CHAPTER III
RESEARCH METHODOLOGY

This chapter explains the details of research methodology used by the researcher to conduct the research. Chapter Methodology includes research design, participant, population and sampling technique, research instrument, research procedure, and data analysis.

3.1 Research Design

This research is a quantitative research with descriptive method. Descriptive method allow the researchers to describe the quantitative information (Kaplan, 2005). In descriptive research, the researchers do not have direct influence to the research and has no manipulation of the factors that influence the research subject (McMillan, 2008). For this research, the descriptive method is used for analyzing the students' misconception in global warming topic gathered through the quantitative information.

The data collected by using two-tier test combined with Certainty of Response Index (CRI). The data gathered in twice with three weeks apart between the first test and the second test. The data from the tests are used for analyzing the current misconception that experienced by the samples of this research. Misconceptions detected when the samples answer the test consistently and they have high certainty of their answers.

3.2 Participants

Participants who involved in this research were students of lower secondary school students. All participants were seventh grade students which were 31 students from school A, 29 students from school B, and 30 students from school C. The total participant for this research is 90 seventh grade students with the range of their age around 13-14 years old. Students' characteristic for being the participants was students who have already learned about global warming topic in the class. Then the participants were explored their misconception in global warming topic by following the test.

3.3 Population and Sample

Population in this research is seventh grade students of three public lower secondary school in Bandung. All schools have already implemented the curriculum 2013 included the global warming topic in the learning process. The schools and samples were determined by using cluster sampling based on the school region in Bandung. The cluster sampling allows the researcher to select randomly one class for each school to be the sample of this research (Kaplan, 2005). Therefore three schools in East Bandung region have been chosen as the location of the research. From those school, one class from each school has been explored the students' misconception in global warming topic.

3.4 Operational Definition

In order to avoid the misinformation in this research, several operational definitions are explained in this study.

1) Students' understanding categories in global warming topic is the categories determined based on students' answer of the two-tier test and CRI. There are 4 levels which are "Understand the concept well", "Understand the concept but lack of confidence/ lucky guess", "Lack of Knowledge", and "Misconception".
2) Students' misconception in global warming topic is the concept that is incompatible with the scientific sense or with the acceptable concepts of global warming by the experts which detected by the two-tier test and CRI.
3) Two-tier test is used for detecting students' misconception in global warming topic. The test is in the form of multiple choice. The first tier consists of the global warming concept while for the second tier consist of the reason for the first tier.
4) Certainty of response index that applied in this research has 6 scales which is from 0-5 index. Students' misconception are identified if students give correct answer and incorrect reason; incorrect answer and correct answer; or incorrect answer and reason, with Certainty of Response Index > 2.5.
5) Students' logical thinking levels measured by scoring the answers of Test of Logical Thinking (TOLT) which developed and validated by Tobin & Capie

(1981). There are 10 questions in the test including two proportional questions, two controlling variable questions, two correlational questions, and two combinatorial questions.

3.5 Research Instrument

There are several instruments have been used in this research in order to collect the data. The instruments were used for detecting students' understanding categories, students' misconception, and students' logical thinking. Those instruments had been given to the research participants as in the Table 3.1

Table 3. 1
Research Instruments

Target	Instrument
Students' understanding categories and misconceptions in global warming topic	Two-tier test and Certainty of Response Index (CRI)
Students' logical thinking levels	Test of Logical Thinking

3.5.1 Two-Tier Test

In this research, the two-tier test consists of first and second tier which contains of the statements related to the global warming topic. The two-tier test is formed in multiple choice questions. For first tier, it consists of one correct statement and two incorrect statements about global warming concept. While the second tier consists of the reasons for the first tier. One of the option is the correct answer, two possible reasons, and the last option is opened response. The last option in second tier is used for students to write down their alternative reason for their answer of the first tier. If students answer both tier correctly, they will get one point. For students who answer one tier or no tiers correctly tier will get zero point. Through this diagnostic test, students have to relate the correct concept at the first tier with the correct reason for the answer at the second tier. If students can answer both statement correctly, it means they understand the concept well.

3.5.1.1 The Development of Two-Tier Test

The concepts of global warming in the two-tier test were developed based on the 2013 curriculum, 2017 version. The questions were developed based on those concepts and finding of the previous research about this topic. Due to the

focus of this research is not to develop the diagnostic test itself, some questions in two-tier test were used from previous research which have been developed and validated by Arslan, Cigdemoglu, & Moseley (2012) in the Atmosphere-Related Environmental Problem Diagnostic Test (AREPDiT). Besides, the other questions had been constructed to cover the concept in global warming for this research. Therefore the instrument was developed and followed some guidance steps for constructing the two-tier test by Treagust (1988). The development of instrument has been followed steps as follows.

1) Defining the content for the test item.

The first step for constructing the test item was to define the global warming concepts included in the test. In this step, the global warming concept identified based on the indicators in seventh grade science curriculum and also from the previous research.

2) Identifying the information of students' misconception in global warming

In this step, some misconceptions on global warming topic had been identified based on the previous studies. For this step, researcher did not directly obtain initial possible misconceptions to students as the Treagust's development step but researcher obtained the information from previous researchers' findings. This step was taken for reduce the time consuming.

3) Composing the draft of the test item.

After the correct concepts and misconceptions collected, researcher had constructed into multiple choice questions. The first tier contained one correct option and two incorrect options. And for the second tier, it contained one correct reason for the first tier, two distractor options, and one open response.

4) Instrument justification.

The next step is expert judgment. For correcting the test item, two natural science lecturers and one earth and space lecturer involved in this step. From the expert judgment, the test item had be revised by the researcher.

5) Try out the test item.

Validating instrument were conducted to students whom were not the research samples after revising the test items. The item tests had been administered to the

students, then it were analyzed by several aspects, which are validity, reliability, discriminating power, and difficulty level.

a. Validity

Kaplan (2004) state that validity can be defined as the agreement between a test score or measure and the quality it is believed to measure. To measure the validity of the test item, it has been examined the Biserial correlation coefficients helped by ANATES v4 software. The interpretation of r_{xy} will be classified into several categories based on Ebel quoted by Crocker & Algina (1986) in Table 3.2.

Table 3. 2
Classification of Validity Coefficient

r_{xy}	Validity interpretation
$0.4 > r_{xy} \leq 1.0$	High
$0.3 > r_{xy} \leq 0.4$	Satisfactory
$0.2 > r_{xy} \leq 0.3$	Low (revised)
$0.00 > r_{xy} \leq 0.2$	invalid

(Source: Ebel in Crocker & Algina, 1986)

b. Reliability

Reliability is the consistency of the test item, it explains consistent the scores are for each individual from one administration of an instrument to another and from one set of test item to another (Fraenkel, Wallen, & Hyun, 2012). For determining the reliability, it used r_{11} and helped by using ANATES v4 software. And the interpretation of the reliability is in Table 3.3.

Table 3. 3
Classification of Reliability Coefficient

Value r_{11}	Validity interpretation
$0.8 > r_{11} \leq 1.00$	Very high
$0.6 > r_{11} \leq 0.8$	High
$0.4 > r_{11} \leq 0.6$	Satisfactory
$0.2 > r_{11} \leq 0.4$	Low
$r_{11} < 0.2$	Very low

(Source: Kirbulut & Geban, 2014)

c. Discriminating Power

Discriminating power is to differ between respondents with a higher score level and a lower score level of respondents (McMillan, 2008). To determine discriminating power, the researcher uses the formula from Yuan, Deng, Zhu, & Li, (2013) with the interpretation in Table 3.4.

$$Pi = \frac{Ai}{Ni}$$

Description:

P_i = difficulty index of item *i*
A_i = Average score to item *i*
N_i = Full score of item *i*

Table 3. 4
Discriminating Power Classification

Value of DP	Discriminating Power Interpretation
DP ≤ 0.00	Very poor
0.00 < DP ≤ 0.20	Poor
0.20 < DP ≤ 0.40	Fair
0.40 < DP ≤ 0.70	Good
0.70 < DP ≤ 1.00	Very good

(Source: Yuan, Deng, Zhu, & Li, 2013)

d. Difficulty Level

Difficulty level is the level of the respondents who have the correct answer (Yuan, Deng, Zhu, & Li, 2013). Difficulty level is determined by the formula from Gronlund (1993) and interpreted based on Table 3.5.

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

Where :
P = the value of students who answered items correctly
R = the number of students who answered items correctly
N = total number of students who conduct the exam

Table 3. 5
Coefficient Classification of Difficulty Level

Value of P	Interpretation
P = 0.00	Very difficult

$0.00 < P \leq 0.30$	Difficult
$0.30 < P \leq 0.70$	Medium
$0.70 < P < 1.00$	Easy
$P = 1.00$	Very easy

(Source: Gronlund, 1993)

3.5.1.2 The Result of Two-Tier Test Development

Firstly, two-tier test item was tested to 31 seventh grade students from one school. The school is different from three schools for collecting the data but in the same region with those three schools. Biserial correlation coefficients, the reliability, discriminating power, and difficulty level were used for item analysis. The result shows this two-tier test instrument has reliability on 0.76, it means the reliability of the instrument is high.

At first, there are 22 questions in the first version of test items. Due to the points of biserial correlation coefficient are very low, 7 questions were eliminated (see Table 3.6) from the test. As the validation result, the blueprint of the item test have been made as follows Table 3.7.

Table 3. 6
Recapitulation of Two-Tier Test Validation

Previous Number	Current Number	DP (%)	DL	Correlation	Significance Correlation	Status
1	1	37.50	**	0.427	Significant	Accepted
2	-	12.50	**	0.089	-	Rejected
3	14	87.50	**	0.673	Very Significant	Accepted
4	-	12.50	***	0.058	-	Rejected
5	10	37.50	**	0.330	-	Revised
6	4	37.50	*	0.364	-	Revised
7	5	37.50	**	0.394	Significant	Accepted
8	-	12.50	**	0.094	-	Rejected
9	11	50.00	*	0.444	Significant	Accepted
10	3	75.00	**	0.530	Very Significant	Accepted
11	-	25.00	**	0.248	-	Rejected
12	8	62.50	**	0.452	Significant	Accepted
13	-	-1.00	***	-0.169	-	Rejected
14	9	75.00	**	0.573	Very Significant	Accepted
15	13	87.50	**	0.595	Very Significant	Accepted
16	-	-2.00	**	-0.084	-	Rejected
17	6	75.00	**	0.655	Very Significant	Accepted
18	-	-1.00	*	-0.069	-	Rejected
19	12	37.50	**	0.413	Significant	Accepted
20	7	62.50	**	0.685	Very Significant	Accepted
21	15	62.50	**	0.469	Significant	Accepted
22	2	62.50	**	0.481	Significant	Accepted

Notes:

DP : Discriminating Power
DL : Difficulty Level (*Easy, **Medium, ***Difficult)

Table 3. 7
The Blueprint of Two-Tier Test

Basic Competence	Main Topic	Indicator	Question Number
3.9 *Menganalisis perubahan iklim dan dampaknya bagi ekosistem.* (Analyzing the climate change and its impacts to ecosystem)	Greenhouse effect	3.9.1 Students are able to describe definition of greenhouse effect.	2,8
	Greenhouse gases	3.9.2 Students are able to explain the greenhouse gases.	3,9,13
	Global warming definition	3.9.3 Students are able to describe the definition of global warming.	1.7
	Cause of global warming	3.9.4 Students are able to describe the cause of global warming.	4,10,14
	Global warming impact	3.9.5 Students are able to describe the impact of global warming.	5,11
	Overcoming global warming	3.9.6 Students are able to describe some efforts to overcome global warming	6,12,15
Total Item Test			15

3.1.1 Certainty of Response Index (CRI)

The answer sheet of the two-tier test is completed by Certainty of Response Index (CRI). The level of students' confidence in answering questions will detect the misconception that might be held by the students. Following the previous research, there are 6-point scale of CRI. The scales are 0 for total guess the answer; 1 almost guess; 2 not sure, 3 sure, 4 certain, 5 totally certain. Students will give the value of CRI for their answers of the questions.

In measuring the score of the CRI, there is an indicator for classifying the students' answers. According to Hasan, Bagayoko & Kelley (1999), the value of CRI below and above 2.5 determines low and high certainty. Students who give low

scale (CRI of 0-2) to their answer identified as the students probably conduct the guesswork for answering the questions. A low CRI scale also indicates students' lack of confidence. In the other hand, students who give high CRI score (CRI of 3-5), then the students have a high degree of confidence of answering the questions. In this situation, students who answer the question correctly indicates they comprehend the concept well. But, students who do not answer the questions correctly and give high CRI determined students have misconception.

3.1.2 Test of Logical Thinking

A set of test about formal reasoning in order to detect students' logical thinking levels in seventh grade. The instrument adapted from the Test of Logical Thinking (TOLT) according to Piaget's theory and developed by Tobin and Capie (1981). The reliability of test was found 0.81. The use of this instrument is to help the researcher to get information about students' logical thinking levels at age 13-14.

This test of logical thinking has ten questions consists of eight multiple choice with reasoning options and two essay questions. Multiple choice questions are used to detect four modes of formal reasoning which are proportional reasoning, controlling variables, probabilistic reasoning, and correlational reasoning. While the essay questions are used to detect combinatorial reasoning. The blueprint of the test is included in Table 3.8.

Table 3. 8
The Blueprint of Test of Logical Thinking (TOLT)

Modes of Formal Reasoning	Question Number
Proportional Reasoning	1, 2
Controlling Variable	3, 4
Probabilistic Reasoning	5, 6
Correlational Reasoning	7, 8
Combinatorial Reasoning	9, 10

For multiple choice scoring, students will get one point as they can answer the both tiers correctly. In the other hand, students who cannot answer both tiers correctly, they will get zero. For the essay questions, students have to answer the

questions briefly and completely. When they can answer the question completely, they will get one point while the ones who do not answer completely, they will get zero. Students who get score 0-1 will be categorized in concrete thinking level and students who get score 2-3 will be categorized in transitional level. Therefore, students who get score 4-10 will be categorized in formal thinking level (Tobin & Capie, 1981).

3.6 Research Procedure

There are several procedures to conduct this research. Those procedure steps are classified into three stages. The stages in this research are preparation stage, implementation stage, and completion stage. Research plot is also made by the researcher to simplify the research procedure to be explained (Figure 3.1).

3.1.3 Preparation Stage

In this stage, there are some steps which are:

1) After the research problem had been formulated, then continue with reviewing and analyzing some relevant research with this research.
2) Constructing a concept map about global warming topic adapted from some journal and students' book reference.
3) Identifying the students' misconception on global warming topic in previous research.
4) Constructing the research instruments. For some instruments, it was adapted from previous research.
5) Conducting the expert judgment and revise the instrument.
6) Validating the instrument to students and revising it.
7) Determining the schools as the place to conduct the research.
8) Determining the research samples and arranging the test date.

3.1.4 Implementation Stage

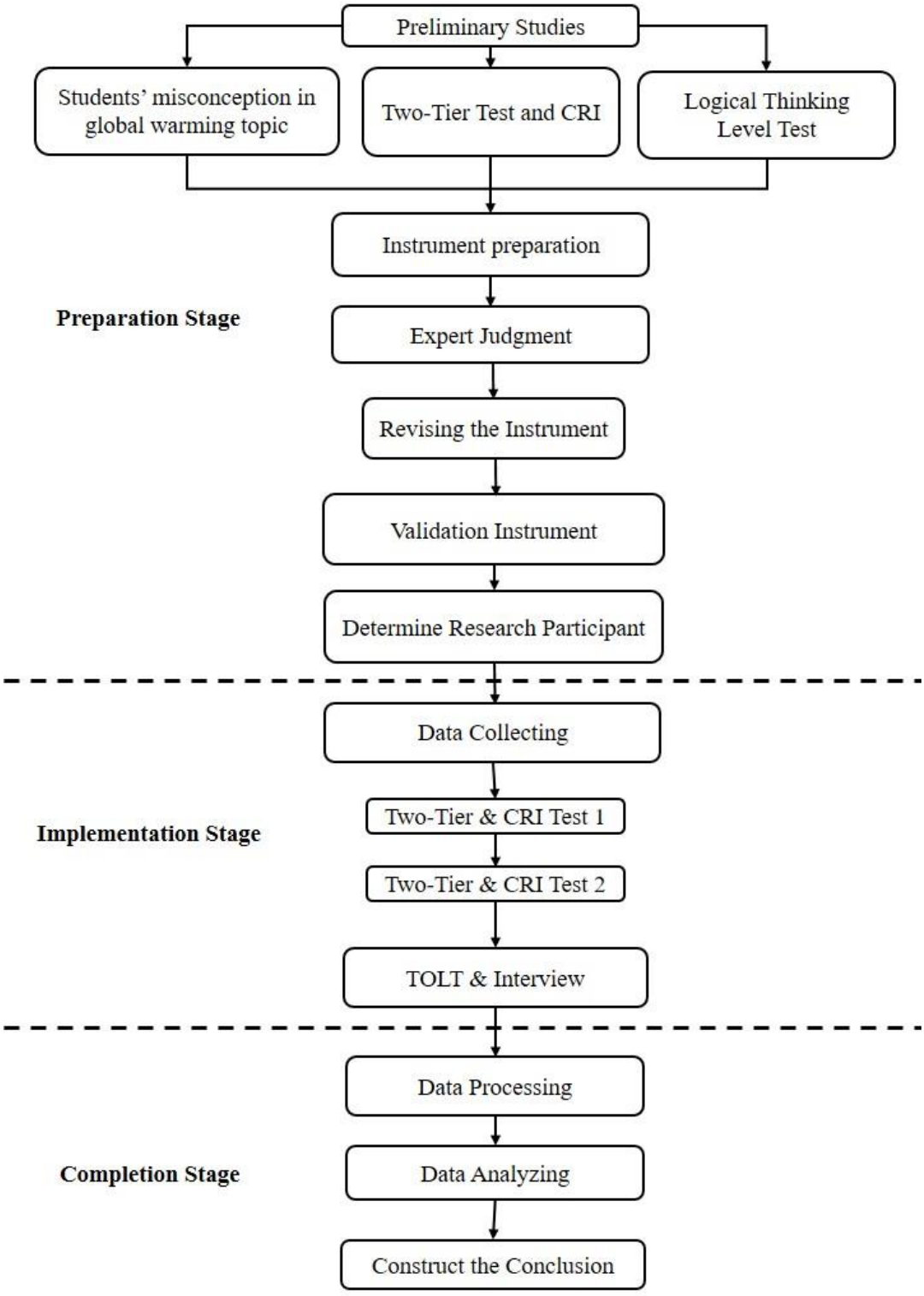
In this stage, there are some steps which are:

1) Conducting the first and second test.
2) Identifying and analyzing the misconception held by student.
3) Conducting TOLT and short interview to students for additional data.

3.1.5 Completion Stage

In this stage, there are some steps which are:

1) Analyzing the data collection.
2) Constructing the research conclusion.
3) Completing the research paper and reporting it.



4) Figure 3. 1 Research Plot

3.7 Data Processing

Data obtained from two-tier test and Certainty of Response Index (CRI) for detecting students' misconception, while the additional data obtained from Test of Logical Thinking and interview. The detail description of data processing technique is explained as follows:

3.1.6 Two-Tier Test and CRI Data Processing

The two-tier test is for analyzing students' answer for the questions. The multiple choice in the first tier has 3 options which are A, B, and C while the second tier has 3 closed responses (1, 2, and 3) and one opened response in option 4. From the students' answer, it shows 12 answer patterns for the data analysis for those who answer the opened response (Table 3.9).

Table 3. 9
The Students' Answer Pattern in Two-Tier Test

First Tier	Second Tier			
	1	2	3	4
A	A1	A2	A3	A4
B	B1	B2	B3	B4
C	C1	C2	C3	C4

After the data of the two-tier answer pattern had been collected, students' answers were categorized into Table 3.10.

Table 3. 10
Two-Tier Test Response Categories

Topic	Item Test	Correct Responses (%)			
		Both Tiers	Only First Tier	Only Second Tier	Both Incorrect

The data from CRI also are used for analyzing students' understanding categories. The analysis of the two-tier test and CRI will determine categories of students' understanding about the global warming concept. The description of the answers stated in Table 3.11

Table 3. 11
Description of Students' Response

Answer (1st Tier)	Reason (2nd Tier)	Value of CRI	Description
Correct	Correct	> 2.5	Understand the concept well
Correct	Correct	< 2.5	Lucky guess, lack of confidence
Correct	Incorrect	> 2.5	Misconception
Correct	Incorrect	< 2.5	Lack of knowledge
Incorrect	Correct	> 2.5	Misconception
Incorrect	Correct	< 2.5	Lack of knowledge
Incorrect	Incorrect	> 2.5	Misconception
Incorrect	Incorrect	< 2.5	Do not know the concept

(Adapted from Arslan, Cigdemoglu, & Moseley, 2012)

Students were classified into categorizes in Table 3.8 after data have been processed. The result of the categories was calculated using a simple formula to determine the percentage of each categories. The calculation is:

$$P = \frac{s}{N} x\ 100\%$$

Description:
P : Percentage of each category
s : Number of the students for each group
N : Total number of students

3.1.7 Test of Logical Thinking (TOLT)

TOLT is used to analyze the students' logical thinking level. The data are used for categorizing students into concrete thinking, transition, formal thinking (Table 3.12).

Table 3. 12
Logical Thinking Levels

Total Correct Answer	Logical Thinking Level
0-1	Concrete Thinking
2-3	Transitional
4-10	Formal Thinking

(Source: Tobin & Capie, 1981)

The result of the categories is calculated using a simple formula to determine the percentage of each categories. The formula is:

$$P = \frac{s}{N} x\ 100\%$$

Description:
P : Percentage of each category
s : Number of the students for each group
N : Total number of students