CHAPTER III
RESEARCH METHODOLOGY

This chapter presents the method that is used in order to find the answers to the research questions stated in Chapter I. The research design includes the participants, the data collection, and the data analysis. The last section involves the concluding remarks of this chapter.

3.1 Statement of the Problems

This research is conducted to answer the following research questions:

1. At what level of self-efficacy did the students belong to?
2. What is the correlation between the students' self-efficacy and their English-speaking performance?

3.2 Research Design

This current research used quantitative approach and descriptive method to analyze the data that have been collected. A quantitative approach and descriptive method were used in this research because the main purpose of this research was to find out whether there was any correlation between students' self-efficacy level and their English-speaking performance or not. In order to find out the correlation, a correlation research design was used in this study.

Creswell (2012, p. 21) explains that a correlation research design is procedures in quantitative research in which the researcher measures the degree of association (or relation) between two or more variables using the statistical procedure of correlation analysis. It can be concluded that a correlation research design allows the researcher to relate two or more variables (students' self-efficacy level and students' English-speaking performance) to see if they influence each other or not. In addition, Privitera (2014, p. 240) has defined a correlation research design as "the measurement of two or more variables to determine or estimate how far the variables are related or change in an identifiable pattern". Spearman's Rank-Order Correlation coefficient was used in this study to find how far the variables are

related or changed. Further, the descriptive method is used to describe the calculation results.

3.3   Participants

The participants of this study were one class of 8th grade students at one of junior high schools in Bekasi. The class of participants consisted of 42 students. Nonetheless, 6 students were not included in this study because they did not fulfill the requirements. Since they did not take all the administered tests. As a result, 36 students were listed as the sample of this research. According to Creswell (2012, p. 146), estimation number of participants for a correlational study is approximately 30 participants.

3.4   Research Instruments

In this research, a research was employed through the questionnaire and speaking test. Each technique will be elaborated thoroughly below.

3.4.1 Questionnaire

According to Creswell (2012), a questionnaire is a form used in a research to collect information from the participants. Regarding this, Mackey & Gass (2005) state that questionnaire is commonly used to collect data related to attitudes or opinions from a group of people. Besides, questionnaire can also be very advantageous. It is because the questionnaire can be given to a large number of participants at the same time. Thus, it could help the researcher to save more time while conducting the research.

The questionnaire that is used in this study was "Children's Perceived Academic Self-Efficacy Questionnaire" adapted from Pastorelli et al. (2001). The questionnaire was employed to find out at what level of self-efficacy the students belonged to. The questionnaire also described students' beliefs and confidence while using English in the classroom.

The questionnaire used a Likert-type scale ranged from 1 (not quite sure) to 7 (completely sure) in terms of students' behavior in English class. The questionnaire used the Likert scale because "the Likert scale is the universal

method for survey data collection, which is easily understood" LaMarca (cited in Yunita, 2014, p. 50). Besides that, according to Bertram (2007, p. 8), the responses from the questionnaire that used a Likert scale were easily quantifiable.

Table 3.1 Likert Scale

not quite sure	a little sure	somewhat sure	Sure	very sure	really sure	completely sure
1	2	3	4	5	6	7

The participants were asked to put a checklist in the box with a number that represented their estimation of their English-speaking skill. After they finished filling up the questionnaire, the total score of all items were processed by using ordinal category formula. Ordinal classification divided the scores into five ranges; very high self-efficacy, high self-efficacy, medium self-efficacy, low self-efficacy, and very low self-efficacy.

3.4.1.1 Validity of Self-Efficacy Questionnaire

According to Hughes (1989, p. 22), a test is said to be valid if it measures accurately what is intended to measure. Therefore, a questionnaire used in this research should measure what was supposed to be measured. In order to check the validity of the questionnaire, according to Masrun (in Michelle, 2013, p. 26), a correlation technique is mostly used. Therefore, a pilot test was conducted in order to check the validity of questionnaire. Participants for this pilot test consisted of 38 students. The formula is as follows.

$$r_{xy} = \frac{\sum xy}{\sqrt{\{(\sum x^2)\}\{(\sum y^2)\}}}$$

r_{xy} : correlation coefficient of a sample

x : ($x_i - \bar{x}$) score item

y : ($y_i - \bar{y}$) total score

(Sugiyono, 2007, p. 228)

The formula was used to correlate between variable x and y. The data were calculated based on the formula above using SPSS 21.0. The data from questionnaire were calculated to find the *r*-value. The *r*-value was obtained from comparing the *r*-result with *r*-table. The instrument was considered as a valid instrument if the *r*-result > *r*-table at 95% confidence level. In contrast, the item was considered as invalid if the *r*-result < *r*-table. If the item was considered as invalid, the item was being dropped or needed to be revised.

In order to calculate the *r*-table, it was also necessary to find the degree of freedom. The degree of freedom calculation formulated as (df = n − 2). Since the sample *(n)* of the pilot test was 38, so 38 - 2 = 36. After that, the value of *r*-result and *r*-table can be seen in the table below.

Table 3.2 The Result of Validity Test on Students' Self-Efficacy Questionnaire

Question	*r*-result	*r*-table	Description
1	.81	.32	Valid
2	.84	.32	Valid
3	.82	.32	Valid
4	.24	.32	Invalid
5	.76	.32	Valid
6	.81	.32	Valid
7	.81	.32	Valid
8	.59	.32	Valid
9	.66	.32	Valid
10	.84	.32	Valid
11	.86	.32	Valid
12	.88	.32	Valid
13	.81	.32	Valid
14	.83	.32	Valid
15	.59	.32	Valid

Based on the results of the questionnaire, 14 items were considered as valid and only one item was considered as invalid. The question number 4 was considered as invalid because the *r*-result (.24) < *r*-table (.32). Therefore, the invalid item was removed from the questionnaire. For further details see in Appendix 2.

3.4.1.2 Reliability of Self-Efficacy Questionnaire

In terms of reliability, Underhill (1987, p. 105) stated that reliability is usually seen as a completely different concept of validity, and the two terms are presented in terms of mutual incompatibility: highly reliable tests are less valid, and vice versa. However, the instrument that was used in this research must test it reliability. By testing the reliability of the instrument, was expected that the instrument would be convincing to provide the same results even though it is carried out in the different situation.

Further, the method to check the reliability of the questionnaire was Cronbach's Alpha method. According to Sugiyono (2007, p. 365), Cronbach's Alpha was used for the interval data or essay. The formula is as follows.

$$r_i = \frac{k}{(k-1)}\left\{1 - \frac{\sum \sigma_i^2}{\sigma_t^2}\right\}$$

r   :   instrument reliability

k   :   number of questions

$\sum \sigma_i^2$:   variance of scores on each question

σ_t^2 :   total variants

(Sugiyono, 2007, p. 365)

The result of reliability test on students' self-efficacy questionnaire using Cronbach's Alpha method can be seen in the table below.

Table 3.3 The Result of Reliability Test on Students' Self-Efficacy Questionnaire

Variable	K	R	Criteria
Students' self-efficacy in speaking English	14	.944	Highly reliable

Based on the table above, the Cronbach's Alpha value for 14 valid items is .944. In order to describe level of reliability of instruments, George & Mallery (2003) has suggested the rule that is commonly used for describing internal consistency of the data. See table 3.4 below.

Table 3.4 The Reliability of the Data Interpretation

Cronbach's Alpha	Interpretation
$\alpha \geq 0.9$	Excellent
$0.9 > \alpha \geq 0.8$	Good
$0.8 > \alpha \geq 0.7$	Acceptable
$0.7 > \alpha \geq 0.6$	Questionable
$0.6 > \alpha \geq 0.5$	Poor
$0.5 > \alpha$	Unacceptable

(George & Mallery, 2003)

According to the table above, it can be concluded that the reliability of the questionnaire was good. Thus, it can be concluded that the instrument can be used in this research.

3.4.2 Speaking Test

The speaking tests were chosen by the researcher based on the standard and basic competency in the syllabus of Education Unit Level Curriculum (KTSP) 2006. After looking at the KTSP 2006, the researcher further read the syllabus for the first semester of 8th grade students. The standard competency for speaking skill states, *"2. Mengungkapkan makna dalam teks lisan fungsional dan monolog pendek sederhana yang berbentuk* descriptive *dan* recount *untuk berinteraksi*

dengan lingkungan sekitar". Further, the basic competency in this section states, *"4.1 Mengungkapkan makna dalam bentuk teks lisan fungsional pendek sederhana dengan menggunakan ragam bahasa lisan secara akurat, lancar, dan berterima untuk berinteraksi dengan lingkungan sekitar"*. The syllabus shows that junior high school students are expected to receive and produce spoken language in form of descriptive text. Therefore, a short dialogue taken from "English in Focus" by Wardiman et al. (2008) was chosen as the instrument for the first speaking test. In the first speaking test, students were given the script before the test to allow them to read it and prepare. After that, students performed it in front of the classroom.

For the second speaking test, the standard and basic competency were used is similar to the first ones. According to the standard and basic competency stated above, junior high school students are expected to produce spoken language in form of descriptive text. For that reason, a series of pictures about hobbies and daily activities were chosen as the instrument media for the second speaking test. In the second speaking test, students were asked to describe the picture that was given by the researcher. Students were also allowed to speak freely. When the students had finished speaking, or if they faltered, the researcher may ask questions designed to elicit particular information, perhaps about a point the student had missed or not made clear. Further, according to Underhill (1987, p. 66), visual stimuli are an economic and effective way of providing a topic of conversation without giving the student words and phrases to manipulate and give back.

3.4.2.1 Validity and Reliability of Speaking Test

Speaking test implemented in this research should be valid and reliable. Underhill (1987, p. 104) states that validity has a variety of meanings. It depends on the situation in which the test is used as much as the test itself. Further, he states that validation is relative, and not an absolute process. A test relates only to the particular circumstances in which a test was established. On the other side, reliability is usually seen as a completely different concept of validity, and the two

terms are presented in terms of mutual incompatibility: highly reliable tests are less valid, and vice versa.

According to Underhill (1987, p. 105), an oral test is a personal encounter between two human beings; it is designed by humans, administered by humans, taken by humans and marked by humans. However, in the end, the evaluation and development of this human activity are surrendered by the statistical machine. Therefore, the test maker should do the evaluation and development of the test. It is because the validity of test relatively depends on the situation in which a test is established, and it cannot be generalized by a statistical machine.

3.5   Scoring Procedure

The procedures of scoring for the Self-Efficacy questionnaire and speaking tests are as it follows.

3.5.1 Questionnaire

The participants were assessed on their level of self-efficacy by using the Self-Efficacy questionnaire. They were asked to put a checklist in the box with number ranged from 'not quite sure' (0) to 'completely sure' (7). They filled the questionnaire based on their estimation of their beliefs on English-speaking skill. The questionnaire consisted of 14 items. Therefore, the maximum score for Self-Efficacy questionnaire was 98 and the minimum score was 14.

3.5.2 Speaking Tests

According to Mertler (2001), rubrics are rating scales that are used with performance assessments. Rubrics are formally formed as scoring guides, consisting of specific pre-established performance criteria, used in evaluating student work on performance assessments. He adds that rubrics are typically the specific form of scoring instrument used when evaluating student performances or products resulting from a performance task. Moreover, Underhill (1987, p. 97) explains that the use of scoring rubric makes the scoring of oral tests easier and more consistent. Therefore, this study needs to use a scoring rubric to assess students' speaking performance.

Linse & Nunan (2005) state that there are two types of rubrics; a holistic rubric and an analytic rubric. According to Nitko (2001), a holistic rubric requires the teacher to score the overall process or product as a whole, without judging the component parts separately. In contrast, an analytic rubric provides detailed information broken down into different categories and useful feedback on areas of strength and weakness. Besides, an analytic rubric also makes scoring more consistent across students and provides more guidance for instructional planning. Moskal (cited in Mertler, 2001) explains that in an analytic rubric, the teacher scores separate, individual parts of the product or performance first. After that the teacher sums the individual scores to obtain a total score.

Linse & Nunan (2005) has suggested Student Oral Language Observation Matrix (SOLOM) by Daniels & Bizar (2004) for assessing students' speaking tests in more detail. It was because SOLOM is an analytic rubric that provides more categories to assess students' speaking tests. The score on SOLOM ranged from 1-5. The lowest score was 1 and the highest score was 5. Students performances are determine using the criteria that is written on the scoring rubric. There are four elements of language that assessed in this study, which are fluency, pronunciation, grammar and vocabulary. Besides, confidence and volume were also included in the scoring rubric.

Table 3.5 Speaking Scoring Rubric 1

No.	Elements of Language	1	2	3	4	5
1.	Pronunciation	Frequent phonemic errors and foreign stress and intonation patterns that cause the speaker to be unintelligible.	Frequent phonemic errors and foreign stress and intonation patterns that cause the speaker to be occasionally unintelligible.	Some consistent phonemic errors and foreign stress and intonation patterns, but the speaker is intelligible.	Occasional pronunciation errors with inappropriate intonation patterns, but the speaker is intelligible.	Occasional pronunciation errors, but the speaker is always intelligible.
2.	Fluency	Speech is so halting and fragmentary that intelligibility is virtually impossible.	Numerous pauses that interferes with intelligibility.	Some pauses but do not interfere with intelligibility.	Speech is clear with occasional lapses.	Speech is smooth and effortless.
3.	Volume	Speaks cannot be understood. Volume is too soft to be heard by all audience members.	Often mumbles or cannot be understood. Volume often too soft to be heard by all audience members.	Speaks clearly and distinctly most of the time (85-94%). Volume is loud enough to be heard by all audience members at least 60% of the time.	Speaks clearly and distinctly all the time (95-100%). Volume is loud enough to be heard by all audience members at least 80% of the time.	Speaks clearly and distinctly all the time (95-100%). Volume is loud enough to be heard by all audience members throughout the presentation.
4.	Confidence	Sometimes stands up straight, leans, slouches, rocks back and forth, doing inappropriate moves. Never establishes eye contact with interlocutor.	Sometimes stands up straight, leans, slouches, rocks back and forth, etc. Establishes eye contact with interlocutor once or twice.	Stands up straight. Doesn't ever rock back and forth, pace, lean-on, board, etc. Establishes eye contact with interlocutor but it is sporadic.	Stands up straight, looks relaxed and confident. Establishes eye contact with interlocutor but it is frequent.	Stands up straight, looks relaxed and confident. Move around the room in a meaningful fashion. Establishes eye contact with interlocutor.

(Adapted from Student Oral Language Observation Matrix by Daniels & Bizar (2004) cited in Linse & Nunan (2005))

Table 3.6 Speaking Scoring Rubric 2

No.	Elements of Language	1	2	3	4	5
1.	Grammar	Errors in grammar and word order so severe as to make speech virtually unitelligible.	Grammar and word order errors make comprehension difficult. Must often rephrase and/or restrict him/herself to basic patterns.	Makes frequent errors of grammar and word order that occasionally obscure meaning.	Occasionally makes grammatical and/or word order errors that do not obscure meaning.	Grammar is always correct.
2.	Vocabulary	Vocabulary limitations so extreme as to make conversation virtually impossible.	Misuse of words and very limited vocabulary, comprehension quite difficult.	Student frequently uses wrong words, conversation somewhat limited because of inadequate vocabulary.	Student occasionally uses inappropriate terms and/or must rephrase ideas because of lexical inadequacies.	Use extensive vocabulary.
3.	Pronunciation	Frequent phonemic errors and foreign stress and intonation patterns that cause the speaker to be unintelligible.	Frequent phonemic errors and foreign stress and intonation patterns that cause the speaker to be occasionally unintelligible.	Some consistent phonemic errors and foreign stress and intonation patterns, but the speaker is intelligible.	Occasional pronunciation errors with inappropriate intonation patterns, but the speaker is intelligible.	Occasional pronunciation errors, but the speaker is always intelligible.
4.	Fluency	Speech is so halting and fragmentary that intelligibility is virtually impossible.	Numerous pauses that interferes with intelligibility.	Some pauses but do not interfere with intelligibility.	Speech is clear with occasional lapses.	Speech is smooth and effortless.

5.	Volume	Speaks cannot be understood. Volume is too soft to be heard by all audience members.	Often mumbles or cannot be understood. Volume often too soft to be heard by all audience members.	Speaks clearly and distinctly most of the time (94-85%). Volume is loud enough to be heard by all audience members at least 60% of the time.	Speaks clearly and distinctly all the time (95-100%). Volume is loud enough to be heard by all audience members at least 80% of the time.	Speaks clearly and distinctly all the time (95-100&). Volume is loud enough to be heard by all audience members throughout the presentation.
6.	Confidence	Sometimes stands up straight, leans, slouches, rocks back and forth, doing inappropriate moves. Never establishes eye contact with interlocutor.	Sometimes stands up straight, leans, slouches, rocks back and forth, etc. Establishes eye contact with interlocutor once or twice.	Stands up straight. Doesn't ever rock back and forth, pace, lean-on, board, etc. Establishes eye contact with interlocutor but it is sporadic.	Stands up straight, looks relaxed and confident. Establishes eye contact with interlocutor but it is frequent.	Stands up straight, looks relaxed and confident. Move around the room in a meaningful fashion. Establishes eye contact with interlocutor.

(Adapted from Student Oral Language Observation Matrix by Daniels & Bizar (2004) cited in Linse & Nunan (2005))

Based on the scoring criteria above, the first scoring rubric consisted of pronunciation, fluency, volume and confidence. This was due to the speaking test that was scripted. Participants were only memorizing the same script that was given to them earlier. Therefore, the scoring rubric did not score the grammatical and vocabulary aspects.

In addition, the second scoring rubric consisted of six items, namely grammar, vocabulary, pronunciation, fluency, volume and confidence. The researcher wanted to score participants' grammatical and vocabulary ability because on the second speaking test, the participants were asked to produce sentences which require them to use grammatical and vocabulary ability.

3.6   Research Procedure

In this research, the data were collected in two sections; through the questionnaire and speaking test. The further explanation will be explained below.

1. The supervisors were checked whether the statements of the questionnaire were really designed in order to answer the research questions or not.
2. The instruction of the questionnaire was given both in oral and written. In order to avoid the misunderstanding, all the instructions were given in *Bahasa Indonesia*. The participants were asked to read and answer all questionnaire items carefully.
3. After the questionnaire had been completed, the English-speaking test were administered by the researcher. The speaking test was conducted twice; role-play dialogue and interview using a picture.
4. For the speaking test, it was held in the classroom because according to Underhill (1987, p. 17), a test that is to be a regular part of teaching program should be held in the familiar surroundings, such as an ordinary classroom.
5. In the first speaking test, participants were given the script before the test to allow them to read it and prepare. After that, they performed it to the researcher.

6. For the second speaking test, the chairs of the researcher and the participants were arranged side by at an angle in order to make the participants feel that the researcher was investigating them. Two or three participants were interviewed by the researcher at the same time.

7. In the interview using a picture, participants were asked to describe the picture that was given by the interviewer. They were also allowed to speak freely. When they had finished speaking, or if they faltered, the interviewer may ask questions designed to elicit particular information, perhaps about a point the participant had missed or not made clear.

3.7 Data Analysis

To answer the first research question, which was *At what level of self-efficacy did the students belong to?,* the researcher used a questionnaire. The questionnaire that was used is Self-Efficacy Scale taken from Children's Perceived Academic Self-Efficacy Questionnaire by Pastorelli, et al. (2001). Further, the questionnaire was analyzed using SPSS 21.0 to interpret the data. SPSS was used in this study because this tool was considered to be much more practical and efficient, and provide more accurate analysis results. After the data were analyzed, the questionnaire was processed through the ordinal classification to determine the range of the students' self-efficacy level, from very high to very low.

For the second question, which was *What is the correlation between the students' self-efficacy and their English-speaking performance?*, a correlation design was used to find the correlation between two variables. The data were also processed using SPSS 21.0. First, speaking tests were administered to get students scores on the speaking tests. After that, the data from both the questionnaire and the speaking tests were calculated to find out whether the data were normally distributed or not. The data further were correlated to find the correlation values between two variables. Further, the correlation value from the calculation was interpreted using the guidelines proposed by Evans (1996) to see the strength of correlation between two variables in this research.

After correlating the two variables, the test of hypotheses was assessed. The test of hypotheses was conducted to determine whether the null hypothesis should be accepted or rejected based on the certain level of significance. The level of statistical significance indicated whether the results were statistically significant and not replicable. It can be formulated as it follows.

H_0 = There is no statistically significant correlation between students' self-efficacy level and their speaking performance for 8th grade students of junior high school in Bekasi.

H_1 = There is a statistically significant correlation between students' self-efficacy level and their speaking performance for 8th grade students of junior high school in Bekasi.

3.8 Concluding Remarks

This chapter has presented how the data in this research are collected and analyzed. It includes research design, participants, research instruments, research procedure, and data analysis. The primary data were taken from a Self-Efficacy Questionnaire and speaking tests. The data were taken from the 8th grade students. Nevertheless, a correlation analysis was used to identify the degree of correlation between students' self-efficacy level and their speaking English performance. In this regard, the SPSS 21.0 was used to analyze and calculate the data. After correlating the two variables, the test of hypotheses was assessed to determine whether the null hypotheses should be accepted or rejected. The results from this research will be elaborated in the next chapter.