

III. RESEARCH METHOD

This chapter illustrates how the research was done; what design of the research was, who the population and the sample were, and how the data were gathered. It also covers the validity and reliability of the instrument, treatment of the data, and data analysis.

3.1 Research Design

This research was quantitative design. The writer used factorial research design. A factor was a discrete variable used to classify experimental units. In this case, there were two factors; they were extrovert and introvert. A factorial design was the most common way to study the effect of two or more independent variables, although it would focus on designs that had only two independent variables for simplicity. The design of this research was as follow:

Personality	Pretest	Posttest	Gain of Listening Achievement
Introvert			
Extrovert			

It compared the variable with the score which was got by the pretest and posttest.

The design was as follow:

$$T_1 \text{ X } T_2$$

Where :

T_1 : Pretest

X : Treatment

T_2 : Posttest

(Setiyadi, 2006:131)

There were two variables that were organized in this research: they were dependent and independent variables. It was a “product” as a result of interaction between variable involved in that particular research while independent variable was the variable whose function was to influence the dependent variable (Setiyadi, 2006:107). From the explanation above, the writer determined the variables as follow:

1. The introvert students were as independent variable. (x_1)
2. The extrovert students were as independent variable. (x_2)
3. Students listening achievements were as the dependent variable. (y)

In order to find students who pose the independent variables, questionnaire was given to the students to be answered. Based on the result of the questionnaire, the writer classified the students into three groups; introvert, mediocre, and extrovert. The introvert and the extrovert groups were taken as the independent variable. Meanwhile, the dependent variable of the research was obtained from the students' result of listening test.

There would be a pre-test before teaching listening and post-test after teaching listening to see the gain of listening score. The reason of choosing pre-test and post-test was in order to get primary data which were more reliable than just conducting listening test in one meeting. The writer assumed that conducting test in one meeting without any treatment was not really reliable because the result of the test might be influenced by other factors at that time. But, by conducting the pre-test, treatment and post-test the score was more reliable.

3.2 Population and Sample

The population of this research was students in the first grade of SMA Kartikatama Metro. The writer decided to take two classes that have same characteristics as the sample of the research. It was called purposive sample. The sample class was selected based on the purpose of study. To know whether the class had the same characteristics or not, the writer would took the data of students' score at the first semester from the teacher.

3.3 Research Procedure

The procedures of the research was as follows:

1. Determining the Population and Selecting the Samples

The population of this research was the first grade students of SMA Kartikatama Metro in the 2012/2013 learning year. The sample class was selected using purposive sampling. It means the sample class was selected based on the purpose of the study. It was two classes as the sample which had the same characteristics. To determine the class was the sample class, the writer took the data of students' score at the first semester of the first grade in the 2012/2013 learning year.

2. Selecting Listening Test

In selecting the listening test, the writer took a look at the syllabus used by the teacher of the sample class. Any material being taught which was corresponding with listening could be taken into the listening test to see their gain in listening skill.

3. Distributing Questionnaire

The writer gave the students questionnaire to the students to be answered. Students were given 10 minutes to answer the questionnaire. The result of this questionnaire was used to group the students based on their type of personality.

4. Conducting Try-out

After distributing questionnaire, the writer decided to conduct try out test to examine whether the item test were good enough and appropriate for the students.

5. Conducting Pre-test

The next step was administering the pre-test to the students to see their score before getting treatment from the researcher.

6. Teaching Listening

After conducting pre-test, the researcher gave treatment to the students by teaching listening. The researcher taught listening based on the syllabus that used by the teacher of sample class. The writer decided to choose listening to narrative text because it referred to the macro skills which become the focus of the research.

7. Conducting Post-test

After giving treatment to the students, the researcher conducted post-test to see the students' achievement after getting treatment from the researcher.

8. Analyzing, Interpreting and Concluding the Data

After collecting the data, the analyzing, interpreting, and concluding the data gained was done. First, the data gained from the test were tabulated and calculated. Next, the data were divided into two group based on the

students type of personality. Independent group t-test was then used to see if the hypothesis were accepted or rejected.

3.4 Research Instrument

The instrument used for collecting data were as follows:

1. Questionnaire

In order to collect the data, the writer used questionnaire as the tool of measurement. Questionnaire was an instrument which was very effective to measure aspects and variables in associated with personality, psychology aspect or sociology (Setiyadi, 2006). The questionnaire for personality test was taken from the test that was used many times to measure personality. There was no doubt of this test because it was taken from the expert namely Eysenck Personality Inventory (EPI). The questionnaire was taken and translated into Bahasa Indonesia in order to minimize the misinterpretation by the students. The questionnaire consisted of 20 items which has 4 options in each question. The scoring system will be : a = 4, b = 3, c = 2, d = 1

Further, the following table was a table of specification of questionnaire items.

Table 3.1 Table of specification (Questionnaire)

Items number	Personality				Items	Percentage
	Introversion		Extroversion			
	Melancholic	Phlegmatic	Choleric	Sanguine		
1,2,3,4,5	x				5	25%
6,7,8,9,10		x			5	25%
11,12,13,14,15			x		5	25%
16,17,18,19,20				x	5	25%

By using the result of the questionnaire, the writer classified the students personality. Since the items of introvert were 10 items and the items of extrovert were 10 items, the score would be compare. If the score of introvert items were higher and the score of extrovert items were lower, the participants would be classified as introvert group. If the score of introvert items were lower and the score of extrovert items were higher, the participants would be classified as extrovert group. If the score of introvert as higher as extrovert, the participants would be classified as mediocre group.

2. Listening Test

The writer decided to administer listening test in order to get primary data from the students. The test was in the form of pre-test before the researcher gave treatment to the students and post-test after the researcher gave treatment to the students. The material to be tested was listening to the narrative text.

In scoring student's test, Arikunto's formula was used. The ideal high score was 100. So, the formula which will be used was as follow :

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

Where :

S : The score of the test

R : The total of the right answer

N : The total items

(Arikunto, 1996: 212)

The writer decided to conduct pre-test and post-test because it was more reliable than simply conducting listening test once in one meeting because the result of the test might be influenced by other factors at that time. If the test conducted two times, it would be more reliable.

3.5 Validity and Realibility of the Instrument

1. Validity of the Instrument

- **Validity of Questionnaire**

Validity was a matter of relevance; it means that the test measures what was claimed to measure. To measure whether the test has a good validity,

it could be analyzed from its content validity and construct validity. Content validity was concerned whether the test was sufficiently representative for the rest of the test or not. While construct validity focuses on the relationship between indicators within the test. Since purpose of the test was to measure as well as to investigate students' personality, the writer applied a test that deals with the students' personality test developed by Eysenck (1961) namely Eysenck Personality Inventory. This was used to measure or classify the respondent to the type of extrovert and introvert. There was no doubt feeling to this standard test, because it was already constructed by the expert and it measured about personality which had been tested many times. To measure the validity of questionnaire, the researcher used inter-rater reliability where there were 3 raters or judges to make sure that the questionnaire was valid.

- **Validity of Listening Test**

As well as for listening test, to claim it was valid, the writer took the topic from the syllabus that had been taught to the sample class, in this case the writer took the topic of narrative text. Then, to ensure that the test was valid the test items should fulfill the construct validity. The table below was table of specification of listening test.

Tabel 3.2 Table of Specification (Listening Test)

No.	Macro aspects of listening	Item	Percentage
1	Determing main idea	7	23.33%
2	Finding specific information	14	46.67%
3	Inference	6	20%
4	Vocabulary	3	10%
Total		30	100%

Since the writer put focus on macro skills, the item test would be the macro aspects of listening which consisted of determining main idea, finding specific information, inference, and vocabulary.

2. Reliability of the Instrument

- **Reliability of Questionnaire**

Reliability refers to the consistency of the measure. A test was defined to be reliable if its scores remain relatively stable from one administration to another (Hatch and Farhady, 1982:144). First of all, the result of the questionnaire was scores based on Likert scale with range of score was 1 to 4. In order to measure the consistency of items in the questionnaire, the writer used *Cronbach Alpha Coefficient* since it was the most commonly used one. The alpha ranges between 0 and 1.

- **Reliability of Listening Test**

For the listening test, reliability of the test can be defined as the extent to which a test produces consistent result when administered under similar conditions (Hatch and Farhady, 1982:243). Pearson Product Moment formula that was used as follows:

$$r_{11} = \frac{\sum xy}{\sqrt{[\sum x^2][\sum y^2]}}$$

Where:

r_{11} : 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

The criteria of reliability were:

0.80 – 1.00 = very high

0.50 – 0.79 = moderate

0.00 – 0.49 = low

(Hatch and Farhady, 1982: 247)

To know the coefficient correlation of whole items, “Spearman Brown’s prophecy formula” was used. The formula was as follows:

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

Where:

r_k : The reliability of the whole test

r_{xy} : The reliability of the half test

(Hatch and Farhady, 1982:246)

3.6 Level of Difficulty

Level of difficulty was related to how easy or difficult the item was from point of view of the students who take the test. To know the level of difficulty, the researcher used the following formula:

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

Where:

LD: Level of difficulty

R: The number of students who answer correctly

N: The total number of students following the test

(Heaton, 1991: 182)

The criteria were:

<0.30 = difficult

0.30-0.70 = average

>0.70 = easy

3.7 Discrimination Power

The discrimination power refers to the extent to which the item differentiates between high and low level students on the test. A good item according to the criteria was one which good students will do well and bad students will fail.

To know the discrimination power of the test, the formula that was used:

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

Where:

DP: Discrimination power

U: The proportion of upper group students

L: The proportion of lower group students

N: Total number of the students

The criteria were:

0.00 – 0.20 = poor

0.21 – 0.40 = satisfied

0.41 – 0.70 = good

0.71 – 1.00 = excellent

(Negative) = bad items (should be omitted)

(Heaton, 1991: 182)

3.8 Treatment of the Data

There were three underlying assumptions that need to be fulfilled if we were going to use *Independent Group T-test*, namely :

1. The data was interval or ratio
2. The data was taken from random sample in a population
3. The data was distributed normally

(Setiyadi, 2006:170)

Although the sample were not taken by randomly, the *Independent group t-test* was able to use in order to see whether the hypothesis were accepted or rejected as long as the data was distributed normally. If the data was not distributed normally, the *Mann-Whitney U Test* would be used to see whether the hypothesis were accepted or rejected.

Therefore, the writer used the following procedures to treat the data:

1. Normality Test

The normality test was used to measure whether the data from students score were normally distributed or not. The writer used SPSS 17 to analyze the data. The hypothesis for the normality test were as follow:

H_0 : the data was not distributed normally

H_1 : the data was distributed normally

The criteria for the hypothesis was H_1 was accepted if $\text{sign} > \alpha$, with the level of significance 0.05.

2. Hypothesis Test

Last, the writer tested the hypothesis whether it was accepted or rejected. First, the writer analyzed the data from questionnaire to categorize the students into three groups. Two groups (Introvert and Extrovert) were analyzed further. Their data from the listening test was analyzed to find out if the hypothesis was accepted or rejected by using the statistical analysis t-test with the level of significance $\alpha = 0.05$.

The formulation was as follow:

$$t = \frac{x_1 - x_2}{S \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

With :

$$S^2 = \frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2}$$

x_1 : the arithmetical mean of the introvert group

x_2 : the arithmetical mean of extrovert group

S : standard deviation

n_1 : the number of students in extrovert group

n_2 : the number of students in introvert group

The proposed hypothesis were:

H_0 : Students with introvert personality do not have better achievement in listening than the extrovert ones.

H_1 : Students with introvert personality have better achievement in listening than the introvert ones.

The writer used one-tailed t-test formula in SPSS 17 to make it easier in doing the calculation, with the level significant 0.05.

The criteria were:

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

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