

CHAPTER III

RESEARCH METHODOLOGY

This chapter explains the research methods used. The methodology applied in this study is the quantitative method. The quantitative data obtained provides information on the student's listening skill level. This chapter also discusses the research design, population and sample, research instruments, data collection procedure, and data analysis.

3.1 Research Design

This research aims to determine the degree of enhancement in students' comprehension and retention of auditory information due to exposure to story-based interactive videos and to know why such cases improve their listening skills. That being said, the data that would be collected are in mixed method with primary quantitative data from students' scores in listening tests and secondary qualitative data obtained from observing students' behavior during four-week treatments. A quasi-experimental study design is implemented in this study. Tests related to the title were used to accomplish the quantitative method. The tests consist of students' listening skills and comprehension. Furthermore, an observation will be held during the treatment weeks to find students' behavior changes during the treatment of the media learning as a qualitative method.

The quasi-experimental study was designed to examine two groups of students' listening scores from two different classroom improvements, each receiving a different treatment. One benefit of using a quasi-experimental research design is that it allows the researcher to study variables that cannot be manipulated or controlled in an experimental setting, such as age, gender, or socioeconomic status. Quasi-experimental research design can also help evaluate the effectiveness of interventions or programs in naturalistic settings, such as schools or community organizations. This method can provide valuable insights into the real-world impact of the intervention and help inform policy and practice (Shadish et al., 2002).

Table 3.1
Quasi-Experimental Research Design

Groups	Pre-Test	Treatment	PosTest
Experimental	T_1E	X_1	T_2E
Control	T_1C	X_2	T_2C

Where;

T_1E = students' listening scores of the experimental group on the pre-test.

T_1C = students' listening scores of the control group on the pre-test.

X_1 = treatments using interactive video for the experimental group.

X_2 = treatments using weekly listening practices for the control group.

T_2E = students' listening scores of the experimental group on the post-test.

T_2C = students' listening scores of the control group on the post-test.

3.2 Population and Sample

The population of this study is first-year students from one of the vocational schools in Bandung. The sample of this study is 70 students from two classes. Researchers chose Vocational students as participants in this study is because, in their age of adolescence, high school students are attracted to any form of entertainment, especially games. Pritchard and Koski (2020) stated that primary and secondary education students (high school is included in secondary education) have yet to reach adult maturity. This limited view of the world may be valuable data, especially in this field of study involving games and how it affects students' academic performance. Furthermore, the researcher chose this vocational school mainly because of the ease of access since, at the time, the researcher is undergoing practical field experience as a teacher in one of the vocational high schools in Bandung. After the meeting with the English teacher at the school, the researcher is permitted to have two classes for them to teach and also become this research data.

Table 3.2
Population of students that will be used in the research

Class	Students			Class Speciality
	Total Students	Gender		
		Male	Female	
Class A	35	3	32	Tourism
Class B	35	3	32	Accountant

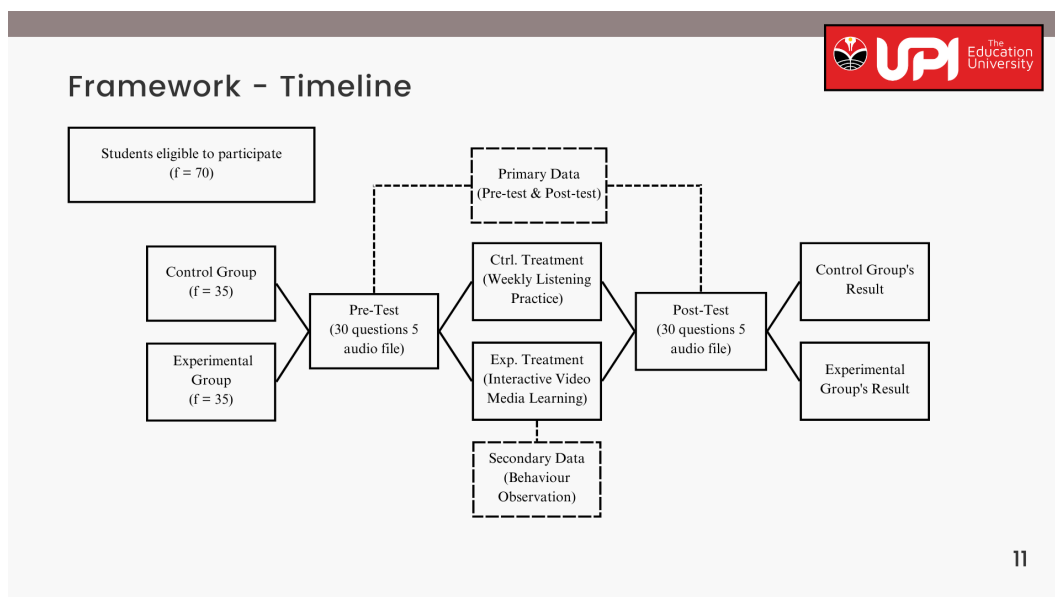
3.3 Research Instrument

In this study, the researcher will use tests and observation checklists. The tests will be a two-set of listening tests in the A1 and A2 English listening skill categories consisting of thirty questions with six parts of audio file provided by English Learning Hub on YouTube. This set of test are prepared to determine the improvement of student' listening skills after they experience the controlled treatment.

The observation checklists will be a set of questions that the researcher needs to answer according to the situation of the classroom whenever the treatment is being done. These consist of twenty observation questions focusing on student retention span, student participation, and metacognitive activity. The questions inside the observation checklists are developed by the researcher by referencing previous studies such as metacognitive knowledge, ability and theory (Helgesen and Brown, 2007; Flavell, 1979; Dweck et al., 1995; Goos et al., 2002) and Student activity, retention, and participation (Christensson, 2011; Kruk & Zawodniak, 2018; Apperley, 2010; Sun et al., 2014; Gedera & Zalipour, 2018).

3.4 Research Procedure

The procedure of this study follows four stages that will span around two to three months during practical field experience. The four stages are Survey, Planning, Treatment and Observation, Follow-up, and Countermeasures. This stage was developed by referencing previous studies that use quasi-experimental research design as their base. For more visual explanation are shown in Figure 3.1 below.



Source: Purnama (2023)

Figure 3.1
This study research procedure

3.4.1 Survey

This stage started around February 2023, while the researcher was conducting Practical Field Experience or PPL in one of the vocational high schools in Bandung. The researcher then had a meeting with the English teacher in the school regarding the class that was chosen to be the researcher's class while also for the research population and sample. The researcher also gets to know the students better to prepare for the study by following the English teacher's last two lectures before handing the class down to the researcher for the next two and a half months.

In this stage, the researcher conducts a pilot test to measure the potential success of this research. The instrument's reliability for this study was established using a pilot test through the pilot test that was conducted on the first week of May and 10 students were chosen randomly outside the sample of the population. The chosen 10 students conducted a mini-listening practice with the researcher to measure the correct implementation of listening practice and to give the researcher feedback on how they should approach the student with listening practice. Assessing a listening test's reliability is essential to ensure that the test scores are

accurate and trustworthy and can be used confidently to make decisions or draw conclusions (Ary et al., 2002)

3.4.2 Planning

A total of six treatments are planned to be held during the lesson in two and a half months. The first meeting was used for the pre-test to gauge the student's ability on their Listening skills. This pre-test used English Listening Test A1 as a starting point to measure the baseline of student listening skills.

The second through the fifth meeting was used for the treatment. A Class was practicing standard listening using listening audio, making them a Control Group. At the same time, B Class received interactive video treatment for their listening practice, making them an Experimental Group.

Finally, in the sixth meeting, the researcher conducted a post-test to determine what changes these students received from different treatments. The post-test used English Listening Test A2 as a measuring point to find improvement in students' listening skills after four weeks of treatment.

3.4.3 Treatment and Observation

With the help of the school's English teacher as a collaborator, the researcher conducted the first meeting to pre-test both classes. The collaborator helps the researcher brief the students beforehand regarding the researcher's plan to use both classes as samples for the researcher's study. For the treatment weeks, the researcher acted solo, giving out treatment and observing students' reactions towards both treatments, noting all changes and sometimes questions that helped the researcher get the needed data.

For the first through fourth treatment, the students in the experimental group started to watch the first installment of the Markiplier interactive video series, which took about forty minutes or one academic hour. The same can be said for control group students who started to do their weekly listening practice as ordered by the researcher, which took them forty minutes to complete and discuss the listening practice. While both groups are busy observing and listening to the material given in the form of videos and audio. The researcher began to observe and note some changes in the observational checklist that were prepared

beforehand. A handful of events that transpired in the class are noted in the weekly treatment findings.

After the treatment had been done, the researcher then went to the school English teacher for a daily teaching report and discussed the findings with them. Some advice was thrown regarding the researcher's method of approach during the treatment and how to make it more effective.

3.4.4 Follow-Up and Countermeasures

Although the researcher has already allocated the time and on-site preparation for the treatment, external interference and human errors may occur during the research. These unforeseen challenges can impact the integrity of the study's outcomes and conclusions. Researchers must remain vigilant and implement robust measures to mitigate the effects of such interferences and errors. Fortunately, there was no significant problem that the researcher and participant had throughout the treatment weeks.

However, there is a countermeasure when there is a significant problem that the researcher and participants might had. If a significant problem comes from the internal (school), the research must halt until further notice from the school, and there is no roundabout solution for it. Furthermore, to ensure the data collected are all reliable and concrete, the researcher must lay off a few participants if it is deemed unavoidable or impossible to mitigate.

3.5 Data Analysis

The data of this study are to be analyzed using IBM SPSS 25 for Windows to compute all data variables easily. The data that use these methods are primarily the test scores for both pre-test and post-test in both experimental and control groups since SPSS enables the management of huge and complicated data sets and the execution of advanced statistical analyses effectively and structured.

3.5.1 Data Analysis on Pre-Test and Post-Test

In analyzing the pre-test and post-test data given to both groups, a dependent t-test was used to compare the mean difference between both tests. Hatch and Farhady (1982) stated that a dependent t-test is used to investigate the significance of the mean between the pre-test and post-test. All hypotheses were

started with the alpha level at 0.05. The data gained through all the above were then computed using IBM SPSS 25 for Windows to compute all data variables easily. However, before the output data are analyzed in a t-test, it should fulfill the following criteria:

1. The data should have a normal distribution (Normality Test).
2. The variance of the two groups must be homogenous (Homogeneity test).
3. The participants must be different in each group. (Basic needs for Quasi-Experimental).

Four steps were accomplished by covering the normality test, homogeneity variance, independent test, and dependent t-test. The details of statistical procedures were as follows:

3.5.1.1 Normal Distribution Test

The Normal distribution test was calculated before the t-test. It intended to investigate whether pre-test and post-test scores were normally distributed among the groups. The statistical calculation for the normality test is used by Shapiro-Wilk as shown below:

1. Setting the hypothesis, H_0 = the score between experimental and control groups are normally distributed.
2. Setting the level of significance (p) at 0.05
3. Analyzing the normality distribution using the Shapiro-Wilk test.
4. Comparing the score between the test result and level of significance with these values:
 - a. If Asymp.Sig. > 0.05, the null hypothesis is rejected, which means the sample for both groups is normally distributed.
 - b. If Asymp.Sig. < 0.05, the null hypothesis is rejected, which means the sample for both groups is not normal.

In addition, if the data is not normally distributed, the Mann-Whitney test will be administered to test the hypothesis with IBM SPSS 25 to calculate the result.

3.5.1.2 Homogeneity Variance

The homogeneity variance test used the Levene test in IBM SPSS 25. This is the step after finding out that pre-test and post-test scores are normally distributed. Then, the steps are as follows:

1. Setting the hypothesis, H_0 = data between the two groups are homogenous.
2. Setting the level of significance (α) = 0.05
3. Measuring the homogeneity variance using Levene's test.
4. Comparing the result of Leven's test and alpha significance level:
 - a. If Asymp.Sig. < 0.05, the null hypothesis is rejected, which means the two groups are not equal.
 - b. If Asymp.Sig. > 0.05, the null hypothesis is not rejected, which means the two groups are homogeneous.

3.5.1.3 Independent T-Test

Analyzing the difference between the means of the experimental and control group, the independent t-test was then used to reveal its result. This research analyzed the independent sample test using computation with IBM SPSS 25.

There are some significant criteria in analyzing the independent t-test, as follows:

1. Setting the level of significance (α) = 0.05.
2. Analyzing the independent t-test using SPSS 25
3. Determining t-test hypothesis:
 - a. $t\text{-obtained} < t\text{-critical} = H_0$ is accepted, and there is no significant difference between both groups in the pre-test mean.
 - b. $t\text{-obtained} > t\text{-critical} = H_0$ is rejected, and both groups have a significant difference in the pre-test mean.

3.5.1.4 Paired-Sample T-Test

Paired t-test was used to determine the differences between the pre-test and post-test in each group. In this research, the independent sample was analyzed using IBM SPSS 25.

The steps are as follows:

1. Setting up the hypothesis, H_0 = there is no significant difference between students listening test in pre-test and post-test.

2. Setting the level of significance (α) at 0.05
3. Calculating t-test score using IBM SPSS 25
4. Comparing t-obtained and t-critical:
 - a. If $t\text{-obtained} > t\text{-critical}$, it means that the null hypothesis is rejected, there is a significant difference between the scores before and after the treatment.
 - b. If $t\text{-obtained} < t\text{-critical}$, the hypothesis is not rejected because there is no significant difference between pre-treatment and post-treatment.

3.5.1.5 Effect Size

The effect size computation is conducted to check the level of effect of the treatment after t-test calculation by using IBM SPSS 25 from the independent t-test of the post-test. The effect size is used to determine how significant the impact of treatment is on the experimental group's score.

To measure the value of effect sizes, the researcher uses Cohen's effect size specifications:

Table 3.3
Cohen's effect size classification

Effect Size (ES)	Interpretation
$0.00 \leq ES < 0.20$	Ignored
$0.20 \leq ES < 0.50$	Small
$0.50 \leq ES < 0.80$	Moderate
$0.80 \leq ES < 1.30$	Large
$1.30 \leq ES$	Very Large

3.5.2 Data Analysis on Student's Behaviour

For this data that are collected from observing students using observational checklists will be analyzed using qualitative method. The researcher must take into account the observation's context, including the location, task, goal, participants, and environment in order to interpret it successfully.

3.5.2.1 Retention Span

Analyzing students' retention spans begins with gathering qualitative data through observation that was conducted within the four-week treatment for the experimental group. The data used are the duration students concentrate, at what point students are distracted, whether the student continues paying attention after being distracted, and how often the student gets distracted. This data provides insights into students' attention span during the treatment and how long it lasted until students got distracted.

3.5.2.2 Participations

Analyzing participation for students in one class requires additional help. The researcher asked their friend, who also enrolled in the same vocational high school as an internship teacher, to record the treatment using their mobile phone. The data used are based on the activity level of the class during each treatment meeting, for instance, how many times students struck a question, how many students were passive throughout the treatment, and how many students were playing along with the interactive video. The data was then converted to the student status list as active or passive in the class.

3.5.2.3 Metacognition

Analyzing the metacognitive activity of the students is mentioned in chapter two about metacognition, where students' ability to think about thinking in the school lessons is observed. These data are based on students' performance in problem-solving both live in class and in their assignments. Hence, this measures the “passive students” better whether they are focused on thinking and absorbing new information or vice versa.

3.6 Hypothesis

The researcher hypothesized that there would be two outcomes at the end of this experiment. The first one is the Null hypothesis (H_0), where there are no significant changes for the students after the researcher administered a controlled learning experiment, proving no significant differences between interactive videos and regular listening audio practice. The second one is the Alternate hypothesis (H_a), where there were significant changes for the students after the researcher

administered the controlled learning experiment, proving that there are significant differences between using interactive videos and regular listening audio practice.

Other hypotheses other than the outcomes of this study also developed inside the subchapter of analyzing the quantitative data. A handful of hypotheses are made as indicators of how the data translated in the collective statement in front of he hypothesized outcome.