## CHAPTER III

## RESEARCH METHODOLOGY

## A. Research Design

This research is an experimental research. Based on what Cresswell said that we use experimental research when we want to establish possible cause and effect between our independent and dependent variables (Cresswell, 2012). A quasi experimental study might compare outcomes for individuals receiving program activities with outcomes for a similar group of individuals not receiving program activities. The type of this research also might compare outcomes for one group of individuals before and after the group's involvement in a program (pre-test/post-test design). The design of this research is a quasiexperimental research, to know the effect of using Crazy Story Game on speaking ability.

This research design used two groups. The first group was an experimental group treated by using Crazy Story Game and symbolized as (X) and the second group was control group treated without using Crazy Story Game and symbolized as (Y). In conducting this research, the researcher involved the eleventh grade students of Vocational High School Pharmacy Ikasari Pekanbaru, the classes were divided into an experimental class and a control class. Both of the classes were given a pre-test to know students' ability in speaking.

After that, the experimental class was given the treatment by using Crazy Story Game for 6 meetings while the control class was given conventional teaching strategy.

## Table III. 1 <br> The Research Design

| Group | Pre-test | Treatment | Post-test |
| :---: | :---: | :---: | :---: |
| Experiment | $\mathrm{X}_{1}$ | T | $\mathrm{Y}_{1}$ |
| Control | $\mathrm{X}_{2}$ | - | $\mathrm{Y}_{2}$ |

Note:

$$
\begin{array}{ll}
\mathrm{X}_{1} & =\text { Pre-test for experimental group } \\
\mathrm{X}_{2} & =\text { Pre-test for control group } \\
\mathrm{T} & =\text { Treatment for experimental group by using CSG } \\
- & \text { The using of traditional strategy } \\
\mathrm{Y}_{1} & =\text { Post-test for experimental group } \\
\mathrm{Y}_{2} & =\text { Post-test for control group }
\end{array}
$$

## B. Location and Time of the Research

This research was conducted at the eleventh grade students of Vocational High School Pharmacy Ikasari Pekanbaru. This research was conducted from $27^{\text {th }}$ July to $30^{\text {th }}$ Auguts 2016 .

## C. Subject and Object of the Research

1. The subject of the research

The subject of this research was the eleventh grade students of VocationalHigh School Pharmacy Ikasari Pekanbaru in (2016-2017) academic year.
2. The object of the research

The object of the research was the effect of using crazy story game on students' speaking ability.

## D. The Population and Sample of the Research

## 1. Population

The population of this research was the eleventh grade students at Vocational High School Pharmacy Ikasari Pekanbaru. The number of the students was 263 in 8 classes. They are 5 pharmacy classes, 1 chemical industry class, and 2 health analysis classes.

Table III. 2
The Total Population of the Eleventh Grade of Vocational High School Pharmacy Ikasari Pekanbaru

| No | Classes | Total |
| :---: | :---: | :---: |
| 1 | XI Pharmacy 1 | 35 |
| 2 | XI Pharmacy 2 | 34 |
| 3 | XI Pharmacy 3 | 30 |
| 4 | XI Pharmacy 4 | 34 |
| 5 | XI Pharmacy 5 | 35 |
| 6 | XI Chemical Industry 1 | 30 |
| 7 | XI Health Analysis 1 | 35 |
| 8 | XI Health Analysis 2 | 30 |
|  | Total | 263 |

2. Sample

The population was large enough to be taken all as sample of the research. Based on the total population above, the researcher took two classes for the samples by using Cluster Random Sampling

Technique. Cluster sampling randomly selects groups, not individuals; Sugiono, (2008) states that all the members of selected groups have similar characteristics. Cluster random sampling is most useful when the population is very large or spread out over a wide geographic area.

Based on explanation above, to find the sample, the researcher used lottery by passing out small rolled paper marked by sequence name of the class. Then after passing out the paper,the samples of this research are XI 6 as experimental class and XI 8 as control class. The data can be seen in the table as follows:

Table III. 3
Sample of the Research

| No | Class | Type | Total |
| :---: | :---: | :---: | :---: |
| 1 | XI.6 Chemical Industry 1 | Experimental Class | 30 |
| 2 | XI.8 Health Analysis 2 | Control Class | 30 |
| Total |  |  | $\mathbf{6 0}$ |

## E. Validity and Reliability of the Instrument

## 1. Validity of the Test

Validity refers to appropriatness of a given test or any of its components parts as a measure of what it is purposed to measure. According to Scarvia B. Anderson in Arikunto (2008: 65), a test is said to be valid if it measures what to be measured. It is supported by Porte (2002: 232, 233), there are several type of validity namely; face validity, content validity, external validity, internal validity, and construct validity. Face validity relates to content validity but assesses
informally and/or intuitively whether the instrument appears to measure what it purposed to measure. Content validity considers formally the extent to which a particular instrument measures accurately what it is claimed to measure. A group of expert would normally decide on this, focusing on the instrument's representativeness and comprehensiveness. External validity is of little value unless it has been preceded by adequate address of internal validity concern, which give us confidence in the basic descriptive conclusion drawn from the data themselves. Internal validity is the extent to which the result of the study can be put down to the treatment applied rather than to the design of the study. It also reflects on the degree to which sound conclusion can be drawn about the result of the study. Construct validity describes the extent to which a particular instrument measures accurately construct of interest that have been obtained theoretically.

The researcher concluded that this research belonged to the content validity, because the test reflected to what the student had learned the content of the curriculum. And also Gay (2000: 164) stated that there is no formula used to calculate the content validity and there is no way how to express it quantitatively. So, it means tests of content validity were given based on the material they had learned. The material of the test was taken from the textbook used by eleventh grade students at Vocational High School Pharmacy Ikasari Pekanbaru.

## 2. Reliability of the Test

According to Gay (2000: 169), reliability is the degree to which a test consistently measures whatever it is measuring.A test is said to be reliable if it can produce stable or consistent scores although the test is administered at different time. Reliability is a very important characteristic of a test. A test is not valid unless it is reliable. There are some factors which affect the reliability of a test, namely; the sample performance, the number of items, the administration of the test, the students' motivation and other factors beyond the control of the tester (such as students sickness, etc).

According to Cresswell (2008: 169), there are five types of reliability. They are test-retest reliability, alternate forms reliability, alternate forms and test retest reliability, inter-rater reliability and internal consistency reliability.In this research, to know the reliability of the speaking test the researcher used inter-rater reliability because the researcher had two raters in order to score the students' speaking ability. Porte (2002: 237) says that Inter-rater reliability of an instrument measures the degree of agreement between two or more raters, and indicates the extent to which the raters assess by using the instrument in the same way. Then, the scores of rater 1 were correlated with the scores of rater 2 . The higher correlation, the higher the inter judge reliability will be.

The researcher used Pearson Product Moment formula by using SPSS 17 version to obtain the correlation between scores from rater 1 and rater 2. Then,to know the level of the correlation, the $r_{0}$ is process through Spearman-Brown Prophecy formula.In analyzing the reliability of the test, the researcher used formula as follows:
$\mathrm{r}_{\mathrm{tt}}=\frac{n r_{A, B}}{1+(n-1) r_{A, B}}$

Where,
$\mathrm{r}_{\mathrm{tt}}=$ inter-rater reliability
$\mathrm{n}=$ the number of raters whose combined estimates of the final mark
$\mathrm{r}_{\mathrm{A} \cdot \mathrm{B}}=$ the correlation between raters, or the average correlation among all raters if there are more than two

The researcher used the categories of reliability that can be seen from the following table.

Table III. 4

## Categories of Reliability

| No | Reliability | Level of Reliability |
| :---: | :---: | :---: |
| 1 | $0.0-0.20$ | Low |
| 2 | $0.21-0.40$ | Sufficient |
| 3 | $0.41-0.70$ | High |
| 4 | $0.71-1.0$ | Very high |

(Tinambunan in Putriani (2011: 35))

Table III. 5
Correlations

|  |  | Rater 1 | Rater 2 |
| :--- | :--- | ---: | ---: |
| rater 1 | Pearson Correlation | 1 | $.472^{* *}$ |
|  | Sig. (2-tailed) | 30 | .008 |
|  | N | $.472^{* *}$ | 30 |
| rater 2 | Pearson Correlation | .008 | 1 |
|  | Sig. (2-tailed) | 30 | 30 |
|  | N |  |  |

**. Correlation is significant at the 0.01 level (2-tailed).

From the output above, it can be seen that $\mathrm{r}_{\mathrm{o}}\left(\mathrm{r}_{\text {obtained }}\right)$ is 0.472 will be correlated to $r_{t}\left(r_{\text {table }}\right.$. It is necessary to find the df (degree of freedom).
$\mathrm{df}=\mathrm{N}-\mathrm{nr}$
df : degree of freedom
N : Number of cases
nr : number of correlated variable
$\mathrm{df}=30-2=28$
The researcher took $\mathrm{df}=28$ to be correlated either at level $5 \%$ or $1 \%$. At level $5 \%, \mathrm{r}_{\text {table }}$ is 0.361 ; while at level $1 \% \mathrm{r}_{\text {table }}$ is 0.463 . Thus, the $\mathrm{r}_{\text {obtained }}$ is obtained higherthan $\mathrm{r}_{\text {table }}$, either at level $5 \%$ or $1 \%$. So the researcher concluded that there is a significant correlation between score given by rater 1 and score given by rater 2 . In the other words, the speaking test is reliable.

## F. The Technique of Collecting Data

## 1. Observation

According to Arikunto (2006: 156), observation is an activity that is concerned on some objects by using the five senses. Observation is the way to get some data, by observing the object of the research. In this research, observation was used to collect data on the application crazy story game in teaching process. While the researcher did the treatment, one person (teacher) contributed as observer. He followed the treatment process and check whether the observation list ran or not. The observation list was collected and analyzed in order to consider how far the implementation of the treatment in the classroom.

## Table III. 6 Observation Checklist

| No | Indicators of Crazy Story Game | Yes | No |
| :---: | :--- | :--- | :--- |
| 1 | Teacher divides students in group, Each group <br> consists of 7-8 person, |  |  |
| 2 | Teacher asks every group to sit in their group, |  |  |
| 3 | Teacher explains the story orally |  |  |
| 4 | Teacher asks students to write word or verb on <br> piece of paper and do not tell anyone, |  |  |
| 6 | Teacher starts the story 1-2 sentences and asks the <br> student one to continue the story using his/her <br> word, | After the student one continue the story 1 or 2 <br> sentences he choose next student to continue the <br> story and continues until the last students in their <br> group. |  |
| 7 | Teacher asks the student to guess what his friend <br> word |  |  |
| 8 | The most correct in guessing his friend word is the <br> winner in this game, and then teacher gives some <br> correction and conclusion in the end of the game. | Total |  |

## 2. Oral Presentation Test

Oral presentation test has been used to collect students' data. The studentsexpressed their ideas based on the topic given by asking, and answering the information orally.

In this research, researche gave pre-test and post-test to every groups. The test would be explained as follows:

## 1. Procedures of collecting data for experimental class:

a. Pre-test

Pre-tes was given to measure students' speaking ability before they were taugh by using crazy story game technique. This test is given to both experimental and control group.
b. Treatment

The treatment was conducted for the experimental class. This research used Crazy Story Game Technique for six meetings.
c. Post-test

After conducting the treatment, the post-test was administered and it was analyzed as final data for this research. The test was given the same test as given in the pre-test.

## 2. Procedures of collecting data for control class:

a. Pre-test

The control class was given pre-test to know their speaking ability. The test was the same as experimental class.
b. No treatment
c. Post-test

Post-test was given to control class and the result was analyzed and used as final data for this research.

Then, the researcher took the total score from the result of the speaking ability test. The classification of the students' score was shown below:

## Table III. 7 <br> The Classification of Students Scores

| Score | Categories |
| :---: | :---: |
| $80-100$ | Very Good |
| $66-79$ | Good |
| $56-65$ | Enough |
| $40-55$ | Less |
| $30-39$ | Fail |

## G. Technique of Data Analysis

In analyzing the students' speaking ability, the researcher used minimum passing grade of English lesson in Vocational High School Pharmacy Ikasari Pekanbaru. It was 72 for the students' speaking ability.

It means that for those who get score $>72$, they pass the passing grade; while those get score $<72$ they don't pass the passing grade.

## 1. Normality of the Data

Before analyzing the data by using T -test formula, the researcher had to find out the normality test of the data. According to Priyatno, the normality of the data test can be analyzed by using lilliefors and One Sample Kolmogorov Smirnov. In this research, the researcherused lilliefors through SPSS. 17 versions.

In analyzing the normality of the data, the researcher used students' post-test score of experimental class and control class.

Analysis:
If the probability $>0.05 \mathrm{H}_{0}$ is accepted
If the probability $<0.05 \mathrm{H}_{0}$ is rejected

## 2. Homogenity of the Data

According to Siregar (2013:167), the purpose of homogeneity test is to know whether the object of the research has the same variance or not. The method used in this test was comparing the biggest variance with the smallest one. In analyzing the homogeneity of the data, the writer used students' pre-test scores of experimental class and control class.

## 3. Analysis of the Data

In analyzing the data, the researcher used students' post-test score in experimental and control classes. This score was analyzed statistically. In this research the researcherr used these formulas:

## a. Independent Sample T-test

Hartono (2011: 207) said that to find out whether or not there is a significant difference between two or more variables independent sample $t$-test was used. In this research, the data were analyzed by using SPSS 17.0 Version. The T-table was employed to see whether or not there is a significant difference among the mean scores both of experimental and control classes. Statistical hypothesis:

1. $\mathrm{H}_{0}=\mathrm{t}_{0}<\mathrm{t}$-table
2. $\mathrm{H}_{\mathrm{a}}=\mathrm{t}_{0}>\mathrm{t}$-table

## b. Effect Size

According to Pallant (2005:199),Effectsize was the strength of the difference between groups or the influence of independent variable Pallant. There are a number of different effect size statistics, the most commonly used is eta squared. Eta squared can range from 0 to 1 and represents the proportion of variance in the dependent variable that is explained by the independent (group) variable. The formula for eta squared is as follows:

Eta squared $^{2}=\frac{\mathrm{t}^{2}}{\mathrm{t} 2+(\mathrm{N} 1+\mathrm{N} 2-2)}$
Where:

$$
\begin{aligned}
& \mathrm{t}^{2}=\mathrm{t}_{\mathrm{o}} \\
& \mathrm{~N}_{1}=\text { Number of students }
\end{aligned}
$$

## 4. Assessment

In this research, the speaking was evaluated by concerning five components that should be considered in giving students' speaking ability score. According to Hughes (2003: 131), there are some components that should be considered in giving students' speaking ability score in assessing speaking; there are five components that should be considered in giving students' speaking ability score. Every component has five levels and criterias. It can be seen in the following table below:

Table III. 8
Speaking Assessment
a. Vocabulary

| Level | Criteria |
| :---: | :--- |
| 1 | Speaking vocabulary inadequate to express anything but the most <br> elementary needs. |
| 2 | Has speaking vocabulary sufficient to express himself simply with <br> somecircumlocutions. |
| 3 | Ableto speak the language with sufficient vocabulary to participate <br> effectively in most formal and informal conversationson practical, <br> social, and professional topics.Vocabulary is broad enough that he <br> rarely has to grope for a word. |
| 4 | Can understand and participate in anyconversation within the range of <br> hisexperiencewith a high degree of precision of vocabulary. |
| 5 | Speech on all levels is fully accepted by educated native speakers in <br> all its features including breadth of vocabulary and idioms, <br> colloquialisms, and pertinent cultural references. |

## b. Grammar

| Level | Criteria |
| :---: | :--- |
| 1 | Errors in grammar are frequent, but speaker can be understood by a <br> nativespeaker used to dealing with foreigners attempting to speak his <br> language. |
| 2 | Can usually handle elementary constructions quite accurately but does <br> not have thorough or confident control of the grammar. |
| 3 | Control of grammar is good. Able to speak the language with sufficient <br> structuralaccuracy to participate effectively in mostformal and informal <br> conversations on practical,social, and professional topics. |
| 4 | Ableto use the language accuratelyon all levels normally pertinent to <br> professional needs. Errors in grammar are quite rare. |
| 5 | Equivalent to that of an educated native speaker. |

## c. Fluency

| Level | Criteria |
| :---: | :--- |
| 1 | No specific fluency description. Refer to other four language are asfor <br> implied level of fluency. |
| 2 | Can handle with confidence but not with facility most social situations, <br> incliding introductions and casual conversations about current events, <br> as well as work, family, and autobiographical information. |
| 3 | Can discuss particular interests of competence with reasonable ease. <br> Rarely has grope for words. |
| 4 | Able to use the language fluentlyon all levels normally pertinent to <br> professional needs. Can participate in any conversation within the range <br> of this experience with a high degree of fluency. |
| 5 | Has complete fluency in the language such that his speech is fully <br> accepted by educated native speakers. |

## d. Pronunciation

| Level | Criteria |
| :---: | :--- |
| 1 | Errors in pronunciation are frequent butcan be understood by a native <br> speakerused to dealing with foreigners attempting to speak his <br> language. |
| 2 | Accent is intelligible though often quite faulty. |
| 3 | Errors never interfere with understanding and rarely disturb the native <br> speaker. Accent may be obviously foreign. |
| 4 | Errors in pronunciation are quite rare. |
| 5 | Equivalent to and fully accepted native speakers. |

e. Comprehension

| Level | Criteria |
| :---: | :--- |
| 1 | Within the scope of his very limited language experience,can <br> understand simple questions and statements if delivered with slowed <br> speech. Repetition,or paraphrase. |
| 2 | Can get the gist of most conversations of non-technical subjects. |
| 3 | Comprehension is quite complete at a normal rate of speech. |
| 4 | Can understand any conversation within the range of his experience. |
| 5 | Equivalent to that of an educated native speaker. |

The speaking result was evaluated by concerning five components of speaking assessment above. The score of each level in every component is 4 . The highest score for all calculated components is 100 . So, the highest score for each components is 20 . It can be seen in the following table below:

Table III. 9
The Specification of the Test

| No | Components of Speaking | The Highest Score |
| :---: | :---: | :---: |
| 1 | Accent | 20 |
| 2 | Grammatical | 20 |
| 3 | Vocabulary | 20 |
| 4 | Fluency | 20 |
| 5 | Comprehension | 20 |
|  | Total | 100 |

Next, to know whether the students' score of speaking ability is higher or lower, the researcher used the score 72 based on the minimum of Passing Grade score (PG) at Vocational High School Pharmacy Ikasari Pekanbaru. Thus, the students whogot score $\geq 72$ passed the minimum of Passing Grade score (PG), while the students got $<72$, they did not pass the minimum of Passing Grade score (PG).The Classification of students' score can be seen in the following table below:

Table III. 10
Classification of Students' Score

| No | The Score Level | Category |
| :---: | :---: | :---: |
| 1 | $80-100$ | Excellent |
| 2 | $66-79$ | Good |
| 3 | $56-65$ | Average |
| 4 | $40-55$ | Poor |
| 5 | $30-39$ | Fail |

(Arikunto, 2008: 245)
From the table aboveit can be seen that the high score $80-100$ is excellent, $66-79$ is good, $56-65$ is average, $40-55$ is poor and $30-39$ is fail.

