

CHAPTER III

RESEARCH METHODOLOGY

A. Research Design

This research uses a quantitative approach with a quasi-experimental design to examine the effect of using Webtoon as a narrative text media on ninth-grade students' reading comprehension at SMP Negeri 1 Danau Kembar. The study involves two groups: an experimental group that receives treatment using Webtoon and a control group that is taught using conventional media.

Quasi-experimental design is chosen because the researcher cannot fully control all external variables that may influence the implementation of the experiment, especially in an educational setting where random assignment is limited. Therefore, the existing classes are used as groups without randomization.

According to Creswell (2014), research design is a structured plan that guides the procedures for collecting, analyzing, and interpreting data in a study. In quantitative research, the design focuses on numerical measurement and statistical analysis to test hypotheses and examine relationships between variables.

Furthermore, quasi-experimental research involves the use of comparison groups but lacks full random assignment, which makes it suitable for real classroom situations. As stated by Campbell and Stanley (1963), quasi-experimental designs are applied when researchers cannot fully control

variables but still aim to investigate causal relationships through treatment and comparison groups. Similarly, Cohen, Manion, and Morrison (2018) explain that quasi-experimental designs are commonly used in educational research where intact groups are utilized, and full experimental control is not possible.

Based on these perspectives, this study is categorized as quasi-experimental research because it aims to identify the causal effect of Webtoon media on students' reading comprehension by comparing the outcomes between the experimental and control groups.

B. Setting of The Research

The research was conducted at SMP Negeri 1 Danau Kembar, where the researcher conducted the study directly at the school. The school is located in Jorong Gurun Data, Nagari Simpang Tanjung Nan IV, Danau Kembar District, Solok Regency, West Sumatra. The research was conducted during the first semester of the 2026/2027 academic year.

C. Data and Source of Data

1. Population

The population of this research consisted of all ninth-grade students at SMP Negeri 1 Danau Kembar. This population was selected because it is directly related to the focus of the study, which examined students' reading comprehension in learning narrative texts using Webtoon media.

In research, a population refers to a group of individuals or objects that share certain characteristics and become the main focus of a study. According to Creswell (2014), a population is a group of individuals or elements that the researcher intends to study and generalize the results. Similarly, Sekaran and Bougie (2016) define a population as the entire group of people, events, or things of interest that the researcher wishes to investigate.

Based on these definitions, the population in this study include all ninth grade at SMP Negeri 1 Danau Kembar. In this study, the researcher selected ninth-grade students as the population, consisiting of 62 students at SMP Negeri 1 Danau Kembar. They represented the target group for investigating the effect of Webtoon media on reading comprehension.

2. Sample

The sample of this research was taken from two classes of ninth-grade students at SMP Negeri 1 Danau Kembar, namely class IX A and IX B. One class was assigned as the experimental group, while the other will served as the control group.

A sample is a subset of a population that is selected to represent the characteristics of the entire population. Creswell (2014) states that a sample is a subgroup of the target population that the researcher plans to study for generalizing about the population. In line with this, Sekaran and Bougie (2016) explain that a sample is a portion of the population chosen to provide information about the population.

In this study, the sampling technique used was purposive sampling, where the researcher selected participants based on specific considerations related to the research objectives. According to Etikan, Musa, and Alkassim (2016), purposive sampling is a non-probability sampling technique in which participants are selected based on particular characteristics and the purpose of the study.

The use of purposive sampling in this research was based on the need to select two intact classes as the experimental class and the control class. The selection of the classes also considered the English teacher's suggestions and the characteristics of the classes. Therefore, class IX A and IX B were chosen as the sample of 41 students because they were suitable for this research and quasi-experimental design.

D. Research Variables

A research variable refers to a characteristic or factor that is studied by a researcher to obtain information and draw conclusions. In this study, there were two variables: an independent variable and a dependent variable.

1. Independent Variable (X)

The independent variable is a variable that influences or affects the dependent variable. In this study, the independent variable (X) was the use of Webtoon as a learning medium in teaching narrative texts. Webtoon was used as a digital and visual media to support students in understanding narrative texts.

2. Dependent Variable (Y)

The dependent variable is a variable that is affected by the independent variable. In this study, the dependent variable (Y) was students' reading comprehension of narrative texts. Reading comprehension refers to students' ability to understand and obtain meaning from the texts they read. Grabe and Stoller (2020) explain reading comprehension as an important part of reading ability in second-language learning.

Therefore, the independent variable in this study was the use of Webtoon as a learning media, while the dependent variable was students' reading comprehension of narrative texts.

E. Instrumentation

Instrumentation refers to the tools used by a researcher to collect and measure data in a study. In this research, the instrument was used to measure students' reading comprehension test in the form of multiple-choice questions. The test was used to measure students' reading comprehension before and after the treatment.

To explain how the data were collected and how the instrument was prepared and analyzed, this section is divided into two parts: data collection technique and data collection instrument.

1. Data Collection Technique

The data were collected through a pre-test and a post-test. The pretest was administered before the students received the treatment. It was intended to measure the students' initial reading comprehension of

narrative texts before Webtoon was used as a learning media. The posttest was administered after the students received the treatment. It was intended to measure the students' reading comprehension of narrative texts after learning through Webtoon.

The tests were used to obtain data on students' reading comprehension of narrative texts at SMP Negeri 1 Danau Kembar.

2. Data collection Instrument

The test consisted of multiple-choice questions with four answers options: a, b, c, and d. The questions measured five aspects of reading comprehension. The test initially consisted of 25 questions for the try-out. After the item analysis, 20 questions were selected for the final test.

Brown and Abeywickrama (2019) explain that language assessment involves designing appropriate tests and test items based on the abilities that need to be assessed. Therefore, the test items in this study were designed based on reading comprehension aspects being measured.

a. Item Difficulty Index

The item difficulty analysis was conducted to determine the difficulty level of each item in the reading comprehension test. The item difficulty index refers to the proportion of test-takers who answer a particular test item correctly. The item difficulty index is represented by P and can be calculated using the following formula:

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

Where:

P = item difficulty index

R = number of students who answered the item correctly

N = total number of students who answered the item

The criteria for the item difficulty index can be seen in Table 3.1

Table 3.1
Classification of the Difficulty Level of Test Items

Difficulty Level	Interpretation
P = 0,00-0,29	Difficult
P = 0,30-0,69	Moderate
P = 0,70-1,00	Easy

Source: Arikunto, 2012: 223

Based on the calculation of the item difficulty index for 25 reading comprehension test items, it can be conclude as prensented in Table 3.2.

Table 3.2
Results of the Item Difficuly Index Analysis of the Test Items

Item Category	Item Number	Total
Difficult	13, 25	2
Moderate	3,6,8,10,11,12,14,16,17,19,22,23	12
Easy	1,2,4,5,7,9,15,18,20,21,24	11
Total		25

Based on the analysis of the 25 test items that were tried out, the item difficulty index showed that 2 items were classified as

difficult, 12 item as moderate, and 11 items as easy. In this study, the difficulty level of test items was determined using Microsoft Excel.

b. Item Discrimination

Item discrimination analysis was conducted to determine the extent to which each test item could distinguish between students with higher and lower levels. Item discrimination can be determined by comparing the item facility of the upper-scoring group with that of the lower-scoring group. With the formula:

$$ID = \frac{R_{upper}}{N_{upper}} - \frac{R_{lower}}{N_{lower}}$$

Where:

ID = item discrimination index

R_{upper} = number of students in the upper group who answered the item correctly

N_{upper} = total number of students in the upper group

R_{lower} = number of students in the lower group who answered the item correctly

N_{lower} = total number of students in the lower group

Table 3.3
Classification of the Item Discrimination Index

Discrimination Index	Classification
0,00-0,19	Poor
0,20-0,39	Good
0,40-0,69	Fair
0.70-1,00	Very Good

Sourcer: Arikunto, 2012

The results of the item discrimination analysis for the trial test can be seen in the table 3.4.

Table 3.4
Results of the Item Discrimination Analysis

Item Category	Item Number	Total
Poor	16,25	2
Fair	1,2,3,4,7,13	6
Good	5,8,9,10,11,14,15,17,18,20,21,22,23,24	14
Very Good	6,12,19	3
Total		25

Based on the item discrimination results above, it can be concluded that, out of the 25 test items, 2 items were classified as poor, 6 items as fair, 14 items as good, and 3 items as very good. In this study, the item discrimination index was determined with the assistance of Microsoft Excel.

F. Validity and Reliability of the Instrument

1. Validity Test

An instrument is considered valid if it is able to measure what it is intended to measure. In other words, a valid instrument can accurately obtain data related to the variable being studied. Validity is an important aspect of language assessment because it concerns the extent to which the interpretation and use of test scores are supported by appropriate evidence. Brown and Abeywickrama (2019) explain that validity is one of the essential principles in evaluating a language assessment.

In this study, the validity test was conducted to determine whether each test item was appropriate for measuring students' reading comprehension of narrative texts. The instrument consisted of 20 multiple-choice questions. Each item was analyzed to determine whether it had a significant relationship with the total test score.

For the multiple-choice test used in this study, each item was scored using a two point scoring system, with 1 for a correct answer and 0 for an incorrect answer. Therefore, the validity of each test item was analyzed using the point-biserial correlation coefficient ($r_{p\beta i}$). The point-biserial correlation was used to determine the strength of the relationship between the score obtained on an individual test item and the students' total test score.

The steps for analyzing the validity of the test items were as follows:

- a. Preparing a calculation table to analyze the validity of test items from item number 1 to item number 20.
- b. Calculating the mean of the total scores (M_t).

$$M_t = \frac{\sum X_t}{N}$$

- c. Calculating the standard deviation of the total scores (SD_t).

$$SD_t = \sqrt{\frac{\sum X^2}{N} - \left(\frac{\sum X_t^2}{N}\right)}$$

- d. Calculating the mean score (M_p) of the students who answered each test item correctly.

$$Mp = \frac{\text{total score of testees who answered the test item correctly}}{\text{number of testees who answered the test item correctly}}$$

- e. Calculating the point-biserial correlation coefficient (r_{pbi}) for each test item using the following formula:

$$r_{pbi} = \frac{M_p - M_t}{SD_t} \sqrt{\frac{p}{q}}$$

Where :

r_{pbi} = The point-biserial correlation coefficient, representing the strength of correlation between variable 1 and variable 2, which in this case is considered the item validity coefficient.

M = The mean score obtained by the students for the test item that was answered correctly.

M_t = The mean of the total scores.

SD_t = The standard deviation of the total scores.

P = The proportion of students who answered the item correctly.

q = The proportion of students who answered the item incorrectly.

In the instrument try-out conducted with 15 students, the critical value of the “r” product-moment correlation was determined from the correlation table. The results were as follows:

- a. At 5% significance level (r_t) = 0,514
- b. At 1% significance level (r_t) = 0,641

If the obtained (r_{pbi}) correlation coefficient is greater ($>$) than the table value (r) = 0,514 at the 5% significance level, then the result obtained is significant, meaning that the test item is declared valid. If the value (r_{pbi}) of the correlation coefficient obtained is smaller ($<$) than the table value (r) = 0,514 at the 5% significance level, then the result obtained is non-significant, meaning that the test item is declared invalid.

From the results of the validity analysis of the 25 test items, it was found that 21 test items were categorized as valid and 4 test items were categorized as invalid. For more clarity, it can be seen in the table below:

Table 3.5
Test Instrument Validity

Item Category	Item Number	Total
Valid	1,2,3,4,5,6,7,8,9,10,11,12,14,15,17,19,20,21,22,23,24	21
Invalid	13,16,18,25	4

2. Reliability Test

Reliability refers to the consistency of a test in measuring the ability or construct for which it is designed. A test is considered reliable when it produces consistent results under similar conditions. According to Brown and Abeywickrama (2019), reliability is one of the important principles of language assessment and refers to the consistency of test scores.

The criteria used to determine the reliability of the test are presented in the following table:

Table 3.6
Reliability Estimation Criteria

Reliability	category
>0,80	High Reliability
0,40-0,80	Moderate Reliability
<0,40	Low Reliability

For the multiple-choice test, internal consistency reliability was calculated using the Kuder-Richardson 20 (KR-20) technique. KR-20 is appropriate for tests in which items are scored dichotomously, such as 1 for a correct answer and 0 for an incorrect answer. The formula used was:

$$r = \left(\frac{n}{n-1} \right) \left(\frac{S^2 - \sum pq}{S^2} \right)$$

Where:

- r : reliability coefficient of the entire test
- p : proportion of students who answered the item correctly
- q : proportion of students who answered the item incorrectly
- $\sum pq$: sum of the products of p and q
- n : number of test items
- S : standard deviation of the test scores

The reliability test was conducted during the instrument try-out. The calculation resulted in a reliability coefficient of 0.90561, indicating that the test had very high reliability.

G. Data Analysis Technique

One of the research activities is data analysis, which is very important and becomes a determinant of the research. According Patton, data analysis is the process of organizing data, arranging it into a pattern, category, and basic description of the analysis.

1. Normality Test

The normality test aims to determine whether the data from the two sample classes are normally distributed or not. In this study, the normality test was conducted with the assistance of SPSS version 23.

Table 3.7
Results of the Normality Test Analysis for the Experimental and Control Classes
Tests of Normality

	Class	Kolmogorov-Smirnov ^a			Shapiro-Wilk		
		Statistic	df	Sig.	Statistic	df	Sig.
Reading Comprehension	Posttest Experimental class	,166	21	,134	,955	21	,421
	Posttest Control class	,162	20	,175	,952	20	,400

Based on the data above, it was found that the sample classes were normally distributed, as the significance value (α) was greater than 0.05, namely $0.421 > 0.05$. This means that the data from both the experimental class and the control class were normally distributed. The normality test results can also be seen in the following histogram.

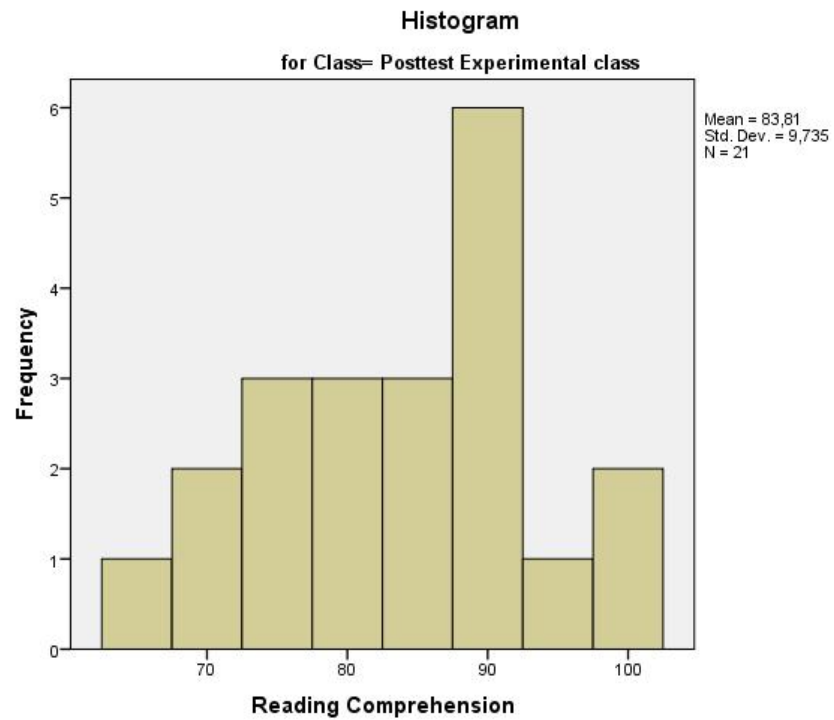


Figure 3.1
Results of the Normality Analysis of the Experimental Class

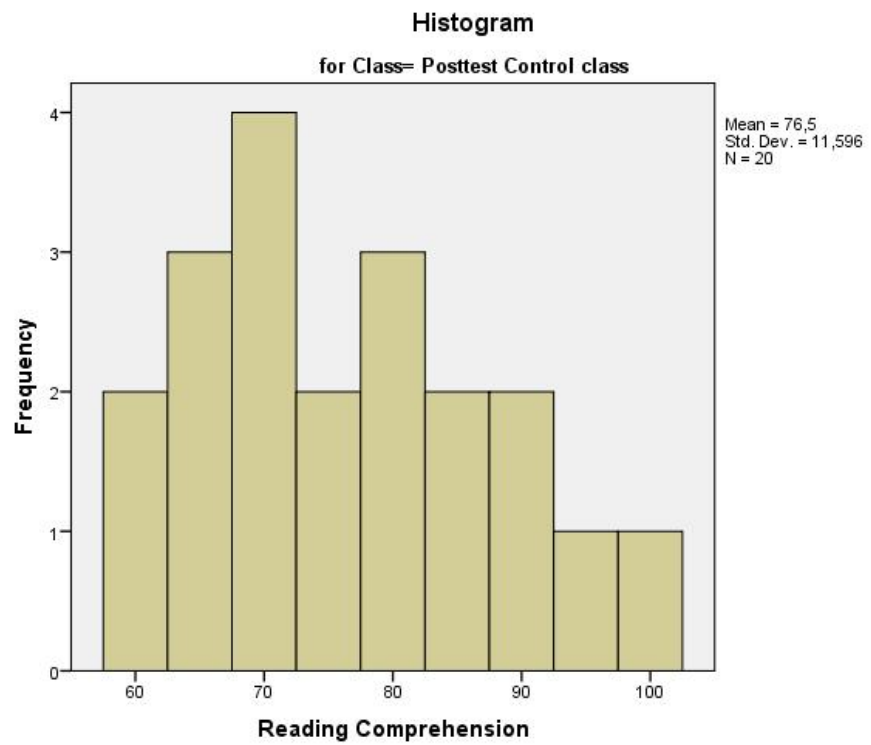


Figure 3.2
Results of the Normality Test Analysis of the Control Class

The normality test can also be further illustrated using the following Normal Q-Q Plot:

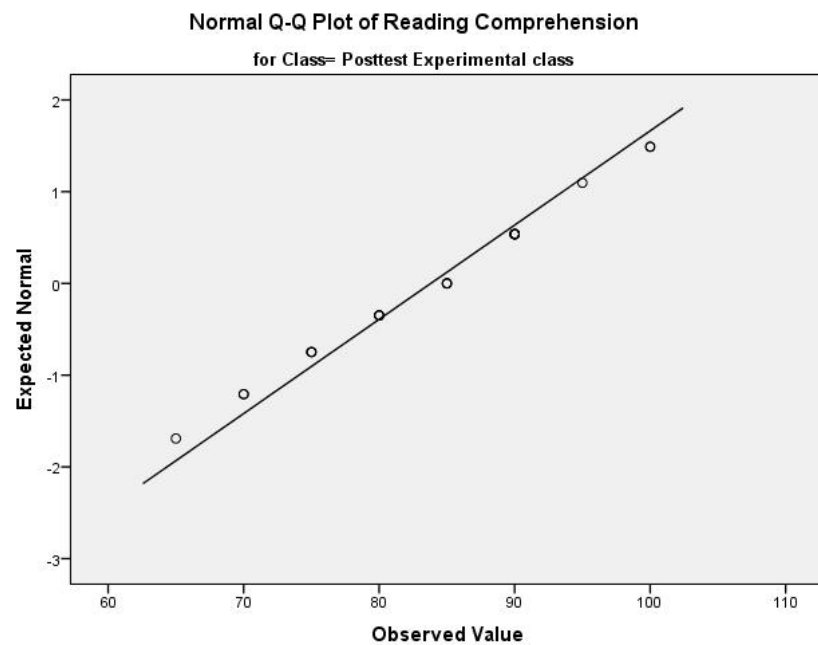


Figure 3.3
Results of the Normality Analysis of the Experimental Class

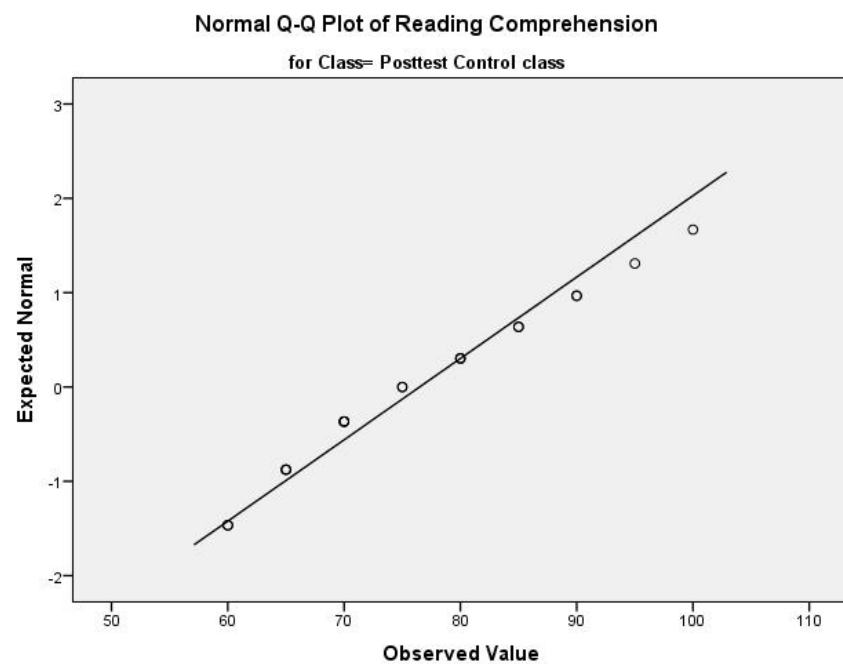


Figure 3.4
Results of the Normality Analysis of the Control Class

2. Homogeneity Test

After obtaining the results of the data normality test, the analysis was continued with a homogeneity test. The homogeneity test was conducted using the Fisher test. According to Sugiyono, the formula for the Fisher test is as follows:

$$F = \frac{S_1^2}{S_2^2}$$

Description:

F : Homogeneity

S_1^2 : Variance of the learning outcomes of the experimental class

S_2^2 : Variance of the learning outcomes of the control class

Decision criteria:

- a. If F-calculated < F-table, H_0 is accepted.
- b. If F-calculated > F-table, H_a is accepted.

The homogeneity test was conducted with the assistance of SPSS version 23, as follows:

Table 3.8
Results of the Homogeneity Analysis of the Experimental and Control Classes

Test of Homogeneity of Variances			
Reading Comprehension			
Levene Statistic	df1	df2	Sig.
,852	1	39	,362

Based on the table above, it was found that the two class groups had equal or homogeneous variances, with a significance value of $0,362 > 0,05$. Since this was a two-tailed test, the data were considered homogeneous. Therefore, it can be concluded that the posttest data of the experimental and control classes were homogeneous, allowing the hypothesis test to be conducted.

3. Hypothesis Test

After conducting the normality and homogeneity tests on the sample groups, the hypothesis test was conducted. The hypothesis test aims to determine whether the research hypothesis is accepted or rejected.

The criteria for testing H_0 are that H_0 is accepted if $t_{\text{calculated}} < t_{\text{table}}$, which can be determined from the t-distribution table with degrees of freedom $df = n_1 + n_2 - 2$ and probability $(1-\alpha)$. The hypothesis is rejected if $t_{\text{calculated}} > t_{\text{table}}$, which can be determined from the distribution table with degrees of freedom $df = n_1 + n_2 - 2$ at a significance level of 0,05. The t-test is used to determine the validity of a statement of hypothesis.

Hypothesis testing is a procedure used to make a decision regarding whether a hypothesis should be accepted or rejected. A hypothesis is an important component of research because it provides direction for further investigation. Therefore, the hypothesis must be tested statistically. In this study, hypothesis testing was conducted using SPSS version 23.

Decision-making criteria:

- a. If $\text{sig } \alpha$ (2-tailed) $> 0,05$, H_0 is accepted and H_a is rejected.
- b. If $\text{sig } \alpha$ (2-tailed) $< 0,05$, H_0 is rejected and H_a accepted.

