III. METHOD

In this chapter the researcher explains some information including the design, data, subject, instrument, data collecting technique, research procedure, scoring criteria, validity and reliability, data analysis, treatment of the data, and hypothesis testing.

3.1 Design

This research was quantitative in which the researcher tried to find out whether there was increase of recount text writing ability of the first year students of senior high school taught by using guiding questions technique. In this research, the writer used the one-group pretest-post test design, dealing with one group which received treatment. The function of pretest given to the group was to find out the students' initial ability in writing as the starting point for teacher before doing the treatment. And to see whether the treatment could be applied or not, the writer saw it from the result of the post test given to the students after treatment. The treatment was given one time which consisted of three meetings with different activities in each meeting. The criteria whether guiding questions could increase the students' ability in writing recount text was determined by the differences between the scores of the pre-test and post-test.

The research design is illustrated as follow:

T1 X T2

where

T1 : Pretest (Test before treatment)

X : Treatment by using guiding questions technique

T2 : Post test (Test after treatment)

(Setiyadi 2006: 131)

3.2 Data

The data in this research was in the form of score. The score was collected from the test which was given to the students. There were two types of test: pretest and post test. Pretest was given before the treatment to know the score before giving the treatment, and post test was given to know the score after treatment.

3.3 Population and Sample

This research was conducted at SMAN 1 Terbanggi Besar in even semester of 2011/2012 school year. The population of the research was the first grade of SMAN 1 Terbanggi Besar. There were nine classes of the first grade in that school which consisted of 30-40 students for each class. There was only one class used as the sample chosen by using purposive sampling. The sample was class XH, consisting of 32 students as the source of the data.

3.4 Instrument

In this research the researcher used writing test as the instrument. The researcher chose the writing test because the researcher wanted to find out students' ability in writing recount text. There were two types of test, pretest and post test. Pretest was given to see the students' ability before treatment as the starting point for the researcher to give the treatment. And the post test was given to see the result of the treatment. In pretest, the researcher asked the students to write their own recount text by using their own technique based on the topic which is given by the researcher. And in post test the researcher asked them to write by using guiding question technique.

3.5 Data Collecting Technique

In collecting the data the researcher applied some technique as follow:

1. **Pretest**

   This test was conducted before the treatment. It was used to see the students' ability before the treatment. It also helped the researcher to be the starting point before the researcher applied the treatment. The time given was 90 minutes.

   After doing pretest the researcher made a note consisting of the errors which were made by students and their ability in writing. After getting the data, the researcher evaluated the data. There were two raters who gave the scores (see Appendices 5 and 7). Then the data treated to see whether it was normal, homogenous and random (see Appendices 11, 13 and 14). The researcher also calculated the reliability of the data (see Appendix 9).

2. **Post test**

   This test was conducted after the treatment. It was used to see the score of the students' writing after treatment. The time given was 90 minutes (see Appendix 3). There were two raters in giving the score (see Appendices 6 and 8). After the data have been collected, the data treated to find out whether it was homogenous, normal and random (see Appendices 12, 13 and 14). Then the researcher calculated the reliability of the data (see Appendix 10). After that the researcher compared the result of pretest and post test to find out the increase (see Appendix 15).

3.6 Research Procedure

The procedures of the research were as follow:

1. **Determining the population and sample**

   The writer chose the first grade of SMA N 1 Terbanggi Besar as the population. The researcher chose the first year students because, based on the 2006 English curriculum; recount text had been learned by them at the second semester. There were nine classes. Each class consisted of 30-40 students. The writer took one class by using purposive sampling. The class was XH which consisted of 32 students.

2. **Selecting the materials**

   The materials used in the research were based on the 2006 English curriculum for the first year of SMA students. The researcher used *English for You* an English textbook for second semester of the first year students

of senior high school adapted from the English curriculum of SMA 2006, School Based Curriculum.

3. Administering pre-test

Pre-test was administered to find out the students' basic ability before treatment. The students were asked to write recount text particularly personal recount text that consisted of orientation, series of events, and re-orientation in 200-250 words in about 90 minutes (see Appendix 2).

4. Conducting treatment

In the treatment, which was given one time, the researcher applied the guided writing in the form of guiding questions in teaching recount text. The researcher also explained the objective of the instruction, how to write recount text, and aspects of writing a good writing. The treatment was divided into three meetings with different activities in each meeting. Each meeting lasted for 90 minutes.

5. Administering post-test

Post-test was administered to find out the increase of the students' ability in writing recount text after receiving the treatments. The test was the same as in pre-test (see Appendix 3).

6. Analyzing the test result

After the researcher got the scores of pre-test and post-test, certainly the comparison between pre-test and post test scores in writing recount text by

using guiding questions was appropriately found out (see Appendices 5-15).

3.7 Scoring Criteria

The students' success in writing was determined if their writing included five aspects of writing. Therefore, the researcher used five aspects of writing to evaluate the students' ability in writing recount text:

1. Content refers to substance of recount text (orientation, series of events and re-orientation), the experience of the main idea,
2. Organization refers to the logical organization of the recount text content (the coherence of events series),
3. Vocabulary refers to the selection of words those are suitable with the content,
4. Language use refers to the use of the correct grammatical and syntactic pattern,
5. Mechanic refers to the use graphic conventional of the language.

(Jacobs et. al., (1981:90).

The score of the test in writing recount text would be derived as follows:

1. Content : 30%
2. Organization : 20%
3. Vocabulary : 20%
4. Language use : 25%
5. Mechanics : 5%

The criteria of scoring are also devised from Jacobs et al (1981: 90) as follows:

Content

30 – 27	Excellent to very good: knowledgeable substantive, development of thesis/topic, relevant to assign topic.
26 – 22	Good to average: some knowledge of subject, adequate range, limited development thesis, mostly relevant to topic but lack detail.
21 – 17	Fair to poor: limited knowledge of subject, little substances, inadequate development, of topic.
16 – 13	Very poor: doesn't show knowledge, not pertinent, or not enough to evaluate.

Organization

20 – 18	Excellent to very good: fluent expression, ideas clearly stated/supported, succinct, well organized, logical sequencing, cohesive.
17 – 14	Good to average: somewhat choppy, loosely organized, but main idea stand out, limited support, logical but incomplete sequencing.
13 - 10	Fair to poor: not fluent, ideas confused or disconnect, lacks logical sequencing and development.
9 – 7	Very poor: doesn't communicate, no organization, or not enough to evaluate.

Vocabulary

20 – 18	Excellent to very good: sophisticated range, effective word or idiom choice and usage, word form mastery, appropriate register.
17 – 14	Good to average: adequate range, occasional errors of word or idiom, choice, usage, meaning confused or obscured.
13 – 10	Fair to poor: limited range, frequent errors of word or idioms, choice, usage, meaning confused or obscured.
9 – 7	Very poor: essentially translation, little knowledge of vocabulary, idioms, word form, or not enough to evaluate.

Language use

25 – 22	Excellent to very good: effective complex construction, few errors of agreement, tense number, word order /function, articles, pronoun, preposition.
21 – 18	Good to average: effective but simple construction, minor problems in simple construction, several errors of agreement , tense, word order/function, articles, pronouns, prepositions, but meaning seldom obscure.
17 – 11	Fair to poor: major problem in complex/simple construction, frequent errors of negation, agreement, tense, number, word order/function, articles, pronouns, prepositions, and/or fragments, run-ons, deletions, meaning confused, or obscured.
10 – 5	Very poor: virtually no mastery of sentence construction rules, dominated by errors, does not communicate, or not enough to evaluate.

Mechanics

5	Excellent to very good: demonstrated mastery of conventions, few errors spelling, punctuation, capitalization, paragraphing.
4	Good to average: occasional errors of spelling, punctuation, capitalization, paragraphing, but meaning not obscured.
3	Fair to poor: frequent errors of spelling, punctuation, capitalization, paragraphing, poor hand writing, meaning confused or obscured.
2	very poor: no mastery of conventions, dominated by errors of spelling, punctuations, capitalization, paragraphing, handwriting illegible, or not enough to evaluate.

3.8 Validity and Reliability

3.8.1. Validity of the Test

Validity is a matter of relevance. It means that the test measures what is supposed to be measured. To measure whether the test has a good validity, the researcher analyzed the test from content validity and constructs validity.

Content validity is concerned with whether the test is sufficiently representative and comprehensive for the test. In the content validity the material given was suitable with the curriculum used. In this research, it could be seen that the instrument which was used in this research is valid in content because in the research, the researcher used recount text writing test that is supposed to be comprehended by the first grade of SMA students.

Beside that, the instrument is also valid in construct. The test was considered as valid in content since the test of writing constitutes a representatives sample of the language skill and structure and also the material was chosen based on 2006 English Curriculum for first year of SMA.

Construct validity is the process of determining the extent to which test performance can be interpreted in terms of one or more constructs (Gronlund, 1985:83). In this research, the researcher administered a writing test and the technique of scoring the students' writing is based on the five aspects of writing; they are content, organization, vocabulary, language use, and mechanic. Therefore it can be concluded that it is valid in construct.

3.8.2. Reliability of the Test

Reliability has to deal with the quality of measurements, the consistency of the measurement device, or the degree to which an instrument measures the same way each time it is used under the same condition with the same subjects. In short, it is the repeatability of the measurements. A measure is considered reliable if it gives the same result repeatedly.

In this research, the writer used inter-rater reliability. It referred to the concern that students' score may vary from rater to rater. There was another person who gave the score besides the writer. In this case the researcher was the first rater and then the writer asked Supriyanto, S.Pd, an English teacher who taught XH in SMA N 1 Terbanggi Besar, to be the second rater.

After calculating the students' writing scores, the researcher calculated the data by using rank order formulation to test the reliability of the device. Reliability is the measure of how stable, dependable, trustworthy, and consistent a test is in measuring the same thing each time (Worthen et al.,

1993). In this case, the researcher used two raters in scoring the students' writing test. The formula is as follows:

$$r = 1 - \frac{6\ x\ \sum D^2}{N\ (\ N^2 - 1\ )}$$

Where:

R : Rank – difference

∑D : The sum of difference between each pairs of ranks

(Harris, 1974: 142)

In this case, the writer also uses the standard of reliability (Arikunto, 1998:260) below:

0.81 – 1.0 = very high

0.61 – 0.8 = high

0.41 – 0.6 = medium

0.21 – 0.4 = low

0 – 0.2 = very low

After calculating the result of the students' recount text writing, the researcher calculated the data by using the formula above (see Appendices 9 and 10). The result of the reliability could be seen in the following tables:

Table 3.1. Reliability of the Test in Experimental Class

	Pre Test	Post Test	Criteria
Reliability	0.99578446	0.99221041	Very high reliability

From the criteria of reliability and calculation, it can be concluded that the reliability of the raters in the experimental class is very high. It means that the researcher's way of scoring the data is similar to the second rater. They had almost the same scoring system therefore there is no subjectivity in scoring the students' writing. Beside that the scoring criteria helps the raters in scoring the students' writing accurately. In addition, the result shows that both raters scored the students' writing consistently and fairly.

3.9 Data Analysis

To measure the students' progress in writing, the researcher analyzed the students score through this activity:

- Scoring the pre-test and post test
- Finding the mean of the pre-test and post test using this formula:

$$Md = \frac{\sum d}{N}$$

M*d* = mean

$\sum$ = total score of the students

N = number of students

- Drawing conclusion by comparing the means of the pre-test and post-test

3.10 Treatment of the Data

The writer used the following procedure in treating the data.

1. Random Test

The random test by using statistical formula of descriptive statistics (SPSS 17) is used to determine whether the data of the students' writing in

experimental is taken from the subject at random. The hypotheses for random test were:

H0: The data is random

H1: The data is not random

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

Random test of pre test in experimental class shows that the two tailed significance is 0.590 (see Appendix 14). Seeing the result, it can be inferred that the data is random since Sign > α (0.590 > 0.05). Meanwhile, the result of the random test of post test shows in the number of 0.604 (see Appendix 14). The value is also > α, in which 0.604 > 0.05. Beside that we could see that n1 < r < n2 in which n1 is cases < test value, r is number of runs, n2 is cases >= test value. It can be seen that in pre test 15 < 16 < 17 and in post test 14 < 15 < 18. It can be stated that the data is random.

In short, the data in experimental class shows that the value of two tailed significance is higher than alpha (sign > α). It is concluded that the data are random.

2. Normality test

The normality test is used to determine whether the data in experimental class are normally distributed. The writer used One-Sample Kolmogorov-Smirnov

Formula (SPSS 17) to test the normality of the data. In this research, the significant level of 0.05 is used. The hypotheses of normal distribution were:

H0: The distribution of the data is normal

H1: The distribution of the data in not normal

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

The result of normality test of pre test in experimental class shows that the value of two tailed significance is 0.200 (see Appendix 11). It means that H0 is accepted and H1 is rejected since 0.200 > 0.05. It implied that the distribution of the test was normal. The result of normality test of post test shows that the value of two tailed significance is 0.200 (see Appendix 12). Since 0.200 > 0.05, it can be stated that the data of post test is normally distributed.

From the result of normality test above, it can be concluded that the hypotheses was accepted which means that the data is normally distributed.

3. Homogeneity Test

The homogeneity test is used to determine whether the data in experimental are homogenous. The researcher used Paired Sample Test to know the homogeneity of the test. The hypotheses for the homogeneity test were:

H0: There is no significant difference in the level of ability (equal)

H1: There is a significant difference in the level of ability (not equal)

The criterion for the hypotheses is H0 is accepted if the result of homogeneity test is higher than 0.05 (sign > α).

The result of homogeneity test of pretest shows that the value of two tailed significance is 0.176 and for post test 0.439 (see Appendix 13). In this case H1 is rejected since 0.176 > 0.05 and 0.439 > 0.05. It means that the XH class has similar ability.

3.11 Hypothesis Testing

To know the increase, the researcher compared the result of pre-test and post-test. After getting the means scores of pre-test and post-test, the researcher analyzed the data using paired sample test to find the increase of the treatment effect. The hypotheses are as follows:

Ha = There is an increase of recount text writing ability of the first year senior high school students who are taught by using guiding questions technique.

Ho = There is no increase of recount text writing ability of the first year senior high school students who are taught by using guiding questions technique.

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

Ha (alternative hypothesis) is accepted if alpha level (α) is lower than 0.05 (α<0.05).

Ho (null hypotheses) is accepted if alpha level (α) is higher than 0.05 (α>0.05).