

CHAPTER III

RESEARCH METHODOLOGY

A. Research Design

The method of this research applied causal-comparative study with quantitative approach. According to Gay (2012), in causal-comparative research, the researcher aims to identify the source or reason for existing variations in the behavior or status of groups or people. In a causal-comparative design, two groups differ on a variable of interest and are compared based on the dependent variable. The researcher chose two groups of participants, frequently referred to as experimental and control groups, but more appropriately referred to as comparison groups. The groups are divided in two ways: either one has a trait that the other does not, or both have the characteristic to varying degrees or amounts.

The aim of this study is to compare two personalities (extrovert and introvert) and determine which one has higher achievement in speaking fluency and whether there is a significant difference between them. Quantitative research approaches are applied to describe current conditions, investigate relations, and study cause–effect phenomena (Gay, 2012). The researcher focused on comparing students' fluency skill based on their personality and found out the significant difference between them.

Gay (2012) stated that variable means a characteristic, property, or attribute of an individual or an organization that can take on different values or vary among individuals or groups. Variables are central to research as

they represent what is measured, observed, or manipulated in a study. In this study, the researcher conducted two variables based on research design above. The variables following:

1. Independent Variable (X)

Independent variable is the variable that affects or causes the change of the dependent variable or in causal-comparative this called grouping variable. In this research, the independent variable is students' personality that is divided into groups, both are sometimes referred to as experimental and control groups, but should more correctly be referred to as comparison groups. Group A is extrovert students and group B is introvert students.

2. Dependent Variable (Y)

Dependent variable is the variable hypothesized to depend on or to be caused by another variable, the independent variable (Gay, 2012). In this research, the dependent variable is students speaking fluency. The result or final score from the test used for a dependent variable.

B. Population and Sample

1. Population

Population is the complete group of individuals or circumstances that are the subject of the research (Gay, 2012). In this research, all students in class X E2 of MAN 1 Padang in academic year 2025/2026 is population. Researcher selected these population based on observation in MAN 1 Padang before began this research. Population in

this study did not involve numerous students in several classes, but only chose one class in grade 10th.

There are 34 students included in this research. The majority of participants are in A1 and A2 level in CEFR, based on discussion from students' English teacher in class X E2. The population consisted of 21 female students and 13 male students in class.

2. Sample and Sampling Technique

According to Gay (2012), sample is a section of the population chosen to take part in a study. Since the sample is selected to be representative of the overall population, researchers can use the data received from the sample to make conclusions about the population.

There are several sampling techniques that can be applied to select what is known as a representative sample. Sampling technique generally divided into two; random and non-random. The researcher used non-random sampling to determine the sample. Nonrandom sampling, also called nonprobability sampling, is the process of selecting a sample using a technique that does not allow the researcher to specify the probability each member of a population has of being selected for the sample.

Nonrandom sampling techniques include the following:

a. Convenience sampling

This technique is selecting individuals who are readily available at the time of the study. The examples are recruiting

volunteers or using pre-existing groups simply because they are easily accessible.

b. Purposive Sampling

Purposive sampling is selecting a sample that is thought to be representative of a specific population based on researcher experience and knowledge about the chosen group.

c. Quota sampling

Quota sampling is selecting a sample based on specific numbers of individuals or groups with diverse characteristics. This method is frequently used in large-scale surveys that cannot include all members of the population of interest.

Based on three sampling techniques above, the researcher chose purposive sampling as sampling technique. This is because the researcher did observation and discussed students in MAN 1 Padang, based on that knowledge the researcher selected one class which is X E2. In this selected class, the researcher assumed different speaking skill might be affected by personality.

According to Etican (2016), purposive sampling is divided into seven categories such as maximum variation sampling, homogenous sampling, typical case sampling, extreme case sampling, total population sampling and expert sampling. The researcher used total population sampling in this research, where the entire population that met the criteria are included in the research. This sampling technique is

chosen because the number of students in this research is relatively small.

Based on the sampling technique which is total population sampling, the sample of this research is the same as population. The sample is all students in class X E2 of MAN 1 Padang in the academic year 2025/2026. There are 34 students included in this research, all of the sample would represent the population.

The majority of sample are never taken English courses and do not have family members or friends who actively use English. Data obtained from surveys filled out by students showed 23 students have never taken English lessons, 6 students have taken English lessons and 4 students chose not to fill out the survey. The data also showed 90% of students from the total of 29 students who filled out the survey do not have family members who can or actively speak English. This data showed less variety of the students, the majority generally learn English at school without English courses or speaking practice with family members or friends.

C. Instruments

In this research, data collected by groups that are already different, which is speaking skill and the researcher sought to determine what is causing the difference. This type of research is sometimes called *ex post facto* in Latin which means “after the fact”. *Ex post facto* study or "after the fact" studies the effect and apparent cause after they have occurred.

Instrument is a tool that is used to measure variables, gather data, or obtain information in research (Gay, 2012). The instruments used in this research are closed-ended questions and set of questions. These instruments are in following:

1. Questionnaire (Closed-Ended Questions)

The instrument of this research used closed-ended questions for questionnaire. This instrument is based on the Extroversion/Introversion in Language Learning Test (EILLT) from Skellet (2017) for determining the student's personality between extrovert and introvert. The researcher met the students directly then gave 26 numbers of close-ended questions and students must answer it based on the likert scale zero until five (0 = never, 1 = very rare, 2 = rare, 3 = frequent, 4 = usually, 5 = always).

The 26 closed-ended questions were asking about students' feelings in context learning English in the classroom. The test mostly asked about boredom in class, working in a group, talking more or quiet, like to go out, nervous to speak loudly and performing in front of people.

All the scores in the questionnaire sheet summed up and determined which personality they are into. Students with scores above 65 are considered extroverts, and students with scores under 50 are considered introverts. Students in the gap between 50 and 65 are considered ambiverts, which was not included in the test.

2. Speaking Test (Set of Questions)

The next instrument is a set of questions based on Aptis Speaking Test to measure speaking fluency and measure with rubric of fluency by Kahng (2014). Aptis Speaking Test is one of some type tests for speaking that is common use for L2 (secondary language) learners, the test is similar to Cambridge General Speaking Test, IELTS, PTE Academic and TOEFL.

This research used Aptis speaking only in task 1 because task 1 is for A1/A2 English level which is the English level of the students in this research. Students have to answer three personal questions about personal information and interests. Test taker gave 3 questions and students had 30 seconds to answer each question. The students should give an extended answer in purpose to make score high.

There are five set questions prepared by the researcher, each set has three questions about personal information and interests. All of the questions in five sets are quite open. For instance, the test taker may ask students questions like: describe, please tell me about or what you like.

The five question sets are in following:

- a. Set A, asking about describing people
- b. Set B, asking about describing and explaining place
- c. Set C, asking about telling activities
- d. Set D, asking about describing and explaining things in school
- e. Set E, asking about describing food

To enable measuring process, individual task performances are recorded in audio files using recorder. Test takers used a recorder application in smartphone to record all the audio. Test takers are also asked to consent to the usage of participants' recordings for the research and assessment process.

Furthermore, all the recordings transcript manually by the researcher using some mark such as parenthesis for length of silent pause, and length of filled pauses was marked next to each filled pause without parentheses. For measuring pauses, digital recorder software used in purpose to achieve length of pause time.

The researcher also used a website to help measure syllables in speaking. Syllable Counter (SyllableCounter.net, n.d.), an online software based on an English syllable dictionary, was used to count words and syllables. The software calculated the amounts of syllables for speech transcripts entered.

The fluency measured by the following rubric evaluation of speaking fluency by Kahng (2013) as bellow:

1) Speed

The formula to measure speed or speech rate is in following table:

Table 3.1 Speed Measurement and Formula

Measurement	Formula
Mean syllable per second	$\frac{\text{total number of syllables}}{\text{speech time excluding silent pause time (s)}}$

Speed, or some expert called speech rate, is how fast the speaker could utter syllables within a second. Mean syllable duration is a pure measure of speed in that it excludes pause time unlike the traditional measure of speech rate. Speed indicated how many syllables the speakers could speak in one second. This aspect revealed how fast the speaker utter syllables in numeric data. The points shown in this aspect are in line with how good the speaker is in fluency.

2) Repair

The formula to measure repair is in following table:

Table 3.2 Repair Measurement and Formula

Measurement	Formula
Number of corrections and repetitions	$\frac{\text{total number of corrections and repetitions}}{\text{spoken time (s)}}$

Repair is when the speaker repeats the utterance. Repair divides into two such as correction and repetition. The difference between these two aspects is repetition when speaker rephrases the same utterance while correction when the speaker rephrases utterance with different words, in other words the speaker tries to correct utterance. Repair should be avoided to keep the fluency good, but in case to keep the flow of speaking, the speaker would recommend used repair rather than pause.

3) Length of run

The formula to measure length of run is in following table:

Table 3.3 Length of Run Measurement and Formula

Measurement	Formula
Mean syllables per run	Average number of syllables produced between two silent pauses.

Run in this context means utterance in syllable between silent pauses. Length of run is measuring mean syllables per two silent pauses. All of the syllables in one run are summed and then divided into the number of runs the speaker spoke. This aspect is similar to speed, this also significantly affects fluency. The difference is this aspect counted in run not spoken time. Same with speed, the bigger points in this aspect means greater fluency the speaker has.

4) Pauses

The formula to pause is in following table:

Table 3.4 Pauses Measurement and Formula

Measurement	Formula
Mean length of all pauses (s)	$\frac{\text{total length of all pause time (s)}}{\text{number of pauses}}$

Pause in speaking means taking a break from uttering words. Pause could be silent or filled with any sound just to keep the flow of utterance. Pause is allowed in speaking but better to avoid in order to keep the speaking flow normal and natural. If in any case the

speaker should use pause, filled pause is a good choice to keep the flow of speaking rather than silent pause. Both filled and silent pause are counted in this aspect. Researcher was not measuring the difference between silent and filled pause since pause is not a significant aspect to measure in this study.

To analyze the mean length of pause, the researcher counted all duration of filled pause and silent pause and then divided it into the number of pauses the speaker used. This would indicate the mean duration of pause. For pause length, if it is equal or longer than 0.2 seconds considered as pause and if shorter than 0.2 second could be ignored and not measured. In this research, pause only counted if the length of pause time greater or same with 0,3 second.

D. Technique of Collecting Data

This research used primer data which means the researcher collected the original data from participants. For causal-comparative research, the data can not be manipulated, therefore the data collected only once both for two comparison group A and group B. Researcher collected data in two different ways as in following:

1. Questionnaire

Questionnaire is a set of written survey questions to be completed by a specific group of research participants (Gay, 2012). Questionnaire is used to determine students' personality, whether they are extrovert or

introvert. The questionnaire is given on paper based, students had to write name and fill the scale beside the questions list.

The researcher used Likert Scale to measure the data. Likert scale is a set of statements used to assess a real or hypothetical condition. Participants rate their level of agreement with offered statements (from strongly disagree to strongly agree) using a metric scale (Joshi et al., 2015). Students had 30 minutes to fill the questions and then collected it to the researcher.

2. Aptis Speaking Test

The Aptis Speaking Test is a test to examine speaking skill developed by the British Council (Shin, 2022). The test is generally for examining the English language proficiency in speaking skill as on the scale of the Common European Framework of Reference for Languages (CEFR).

The Aptis Speaking test consists of four parts targeting different CEFR levels (Tavakoli, 2017), as following:

- a. Task 1: Respond to three questions on personal topics and interests.
Target Level: A1/A2. Response time: 30 seconds/question.
- b. Task 2: Respond to 3 questions, including describing a photo and answering a concrete, familiar topic related to the photo. Target Level: B1. Response time: 45 seconds/question.
- c. Task 3: Respond to 3 questions related to 2 contrasting pictures.
Target Level: B1. Response time: 45 seconds/question.

- d. Task 4: Provide a long turn, integrating responses to a set of 3 questions. Target Level: B2. Response time: 2 minutes and 1 minute for preparation.

The first task in Aptis speaking test means the easiest and the last means the hardest. The first task is for beginner in English which is in A1 or A2 level (the lowest level in CEFR level). The second and the third task are for higher level which is B1. The last task is the hardest for level B2 in CEFR level.

Based on discussion with an English teacher in MAN 1 Padang and observation, the majority participants are in level A1/A2. This is the reason the researcher only used task 1 for the test. Students would get a speaking test and the result recorded to measure their fluency. The researcher measured fluency using rubrics from Kahng (2014) that consist of 4 aspects. Students got one random set question which has three different questions about their personal topics. The questions are read one by one and students have 30 seconds to answer each question.

In order to get an appropriate situation in a test, when a student got test with a task taker, the other students must wait for their turn. Other students can do any activity they would like such as reading, drawing, writing, preparing themselves or any kind of activity as long as it did not bother other people.

E. Technique of Data Analysis

Analysis of data in causal-comparative study involved a variety of descriptive and inferential statistics (Gay, 2012). Descriptive statistics are the average performance of a group on a measure of some variable and the data showed in descriptive statistic in the table

For the normality test, the researcher used the Shapiro-Wilk test for testing if the data were normally distributed or not. The Shapiro-Wilk test is a famous normality assumption diagnostic tool with good power qualities. It is based on correlation within given observations and associated normal scores (Das, 2016). This ratio is both scale and origin invariant, hence the statistic is useful for testing the composite hypothesis of normality (Saphiro-Wilk, 1965).

In a paper conducted by Saphiro and Wilk in 1958, data simulation that used is under 50 samples ($n < 50$). Therefore, this is recommended for this research since participants are 34 students. In the test, data is said to be normally distributed if the significance value is more than 0.05 (sig. > 0.05).

The Saphiro-Wilk formula as follow:

$$T_3 = \frac{1}{D} \left[\sum_{i=1}^k a_i (X_{n-i+1} - X_i) \right]^2$$

Where:

D = Coefficient test Saphiro-Wilk

a_i = constants generated from the means, variances and covariances

X_i = the ordered participant's values

n = amount of participants

For the coefficient test Saphiro-Wilk:

$$D = \sum_{i=1}^n (X_i - \underline{X})^2$$

Where:

X_i = the ordered participant's values

$\underline{X}$ = mean of participants' value

n = amount of participant

The next step after showing data in the table as a descriptive statistic is analyzing data. For comparing the students' speaking fluency final score, the researcher calculated if there is a significant gap between introvert students and extroverted students in speaking fluency. This analysis used inferential statistic. The most commonly used inferential statistic in causal comparative research is the t-test, this test is to determine whether the scores of two groups are significantly different from one another (Gay, 2012).

Definition of t-test is a type of statistical test for comparing the means of two groups. t-test divided into two types; independent and dependent (paired). Independent t-test can be used when two groups are completely different and independent for each other, while paired t-test can be used when the same group with different times, treatment or condition (Kim, 2015). The researcher used an independent t-test in fact that this research used two different groups as extrovert and introvert. According to Zimmerman & Zumbo (1993), if population variances or sample sizes are

unequal, the Welch t-test is recommended to use for independent t-test. In this research, Welch t-test is used because of the different variance of two groups.

The data analyzed using Welch t-test. The t-test formula as follow:

$$t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

Where:

$\bar{x}_1$ and $\bar{x}_2$ = Average values of each of the sample sets

s_1^2 and s_2^2 = Variance of each of the sample sets

n_1 and n_2 = Number of students in each sample set

For the variance formula:

$$s^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2$$

Where:

s^2 = Variance

n = Number of participants in the data set

x_i = Each value in the data set

$\bar{x}$ = Mean of all values in the data set

The researcher compared t-item to t-table at a valuable level of 5% ($\alpha = 0.05$). If the t-item is fewer than t-table (0.05), that means there is a significant gap between two types of groups. And the opposite, if the t-item

is greater than t-table, that means there is no significant gap between two types of groups.

Both normality and Welch t-test applied using SPSS (Statistical Package for the Social Sciences). SPSS is a software that enables you to obtain statistics ranging from simple descriptive numbers to complex analyses of multivariate matrices (Arkkelin, 2014). This research used SPSS 27 to create a descriptive statistic, test the normality data and compare two variables.

