CHAPTER III
RESEARCH METHOD

3.1 Research Design

In this research, the research design is a Cross-sectional survey design. Surveys are a quantitative research method in which a researcher surveys a target sample and population to determine a population's attitudes, opinions, behaviors, or characteristics. In a cross-sectional survey design, the researcher measures data at one point. For instance, when middle school children respond to questionnaires about teasing, they provide information about their impressions at that moment. This design carries an obvious advantage: measuring present attitudes or practices. It also carries the added advantage of ascertaining information within a relatively brief time, including the time it takes to administer the survey and gather the information. (Creswell, 2012) Researchers collected data using questionnaires to determine the causes of students experiencing misconceptions (Ishtiaq, 2019).

This study used a quantitative method to investigate students' misconceptions about science, especially biology. Quantitative methods emphasize statistical, mathematical, or numerical measurement and analysis of data obtained from surveys, polls, or questionnaires or computational manipulation of previously collected statistical data. This research aims to identify student misconceptions and create a practical diagnostic test to analyze the test's validity, reliability, and difficulty level (Tumanggor et al., 2020).

This method is used to collect and test the first round of quantitative data, followed by the collection and examination of the second round of qualitative data, which is developed based on the quantitative results of the first round. As a result, questionnaires and open-ended questions are more suitable for diagnosing misconceptions in small groups. Conventional multiple-choice questions and their derivatives (two-, three-, and four-tier diagnostic tests) are more suitable for large groups. Researcher data can be categorized, ranked, or measured using units of measurement. Quantitative research

allows the design of graphs and tables to facilitate the analysis of raw data (Ahmad et al., 2019).

3.2 Participant

The respondents of this research are the students from grade 8th. The location of this research is a junior high school; for the preliminary study, a sample of 35 respondents was taken on SMP A. For the pilot test, data on SMP B Bandung and SMP C will be taken; the total number of respondents is 120. For the actual research, take data on SMP D; the total number of respondents is 131. All the schools above used the *Merdeka* curriculum. The data was held in the first and second semesters of the 2023/2024 academic year.

3.3 Research Instrument

A four-tier diagnostic test instrument was developed to investigate the digestive system concept and meet the research objectives. The diagnostic test instruments, expert assessment by the test instrument, testing the instrument on several non-sample students, converting the diagnostic test instrument into digital form and publishing it online, expert media review, data collection, data processing and analysis, and concluding are the first steps in the research process. The four-tier diagnostic test instrument is the best for diagnostic application in large groups because the assessment process is relatively more efficient and objective (Jumadi et al., 2023).

This study collected data using a four-tier diagnostic test conducted via a Google form. The focus of this research. Held by students of various ages from 12-14 regarding certain concepts and their refusal to accept new information. The instruments created were distributed to the research sample. Students work on questions according to their talents. A total of four levels of multiple choice are presented with 10 questions. The questions cover three concepts: micronutrients, minerals and water, absorption, and elimination. The distribution of questions and the concept of these questions are described in Table 3.1.

Table 3.1 Question Distribution

Topic	Sub Topic	Test Item
Digestive System	Micronutrient	1, 2, 3, 4, 5
	Mineral and Zinc	6, 7
	Absorption and Disposal	8, 9, 10

The question itself consists of four tiers. The first tier assesses students' knowledge of concepts by asking straightforward questions. This helps identify whether students have a basic knowledge of the topic. The second tier is the trust rating; after answering the content questions, students rate their confidence in their answers, usually from “sure” and “not sure.” The third tier lies in reasoning questions. Students explain the reasoning behind their answers, providing insight into their thought processes and potential misunderstandings. The fourth tier is confidence in reasoning. Students rated their confidence in the reasons they provided, similar to confidence ratings at the second level. An example of the question is shown in Table 3.2.

Table 3.2 Sample Question

No	Tier	Question
	1st	What happens when the body experiences problems with the digestive system, affecting its ability to process food?
		A. The digestive process becomes faster
		B. The feces become harder to pass
		C. Hydration of the skin decreases
		D. Nutrients from food cannot be absorbed properly
7	2nd	State the confidence level of your answer

No	Tier	Question
		A. Sure
		B, Not sure
	3rd	Which of the following options is the reason for your answer?
		A. The body has a lack of water to hydrate
		B. The body digests too quickly, causing diarrhea
		C. The body has difficulty dissolving nutrients
		D. The body has constipation, making defecation difficult
	4th	State the confidence level of your answer
		A. Sure
		B. Not sure
	Scientific Concept	Nutrient absorption is disrupted when the body experiences digestive system problems, such as malabsorption disease or other digestive dysfunction. This means that food's essential vitamins, minerals, and other nutrients must be adequately absorbed into the bloodstream. If the body cannot absorb these nutrients adequately, it can cause various health problems.

The question shown in Table 3.2 is given to the student to determine their understanding of the function of minerals and water during the role of water in the digestive system. In this question, students are given four options to analyze which option is appropriate for answering the causing effect when our digest lacks water. The digestive system's ability to break down and absorb nutrients is impaired when the body experiences digestive system difficulties, such as inflammation, infection, or disease. This

can be caused by damage to the intestinal lining, which limits the surface area accessible for nutrient absorption, or by decreased synthesis of digestive enzymes, which prevents the breakdown of food into absorbable molecules. As a result, essential nutrients, vitamins, and minerals are not absorbed efficiently into the bloodstream, resulting in malnutrition and other health problems. After that, students are asked to give the reason for answering the answer that supports their option. Then, they have to give their confidence rating for their answer, whether it is sure or not sure, for their first and third-tier question.

3.3.1 Preliminary Test

A preliminary or feasibility study is the first examination that evaluates a research project's feasibility, timing, costs, risks, and adverse events. It is a replica of a more extensive study intended to test the methods, procedures, and instruments used in it. The main purpose of this research, based on a preliminary study, is to analyze students' understanding of the digestive system by identifying misconceptions and potential improvements in teaching techniques and models to offer a good learning environment for students (Marviyani et al., 2023). This topic was used in the form of open-ended questions with two levels. The first level consists of questions intended to assess students' understanding of the material, and the second level is the reasoning behind the answers to the questions. The open-ended question of the preliminary study is shown in Table 3.4.

Table 3.3 The Distribution of Preliminary Study Question

Item	Part	Question
1	**1**	Our bodies need carbohydrates, proteins, fats, minerals, and vitamins. Simple carbohydrates are commonly called “sugar.” The type of sugar needed by our body cells is...
	2	Give a brief explanation of the answer you have given...
2	**1**	What causes the protein to function as a nutrient?

Item	Part	Question
	2	Give a brief explanation of the answer you have given...
3	**1**	Fat often has a bad connotation, but is there any good fat for our body?
	2	Give a brief explanation of the answer you have given...
4	**1**	Vitamins are needed to maintain the function of your body systems. They are also commonly found in foods. How do vitamins naturally occur in our bodies?
	2	Give a brief explanation of the answer you have given...
5	**1**	Vitamins and minerals do not produce energy. Why does our body still need them?
	2	Give a brief explanation of the answer you have given...
6	**1**	All the important activities we do require water in our bodies. Why is water so important for our bodies?
	2	Give a brief explanation of the answer you have given...
7	**1**	Candies, sodas, and other sweets are foods that we often encounter in our daily lives. What are the side effects when we consume excessive amounts of sugar-containing foods?
	2	Give a brief explanation of the answer you have given...
8	**1**	Why do we need to use preservatives to preserve food?
	2	Give a brief explanation of the answer you have given...

Item	Part	Question
9	**1**	The processed foods we eat usually contain additives such as thickeners. What do you think are the main benefits of using these additives in our food?
	2	Give a brief explanation of the answer you have given...
10	**1**	Digestion is the process by which the food you eat is broken down into smaller nutrients. List the steps in the digestion process in your body!
	2	Give a brief explanation of the answer you have given...
11	**1**	Chemical digestion occurs when the salivary glands in the mouth produce an enzyme called amylase. What is the role of amylase in chemical digestion?
	2	Give a brief explanation of the answer you have given...
12	**1**	Absorption occurs after digestion. This absorption process is the longest chemical process, why does it happen?
	2	Give a brief explanation of the answer you have given...
13	**1**	After chemical digestion is complete, the nutrient molecules are ready to be absorbed by the body. Briefly, how does the absorption of nutrient molecules occur?
	2	Give a brief explanation of the answer you have given...
14	**1**	In the large intestine, water is absorbed while food debris is ready to be excreted. Why is the function of water absorption by the large intestine important?

Item	Part	Question
	2	Give a brief explanation of the answer you have given...

Before the instrument is developed, the first thing that has to be done is the preliminary test for the students who have already learned this topic, which is given 14 questions to 35 students of grade 9^{th} who have already learned it in their previous class. Using their prior knowledge to answer the open-ended question in Google Form (see Table 3.4) to determine their answer and make the optional multiple-choice four-tier test. After the answers are formed to a multiple choice, the previous questions are developed, so one open-ended question is formed to be two multiple-choice questions in two questions. They have the same option but different questions to make the option possible for answering the first and third tier of two questions. In conclusion, student response results were evaluated using Microsoft Excel and SPSS to determine the validity of the test questions. These questions will be separated into valid and invalid test items. Test questions that are classified as invalid will be excluded from this study. The preliminary test question had 14 questions and was then developed into 28 questions (see Appendix A.2).

3.3.2 Validation Test

Validity describes how effectively the data obtained covers the actual topic of investigation; it means "measuring what you want to measure"**.** Validity refers to the accuracy of measuring ideas in quantitative studies. Validity refers to the accuracy and precision of a measuring instrument. Content validity refers to the correspondence between what is to be measured and the existing instrument. Content validity describes procedures that establish the extent to which a battery of tests is relevant and reflective of each domain in which test scores are to be interpreted. This evaluates the instrument's coverage of relevant variable material. In other words, the instrument measures all domains related to that variable or concept (Heale et al., 2015). The formula for calculating content validity is the Aiken V coefficient, which is determined by giving a

number from 0 to 1 (not representative to very representative) (Rahmawati & Harun, 2019b). The following is the Aiken V formula:

$$V = \frac{\Sigma s}{\mathrm{n(c-1)}}$$

V = Rater's Fit Index
s = Average score – the lowest score in the category
c = Number of Categories
n = Number of Raters

The validation results of the Aiken Formula were used to analyze each item from three experts and determine the content validity of a sophisticated multiple-choice instrument for assessing students' critical thinking abilities. The Aiken formula calculates the content validity coefficient depending on how many people validate an item and how well the item represents the measured indicator (Arthur, 2024).

Table 3.4 Criteria for Validity using the Aiken Index

Index Validity (V)	Interpretation
V > 0.8	High Validity
0.4 < V ≤ 0.8	Moderate Validity
V ≤ 0.4	Invalid

The item score is considered valid if the analysis results show that each question given has a value greater than Table V, which ranges from 0.4 to 1 according to the number of validators. If the data can be considered valid, then it can be used. Meanwhile, the object is thrown away and unused if it is invalid. Therefore, it is crucial to re-study and even recreate the items in each evaluation component and carry out in-depth analysis and validation of the material (Hidayah et al., 2023).

Table 3.5 Aiken Test for Tier 1 Validation Result

Item	**Expert**			**S_1**	**S_2**	**S_3**	**∑s**	**n(c-1)**	**V**	**Validity**
	I	II	III							
1	1	1	0	1	1	0	2	3	0.67	Moderate
2	1	0	0	1	0	0	1	3	0.33	Low
3	1	1	1	1	1	1	3	3	1.00	High
4	1	1	1	1	1	1	3	3	1.00	High

Item	Expert			S_1	S_2	S_3	$\sum s$	n(c-1)	V	Validity
	I	II	III							
5	1	0	1	1	0	1	2	3	0.67	Moderate
6	1	1	1	1	1	1	3	3	1.00	High
7	1	1	1	1	1	1	3	3	1.00	High
8	1	1	1	1	1	1	3	3	1.00	High
9	1	1	1	1	1	1	3	3	1.00	High
10	1	1	1	1	1	1	3	3	1.00	High
11	1	1	1	1	1	1	3	3	1.00	High
12	1	0	1	1	0	1	2	3	0.67	Moderate
13	1	0	1	1	0	1	2	3	0.67	Moderate
14	1	1	1	1	1	1	3	3	1.00	High
15	1	1	1	1	1	1	3	3	1.00	High
16	1	0	1	1	0	1	2	3	0.67	Moderate
17	1	1	1	1	1	1	3	3	1.00	High
18	1	1	1	1	1	1	3	3	1.00	High
19	1	0	1	1	0	1	2	3	0.67	Moderate
20	1	1	1	1	1	1	3	3	1.00	High
21	1	0	1	1	0	1	2	3	0.67	Moderate
22	1	0	1	1	0	1	2	3	0.67	Moderate
23	1	0	1	1	0	1	2	3	0.67	Moderate
24	1	0	1	1	0	1	2	3	0.67	Moderate
25	1	1	1	1	1	1	3	3	1.00	High
26	1	1	1	1	1	1	3	3	1.00	High
27	1	0	1	1	0	1	2	3	0.67	Moderate
28	1	1	1	1	1	1	3	3	1.00	High

In Table 3.4, as seen above, the V coefficient value for each test item ranges from 0.4 to 1, with an average value of 0.85 for all 28 questions at the first and third levels, indicating good validity. These test questions are suitable for measuring the desired results.

Table 3.6 Aiken Test for Tier 3 Validation Result

Item	Expert			S_1	S_2	S_3	∑s	n(c-1)	V	Validity
	I	II	III							
1	1	1	0	1	1	0	2	3	0.67	Moderate
2	1	0	0	1	0	0	1	3	0.33	Low
3	1	1	1	1	1	1	3	3	1.00	High
4	1	1	1	1	1	1	3	3	1.00	High
5	1	0	1	1	0	1	2	3	0.67	Moderate
6	1	1	1	1	1	1	3	3	1.00	High
7	1	1	1	1	1	1	3	3	1.00	High
8	1	1	1	1	1	1	3	3	1.00	High
9	1	1	1	1	1	1	3	3	1.00	High
10	1	1	1	1	1	1	3	3	1.00	High
11	1	1	1	1	1	1	3	3	1.00	High
12	1	0	1	1	0	1	2	3	0.67	Moderate
13	1	0	1	1	0	1	2	3	0.67	Moderate
14	1	1	1	1	1	1	3	3	1.00	High
15	1	1	1	1	1	1	3	3	1.00	High
16	1	0	1	1	0	1	2	3	0.67	Moderate
17	1	1	1	1	1	1	3	3	1.00	High
18	1	1	1	1	1	1	3	3	1.00	High
19	1	0	1	1	0	1	2	3	0.67	Moderate
20	1	1	1	1	1	1	3	3	1.00	High
21	1	0	1	1	0	1	2	3	0.67	Moderate
22	1	0	1	1	0	1	2	3	0.67	Moderate
23	1	0	1	1	0	1	2	3	0.67	Moderate
24	1	0	1	1	0	1	2	3	0.67	Moderate
25	1	1	1	1	1	1	3	3	1.00	High
26	1	1	1	1	1	1	3	3	1.00	High
27	1	0	1	1	0	1	2	3	0.67	Moderate
28	1	1	1	1	1	1	3	3	1.00	High

In Table 3.5, as seen above, the V coefficient value for each test item ranges from 0.4 to 1, with an average value of 0.8 for all 28 questions at the first and third levels, indicating good validity. These test questions are suitable for measuring the desired results.

Table 3.7 Result of Validity Construct Pilot Test

Question	Tier	Pearson Correlation	N of (5%)	Interpretation	Decision	Question	Tier	Pearson Correlation	N of (5%)	Interpretation	Decision
Q1	1	0,287	0.195	Valid	Accepted	Q2	1	0,180	0.195	Not Valid	Rejected
	3	0,269	0.195	Valid			3	-0,027	0.195	Not Valid	
Q3	1	0,302	0.195	Valid	Accepted	Q4	1	0,032	0.195	Not Valid	Rejected
	3	0,207	0.195	Valid			3	-0,058	0.195	Not Valid	
Q5	1	0,314	0.195	Valid	Accepted	Q6	1	0,219	0.195	Valid	Revised
	3	0,208	0.195	Valid			3	0,135	0.195	Not Valid	
Q7	1	0,197	0.195	Valid	Accepted	Q8	1	0,277	0.195	Valid	Accepted
	3	0,235	0.195	Valid			3	0,578	0.195	Valid	
Q9	1	0,414	0.195	Valid	Accepted	Q10	1	0,116	0.195	Not Valid	Rejected
	3	0,572	0.195	Valid			3	-0,134	0.195	Not Valid	

Question	Tier	Pearson Correlation	N of (5%)	Interpretation	Decision	Question	Tier	Pearson Correlation	N of (5%)	Interpretation	Decision
Q11	1	0,174	0.195	Not Valid	Rejected	Q12	1	0,414	0.195	Valid	Accepted
	3	0,114	0.195	Not Valid			3	0,582	0.195	Valid	
Q13	1	0,214	0.195	Valid	Revised	Q14	1	0,102	0.195	Not Valid	Rejected
	3	-0,013	0.195	Not Valid			3	-0,039	0.195	Not Valid	
Q15	1	0,473	0.195	Valid	Accepted	Q16	1	0,433	0.195	Valid	Revised
	3	0,248	0.195	Valid			3	0,127	0.195	Not Valid	
Q17	1	0,069	0.195	Not Valid	Rejected	Q18	1	0,271	0.195	Valid	Accepted
	3	0,162	0.195	Not Valid			3	0,386	0.195	Valid	
Q19	1	0,114	0.195	Not Valid	Rejected	Q20	1	0,109	0.195	Not Valid	Revised
	3	0,109	0.195	Not Valid			3	0,243	0.195	Valid	
Q21	1	0,138	0.195	Not Valid		Q22	1	0,212	0.195	Valid	

Question	Tier	Pearson Correlation	N of (5%)	Interpretation	Decision	Question	Tier	Pearson Correlation	N of (5%)	Interpretation	Decision
	3	-0,010	0.195	Not Valid	Rejected		3	0,005	0.195	Not Valid	Revised
Q23	1	0,199	0.195	Valid	Accepted	Q24	1	0,097	0.195	Not Valid	Not Valid
	3	0,201	0.195	Valid			3	0,162	0.195	Not Valid	
Q25	1	0,230	0.195	Valid	Revised	Q26	1	0,009	0.195	Not Valid	Rejected
	3	-0,030	0.195	Not Valid			3	-0,184	0.195	Not Valid	
VQ27	1	0,123	0.195	Not Valid	Rejected	Q28	1	0,195	0.195	Valid	Rejected
	3	0,016	0.195	Not Valid			3	0,100	0.195	Not Valid	

3.3.3 Reliability Test

Reliability refers to the extent to which measuring a phenomenon produces reliable and consistent results. It is also related to repeatability. For example, a scale or test is reliable if repeated measurements under constant conditions produce the same results (Taherdoost, 2016).

In this research, reliability estimation was carried out using the Product Moment Correlation method using SPSS. Data from the validation of instrument items were

evaluated by calculating the reliability test $h \geq r\ t$ using the Alpha Cronbach formula assisted by SPSS. Specifically, if the Alpha value > r table, the instrument meets the reliability test (Hidayah et al., 2023).

3.4 Data Analysis

The variables option selected for the answer tier, confidence selected for the answer tier, option selected for the reason tier, and confidence selected for the reason tier were ticked by each student (Zhao et al., 2023). For each question, there will be categories for Scientific Knowledge (SK), False Positive (FP), False Negative (FN), Misconception (M), and Lack of Knowledge (LK). Scientific knowledge is an understanding of concepts, principles, and facts that have been established and validated using scientific procedures. This includes not only knowing factual information but also understanding the underlying principles, applying them to various situations, and reasoning rationally. False positives occur when students successfully answer a question or appear to understand a concept, but this apparent understanding is not genuine or based on sound reasoning. This could be the result of guesswork, luck, or superficial knowledge. False negatives occur when students give incorrect answers or appear to lack understanding of a concept even though they know it. This could be due to misreading the question, making a small mistake, or a lack of confidence. A misconception is a false belief or understanding of a scientific idea. Misconceptions are usually ingrained and difficult to overcome because they are based on intuitive but faulty reasoning or flawed prior information. Lack of knowledge refers to a state in which a learner lacks the necessary information or understanding of an idea. In contrast to misunderstanding, which occurs when students have inaccurate knowledge, lack of knowledge indicates that students are unaware or uneducated regarding the ideas. A comparison of the decision of the four-tier test is shown in Table 3.5.

Table 3.8 The Decisions of Four-Tier Test (Kiray & Simsek, 2021)

1. Tier	2. Tier	3. Tier	4 . Tier	The decision of the four-tier test
True	Confident	True	Confident	SK

True	Confident	False	Confident	FP
False	Confident	True	Confident	FN
False	Confident	False	Confident	M
True	Confident	False	Not Confident	LK 1
True	Not Confident	False	Confident	LK 2
True	Not Confident	False	Not Confident	LK 3
True	Confident	False	Not Confident	LK 4
True	Not Confident	False	Confident	LK 5
True	Not Confident	False	Not Confident	LK 6
False	Confident	True	Not Confident	LK 7
False	Not Confident	True	Confident	LK 8
False	Not Confident	True	Not Confident	LK 9
False	Confident	True	Not Confident	LK 10
False	Not Confident	True	Confident	LK 11
False	Not Confident	True	Not Confident	LK 12

Table 3.8 *SK* scientific knowledge, *LK* lack of knowledge, *M* misconception, *FP* false positive, *FN* false negative

The data was obtained with Excel by considering the possibilities in Table 3. If the answer were correct, the score would be "1," and if the answer were incorrect, the score would be "0." This also applies to confidence rating. This score measures the first tier and third tier; if the answer is confident, the score is "1," and if the answer is not confident, the score is "0." It measures the second tier and fourth tier.

1. For example, the categories of scientific knowledge in the first and third tiers should be correct, and the second and fourth tiers should also be confident that the total score is "1-1-1-1".
2. It also happens to false positives. The first tier should have correct answers, but the third tier should have incorrect answers. The second and fourth tiers should have confident answers. The total is "1-1-0-1."

3. The first-tier answer should be incorrect for false negatives, but it should be correct for the third-tier. The second and fourth tiers should be confident so that the total score would be calculated as “0-1-1-1”.
4. The first and third tiers can identify misconceptions that should be incorrect answers; hence, the second and fourth tiers should be confident answers, so the total answer is “0-1-0-1”.
5. The last is lack of knowledge; the rest of the answers are not categorized as scientific knowledge, false positive, false negative, or misconception (Kiray & Simsek, 2021).

3.5 Research Procedure

To organize the research, the research procedure is planned based on the stage of the guided-inquiry learning model using web-based inquiry implementation. The procedure has three stages: the preparation stage, the implementation stage, and the completion stage. Those three stages are explained below.

A. Preparation stage

a. Identify the research question and problem
b. Formulate research objectives
c. Analyze the topic of The Digestive System in the science textbook for grade 8 based on Kurikulum Merdeka to develop the question for preliminary study.
d. Formulate 28 questions, two from each concept, by using the open-ended questions for testing the preliminary study.
e. Validation of research instruments by experts.
f. Revise preliminary instruments.
g. Choose one question for each concept, eliminating 14 questions, and test the 14 questions that have already been accepted by the expert.
h. Distribute the preliminary study by testing the 14 open-ended questions on grade 9^{th} students.
i. Collect the results from the primary study
j. Distribute the questions and make the branches from one question to two questions, and the option for multiple choice based on the students' responses.

k. Validate and revise the question to the supervisor.
l. Validate the question to 3 expert judgments, which are 2 biology lecturers and 1 biology teacher.
m. Distribute the result of instrument judgment to Excel for the Aiken Test and revise the question.
n. Conduct the validity test for 30 students in grade 8.
o. Analyze the result of the validity test by using SPSS.
p. Revise the question to eliminate the sort of question that is valid by SPSS, check the correlation and reliability test of the validation, and take the 10 valid questions for the real research.

B. Implementation stage

a. Contact the school and provide the permission letter to take the data.
b. Make an appointment with the vice principal to conduct the real research.
c. Coordinate with the homeroom teacher to spread the Google form link to the grade 8^{th} students.
d. Take the documentation while the students fill out the questionnaire.
e. Interview with the science teacher and 3 students who have a large of misconceptions.

C. Completion stage

a. Analyze the statistical data that has already been obtained in the real research by using SPSS.
b. Validate and discuss the result with the supervisor.
c. Make a conclusion, implication, and recommendation based on data analysis.
d. Create the report on the research paper thesis.
e. Put all the data evidence in the Appendix.
f. Upload the journal based on the research paper thesis to Sinta 3. The sequence of the research procedure is shown in Figure 3.1.

Figure 3.1 Flowchart of Research Procedure