CHAPTER III

RESEARCH METHOD

This chapter describes the features design of the fundamental research methods which were applied by the researcher. Inside it, there were any cases would be explained, they were research design, population and sample, research instruments, data collection technique, and data analysis technique.

A. Research Design

Research design was a procedure to collect, analyse, interpret, and report on data obtained during the research period. The study conducted by the researcher was classified as a correlational study, because the purposed of this study was to reveal the correlation between one variable to another variable. In this study, the researcher did not alter or modify variables, so this research was categorized as a survey research. The descriptive-quantitative method was used to analyse and report the result.

Correlational research was the method used to investigate variables (two or more) to determine the relationship diversity characteristic of variables. Another research asserted that correlational study means a statistical relationship between a set of variables, and none of them were manipulated.\(^1\) According to Creswell, correlational study was included in the quantitative research.\(^2\) In addition, survey study was the part of

---

correlational study.\textsuperscript{3} It was designed to provide a quantitative or numeric
description of attitude, trend, or opinion of the population by studying a
sample of it.\textsuperscript{4} Babbie (as cited by Creswell) stated that the survey research,
the procedure of the data collection was done by questionnaire or
structured interview.\textsuperscript{5}

According to the explanation above, the researcher investigated
two variables in a condition where they were not manipulated. SES
(Socio-Economic Status) and LLS (Language Learning Strategy) are two
variables that would be investigated.

The first procedure was determining the object (population and
sample). Because this research was universal (the results could be applied
to any person), the sample could be selected randomly. The simple random
sampling was used to determine the sample.

Furthermore, in the data retrieval from the sample was applied by
using questionnaires. Questionnaire was a document that contains some
questions or items used to investigate and collect the data would be
analysed.\textsuperscript{6} According to Babbie (as cited by Heum Lee), A questionnaire
could collect data by asking people with questions or asking them to agree

\textsuperscript{3} Ibid., 14.
\textsuperscript{4} John W. Creswell, \textit{Research Design: Qualitative, Quantitative, and Mixed Method Approaches},
\textsuperscript{5} Creswell, \textit{Research Design}, 14.
\textsuperscript{6} Bidhan Acharya, “Questionnaire Design,” in \textit{A Paper Prepared for a Training Workshop in
Research Methodology Organised by Centre for Post Graduate Studies Nepal Engineering
College in Collaboration with University Grant Commission Nepal, Pulchok, June, 2010}, 2,
%20Research%20Methods/6.4%20Questionnaire%20Design\_Acharya%20Bidhan.pdf.
or disagree with statements representing different points of view.\textsuperscript{7} The questionnaire used in this study consists of two kinds, namely SES (Socio-Economic status) questionnaire and LLS (language learning strategies) questionnaires.

The third procedure was the process of data analysis. The results of the data collection were analysed using descriptive statistic and inferential statistic. In the descriptive statistic the normality test and homogeneity test was acted as the requirement before doing the inferential statistic. in the next procedure, MANOVA was carried out to analyse the hypothesis. Furthermore, Pearson product-moment formula used to determine the relationship between SES and LLS.

\section*{B. Population and Sample}

\subsection*{1. Population}

All item in any field of inquiry was defined as ‘Universe’ or ‘Population.\textsuperscript{8} In this study, the population that was studied was the EFL or ESL learners in English Education Department, Faculty of Tarbiyah and Teacher Training, Sunan Ampel State Islamic University, Surabaya.. The total of the population was around 400 learners.

\textsuperscript{7} Sung Heum Lee, “Constructing Effective Questionnaires,” in Handbook of Human Performance Technology (Hoboken: Pfeiffer Wiley, 2006), 760.

\textsuperscript{8} C. R. Kothari, Research Methodology: Method and Technique, Second Revised. (New Delhi: New Age International (P) Ltd, 2004), 55.
2. Sample

It was defined as the selected respondents was taken from the population.
C. Research Instrument

In this study, researcher used a single type of instrument in the form of questionnaire. Kothari said “A questionnaire consists of a number of questions printed or typed in a definite order on a form or set of forms.”11 The questionnaire was mailed to respondents who were expected to read and understand the questions and write down the reply. The respondents had to answer the questions on their own.

In this research, the researcher uses 2 type of questionnaires. The first questionnaire was purposed to measure the Language Learning Strategy used by students, and the second was a questionnaire that was served to investigate Socio-Economic Status owned by students.

1. Language Learning Strategies Instrument

For a questionnaire that measures LLS, the researcher used a questionnaire namely “Strategy Inventory of Language Learning for EFL/ESL v 0.7 (SILL)” which was designed by Rebecca Oxford (See appendix I). The instrument consisted of 50 items based on taxonomy developed by Rebecca Oxford. They were divided into Memory strategies, cognitive strategies, compensatory strategies, metacognitive strategies, affective strategies and social strategies. for the details explanation please read this following explanation:

a) Memory strategies (items 1–9)

b) Cognitive strategies (items 10–23)

c) Compensatory strategies (items 24–29)

---

11 Kothari, Research Methodology: Method and Technique, 100.
d) Metacognitive strategies (items 30–38)

e) Affective strategies (items 39–44)

f) Social strategies (items 45–50)

The LLS questionnaire was categorized as international or standardised instrument, it was mean that the instrument validity and reliability was absolute, because it was often used as measurement tool without any change.

While for item scoring, every item had 1-5 score. The score was gotten from linear scale which was described as follow:

1. Never or almost never true of me
2. Usually not true of me
3. Somewhat true of me
4. Usually true of me
5. Always or almost always true of me

Every item had been answered got a score according the criterion above, then the result was calculated in each part of strategies of LLS. And the data was taken and proceed from the instrument was categorized as interval data.

2. Socio-Economic Status Instrument

The researcher used a questionnaire from a journal written by Aggarwal et al (See Appendix 2), because there were cultural similarities between their study with this research.
There are many questionnaires for measure SES, but they had no significant correlation with education, most of them use the SES questionnaires to measure the SES has correlation with health or economy. For education SES has several categories as the measurement tool, they are income, occupation, and education which were measured from family background aspect.

The total of question of SES questionnaire was 22 item. Every item was labelled as X (for example item number 1 was labelled as X1). For the data scoring, every answer of the data had been labelled with a number which was started from 1 and so on (it was depended on the total of answer items of multiple choices) Then the researcher calculated total of the score. It was the basic of SES classification.

According to Aggarwal the instrument still needed a test to check the validity and reliability. The researcher used pilot validity to check it. Sudarmanto asserted that the validity test could be applied by using *Pearson Correlation Product Moment (PCPM) technique*. the researcher used IBM SPSS 23 to calculate the data. But in manual it was measured by this following formula.

\[
r = \frac{n \sum_{i=1}^{n} x_i y_i - \sum_{i=1}^{n} x_i \sum_{i=1}^{n} y_i}{\sqrt{\sum_{i=1}^{n} x_i^2 - \left( \sum_{i=1}^{n} x_i \right)^2} \sqrt{\sum_{i=1}^{n} y_i^2 - \left( \sum_{i=1}^{n} y_i \right)^2}}
\]

**a. Pilot Validity 1**

The researcher distributed the questionnaire into 10 persons. After the data had been collected, the validity
investigated by using PCMP. Before performed it, the data was calculated to know about the total score.

After performing PCPM, the result of pilot validity 1 (see appendix 3 (Table 3.1. Pilot Validity 1)) showed there were only 8 items of the questionnaire was decided as valid.

The item could be valid if the r score (could be seen in correlation product moment column) > r table (could be seen on appendix 4 (r table)).

The significant value was decided by the researcher was 5% or 0.05. It was mean that the value of r table was 0.632. According to table 3.1, there only 8 item was decided as valid instrument, because they had a greater value than r table. they are:

1) Item x1=.817** > 0.632
2) Item x2=.683* >0.632
3) Item x3=.776* > 0.632
4) Item x4=.880** > 0.632,
5) Item x11=.726** > 0.632
6) Item x14=.734** > 0.632,
7) Item x19=.881** > 0.632
8) Item x20=.907* > 0.632
b. **Pilot Validity 2**

The fact was support the homogeneity choice and the non-understandable factor was indicated as the reason about non-validity of the items. So the researcher tried to give item 15 and 21 (see *Appendix 5 (SES Instrument phase 2)*) some additional explanation/information/description based on the culture in Indonesia to make the items clear to understand and remove one item that hard to understand. so the total items were 21 from the previous item (22).

To solve the first problem (the answer had been chosen was homogeneity), the researcher tried to add the total of the respondents. So in this Pilot Validity 2, the questionnaire was spread out to the 20 people which was different with the first pilot test. The increment of the respondent was purposed to make a more answer variety, so the data had a higher chance to be valid.

The method used to analyse the data was same with the *Pilot Validity 1*. The result (see *Appendix 6 (Table 3.2. Pilot Validity 2)*) presented the same result as the previous test. However, the significant score of the Pearson Correlation was different, but the valid item was same.

Because the total of respondents was 20, the r table was decided in sig. 5% was 0.444. From the table 3.2, the valid item
was still on the same item but in different value of correlation and significant (see the marking score in Appendix 6 (Table 3.2. Pilot Validity 2)). for more detail see the table 3.3 bellow

Table 3.3 (The conclusion of Pilot Validity 2)

<table>
<thead>
<tr>
<th>Number of Items</th>
<th>Score of the correlation (r score)</th>
<th>&lt; (less than) or &gt; (more than)</th>
<th>r table on level 20 N</th>
</tr>
</thead>
<tbody>
<tr>
<td>X1</td>
<td>871**</td>
<td>more than</td>
<td>0.444</td>
</tr>
<tr>
<td>X2</td>
<td>.687*</td>
<td>more than</td>
<td>0.444</td>
</tr>
<tr>
<td>X3</td>
<td>.598*</td>
<td>more than</td>
<td>0.444</td>
</tr>
<tr>
<td>X4</td>
<td>.889**</td>
<td>more than</td>
<td>0.444</td>
</tr>
<tr>
<td>X11</td>
<td>.766**</td>
<td>more than</td>
<td>0.444</td>
</tr>
<tr>
<td>X14</td>
<td>.843**</td>
<td>more than</td>
<td>0.444</td>
</tr>
<tr>
<td>X19</td>
<td>.471**</td>
<td>more than</td>
<td>0.444</td>
</tr>
<tr>
<td>X20</td>
<td>.867*</td>
<td>more than</td>
<td>0.444</td>
</tr>
</tbody>
</table>

c. Pilot Validity 3

After holding the pilot validity twice, the researcher decided to take the valid item and used it as the instrument for this research. So the questionnaire of SES included 8 item (See Appendix 7 (SES Instrument Used for Research)).

The chosen items were argued had capability to measure the variable, because the three main concept of Socio-Economic status (income, occupation, and education) had been included. For investigates why the data could be invalid, the researcher use and elicitation question which was related to each item, and the reason why the other item was clarified as invalid item was not because the respondents did not
understand about the question. But because they had their own reason such as follow:

1) They did not need it however they could have it (it was related with a question which ask about the proprietary, such as agricultural land, house, vehicle and other)

2) It was a privacy (it was related to the tax)

3) They do not know the fact (some of them did not know the real information of the family)

Before applied it to the real sample, the researcher spread out the chosen item to 10 persons, it was purposed to check the validity and their reliability. The reason why the researcher decrease the sample was the difficulties of the correlation test. one of the factor of it was the total of respondents. The higher of respondents’ total, the higher correlation probability could occur. If this test was passed with only 10 persons, it was mean that if it was applied to the respondents with higher mount, it will more valid.

The result of the Pilot Validity 3 (see Appendix8 (Table 3.4. Pilot Validity 3)) showed that in line Pearson Correlation indicated that all of item was clarified had a correlation each other. And the value of r score was higher than r table in 5% significant level (see table 3.4). So it could be concluded that the instrument was a valid instrument for each item.
Table 3.5 (The conclusion of Pilot Validity 3)

<table>
<thead>
<tr>
<th>Number of item</th>
<th>Pearson correlation score (r score)</th>
<th>&lt; (less than) or &gt; (more than)</th>
<th>R table in 5% significant level in 10 N</th>
</tr>
</thead>
<tbody>
<tr>
<td>X1</td>
<td>0.897</td>
<td>&gt; (more than)</td>
<td>0.632</td>
</tr>
<tr>
<td>X2</td>
<td>0.669</td>
<td>&gt; (more than)</td>
<td>0.632</td>
</tr>
<tr>
<td>X3</td>
<td>0.776</td>
<td>&gt; (more than)</td>
<td>0.632</td>
</tr>
<tr>
<td>X4</td>
<td>0.706</td>
<td>&gt; (more than)</td>
<td>0.632</td>
</tr>
<tr>
<td>X5</td>
<td>0.880</td>
<td>&gt; (more than)</td>
<td>0.632</td>
</tr>
<tr>
<td>X6</td>
<td>0.874</td>
<td>&gt; (more than)</td>
<td>0.632</td>
</tr>
<tr>
<td>X7</td>
<td>0.737</td>
<td>&gt; (more than)</td>
<td>0.632</td>
</tr>
<tr>
<td>X8</td>
<td>0.670</td>
<td>&gt; (more than)</td>
<td>0.632</td>
</tr>
</tbody>
</table>

**d. The Reliability Test**

The validity test was run for three times and showed the same valid item, it was mean that the instrument was used by the researcher was reliable.

**D. Data Collection Technique**

According to Sudjana, data collecting could be applied by two styles, first was census and second was sampling. The technique used by the researcher was a random sampling. 25 learners were chosen randomly from every semester in English education department. the researcher did not take the sample in single semester, because there was prohibition that the distribution to the population could not be effective, remembering that the population was EFL/ESL learners in English education department of Sunan Ampel Surabaya. By taking population from every semester, the distribution of the research result could be more effective. Sudarmanto,

---

asserted that in the random sampling, the sampling cloud be effectively when the population was bounded, then the sample could be taken.\textsuperscript{13}

The research was carried out independently by the researcher. It needed time around 20 days to collect the data. The author spread out the questionnaire via online (by using social media, such as WhatsApp, BBM, and google form) and offline. But the much of information or data was arrived from offline method. 63 persons was filled the offline questionnaire, and 12 persons was filled the online questionnaire.

E. Data Analysis Technique

The data analysis technique contained about the way of the researcher to process the data had been collected. This section described about the statistical analysis used by the researcher. It was divided into two categorizations, first was descriptive statistic and the second was inferential statistic (parametric or non-parametric). The data analysis was performed by using IBM SPSS 23, so the descriptive statistic (included mean, median, mode, z score, and standard deviation) was calculated coincide with the inferential statistic.

This research used CPS (correlation product moment) to investigate the correlation between Socio-Economic status and language learning strategies. But before perform it, there were some requirements had to fulfil,\textsuperscript{14} they are:

\textsuperscript{13} Sudarmanto, Statistik Terapan Berbasis Komputer: Dengan Program IBM SPSS 19, 47.

\textsuperscript{14} Edi Riadi, Metode Statistika Parametrik & Non-Parametrik (Tanggerang: Pustaka Mandiri, 2014), 141.
1. The sample was random
2. The data was categorized as interval or ratio
3. The data had normal distribution characteristic
4. The data had homogeny characteristic

The 1st requirement had been completed because the sample was taken using simple random sampling technique. Then the 2nd requirement also had been fulfilled because the score of the data could be calculated. While mean, median, mode, standard deviation, and z score was needed to calculate before performing 3rd and 4th requirements.

The first step taken by the researcher carried out the descriptive statistic. In this step mean, median, mode, standard deviation, z score, and classifying the learners according SES level was applied.

The second step was applying inferential statistic. the researcher executed the normality test, homogeneity test, MANOVA test, and Pearson Correlation Product Moment test chronologically. For easier understanding see figure 3.1 bellow.
For further information, let’s shift to the explanation below.

1. **Descriptive Statistic**

   It was known as deductive statistic, simple statistic, and descriptive statistic. The purpose of it was to organize and analyse the numerical data, so it could give a description of the data clearly. Sugiyono (as cited by Sudarmanto,) asserted that the function of descriptive statistic was to describe the object of research by using sample or population without carried out an analysis.\(^{15}\)

   Mean, modus, median, standard deviation, table, graphic, and the others were some of the descriptive statistic. They were very important, because the inferential statistic was begun from them. The basic descriptive statistic was mean, median, and modus. After finding them the

---

standard deviation and z score could be found, then they could be used to carry out the normality and homogeneity test.

This was the first series of the data analysis way which must be calculated before applying the inferential statistic. In this part, the researcher had to found mean, median, mode, standard deviation, and z score. After calculated them, the researcher could classify the learners according to their SES level.

a. Mean

Mean could be defined as average score of the data. It could be calculated by using the formula bellow.

\[
\bar{x} = \frac{x_1 + x_2 + x_3 + \cdots + x_n}{n}
\]

\(\bar{x}\) = Mean  
\(x_1\) = Data Frequency  
\(n\) = Total of data/respondent

b. Median

Median was explained as the location of data which was disport it into 2 part with the same portion.\(^{16}\) It could be calculated with:

\[
Me = \frac{n + 1}{2}
\]

\(Me\) = Median  
\(n\) = Total of data

---

c. Modus

While for the modus was explained as the value of the data with the most frequencies. To calculate the modus the researcher only arranged the data form the smallest in to the highest. Then the score with highest frequencies could be found.

d. Standard Deviation (SD)

The next was standard deviation and z score. Standard deviation (SD) was used to measure how the data could be spread out. In general SD could be measured by using this formula.

\[
s^2 = \frac{n\sum f_i x_i^2 - (\sum f_i x_i)^2}{n(n-1)}
\]

\[s^2 = \text{Variance} \]
\[x_i = \ldots \text{ Score} \]
\[n = \text{Sample Value} \]

e. Z Score

After found the SD, another significant step was calculating the Z score. it could be used in almost parametric statistic. It also needed to measure the homogeneity and normality assumption test. In general Z score could be measure by this following formula.

\[17\text{ Ibid., 48.}\]
\[ z = \frac{x - X}{SD} \]

\( x \) = the data would be calculated

\( X \) = Mean

SD = Standard Deviation

f. SES classification

<table>
<thead>
<tr>
<th>Table 3.1 (SES Classification Example)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Score reference</td>
</tr>
<tr>
<td>LOW</td>
</tr>
<tr>
<td>MIDDLE</td>
</tr>
<tr>
<td>HIGH</td>
</tr>
</tbody>
</table>

The researcher divided data into the 3 class, they are high, middle, and low. According to the table 3.1, the classification based on the range of the score had been divided. Learners who got total of score 1-17 was categorized as LOW SES, learners who got score around 18-34 was classified as MIDDLE SES, and they were get more than 34 of Total SES score was categorized as HIGH SES.

2. Inferential Statistic

After committed the descriptive statistic, and all of the component (mean, median, mode, z score, and SD) was found. The homogeny and normality test was carried out. Then if the data passed both of those assumption, MANOVA and Pearson Correlation Product
Moment could be used. For more detail information about the homogeneity, normality, MANOVA, and Pearson Correlation Product Moment and how to interpret the result of them, let see this following information.

a. Normality Test/Normal Distribution Test (Kolmogorov Smirnov Test)

After finding all of the requirement to apply the normality test and homogeneity test, the next step was calculating the normality test using Kolmogorov Smirnov method. Its application was equal with Lilliefors’ method, the differences was located on the comparator table. Lilliefors used the Lilliefors’ table and Kolmogorov Smirnov used Kolmogorov Smirnov’ table. Manually the Kolmogorov Smirnov test was carried out by these following steps.

1. Arrange the data from the smallest into the highest
2. Decide the frequencies of each data
3. Decide the cumulative proportion (CP) = \[
\frac{\text{cumulative frequencies}}{n}
\]
4. Decide the Z score for each data
5. Use the Z table
6. Use cumulative probabilities table (Fz)
7. Find out the value of \( a_1 \) and \( a_2 \) with this following criterion
a. \[ a_2 = |Cp - Fz| \] (the quarrel of CP and Fz)

b. \[ a_1 = |a_2 - \frac{f_i}{n}| \]

8. Looking for the highest value of a1 and a2 (Dh)

9. See Kolmogorov Smirnov’s table (Dt)

10. Compare Dh and Dt. (Dh<Dt = Normal)

But the because the data will be analysed using IBM SPSS 23, the analysis was little differences. According to Sudarmanto, the result of Kolmogorov Smirnov test could be analysed by using the value of Asymp. Sig. (2-tailed) which was compared with Alpha Value.18 For the data interpretation, if the Asymp. Sig. had score more than Alpha score; it was mean that the data was normally distributed.

b. **Homogeneity Test (Levene’s Test)**

While for the Homogeneity test was carried out by using Levene’s test. The homogeneity test had function to check was the data came from the homogeny population or not.19 Levene’s test was one of the method to carried out that assumption test. In General, it could be calculated by using this following formula.

\[
L = \frac{(N - k)\Sigma ni (V - vk)}{(k - 1)\Sigma (Vij - Vi)}
\]

---

18 Sudarmanto, *Statistik Terapan Berbasis Komputer: Dengan Program IBM SPSS Statistik* 19, 130.
\(L = \text{Levene’s Value}\)

\(X = \text{Value of Residual Data}\)

\(x = \text{Mean}\)

\(N = \text{Total of sample}\)

\(K = \text{Total of Population}\)

To know the result of the data was homogeneity or not, it could be seen from the significant value of each variable. When the score of \(\text{sig.}\) was more than 0.05 or 5\% (\(\text{Alpha}\)), it could be concluded that the data was homogeneity.

c. MANOVA

After both of the assumption was complete, the hypothesis test was carried out using MANOVA. It was equal with ANOVA, the differences was located on the independent variable.\(^{20}\) The researcher chose this method because the total of variable was more than 2. If the variable only 2, it could be analyzed using T-test. While if the variable was 3, it could be analyzed by ANOVA. It was technique purposed to know about was the differences between variance score of one dependent variable was caused by the differences of each variable of independent variable.\(^{21}\) While for MANOVA the total of dependent variable was more than two. According to


Sarwono, MANOVA was used to check the significant of mean between two groups with two dependent variable or more. \(^{22}\) The purposed of MANOVA was to investigate the hypothesis (by using MANOVA analyzation) and find out the differences of independent variable in dependent variable (by using *Between Subject-Effect Test* which had included in MANOVA). \(^{23}\) The analysis explanation of the first and second purposed of MANOVA was carried out by using the *sig. value* of each variable or variance. When the *sig. value* was less than *Alpha score*, it was mean that the null hypothesis was rejected.

d. **Pearson Correlation Product Moment**

It was purposed to investigate the kind of correlation between SES and LLS (it could be negative or positive correlation). The CPP (Correlation Product Moment) was used. It was defined as the test device was purposed to investigate the correlation between two random sample which was homogeny and normal. \(^{24}\) The analysis of data result was if the data had (*) or (**) sign, it was mean that there was significant correlation.

\(^{23}\) Ibid.
and the sign “- or +” in the score indicated the kind of correlation.\textsuperscript{25}

\textsuperscript{25} Sami’an, “SPSS Korelasi” (presented at the Penggunaan SPSS Dalam Statistik, Universitas Gadjah Mada, 2010), 10.