 1 Group

E_2 = experimental 2 Group

T ₁	=	the pretest
T ₂	=	the posttest
X ₁	=	treatment 1 (Teaching reading through narrative text by using authentic material)
X ₂	=	treatment 2 (Teaching reading through narrative text by using non-authentic material)

3.2. Variable

To find out the answer of this research, the researcher used the reading score of the student. The score that used are from student's pre-test score and student's post-test score in experimental class 1 and student's pre-test score and student's post-test score in experimental class 2. Those score used as a data of the research.

3.3. Population and Samples

3.3.1. Population

The population of the research was the second grade students of SMP N 1 Metro. There were seven classes in the second grade of SMP N 1 Metro and consisted of 26 students in each class (VIII.1 - VIII.10). The samples of this research were two classes of the second year students.

3.3.2. Samples

The sample classes were taken through lottery, because all the classes have the same opportunities to be chosen as the sample of this research and to make sure that the students' abilities were homogeneous or not by seeing the data of the teacher in the school. The researcher took one class as the experimental class 1, and the other one as the experimental class 2. In this case, the researcher asked the leader of the each class to take a small piece of paper in order to know the class will be as experimental class 1 or experimental class 2.

3.4. Instrument

The data of the research were collected by researcher used reading test and treatments as the instrument. There were two kinds of test, pre-test and post-test. Pre-test was administered in order to measure the student's writing ability achievement before the treatment, and post-test was administered after presenting the treatment in order to know the achievement of reading ability.

1. Try Out

Try out test was to measure that the research instrument is good for testing students' reading comprehension. The quality measurement is based on the calculation of its reliability, level of difficulty and discrimination power. In the reliability of the try-out, the Split-Half Method is used in order to analyze the odd (x) and even (y) of the test items. To measure the coefficient of the reliability between odd and even group, the researcher used Spearman-Brown formula.

2. Pretest

The pretest was given before the treatment has applied. The pretest was done to know the basic of students' reading ability in comprehending texts before getting treatments. It was also needed to know whether both the two classes were equal or not in the terms of their reading comprehension achievement. The test was in multiple choice forms with four options a, b, c, d.

Table 1. Specification of Pretest

No.	Reading Sub-skills	Item Number	Percentage of Items
1.	Identifying Main Idea	1, 7, 9, 14	20%
2.	Identifying Details	2, 11, 13, 17	20%
3.	Making Inferences	4, 10, 19	15%
4.	Identifying References	3, 8, 15, 18	20%
5.	Understanding Vocabulary	5, 6, 12, 16, 20	25%

3. Treatment

In the treatment, the researcher teaches reading comprehension by using authentic materials in class 1 and non-authentic materials in class 2.

4. Post test

The post test was administered once after treatment. It is given to measure students' reading ability in comprehending text and also to know the effect of authentic and non-authentic materials in their reading comprehension achievement.

Table 2. Specification of Posttest

No.	Reading Sub-skills	Item Number	Percentage of Items
1.	Identifying Main Idea	6, 8, 13, 18	20%
2.	Identifying Details	1, 9, 15, 17	20%
3.	Making Inferences	3, 11, 14	15%
4.	Identifying References	2, 7, 10, 19	20%
5.	Understanding Vocabulary	4, 5, 12, 16, 20	25%

3.5. Result of Try Out

The research instrument were tried out to measure the quality in terms of validity, reliability, level of difficulty and discrimination power. Here are some elements tested as follows:

1. The Validity

Validity is the extent to which an instrument measures what it is supposed to measure and performs as it is designed to perform. It is rare, if nearly impossible, that an instrument be 100% valid, so validity is generally measured in degrees. As a process, validation involves collecting and analyzing data to assess the accuracy of an instrument. A test can be said valid if it measures the object to be measured and suitable for the criteria (Hatch and Farhady, 1982: 251).

There are four basic types of validity: content validity, criterion-related validity, face validity, and construct validity, (Hatch and Farhady, 1982:251). To determine the validity of the test, the researcher used content validity construct validity since the other are considered to be less needed.

Content Validity

According to Hatch and Farhady (1982:251) content validity is the extent to which the test measures a representative sample of the subject matter content. Content validity concerns whether the test are good reflection of the materials that need to be tested. To get content validity of the test, the researcher adopted materials based on the objective of teaching in syllabus (Curriculum 2013) for second grade of junior high school students, and represented the materials taught in the class. In line with the syllabus for the second grade of junior high school at the second semester, the students are required to be able to comprehend the short functional texts of narrative text.

Construct Validity

A test is said to have construct validity if it can be demonstrated that it measure. The word 'construct' refers to any underlying ability or trait which is hypothesized in a theory of language ability (Hughes, 1991:26). In this case, what means by reading is to understanding the ideas and information explicitly states in the passage (Suparman, 2012). Based on that theory, some of the reading comprehension aspects that should be mastered are; identifying main idea, identifying details, making inferences, making references, and understanding vocabulary. Therefore, to make sure that the items of the test already good in the term of construct validity, the researcher specify them into table of specification as follows:

Table 3. Table specification of TRY OUT TEST

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

2. The Reliability

According to Hatch and Farhady (1982:243), the reliability of a test can be defined as the extent to which a test procedures consistent result when administered under similar conditions. To estimate the reliability of the test, the split-half method is used in order to analyze the odd (x) and even (y) of the test items. To measure the coefficient of the reliability between odd and even group, the researcher will used this following formula:

$$R_{XY} = \frac{N \sum XY - (\sum X)(\sum Y)}{\sqrt{(N \sum X^2 - (\sum X)^2)(N \sum Y^2 - (\sum Y)^2)}}$$

Notes:

R = coefficient of reliability between odd and even numbers

N = number of the students

X = square of X

Y = square of Y

$\sum X$ = total score of odd number

$\sum Y$ = total score of even number

(Hatch and Farhady, 1982:199)

After getting the reliability of half test, the researcher used Spearman Brown to determine the reliability of the whole test, as follows:

$$r_k = \frac{2 r_{xy}}{1 + r_{xy}}$$

The criteria of reliability are:

0.90 – 1.00 : high

0.50 – 0.89 : moderate

0.0 – 0.49 : low

(Hatch and Farhady, 1982:247)

3. The Level of Difficulty

The difficulty level of an item shows how easy or difficult that particular item done by the participants, (Heaton, 1975: 182). Level of difficulty is generally expressed the percentage of the students who answered the item correctly. It is calculated by the following formula:

$$LD = \frac{U + L}{N}$$

Notes:

LD : the level of difficulty

U : the number of upper group who answer correctly

L : the number of lower group who answer correctly

N : the total number of students in upper and lower groups

(Arikunto, 1993:221)

The criteria of the difficulty level are:

< 0.30 = difficult

0.30-0.70 = average

> 0.70 = easy

4. The Discrimination Power

Discrimination power is the proportion of the high group students getting the items correct minus the proportion of the low-level students who getting the items correct. The formula of the discrimination power is:

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

Notes:

D : discrimination power

U : the number of students from upper group who answer correctly

L : the number of students from lower group who answer correctly

N : the number of students

The criteria are:

DP = 0.00 – 0.19 = poor

DP = 0.20 – 0.39 = satisfactory

DP = 0.40 – 0.69 = good

DP = 0.70 – 1.00 = excellent

DP = negative/minus (-), all poor

(Arikunto, 1993: 221)

In this research, the instrument to be tried was administered to Class VIII.1 that consists of 20 students who has the same characteristic as the samples that was chosen. The material was based on the objective of teaching syllabus for the students at the second grade of junior high school. It means that, this test was completely proved the criteria of content validity. Then, to find the reliability of the test, the researcher used the data from tryout test (see Appendix 5). If the reliability test reaches 0.50, the researcher will consider that the test has been reliable.

Meanwhile, in this research the reliability was 0,98. So, it could be stated that the test had a very high reliability since 0,98 include to 0,80 – 1,00 range of very high reliability.

In this test, there were 50 items of multiple choices with four options (a, b, c, and d), one of these as the correct answer and the rest were distracters. This was conducted in 80 minutes. After analyzing data in level of difficulty, the writer found that the 10 items were difficult, 21 items were average. Meanwhile, in the discrimination power, there were 22 good items and 28 poor items. In this case, there were 30 items that was dropped because it had easy or difficult in level of difficulty, and poor in the discrimination power. Therefore, there were 20 items that possible to use for pretest and posttest.

5. Scoring System

In scoring the result of the students' test, the researcher used Percentage Correct (Lyman. 1971:95). The percentage correct score was used in reporting the result of achievement test. The research calculated the average of the pre-test and post-test by using this formula:

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

Where:

S = the score of the test

R = number of right answer

N = number of items on test

3.6. Procedure

There were some procedures that were applied for taking the data, they were:

1. Determining the population and the sample.

The researcher took two classes to determine the experimental class 1 and experimental class 2.

2. Administering the try-out test

The researcher prepared the try out materials and gave the try out test to another class which has the same characteristic as the samples that was chosen, it is VIII.1.

3. Administering pre-test.

The researcher administered the pre-test on both groups' experimental class 1 and experimental class 2.

4. Conducting treatment.

The researcher taught the student in experimental class 1 by using authentic materials and taught the students in experimental class 2 by using non-authentic materials.

5. Administering the post-test.

The researcher administered the post-test, in both of experimental class 1 and experimental class 2.

6. Scoring the student's work.

The researcher scores the learner's work in order to get the data.

7. Analyzing the data.

After collecting the data, the researcher analyzed the data.

8. Testing hypothesis.

After analyzing the data, the researcher tested the hypothesis.

3.7. Data Analysis

The writer computed the students using the following steps:

- Scoring the pre-test and post-test.
- Tabulating the results of the test and calculating the score of the pre-test and post-test.
- Drawing conclusion from the tabulated results of the pre-test and post-test administered, that is by statistically analyzing the data using statistical computerization i.e. *Independent Groups t-Test of Statistical Package for Social Science (SPSS) version 16.0 for windows* to test whether the increase of students' gain is significant or not, in which the significance was determined by $p < 0.05$. It was used as the data come from the two samples (Hatch and Farhady, 1982:111).

3.8. Data Treatment

The aim of the data treatment was to determine whether the students' reading comprehension achievement is increase or not. The data of the research was examined by using independent group t-test, because the independent variable has more than one group, those are; authentic material and non-authentic material, which means that two different groups (experimental class 1 and experimental class 2) were compared. The data was statically computed through the Statically Package for Social Science (SPSS) version 16.0. In doing so, the researcher analyzed the data statically by administering the normality test, homogeneity test, and random test.

1. Normality Test

Normality Test was used to measure whether the data of the test have normal distribution or not. It caused the students' score of pretest and posttest was analyzed to gain the normality test. The hypotheses for the normality are as follows:

H_0 : The data is distribute normally

H_1 : The data is not distribute normally

The data will be determined normal if it met the criterion, and the criterion for the hypothesis is:

The hypothesis will be accepted if $\text{sign} > \alpha$. In this case, the researcher will be used the level of significant of 0.05.

2. Homogeneity Test

The homogeneity testing is intended to test whether the variance of the data is equal or not. The data is statically will be computed through SPSS (Independent Sample Test). And the hypothesis for the homogeneity test is as follows:

H_0 : The data is not homogeneous

H_1 : The data is homogeneous

In which, " H_0 is accepted if significant value does not exceed level of significance at 0.05, meanwhile, H_1 is rejected if significant value exceeds level of significance at 0.05".

3. Random Test

In this research, random test was used to see whether the data in experimental class 1 and experimental class 2 was random or not. The hypothesis for random test is as follows:

H_0 : The data is random

H_1 : The data is not random

3.9. Hypothesis Testing

After collecting the data, the writer recorded and analyzed them in order to find out whether there were any difference of student's reading comprehension achievement between the students who are taught using authentic material and those who are taught using non-authentic material and which of the two material is more effective for teaching reading. The writer used Independent Group T-test to know the level of significance of the treatment effect. The formulation is:

$$t_{obs} = \frac{\bar{X}_e - \bar{X}_c}{S_{(\bar{X}_e - \bar{X}_c)}}$$

With:

$$S_{(\bar{X}_e - \bar{X}_c)} = \sqrt{\left(\frac{S_e}{\sqrt{n_1}}\right)^2 + \left(\frac{S_c}{\sqrt{n_2}}\right)^2}$$

\bar{X}_e : Mean from the difference pre-test and post-test of experimental class 1 and 2

\bar{X}_c : Mean from the difference pre-test and post-test of experimental class 1 and 2

$S_{(\bar{X}_e - \bar{X}_c)}$: Standard error of differences between means

N : Subjects on sample

(Hatch and Farhady, 1982:111)

The criteria are:

If the t-ratio is higher than t-table : H_1 is accepted

If the t-ratio is lower than t-table : H_0 is accepted

In this chapter, the researcher explains about methods deals with design, variable, and source of data, instrument, procedure, data analysis, data treatment and hypothesis testing.