

III. RESEARCH METHOD

This chapter discusses the methods of research used in this study, they are: research design, population and sample, data collecting technique, research procedure, scoring system, and data analysis.

3.1 Research Design

This research was aimed at finding out whether RTT in narrative text can increase students' reading comprehension achievement. This research was conducted in one class as experimental class which was chosen purposively. The researcher used one-group pre-test post-test design. Pre-test was administered to find out students' reading ability in narrative text, before the treatments and post-test was administered to find out students' reading ability in narrative text, after the treatments.

The design can be illustrated as follows:

T1 X T2

Which:

X : Treatment (using RTT)

T1 : Pre-test (before treatment)

T2 : Post-test (after treatment)

(Hatch and Farhady, 1982: 20 in Setiyadi, 2006: 131)

Besides, the observation and the interview were also conducted in this research. The observation was conducted to observe the teaching-learning process during the treatment of teaching reading comprehension through RTT. The interview was conducted in the form of open questions and formal types. It was conducted toward some representatives of the students as the interviewees, which were chosen from low and high scores based on the mean score of the post-test, to find out the problems faced by the students in learning reading comprehension through RTT, in which their answers were classified and generalized as the resource.

3.2 Population and Sample

The population of this research was the 2nd year students of SMAN 1 Kalirejo. There were seven classes of the second year students and each class consists of about 30-33 students. This research was conducted purposively in XI IPS 3 as the subject of the research because based on the observation done by the researcher, the students in this class had mainly problem of reading comprehension relying on their reading quiz score.

3.3 Data Collecting Technique

In collecting the data, the research procedure used these following steps:

1. Giving the Pre-test

The pre-test was conducted before the treatment of teaching reading comprehension through RTT, in order to find out the students' basic reading comprehension. Narrative text was the type of reading text that researcher was used in this research based on KTSP 2006 curriculum of senior high school.

2. Giving the Post-test

The post-test was given to the students after the treatment of teaching reading comprehension through RTT by using three times treatments. The result of the post-test was compared with pre-test in order to know whether RTT can increase students' reading comprehension achievement or not, especially in narrative text.

3. Conducting the Observation

The observation was conducted in XI IPS 3 class to observe the teaching-learning process during the treatment of teaching reading comprehension through RTT.

4. Conducting the Interview

The interview was conducted in XI IPS 3 class, in which the students' answers were classified and generalized as the resource. The interview was conducted to find out the problems faced by the students in learning reading comprehension through RTT.

3.4 Research Procedure

In constructing the research, the research procedure used these following steps:

1. Determining the population and sample

The population of this research was the 2nd year students of SMAN 1 Kalirejo. In this research, the class of XI IPS 3 was taken as the sample, the students are taught reading narrative text by using RTT.

2. Determining the research instrument

The instrument of this research was objective reading text of multiple choice test. According to Henning (1987), to measure reading comprehension, requesting students to write short sentence answer to written questions was less valid a procedure than multiple-choice selection (as cited in Henning, 1987: 48). The numbers of the items in the test were 25 items which had five options of answers, (A, B, C, D and E), in which one of them is the correct answer and the rests are the distracters. Narrative text was the type of reading text that researcher was used in the research and the most of materials were taken from students' English textbook, students' task sheet and internet.

3. Administering the try-out test

The try-out test was conducted in the first meeting in another class to know the quality of the test as the instrument of the research. The try-out test was conducted in one class as the tryout class of this research. The numbers of the tests items were 40 items contain five options of answer for each (A, B, C, D and E). Time allocated was 90 minutes. The result of try-out test was used to measure the level of difficulty and discrimination power, to find out the validity and reliability of the test.

4. Administering the Pre-test

The pre-test was conducted before the treatment of teaching reading comprehension through RTT, to see the students' base reading comprehension. The students in experimental classes were given the pre-test in the second meeting before the treatment. The pre-test was given as an objective test in the form of multiple choices. The numbers of the items in the test were 25 items which have five options of answers, (A, B, C, D and E), in which one of them was the correct answer and the rests are the distracters. Narrative text was the type of reading text that researcher was used in the research based on KTSP 2006 curriculum of senior high school.

5. Conducting the Treatment

After having the pre-test, the students were given the reading teaching-learning treatment through RTT. The researcher gave the three-time treatments in 90 minutes for each treatments. The topics of the materials about *The Slippers of Buffalo Skin*, *The Legend of Prambanan Temple*, and *The Jackal who Saved the Lion*.

6. Administering the Post-test

The post-test was given to the students after the treatment of teaching reading comprehension through RTT by using the same topic and three times repetition of one question like in the pre-test. The result of post-test was compared with the result of pre-test to analyze whether there is any increase of the students' reading comprehension achievement in narrative text in learning through RTT.

7. Conducting the Observation

The observation was conducted in XI IPS 3 class to observe the teaching-learning process during the treatment of teaching reading comprehension through RTT. Observation sheet (note where the treatment process is reported) was used to observe teaching-learning activity and to note the classroom events during the treatment process. The observation sheet in the form of a check list was established by Frank Serafini (2010). It was used to observe the teaching-learning process of teaching reading comprehension.

8. Conducting the Interview

The interview was conducted in XI IPS 3 class, in which the students' answers were classified and generalized as the resource. Some representatives of the students as the interviewees were chosen from low and high scores based on the mean score of the post-test. The interview was in the form of open and formal questions (the questions must be in the form of explanation or description rather than "yes" or "no" answers, to avoid the students from being reluctant to answer the questions given). The interview was conducted to find out the problems faced by the students in learning reading comprehension through RTT.

9. Analyzing the Result of the Test

Both of the pre-test and post-test result of the class were analyzed using Repeated Measures T-test to compare the data average score (mean) of both pre-test and post-test in one sample (Hatch and Farhady, 1982: 114). It was tested in order to compare the two means of the same students and to find out whether there is any

increase of students' reading comprehension achievement in narrative text after thought by using RTT.

3.5 Scoring System

The reseacher used Arikunto's formula to score the students' work of the test. The ideal score of the test is 100. The pre-test and post-test score were calculated by using the following formula:

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

where:

S : the score of the test
 R : the total of the highest answers
 N : the total items

(Arikunto, 2005: 236)

3.6 Criteria of Good Test

As mentioned previously, a good test should meet some criteria, they are: a good validity, reliability, level of difficulty and descrimination power.

3.6.1. Validity

Validity is important to find out the instrument of validity. According to Hatch and Farhady (1982: 250), validity is the extent to which an instrument really measures the objective to be measured and suitable with the criteria. If the test can

shows precisely measure the quality of the test, it means that it can be considered to be valid.

According to the Hatch and Farhady (1982:281), there are two basic types of validity, such as content validity and construct validity. Therefore, to measure whether the test has a good validity, this research used content and construct validity.

3.6.1.1. Content Validity

Content validity is concerned with whether the test is sufficiently representative and comprehensive for the test. According to Hatch and Farhady (1982:251), since content validity is the extent to which a test measures a representative sample of the subject matter, the focus of content validity is adequacy of the sample of the appearance of the test. 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 are arranged based on the materials already given and it was suitable based on KTSP 2006 curriculum of senior high school. Thus, if the measuring instrument has represented all the ideas that connected with the materials that will be measured, that measuring instrument has fulfilled the aspect of content validity.

3.6.1.2. 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 that is being measured. Nuttall (1985) states that the construct being measured was students' skill in reading comprehension, i.e. determining main idea, finding the detail information,

reference, inference, and understanding vocabulary. Hence, this research was more focused on increasing the students' reading comprehension achievement in five aspects of reading. They was determining main idea, finding the detail information, reference, inference, and understanding vocabulary.

In order to fulfill the criteria of construct validity, the test items would be presented in the table specification below:

Table 3.1. Specification of Data Collecting Instrument (Try-out Test)

No.	Reading Skills	Item Number	Number of Item	Percentage of Items
1.	Determining Main Idea	1, 4, 6, 16, 23, 30, 31, 36	8	20%
2.	Finding Specific Information	7, 9, 11, 12, 17, 18, 22, 28, 29, 32, 35, 37	12	30%
3.	Reference	5, 10, 34, 40	4	10%
4.	Inference	2, 13, 14, 20, 25, 26, 33,39	8	20%
5.	Understanding Vocabulary	3, 8, 15, 19, 21, 24, 27, 38	8	20%
TOTAL			40	100%

Based on Table 3.1, the percentage of determining main idea, inference and vocabulary took bigger part than reference, because they are fundamental to successful reading comprehension. Hence, the percentage of finding specific information is the biggest because if the students are still confused to determine the main idea and understand vocabulary, it might make the students get difficulty to find the specific information of the text. That was why determining main idea, finding specific information, inference and vocabulary took big part than reference.

In order to measure the content and construct validity, *inter-rater* analysis was used to make the reading test instrument more valid. Thus, three English teachers of SMAN 1 Kalirejo would be the raters, they are Sadikan, S.Pd., M. Uzuddin,

S.Pd., Rismala Dewi, S.Pd. took part in measuring the content and construct validity of the test instrument. If the percentage of one item was >50%, it meant that the item test would be taken. The result of *inter-rater* analysis can be seen on table 2 below:

Table 3.2. Inter-rater Analysis of the Try-out

Item Numbers	Micro Skill Types of Listening Comprehension	Raters			Total Percentage
		R1	R2	R3	
1.	Determining main idea	33.33%	33.33%	33.33%	100%
2.	Inference	33.33%	33.33%	33.33%	100%
3.	Understanding Vocabulary	33.33%	33.33%	33.33%	100%
4.	Determining main idea	33.33%	33.33%	33.33%	100%
5.	Reference	33.33%	33.33%	33.33%	100%
6.	Determining main idea	33.33%	0%	33.33%	66.66%
7.	Finding Specific Information	33.33%	33.33%	33.33%	100%
8.	Understanding Vocabulary	0%	33.33%	33.33%	66.66%
9.	Finding Specific Information	33.33%	33.33%	33.33%	100%
10.	Reference	33.33%	33.33%	33.33%	100%
11.	Finding Specific Information	33.33%	33.33%	33.33%	100%
12.	Finding Specific Information	0%	33.33%	33.33%	66.66%
13.	Inference	33.33%	33.33%	33.33%	100%
14.	Inference	33.33%	33.33%	33.33%	100%
15.	Understanding Vocabulary	33.33%	33.33%	33.33%	100%
16.	Determining main idea	33.33%	33.33%	33.33%	100%
17.	Finding Specific Information	33.33%	33.33%	33.33%	100%
18.	Finding Specific Information	33.33%	33.33%	33.33%	100%
19.	Understanding Vocabulary	33.33%	33.33%	33.33%	100%
20.	Inference	33.33%	33.33%	33.33%	100%
21.	Understanding Vocabulary	33.33%	33.33%	33.33%	100%
22.	Finding Specific Information	33.33%	33.33%	33.33%	100%
23.	Determining main idea	33.33%	33.33%	33.33%	100%
24.	Understanding Vocabulary	33.33%	33.33%	33.33%	100%
25.	Inference	33.33%	33.33%	33.33%	100%
26.	Inference	33.33%	33.33%	33.33%	100%
27.	Understanding Vocabulary	33.33%	33.33%	33.33%	100%
28.	Finding Specific Information	33.33%	33.33%	33.33%	100%
29.	Finding Specific Information	33.33%	33.33%	33.33%	100%
30.	Determining main idea	33.33%	33.33%	33.33%	100%
31.	Determining main idea	33.33%	33.33%	33.33%	100%
32.	Finding Specific Information	33.33%	33.33%	33.33%	100%
33.	Inference	33.33%	33.33%	33.33%	100%
34.	Reference	33.33%	33.33%	33.33%	100%
35.	Finding Specific Information	33.33%	33.33%	33.33%	100%
36.	Determining main idea	33.33%	33.33%	33.33%	100%
37.	Finding Specific Information	33.33%	33.33%	33.33%	100%
38.	Understanding Vocabulary	0%	33.33%	33.33%	66.66%
39.	Inference	33.33%	33.33%	33.33%	100%
40.	Reference	33.33%	33.33%	33.33%	100%

Based on Table 3.2 above, the 3 raters agreed with all item numbers of the try-out. Though one of the rater disagreed with some item numbers, such as in item number 6, 8, 12, and 38; but since the 2 other raters agreed with those number items and the total percentage of those number items were 66.66% (>50%); thus, they were taken and considered valid.

3.6.2 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). Meanwhile, Heaton (1991: 162) also states that reliability is a necessary characteristic of any good test, if the tests are given to the same person in other time without any treatment or language learning then it produce different significance results it is no where reliable.

The reseacher used split-half technique to estimate the reliability of the test since this formula is simple to use since 1) it avoids troublesome correlations and 2) in addition to the number of item in the test, it involves only the test mean, and standard deviation, both of which are normally calculated (Heaton, 1991: 164). To used spilt-half method, the researcher classified the test items into two similar parts, i.e. odd and even numbered. By splitting the test into two equal parts, it was made as if the whole tests had been taken twice. The correlation between the those two parts encountered the reliability of half test by using Pearson Product Moment (Henning, 1987: 60). The researcher used Pearson Product Moment to measure the correlation coefficient of the reliability between odd and even number (Henning, 1987: 60) in following formula:

$$r_{xy} = \frac{N \sum xy - (\sum X)(\sum Y)}{\sqrt{[N \sum X^2 - (\sum x)^2][N \sum Y^2 - (\sum Y)^2]}}$$

where:

- r_{xy} : the correlation coefficient of reliability between odd and even
 N : the number of students who take part in the test
 x : the total numbers of odd number items
 y : the total number of even number items
 x^2 : the square of x
 y^2 : the square of y
 $\sum x$: the total score of odd number items
 $\sum y$: the total score of even number items

(Henning,1987: 60)

To measure the coefficient reliability between odd and even number of calculation, the researcher used formula:

$$r_{xy} = \frac{N \sum xy - (\sum X)(\sum Y)}{\sqrt{[N \sum X^2 - (\sum x)^2][N \sum Y^2 - (\sum Y)^2]}}$$

$$r_{xy} = \frac{(30)(5307) - (393)(346)}{\sqrt{[(30 \times 6075) - (393)^2][(30 \times 4710) - (346)^2]}}$$

$$r_{xy} = \frac{159210 - 135978}{\sqrt{[182250 - 154449][141300 - 119716]}}$$

$$r_{xy} = \frac{23232}{\sqrt{[27801][21548]}}$$

$$r_{xy} = \frac{23232}{\sqrt{599055948}}$$

$$r_{xy} = \frac{23232}{24475.61}$$

$$r_{xy} = 0.95$$

After getting the reliability, the researcher then used Spearman Brown's Prophecy formula (Hatch and Farhady, 1982: 246) to determine the reliability of the whole test as follows:

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

where:

r_k : the reliability of the whole test
 r_y : the reliability of half test

(Hatch and Farhady, 1982: 247)

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

$$r_k = \frac{2(0.95)}{1 + 0.95}$$

$$r_k = \frac{1.9}{1.95}$$

$$r_k = 0.97$$

The criteria of reliability are:

0.90-1.00 : high
 0.50-0.89 : moderate
 0.0-0.49 : low

The result computation showed that r_{xy} was 0.95 (see Appendix 16). The result of the whole reliability showed that r_k was 0.97 (see Appendix 16). Based on the criteria of reliability, it can be seen that reliability of the test was high.

3.6.3 Level of difficulty

In order to see the level of difficulty. The researcher used the following formula:

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

where:

FV : Index of difficulty
 R : Number of difficulty
 N : Total number of the students

According to Heaton (1975: 182) states that the criteria of the index of difficulty.

Here are the criteria:

FV= <0.30 : difficult
 FV= 0.30-0.70: average
 FV= >0.70 : easy

After analyzing the criteria of good test by using level of difficulty and discrimination power (see Appendix 12), it can be seen that 15 items were dropped, such as item number 2, 4, 7, 8, 10, 13, 14, 17, 19, 20, 24, 27, 31, 32, and 37. The try-out test consisted of 6 difficulty items (4, 7, 8, 20, 34 and 37); 12 average items (6, 9, 14, 18, 21, 26, 30, 31, 32, 33, 35, and 36); and 22 easy items (1, 2, 3, 5, 10, 11, 12, 13, 15, 16, 17, 19, 22, 23, 24, 25, 27, 28, 29, 38, 39, and 40). The poor and difficult items were revised, while the average and satisfactory items were administrated in the pre-test and post-test.

3.6.4 Discrimination Power

Discrimination power is another important characteristic of a test item which discriminates between weak and strong examines in the ability being tested. The

result of student' test was divided into two groups, upper and lower students. The students in the upper groups is the students who can answer the questions correctly more than the lower students who answer correctly (upper students > lower students). Heaton (1975: 182) states that the formula to determine of discrimination as follow:

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

where:

DP : discrimination power
 U : number of upper group who answer correctly
 L : number of lower group who answer correctly
 N : the total number of the students

The criteria:

DP= 0.00- 0.19 : poor
 DP= 0.20-0.39 : satisfactory
 DP= 0.40-0.60 : good
 DP= 0.70-100 : excellent
 DP= - (Negative) : bad items (be omitted)

Dicrimination power will have positive result, when the upper students get correct answer than lower students. The result is zero, it means that there is no discrimination. The result have negative discrimination power when the lower students get correct answer than upper lower.

In analyzing discrimination power (see Appendix 12), there were 15 poor items (2, 4, 7, 8, 10, 13, 14, 17, 19, 20, 24, 27, 31, 32, and 37), 21 satisfactory items (1, 3, 9, 11, 12, 15, 16, 18, 21, 22, 23, 25, 26, 29, 30, 34, 35, 36, 38, 39, and 40) and 4

good items (5, 6, 29, 33). The items which had negative discrimination were revised, in which the items which had satisfactory level of difficulty and good discrimination indexes were administered for the pre-test and post-test.

3.7 Data Analysis

In order to find out whether implementing RTT in narrative text can increase the students' reading comprehension achievement, the data analyzed by these following procedures:

1. Scoring the pre-test and post-test.
2. Tabulating the results of the test and calculating the scores of the pre-test and post-test.
3. Drawing conclusion by tabulated-result of the pre-test and post-test which statistically analyzed using Repeated Measures T-test computed through SPSS version 17.0 windows to test how significant the difference between the score of pre-test and post-test, in which the significance will be determined by $p < 0.05$ (Hatch & Farhady, 1982:114).

Furthermore, in order to find out the problems faced by the students in learning of reading comprehension through RTT, triangulation method was used. The first instrument was used is observation, which was conducted in XI IPS 3 class to observe the teaching-learning process during the treatment of teaching reading comprehension through RTT. The second instrument was used is interview, which was conducted in XI IPS 3 class in the form of open and formal questions (the

questions must be in the form of explanation or description rather than “yes” or “no” answers, to avoid the students from being reluctant to answer the questions given). To analyze its qualitative data, matrix analysis, in this case description analysis was used, since the researcher used her own idea, including her own interpretation toward the data. (Setiyadi, 2006:262).

3.8 Normality Test

The purpose of computing the normality test was find out whether the data were distributed normally, since it is one of the perquisites to compute data using T-test. In this research, the significant level of 0.05 was used in determining the normality of the data. The hypothesis of the normal distribution test were:

H_0 : the distribution of the data is normal

H_1 : the distribution of data is not normal

The hypothesis is accepted if the result of normality test is higher than 0.05 ($\text{sig} > \alpha$). In this case, level of significance of 0.05 was used. From the result of normality test, see Appendix 20, it can be determined that $\text{sig} > \alpha$ ($0.28 > 0.05$) in pretest and $\text{sig} > \alpha$ ($0.50 > 0.05$) in posttest of experimental class. It proved that H_0 was accepted and the data were normally.

3.9 Hypotesis Testing

The hypothesis was used to prove whether the hypothesis in this research was accepted or not. The hypothesis of this research was there is any increase of students' reading comprehension achievement after being taught through RTT.

The hypothesis was statistically analyzed using Repeated Measures T-test that was used to draw conclusion in significant level of 0.05 ($P < 0.05$).

The formulation is:

$$SD = \frac{\sqrt{\sum d^2 - \frac{(\sum d)^2}{n}}}{n-1} \quad Sd = \frac{SD}{\sqrt{n}} \quad r = \frac{T_1 - T_2}{Sd} \quad df = n - 1$$

Notes:

- r : Ratio
- T_1 : Mean from pre-test
- T_2 : Mean of post-test
- S_d : Standard error of differences between means
- d : Error of differences between mean
- n : Subjects on sample
- SD : Standard Deviation
- df : Degree of freedom

(Hatch and Farhady, 1982:114)

To determine whether the first hypothesis is accepted or rejected, the following criteria acceptance is used:

- H_0 : There is no increase of students' reading comprehension achievement in narrative text through RTT.
- H_1 : There is an increase of students' reading comprehension achievement in narrative text through RTT.

The criteria are:

1. If the t-ratio is higher than t-table: H_1 is accepted
2. If the t-ratio is lower than t-table: H_0 is accepted

Besides, the observation and the interview were conducted in this research. The interview in XI IPS 3 class was in the form of open and formal questions (the questions must be in the form of explanation or description rather than “yes” or “no” answers, to avoid the students from being reluctant to answer the questions given); to analyze its qualitative data in order to find out the problems faced by the students in learning reading comprehension through RTT, matrix analysis, in this case description analysis was used, since the researcher used her own idea, including her own interpretation toward the data (Setiyadi, 2006:262).