

## CHAPTER III

### RESEARCH METHODOLOGY

#### III.1. Design of the Research

In this study, the data were collected through their reading comprehension after giving a treatment. This study used quantitative method with quasi experimental design. According to L. R Gay (2000: 364), the quasi-experimental design involves selecting two groups or more differing on some independent variables and comparing them to some dependent variables. The experiment was treated to two groups. One group was taught by using DRTA and one group was taught by using QAR strategy. Experimental design was used in this research aimed to compare the effect of using DRTA and using QAR on students' reading comprehension. It consisted of two independent variables and one dependent variable. They can be drawn in the following table:

**Table 3.1 Research Design**

E1	01	X1	02
E2	03	X2	04

Figure (Gay, 2000: 353)

E1 : Experimental Group1

E2 : Experimental Group2

X1: Independent variable 1 (DRTA)

X2 : Independent variable 2 (QAR)

01,03 : pre-test

02,04 : post-test

Gay (2000: 354) states that the definition and selection of comparison group very important part of the quasi-experimental design procedure. The independent variable differentiating the groups must be clearly and operationally defined, since each group represents a different population.

### **III.2. Location and Time of the Research**

The location of this research was at Language Development Center of UIN SUSKA RIAU. It is located at Jl. KH. Ahmad Dahlan no. Sukajadi. The duration of time to conduct of this research was within two months starting from January up to February 2017.

### **III.3. Subject and Object of the Research**

The subject of the research was the 2<sup>nd</sup> level students' at Language Development Center of UIN SUSKA RIAU, and the object of this study was a comparison between the effect of using DRTA and QAR strategies toward students' reading comprehension.

### **III.4. Population and Sample**

#### **a. Population**

The population of this research was the students of the first year at Language Development Center of UIN SUSKA RIAU. The total number of the

population was 120 students. The target population was the students of accounting of the Second Level of Language Development Center of State Islamic University of Sultan Syarif Kasim Riau in Academic Year 2015/2016. Based on the population of this research, the sample was selected by using cluster sampling. According to (Gay and Airasian, 2000), cluster sampling randomly selects groups not individuals. Based on teachers' information, all the members of selected groups have similar characteristics, and three classes are chosen by using cluster sampling in this research. The population can be shown from the table below:

**Table 3.2**  
**Population**

No	Class	Male	Female	Number of Students
1	PB 19	12	18	30
2	PB 20	10	20	30
3	PB 21	17	13	30
4	PB 22	13	27	30
<b>Total</b>	<b>52</b>	<b>78</b>	<b>120</b>	

**b. Sample**

The homogenous characteristics were the consideration. Before giving pre test, the teacher gave information about the students' ability. The teacher stated that PB 19 and PB 20 had same capability. All the members of selected groups had similar characteristics. Because all classes were homogenous, the sample was

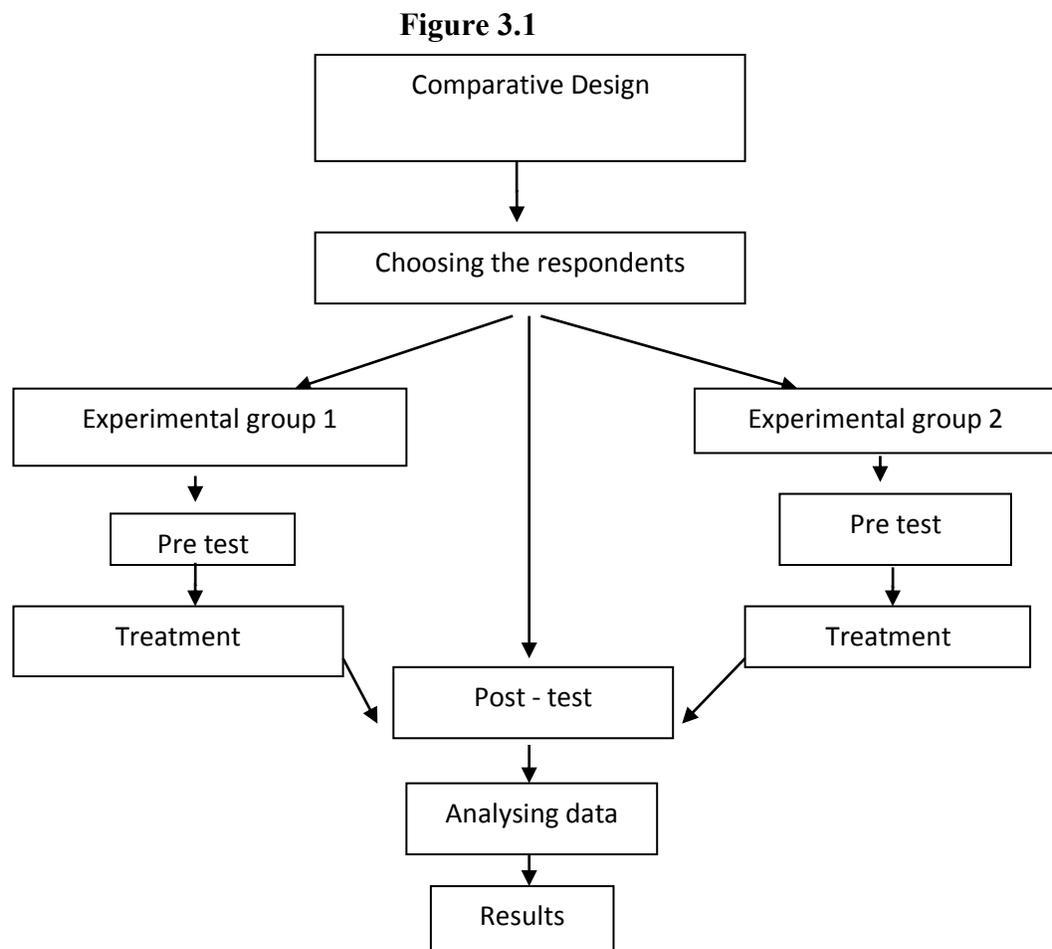
chosen randomly, PB 19 as an Experimental class 1, PB 20 as an experimental class

2. Two classes were taken as the sample of this research as follows:

<b>Class</b>	<b>Male</b>	<b>Female</b>	<b>Total of Students</b>
PB 19	12	18	30
PB 20	10	20	30
<b>Total Participants</b>	<b>22</b>	<b>38</b>	<b>60</b>

Since it was a comparative study with a quasi-experimental research design that had a certain purpose, the researcher used random cluster sample technique. It consisted of two groups taught by using different treatments; two classes was taken as experimental classes; PB 1 for an experimental class 1 and PB 2 for experimental class 2.

### III.5. Research Procedure



### III.6. Research Instruments

To collect the data, reading test was administered as the instrument of this study. The test was applied for pre-test and post-test. It was administered to two classes which consisted of PB 19 and PB 20. The pre-test aimed at finding out the prior reading of the students. While post-test aimed at finding out the students reading comprehension after treatment was given a teaching with DRTA Strategy and QAR Strategy. This activity was also intended to find out whether the students' skill kept holding of the material after the treatment.

After the students did the test, the writer counted the scores by using scoring guidance formula:

$$\text{Total score} = \frac{\text{correct answer}}{\text{Total question}} \times 100$$

By using the formula above, the writer was able to determine that the score of students' reading comprehension included in a certain classification of the score. The classification of the students' scores shown on Suharsimi (2007:245) below:

<b>Score</b>	<b>Categories</b>
80-100	Very Good
66-79	Good
56-65	Enough
40-55	Less
0-39	Fail

In try out test, the total amount of the questions was 25 questions. They were divided into five indicators; *Identifying the main idea*, *Identifying detail information* , *Identifying meaning of vocabulary*, *Identifying Inference*, *Identifying Reference* , *Identifying the moral value*. The total items were 25 questions. All of the questions were distributed to the students to be answered.

### **III.7. Data Collection Technique**

In this research, the data were collected by administrating pre-test and post-test to the students. Each reading text was considered the time and the procedures of DRTA Strategy and QAR Strategy and conventional reading text. The tests was taken from the students' textbook and internet materials.

In order to get the data to support this study, the researcher used the technique as follows:

a. Observation check list

Observation was used to observe directly the students using DRTA and QAR strategies in reading comprehension and to observe the influence of DRTA and QAR strategies toward the students' ability in reading comprehension. In observation technique the researcher had a list of observational items to be observed in the class during teaching and learning process.

b. Test

To find out the comparison of using DRTA and QAR on the students' reading comprehension at the second level of Language Development Center Pekanbaru,

the researcher administered the test to assess students' English ability. The test was administered in two stages. It consisted 25 questions in multiple choice form. The first was a pretest which was done before the treatment. The second was a post-test which was done after the treatment. The researcher measured the total score from the results of the students' English reading comprehension test.

### **III.8. Validity and Reliability Test**

Trying out of the test was administered to determine the quality of the test that was used to collect the data in order to know the test whether the instrument was valid or not. The study of the instrument was measured the validity and reliability. After giving a try out test, this research had found whether the question was significant or not. Furthermore the significant questions had been tested again to find the comparison between DRTA and QAR on students' reading comprehension.

#### **a. The validity of instrument**

In general, validity refers to the appropriateness of the test given or any of its component parts as a measure of what it is purposed to measure. It means that the test is valid to extend the instrument. That is measured what it is supposed to measure. In addition, every tests in this reasearch of the reading text whether it is a short, informal classroom test or public examination should be as valid as test constructor that can make it. The instrument of the test must aim at providing a true measure and useful. Validity is the extent to which inferences were made from assessment results that are appropriate, meaningful, and useful in terms of the purpose of the assessment (Brown, 2003: 22). Validity is important because it can help to determine what

types of tests to use, and help to make sure researchers are using methods that are not only ethical, and cost-effective, but also a method that truly measures the idea or construct in question. Furthermore, Brown says that there are three kinds of validity (2001: 388). They are content validity, face validity, and construct validity. All of them have different usage and function.

While according to Gay (2000: 161), Validity is the appropriateness of the interpretations made from test score. There are also three kinds of validity; content validity, criterion-related validity, and construct validity.

Before collecting the data, each item of the questions had been tested in order to be ideal to try out. The purpose of the try out was to find out the quality of the test items. Brown (2000;22) states that a test is a method of a measuring a person's ability, knowledge, or performance in a given domain. Validity is the extent to which inferences make from assessment results which are appropriate, meaningful, and useful in terms of the purpose of the assessment.

To judge the validity of each item could be seen from the column corrected item-total correlation ( $r_{\text{count}}$ ) compared with  $r_{\text{table}}$ . If  $r_{\text{count}} > r_{\text{table}}$  that item was valid or otherwise using  $\alpha=0.05$ . Based on trying out instrument of students' reading comprehension which was conducted to 35 participants with 30 multiple choice questions related to narrative text, it was obtained that 30 items were valid and 5 items were invalid. The invalid items were removed and students' reading comprehension instrument using 25 valid questions. It could be seen in the followingtable:

<b>Item</b>	<b>r-count</b>	<b>r-table</b>	<b>Status</b>
1	0.587	0.334	Valid
2	0.540	0.334	Valid
3	0.534	0.334	Valid
4	0.616	0.334	Valid
5	0.485	0.334	Valid
6	0.132	0.334	Drop
7	0.421	0.334	Valid
8	0.221	0.334	Drop
9	0.555	0.334	Valid
10	0.444	0.334	Valid
11	0.478	0.334	Valid
12	0.439	0.334	Valid
13	0.452	0.334	Valid
14	0.579	0.334	Valid
15	0.566	0.334	Valid
16	0.591	0.334	Valid
17	0.510	0.334	Valid
18	0.248	0.334	Drop
19	0.574	0.334	Valid
20	0.439	0.334	Valid
21	0.579	0.334	Valid
22	0.444	0.334	Valid
23	0.471	0.334	Valid
24	0.431	0.334	Valid
25	0.573	0.334	Valid
26	0.532	0.334	Valid
27	0.537	0.334	Valid
28	0.470	0.334	Valid
29	0.239	0.334	Drop
30	0.209	0.334	Drop

The points of difficulty level and discrimination index was analyzed by using a formula (Heaton, 1975 : 178).

$$FV = \frac{R}{N} \times 100\%$$

Where :

FV: The index of difficulty

R: The number of correct answers

N: The number of respondents

### **b. The Reliability of Instrument**

A reliability is an important characteristic of a good test. In order to calculate the reliability of the test, the mean of the students' scores and the standard deviation are sought. To find out the reliability of the test, the following formula was used; the discrimination index of an item indicated the extent to which the item discriminated between the students, separating the more able students from the less able. The following formula was taken from Heaton (1975: 164) as follows :

$$r_{ii} = \frac{N}{N-1} \left( 1 - \frac{m(N-m)}{N(X)^2} \right)$$

$$\text{Where : } M = \frac{\sum x}{N} \text{ and } S^2 = \frac{\sum x^2 - \frac{(\sum x)^2}{N}}{N}$$

$r_{ii}$  : Reliability of the test

N : The number of item in the test

M : The mean score of all the test

$S^2$  : The standard deviation of all the test score

**Table 3.5 Criteria Coefficient of Reliability**

<b>Coefficient Reliability</b>	<b>Criteria</b>
$0,80 \leq r_{11} \leq 1,00$	Highest reliability
$0,60 \leq r_{11} \leq 0,79$	High reliability
$0,40 \leq r_{11} \leq 0,59$	Middle reliability
$0,20 \leq r_{11} \leq 0,39$	Low reliability
$0,00 \leq r_{11} \leq 0,19$	Lowest reliability

According to Sudjono (2001), the criteria of instrument reliability ( $r_{11}$ ) is stated as follows:

- If  $r_{11}$  was equal or higher than 0.7, it means that the instrument was reliable.
- If  $r_{11}$  was lower than 0.7, it means that the instrument was unreliable.

Based on trying out reading comprehension instrument which was done to 35 participants, there were 25 items of reading comprehension test were valid and reliable with reliability  $0.919 > 0.7$ . It could be concluded that the test of the instrument in this study was reliable. Based on the coefficient of the reliability table, it was categorized into highest reliability.

### **III.9. Data Analysis Technique**

The scoring guide was chosen as the criteria of scoring representing the basic aspects of reading. There are 6 aspects of reading comprehension, they are; main idea, the detail information, reference, inference and vocabulary. The reading comprehension results were evaluated by considering six aspects and each aspect had a score or a level. In analyzing the data, it used score of post-test of experimental

and control groups. This score was analyzed statistically for both descriptive and inferential statistic.

**a. Independent sample t-test**

To find out whether there is a significant difference or there is no significant difference between two or more variables can be analyzed by using an Independent Sample Ttest. Gay (2001) adds that the t-test for independent sample is used to determine whether there is probably a significant difference. Independent sample t-test was used to find out the results of the hypotheses. They are as follows:

- a. To find out whether there was any significant difference on students reading comprehension before giving the treatment by using DRTA Strategy and QAR strategy for the experimental class and non treatment for the control class.
- b. To find out whether there was significant difference on students' reading comprehension after giving the treatment by using DRTA Strategy and QAR strategy for the experimental class.

To analyze the final-test scores of the experimental group, the following formula is used:

$$t = \frac{M_X - M_Y}{\sqrt{\frac{(SD_X)^2}{N_1 - 1} + \frac{(SD_Y)^2}{N_2 - 1}}}$$

Where:

t = The value of comparing two means

$M_X$  = Mean of the score in pre-test

$\frac{1}{M_Y}$  = Mean of the score in post-test

$SD_X$  = Standard deviation of experimental group

$SD_Y$  = Standard deviation of control group

$N_1$  = Number of the sample in pre-test

$N_2$  = Number of the sample in post-test

1 = the constant number

The t-table has the function to see if there is a significant difference between the mean of the score of both experimental and control groups. The t-obtained value is consulted with the value of the t-table at the degree of freedom (df) =  $(N_1+N_2)-2$  which is hypothesized

$H_a$ :  $t_o > t\text{-table}$

$H_o$ :  $t_o < t\text{-table}$

$H_a$  is accepted if  $t_o > t\text{-table}$  or there is effect after giving the treatment of DRTA Strategy and QAR Strategy on students' reading comprehension.

$H_o$  is accept if  $t_o < t\text{-table}$  or there is no effect after giving the treatment DRTA Strategy and QAR Strategy on reading comprehension.

Afterward, it is better to find the effect size of T-test by following formula:

$$\tilde{\eta}^2 = \frac{t^2}{t^2+n-1}$$

$$\text{eta squared} = \tilde{\eta}^2 \times 100\%$$

Where:

*eta squared* : Coefficient effect

$\tilde{\eta}^2$  : Coefficient

### b. Paired Sample T-Test

Non-independent sample  $t$ -  $t_{\text{test}}$  is known also as Paired-Sample  $t_{\text{test}}$ . The researcher used this formula to obtain the result of the hypotheses that was to find out whether there was significant effect of using DRTA and QAR strategies on students' reading comprehension at the second level students. L.R Gay states that  $t$ -test for non-independent sample is used to compare groups that are formed by some types of matching or to compare a single group's performance on a pre-test and post-test or on two different treatments. (L.R Gay, 2000: 488).

Pre-test and post-test scores were used in the experimental class in order to find the significant effect of using DRTA and QAR on students' reading comprehension of the second level students at Language Development Center Pekanbaru. To obtain the data, SPSS 20 was used.

The formula of paired-sample  $t_{\text{test}}$ :

$$t = \frac{\bar{D}}{\sqrt{\frac{\sum D^2 - \frac{(\sum D)^2}{N}}{N(N-1)}}$$

$D$  : Gain Score ( $D=X_2-X_1$ )

The  $t$ -table has the function to see if there is a significant improvement among the mean of the score of both pretest and posttest. The  $t$ -obtained value is consulted with the value of  $t$ -table at the degree of freedom ( $df$ ) =  $N-1$  which is statistically hypothesis:

Ha:  $t_o > t\text{-table}$

Ho:  $t_o < t\text{-table}$

Ha is accepted if  $t_o > t\text{-table}$  or there is a significant effect after giving the treatment DRTA and QAR toward students' reading comprehension at the second level students at Language Development Center Pekanbaru

Ho is accepted if  $t_o < t\text{-table}$  or there is no significant effect after giving treatment DRTA and QAR toward students' reading comprehension at the second level students at Language Development Center Pekanbaru.

Afterward, it is better to find the coefficient effect of T-test by following formula:

$$\tilde{\eta}^2 = \frac{t^2}{t^2 + n - 1}$$

$$kp = \tilde{\eta}^2 \times 100\%$$

Where:

kp : Coefficient effect

$\tilde{\eta}^2$  : Coefficient

Notation  $\tilde{\eta}^2 = \text{Eta square}$

The effect size can assist between 0 to 1, according to Cohen (Cohen, Manion, and Morrison :2007 p.521) the category of effect size is as follow :

0-0.20 = weak effect

0.21 – 0.50 = Modest effect

0.51- 1.00 = moderate effect

> 1.00 = Strong effect