III. RESEARCH METHODS

The design of this research was organized in a manner that allows composition between pretest and posttest result. To do this, analytical data measuring the caliber of students' writing was collected using pretest and posttest. Using a controlled analysis of the result as channeled through the research design, the transformation of the subjects' writing skills could be adequately revealed and composed across the two- test period.

3.1 Design

This research was a quantitative study which was intended to see the students' descriptive writing improvement after the implementation of the teacher's indirect feedback. The research design was one group pretest-posttest design. The design was used to compare the sudents' ability in pre test and postest after the treatment was given. This design used pretest to find out the students' initial ability before the treatment (Hatch and Farhady, 1982: 21-22). It could be illustrated as follows:

T1	**X**	**T2**

T1 : Pretest

T2 : Posttest

X : Treatment by the researcher (teacher's indirect feedback)

3.2 Population and Sample

The population of this research was the first year of MA Al-Hikmah Bandar Lampung in the academic year of 2013/2014. There were three classes of the first year students. A class was taken as the sample of this research to be the experimental class. The chosen class was XB which consisted 27 students. In determining the experimental class, this study used simple probability sampling by using lottery; so that those all the first year classes in the school got the same chance to be the sample.

3.3 Data Collecting Technique

The aim of this research was to gain the data on the students' descriptive writing ability score before the treatment (pre test) and after the treatment (post test).The text used was descriptive text concerning five aspects of writing: content, organization, language use, vocabulary, and mechanics. Pre test and post test were conducted to see whether there was a significant improvement on sudents' score after the implementation of teacher's indirect feedback. The data was gained from:

1. **Pre test**

   The pretest was conducted before the treatment was administered. It was administered to the experimental class. It was to see the basic quality of students' descriptive writing performances before receiving the treatment. The pretest was a writing test. The pretest was conducted in 90 minutes. Instructions used by the researcher for the pre- test were:

a. Make a descriptive text about one of your family member (father, mother, sister, or brother)
b. Make it in three or more paragraphs. Each paragraph consists of four or more sentences.
c. Becareful in choosing the words, especially adjective.
d. Pay attention to your grammatical structure. Check them carefully before you submit it.

2. **Post test**

The post test was conducted after the treatment was administered. It was used to know the improvement of student's ability in writing simple descriptive text. It was conducted in 90 minutes. The post test was administered once.Instructions used by the researcher for the post- test were:

a. Make a descriptive text about an important person in your life (friend, teacher, etc.)
b. Make it in three or more paragraphs. Each paragraph consists of four or more sentences.
c. Becareful in choosing the words, especially adjective.
d. Pay attention to your grammatical structure. Check them carefully before you submit it.

3.4 Procedures of Data Collecting Technique

In collecting the data, this study usedthe following steps:

1. **Selecting materials for treatment**

   In selecting materials for treatment, the researcher selected some samples of descriptive text from English books and the internet.

2. **Determining the population and selecting sample**

   The population of this research was the first year of MA Al-Hikmah Bandar Lampung. There were three classes, the researcher chose one class as the experimental class randomly by using lottery, since every class has the same opportunity to be chosen.

3. **Administering the pretest**

   The pretestwas conducted to measure students' preliminary ability before treatment. Here, students in experimental class were assigned to write a descriptive text. The topic was about family member and the time allocation was 90 minutes.

4. **Conducting the treatment**

   After giving the pretest to the students, the experimental class was given treatment by using teacher's indirect feedback.The treatment was conducted in 90 minutes, based on the time allocation in the syllabus of the first grade of

SMA. The treatment was conducted in three meetings. In those three meetings the students was guided to write a descriptive text. Indirect feedback was given to the errors on every students' composition until the student make a good composition from the first to the final draft.After the treatment was given, the posttest was given to the students to evaluate their ability in writing descriptive text after the implementation of teacher's indirect feedback.

5. **Administering the posttest**

   In order to see the improvement of student's writing ability, the posttest was conducted in the experimental class after they were being given the treatment. The test was in form of writing. The students were asked to develop their descriptive text writing based on the topic of important person in my life. The posttest was conducted in 90 minutes.

6. **Analyzing the test result (pretest and posttest)**

   After scoring pretest and posttest, the data was analyzed by using SPSS version 17.0 software program. It was used to find out the means of pretest and posttest and how significant the improvement was.

3.5 Scoring Criteria

The students could succeed in writing if their writing includes five aspects of writing. Therefore, the aspects of writing were evaluated in the students' paragraph writing in the form of simple descriptive text. They were content, organization, grammar,

mechanics, and vocabulary. The scoring criteria was modified from ESL composition profile designed by Jacobs et al (1981) as can be seen below:

Table 3.1. The Scoring Criteria

ASPECTS OF WRITING	SCORE	CRITERIA
Content	20	EXCELLENT TO VERY GOOD: knowledgeable. substantive. thorough development of thesis. relevant to assigned topic.
	15	GOOD TO AVERAGE: some knowledge of subject. adequate range. limited development of thesis. mostly relevant to topic, but lacks in detail.
	10	FAIR TO POOR: limited knowledge of subject. little substance. inadequate development of topic.
	5	VERY POOR: limited knowledge of subject. non-substantive. not pertinent. or not enough to evaluate.
Organization	20	EXCELLENT TO VERY GOOD: fluent expression. ideas clearly stated/supported. succint. well-organized. logical sequencing. cohesive.
	15	GOOD TO AVERAGE: somewhat choppy. loosely organized but main ideas stand out. limited support. logical but incomplete sequencing.
	10	FAIR TO POOR: non-fluent. ideas confused or disconnected. lacks logical sequencing and development.
	5	VERY POOR: does not communicate. no organization. Or not enough to evaluate.
Vocabulary	20	EXCELLENT TO VERY GOOD: sophisticated range. effective word/idiom choice and usage. word form mastery. appropriate register
	15	GOOD TO AVERAGE: adequate range. occasional errors of word/idiom form, choice, and usage but meaning not obscured.
	10	FAIR TO POOR: limited range. frequent errors of word/idiom form, choice, and usage. meaning confused or obscured.
	5	VERY POOR: essentialy translation. little knowledge of

		English vocabulary, idioms, word form. or not enough to evaluate.
Language use (grammar)	20	EXCELLENT TO VERY GOOD: effective complex constructions. few errors of agreement, tense, number, word order/function, articles, pronouns, prepositions.
	15	GOOD TO AVERAGE: effective but simple constructions. minor problems in complex constructions. several errors of agreement, tense, number, word order/function, articles, pronouns, prepositions. but meaning seldom obscured.
	10	FAIR TO POOR: major problems in simple/complex constructions. frequent errors of negation, agreement, tense, number, word order/function, articles, pronouns, prepositions and/or fragments, run-ons, deletions. meaning confused or obscured
	5	VERY POOR: virtually no mastery of sentence construction rules. dominated by errors. does not communicate. or not enough to evaluate.
Mechanics	20	EXCELLENT TO VERY GOOD: demonstrates mastery of conventions. few errors of spelling, punctuation, capitalization, and paragraphing.
	15	GOOD TO AVERAGE: occasional errors of spelling, punctuation, capitalization, and paragraphing but meaning not obscured.
	10	FAIR TO POOR: frequent errors of spelling, punctuation, capitalization, and paragraphing. poor handwriting. meaning confused or obscured.
	5	VERY POOR: no mastery of conventions. dominated by errors of spelling, punctuation, capitalization, and paragraphing. handwriting illegible or not enough to evaluate.

Jacobs et al (1981)

3.6 Instruments

Writing test was a device which requires the students to compose their own idea, and extends responses to problem set by the teacher. The instrument of this research was descriptive text writing. The researcher administered writing test to find out whether there was any improvement of students' descriptive text writing ability after the

implementation of teacher's indirect feedback or not. That was why the students were asked to write a descriptive text. The students were given a chance to make writing composition for about 90 minutes.

3.7 Validity

A test could be considered valid if the test measures the objectives to be measured and suitable with the criteria (Hatch and Farhady, 1982: 250). According to Hatch and Farhady (1982: 281) there are two basic types of validity; content validity and construct validity. In order to measure whether the test has a good validity, those two types of validity were analyzed.

Content validity is concerned with whether the test is sufficiently representative and comprehensive for the test. In the content validity, the material given is suitable with the curriculum. Content validity is the extent to which a test measures a representative sample of the subject meter content, the focus of content validity is adequacy of the sample and simply on the appearance of the test (Hatch and Farhady, 1982: 251). This study used descriptive writing test that was supposed to be comprehended by the first year of senior high school students. The test was considered as valid in content validity since the test of writing constitutes a representatives sample of the language skill and structure and also the material used was chosen based on 2006 English Curriculum of KTSP for first year of senior high school.

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, it was examined whether the test given actually reflect what it means to know a language. In this research, scoring criteria was based on the five aspects of writing; content, organization, language use, vocabulary, and mechanics that were suggested by Jacobs et al (1981: 90).

3.8 Reliability

Hatch and Farhady (1982:243) established that the reliability of a test could be defined as the extent to which a test produces consistent result when it administered under similar conditions. A test could be considered reliable if the tests has a consistent result. In order to ensure the reliability of scores and to avoid the subjectivity of the research, there was inter-rater reliability. Inter-rater reliability was used when score on the test is independently estimated by two or more judges or raters. In this case, the first rater was researcher and the secondwas English teacher in MA Al-Hikmah Bandar Lampung. Before scoring the students' descriptive text writing, it was important to make sure that both raters use the same criteria of scoring. Hereby, the first and the second rater used scoring criteria devised from Jacobs et al (1981: 90). To measure how reliable the scoring was, this study used *Rank – order Correlation* with the formula:

$$\boldsymbol{p}\mathbf{= 1 -} \frac{6.\sum d^2.}{N(N^2-1)}$$

p : Coefficient of rank order

d : Difference of rank correlation

N : Number of students

1-6 : Constant number

(Hatch and Farhady, 1982: 206)

In this case, the coefficient of rank correlation was analyzed with the standard of reliability as follows:

1. 0.80000 - 1.0000 : very high reliability
2. 0.60000 - 0.7900 : high reliability
3. 0.40000 - 0.5900 : medium reliability
4. 0.20000 - 0.3900 : low reliability
5. 0.0000 – 0.1900 : very low reliability

Based on the standard of reliability above, it could be concluded that the writing tests would be considered reliable if the tests reached the range of 0.60-0.79 (high reliability).

The reliability of this research could be seen on the explanation below:

1. **Result of Reliability of the Score in Pretest**

$$\boldsymbol{p} = 1 - \frac{6.\Sigma d^2}{N(N^2-1)}$$

$$\boldsymbol{p} = 1 - \frac{6.\ (646.5)}{27(729-1)}$$

$$\boldsymbol{p} = 1 - \frac{4200}{19656}$$

$\boldsymbol{p} = 1 - 0.19734432$

$\boldsymbol{p} = 0.80265568$ (Very high reliability)

2. Result of Reliability of the Score in Posttest

$$\boldsymbol{p} = 1 - \frac{6.\Sigma d^2}{N(N^2-1)}$$

$$\boldsymbol{p} = 1 - \frac{6.\ (543.49)}{27(729-1)}$$

$$\boldsymbol{p} = 1 - \frac{3260.94}{19656}$$

$\boldsymbol{p} = 1 - 0.16590049$

$\boldsymbol{p} = 0.83409951$ (Very High Reliability)

3.9 Data Analysis

The result of student's descriptive writing ability in each test was evaluated based on content, language use, organization, vocabulary, and mechanics. The results of students' performance in pretest then werecompared with the result of their performance in posttest to the impact of the instruction in their writing performance. To analyze the data gained from writing test, the researcher treated the data through the following steps:

1. Sorting the data

Each rater scored the students' writing of pretest and posttest. Then, the average scores between two raters were taken to be the final score that was

analyzed statistically using *Repeated Measured T-test (Paired sample T-test)* that was to show the differences between pretest and posttest of experimental class for answering the hypothesis. The data was computed through SPSS version 17.0.

2. **Drawing conclusion**

   The score of the pretest and posttest of two groups were statistically analyzed using *Repeated Measured T-test (Paired Sample T-test)* to draw a conclusion. It was computed through the Statistical Package for Social Science (SPSS) version 17.0.

3.10. Data Treatment

According to Setiyadi (2006:168-169), using T-Test for hypothesis testing has three basic assumptions that can be described as follows:

a. The data is an interval.

b. The data is taken from random sample in population.

c. The data is distributed normally.

Therefore before testing the hypothesis using T-test, it is necessary to find out whether the data in experimental class was normally distributed or not. Since the objective of this study was only to find out the improvement of students' writing ability,the data was treated by only using normality test.

This test was used to measure whether the data is normally distributed or not. The data was tested by *One-sample Kolmogorov-Smirnov* Formula (SPSS 17.0). The criteria of normal distribution were:

H_0 : the distribution of the data is normal

H_1 : the distribution of the data is not normal

The hypothesis would be accepted if the result of the normality test is higher than 0.05 (sign > α). In this case, the researcher used the level of significance of 0.05.

3.11 Hypothesis Testing

The hypotheses were stated as below:

H_0: "There is no positive effect of teacher's indirect feedback in improving students' ability in terms of: content, organization, vocabulary, language use, and mechanics".

$$\mathbf{H_0: X_1 = X_2}$$

H1: "There is positive effect of teacher's indirect feedback in improving students' ability in terms of: content, organization, vocabulary, language use, and mechanics".

$$\mathbf{H_1: X_1 \neq X_2}$$

Repeated Measured t-test (Paired Sample T-test) was used to test the hypothesis.The formulation was:

1. Paired Sample t-test

$$t = \frac{Md}{\sqrt{\frac{\sum x^2 d}{N(N-1)}}}$$

and

$$\sum x^2 d = \sum d^2 - \frac{(\sum d)^2}{N}$$

Md = mean from the differences pretest and posttest (posttest-pretest)

Xd = deviation of each subject (d – md)

$\sum x^2 d$ = total of quadratic deviation

N = subjects on sample

(Arikunto, 2010: 349-350)

The analysis was computed using SPSS version 17.0. The hypothesis was analyzed at significant level of 0.05 ($p<0.05$) in which H_0 would be approved if Sign>α. It means that the probability of error in the hypotheses is only about 5%; and H_1 would be approved if Sign<α (α<0.05).