CHAPTER III
RESEARCH METHODOLOGY

3.1 Research Method

The research method that is used in this research is the developmental research method. This method is a research method used to produce new products and test the effectiveness of these products (Sugiyono. 2016). The developmental research method is also a method that is carried out deliberately, systematically, to find, improve, develop, produce, test the effectiveness of a product, model or strategy so that the product can be validated testing (Sugiyono, 2016). The method in this research consists of constructing the application, determining the material, determining validity and reliability, normality, homogeneity, and in the end, there is an evaluation (Gravemeijer, 1998).

This method is suitable for making "Captain Lungs 3D" applications in the topic of the respiratory system, because this research focuses on the development stage, followed by pre-test post-test analysis, rather than analyzing data.

3.2 Research Design

The development method in this study uses the Research and Development method. This method is called the development of a developmental design model, to produce a product, then the product is made and refined so that it can support learning using that model. The development model is depicted in Figure 3.1
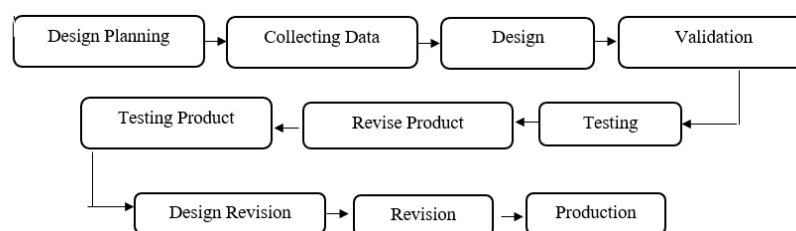

Figure 3.1 Developmental method

(Source : Sugiyono, 2017)

The development method based on Figure 3.1 consists of the planning stage of problem identification, creating a new design, design validation, and product testing. In developmental research method consists of a literature review stage regarding the human respiratory system, then critical thinking skills, making 3D games. At the stage of development, the instrument is developed into an application, evaluating its validity, reliability, normality and homogeneity to analyze data based on pre-test and post-test in the evaluation stage.

3.3 Population and Sample

The subjects in this study were students and also experts. The composition of these experts is divided into three types, including experts on science topics, experts on media, and teachers in computer and science topics. Experts selected in science topics are experts based on recommendations, for example, have achieved a doctoral degree. The second is an expert on media, and an expert from a school teacher., Expert, selected in the teacher is selected from a teacher who has experience in the field of science and computers.

A total of 60 students were selected as the sample in this study, consisting of 33 female students and 27 male students in grade 9 of junior high school. Due to the conditions when doing this research coincided with a pandemic, the research was conducted online via google classroom and using respective smartphones in order to run this game smoothly. For student selection, convenience sampling was used. Students can get the opportunity to be selected. Before doing the research, the response from the expert was first carried out so that the data collection process could take place

3.4 Operational Definition

In order to avoid misconceptions about this research, operational definitions are explained in this research. Those terminologies are explained as follow:

1) Game-based learning using Unity 3D application as a tool to create learning media and can be export into apk files. The smartphone used in this research as a tool for students to learn about biology material given.

2) Game-based learning was created based on the respiratory system concept of respiratory system organs, the mechanism of breathing, and respiratory system diseases. To measure the quality of the game, tested using a questionnaire contained in the game

3) Critical thinking is a method of solving problems and also enhance students' capability in critical thinking. Student's critical thinking skills were judge based on the student's answer in the twenty eight-question related to respiratory system topic. The questions used in the game are adapted from the Facione principle, which consists of interpretation, analysis, evaluation, inference, explanation, and self-regulation.

3.5 Research Instrument

The purpose of the objective test is to measure student learning outcomes regarding the critical thinking aspects of the questions to be tested. The right instrument must meet two conditions, validity and reliability (Sugiyono, 2013). In the objective test, which consists of 28 questions, it is divided into five indicators, which include interpretation, analysis, evaluation, inference, explanation. The instruments used in this study were as follows :

3.5.1 Objective Test

This test is used to measure students' abilities in testing critical thinking skills, which were tested in the Facione (2015) aspect. The test carried out in this study is an objective test in the form of multiple choices, before being tested on students, this test is assessed first by a science expert to assess that this question is suitable for use or must be revised first. The test to be used consists of 28 questions, with the division of 5 questions for interpretation, six questions for analysis, six questions for evaluation, six questions for the explanation, five questions for inference in the form of multiple-choice questions via a google form.

3.5.2 Non Test Instrument

a) **Interview**

At this stage, interviews were conducted with several samples of students to

determine the extent of students' critical thinking skills, while the questions asked of students included understanding and opinions regarding aspects of critical thinking consisting of interpretation, analysis, evaluation, explanation, inference, and regulation. The interview will be conducted online due to the Covid-19 pandemic

b) Expert Judgement of the Rubric

The rubric was used to validate the data in this study. This rubric was adapted from the Multimedia Team at North Carolina State University (Mc Cullen et al., 2015). In the criteria of assessment of the aspects listed already meet the criteria for ready to be assessed by experts, and explanations used in the rubric was very well structured and clear. This rubric was chosen because Media created using game-based learning has indicators, criteria, and aspects of having a scale 1 until 4. Each scale has a different value. Experts can judge which scale suits the media based on their own perspectives. Details about the rubric can be seen in table 3.1 regarding the Expert's Rubric Scale

The result of the scoring of this rubric is the overall average of all scores given by experts. This average is the total of all scores divided by the number of scores **is** depicted in Table 3.1

Table 3.1

Expert Judgement Rubric Scale

No.	Criteria	Assessed Aspects	Experts Assessment	
1	Technical	The game runs perfectly with no technical problems. There are no error messages, all sound, video, or other files are found.	Yes	No
2	Navigation	Users can progress intuitively throughout the entire game in a logical path to find information. All buttons and navigational tools work	Yes	No
3	Spelling and grammar	The game honours all rules of spelling or grammar	Yes	No
4	Completion	Game is completely finished	Yes	No
5	Screen Design	The combination of multimedia elements and content takes communication to a superior level. There is clear attention given to balance, proportion, harmony, and restraint. The synergy reaches the intended audience with style and pizzazz.	Yes	No
6	Use of enhancements	All graphics, video, audio, or other enhancements are used effectively to enrich the learning experience. Enhancements contribute significantly to convey the intended meaning	Yes	No
7	Organization	The sequence of information is logical and intuitive. Menus and paths to all information are clear and direct	Yes	No
8	Branching	Game is genuinely multimedia, rather than linear and contains a significant number of well-designed and age-appropriate choices	Yes	No
9	Citing resources	All sources are properly cited	Yes	No
10	Permissions obtained for resources	All permissions to use text, graphics, audio, video, etc. are available	Yes	No
11	Originality	The game shows significant evidence of originality and inventiveness. The majority of the content and many of ideas are fresh, original, and inventive	Yes	No
12	Curriculum	Clear evidence of a connection to the target curriculum. Frequent and explicit references are made to facts, concepts, and cited resources. Users will learn from this game	Yes	No
13	Evidence that objectives were met	Clear evidence that game content supports stated objectives	Yes	No
14	Depth & Breadth of Game content	Clear evidence that higher-level thinking skills were used in the creation of this game	Yes	No
15	Subject knowledge	Subject knowledge is evident throughout the game. All information is clear, appropriate, and correct	Yes	No

No.	Criteria	Assessed Aspects	Experts Assessment	
16	Items	Items in the instrument are clear enough to take data to the users and have aspects that measurable	Yes	No

(Source : McCullen, et al., 2015)

c) Students Questionnaire

This Student Questionnaire is a student assessment of learning using the "Captain Lungs 3D" application on the topic of the student's respiratory system. This Student Questionnaire will be given to students after carrying out the pretest and also the post-test to evaluate all the learning that has been done. Science subject teachers also have a role in assessing the course of this research. Student's Questionnaire is depicted in Table 3.2.

Table 3.2

Students Questionnaire

No.		Question	Answer	
1	Difficulty level	Do you find difficulties when playing Captain Lungs 3D games?	Yes	No
2	Error navigation	Does the button on the "Captain Lungs 3D" game work well?	Yes	No
3	Language Error	Are the language and writing used in "Captain Lungs 3D" games easy to understand?	Yes	No
4	Clear screen	Does the "Captain Lungs 3D" game have an attractive color and appearance?	Yes	No
5	Challenge	Does each level in the "Captain Lungs 3D" game have different levels of difficulty?	Yes	No
6	Contain material learning	Is there a circulatory system material in humans in the "Captain Lungs 3D" game?	Yes	No
7	Content Understanding	Does the "Captain Lungs 3D" game increase understanding of the material in the circulatory system in humans?	Yes	No
8	Scientific term	Can the biological terms in the "Captain Lungs 3D" game be understood?	Yes	No
9	Delivery of Content	Is the delivery of the material of the respiratory system to humans in a "Captain Lungs 3D" game clearly delivered?	Yes	No
10	Question of Material	Are the questions on "Captain Lungs 3D" this game understandable?	Yes	No

For questionnaire data, the data was tested using a Likert scale to determine the rating scale which can be described as follows Likert Scale for Student's Level on

Agreement shown in Table 3.3

Table 3.3

Criteria For Expert Judgement

Scale Criterion	Point
Strongly Disagree	1
Disagree	2
Agree	4
Strongly Agree	5

After the students fill out the questionnaire based on aspects of the assessment, then we can see the scale of her judgment after summing up all the numbers and process data with a Likert scale. The results of this assessment will later be depicted in a bar chart and also the pie chart

3.6 Objective Test

This test is used to measure students' abilities in testing critical thinking skills, which were tested in the Facione (2015) aspect. The test carried out in this study is an objective test in the form of multiple choices, before being tested on students, this test is assessed first by a science expert to assess that this question is suitable for use or need to be revised. The test to be used consists of 28 questions, with the division of 5 questions for interpretation, six questions for analysis, six questions for evaluation, 6 questions for the explanation, five questions for inference in the form of multiple-choice questions via a google form. To check validity, reliability, distinguishing power, difficulty level, and distractor are tested as follows:

a) Validity

Validity is a measure to show the validity of an instrument. Good validity is the validity that has a high value, whereas low validity is the validity of having a low value. Validity testing is carried out using the

product-moment correlation (Suherman, 2003). The validity formula is described as follows:

$$Rxy = \frac{N\Sigma xy - (\Sigma x)(\Sigma y)}{\sqrt{\{N\Sigma x^2 - (\Sigma x)^2\}\{N\Sigma y^2 - (\Sigma y\}^2}}$$

(Arikunto, 2006)

Description

r_{xy} : Question Validity

n : amount of sample

x : score from each item

ΣX : Total score from each item

ΣY : Total Score

Table 3.4

Interpretation Value of Validation Coefficient

Correlation Coefficient	Criteria
$0,800 < r_{xy} \le 1,00$	Very High
$0,600 < r_{xy} \le 0,800$	High
$0,400 < r_{xy} \le 0,600$	Average
$0,200 < r_{xy} \le 0,400$	Low
$0,00 < r_{xy} \le 0,200$	Very Low
$R_{xy} \le 0,00$	Invalid

(Arikunto, 2006)

b) Reliability

Reliability is one level of reliable instrument. Reliability data is data that has been tested many times. It will remain the same (Arikunto, 2006). In assessing reliability, the formula used uses the Kuder Richanson (KR-21) formula, explained as follows :

$$r1 = \left(\frac{k}{k-1}\right)\left(1 - \frac{M\,(k-M)}{(k)\,(s)2}\right)$$

With

$$M = \frac{\sum X}{N}$$

$$s2 = \frac{N\sum X2 - (X)2}{N\,(N-1)}$$

Description

r_i : Reliability test coefficient

k   : amount of test in instrument

(s)2 : Total Variance

M   : Total Mean

N   : Amount of students

$\sum X$: Total Score

Reliability interpretation refers to the opinion of Guildford (Ruseffendi), with the following criteria shown in Table 3.5

Table 3.5

Interpretation of the reliability coefficient

ri amount	Criteria
$0,90 < r_i \leq 1,00$	Very High
$0,70 < r_i \leq 0,90$	High
$0,40 < r_i \leq 0,70$	Average
$0,20 < r_i \leq 0,40$	Low
$0 < r_i \leq 0,20$	Very Low

(Guildford, 2005)

c) Difficulty Level

A difficulty level is a number that shows the difference between the difficulty and the ease of a problem called the difficulty index. (Arikunto, 2006). The difficulty index ranges from 0.00 to 1.00. If the difficulty value shows a value of 0.00, it means the question is difficult. If the difficulty value is around 1.0, indicating that the problem is easy, the formula for showing difficulty is described as follows:

$$P = \frac{B}{JS}$$

(Arikunto, 2006).

Description

P : Difficulty Index

B : The number of students who answered correctly

JS : Total Number of Sample

The interpretation of the difficulty index is presented depicted in Table 3.6:

Table 3.6

Interpretation Difficulty Index

Difficulty Index	Criteria
P = 0,00	Too difficult
$0,00 < P \leq 0,30$	Difficult
$0,30 < P \leq 0,70$	Average
$0,700 < P \leq 1,00$	Easy
P = 1,00	Too Easy

(Arikunto, 2006)

d) Distractor Effectiveness

Distinguishing power is the ability of a question to distinguish students who are smart (high ability) from students who are less intelligent (Arikunto, 2006). The formula for knowing the power of distinction is as follows :

$$Dp = \frac{BA}{JA} - \frac{BB}{JB} = PA - PB$$

Description

Dp   : Discrimination Index

JA   : Number of upper group

JB   : Number of a lower group

JA   : Number of participants in the upper group who answered correctly

JB   : Number of participants in the lower group who answered

correctly

PA   : Proportion upper group who answered correctly

PB   : Proportion of lower group who answer correctly

In question there also distractor effectiveness to compare between the group of high level and low level, to distinguish her then there is a distractor effectiveness. To distinguish it then seen depicted in Table 3.7

Table 3.7

Distractor Effectiveness

Distractor effectiveness	Criteria
$D_p \leq 0{,}00$	Very Less
$0{,}00 < D_p \leq 0{,}20$	Less
$0{,}20 < D_p \leq 0{,}40$	Average
$0{,}40 < D_p \leq 0{,}70$	Good
$0{,}70 < D_p \leq 1{,}00$	Very Good

(Arikunto, 2006)

3.7 Data Analysis for Student's Critical Thinking

Data processing for instruments on this scale uses the rating scale method. Rating scale is a subjective measure made using a scale. A rating scale is a subjective measure made to measure the respondent's scale to measure socioeconomic status, institutions, knowledge, abilities, and other

activities (Sugiyono, 2013).For measuring rating scale using the following formula as follows :

$$P = \frac{Score\ data\ collection\ results}{ideal\ score}\ x\ 100\%$$

Description

P   : Percentage number

Rating Scale Interpretation

Table 3.8

Interpretation Rating Scale

Percentage Score (%)	Criteria
$0 \le P < 20$	Very Not Good
$20 \le P < 40$	Not Good
$40 \le P < 60$	Average
$60 \le P < 80$	Good
$80 \le P < 100$	Very Good

(Sugiyono, 2013)

Description

P   : Percentage number

a) Normality

The normality test in this study uses the Kolgomorov-Smirnov normality test. The basic concept of this normality test is to compare the distribution of data with standard normal. This normality test uses the SPSS version 25 application

b) Homogeneity

The homogeneity of this test was carried out on the pretest and pre-test and post-test results data that were normally distributed with the aim of knowing whether the variance in the class had the same variance or not. The standard for testing homogeneity uses Fisher's test with a significance of α = 5% or α = 0.05. This homogeneity test uses Microsoft Excel. The homogeneity test is calculated by the equation:

$$F = \frac{Large\ Variant}{Small\ Variant}$$

In the homogeneity test using the *Fisher* test method, the scoring in this method uses the right only method, where correct answers are given a score of one (1), wrong answers are given a score (0), and unanswered answers are given a score (0). Here's the formula for calculating the score :

$$S = \frac{Right\ Answer}{False\ score}\ x\ 100$$

c) N Gain Score

The N gain test was conducted to determine the results of the pre-test and post-test score, then divided by the maximum score and subtracting the pre-test score. The N gain test aims to improve students' understanding after using "Captain Lungs 3D". This N Gain test is to calculate the gain test formula. Criteria for Gain Index is shown in Table 3.9

Table 3.9

Criteria for N Gain

g value	Criteria
$0,7 \leq g < 1$	High
$0,3 \leq g < 0,7$	Average
$0 \leq g < 0,3$	Low

d) Score for Questionnaire

For measuring questionnaire score, is depicted in Table 3.10 and 3.11

Table 3.10

Criteria for Questionnaire

Criteria	Quantity
Very Good	4
Good	3
Enough	2
Less	1

Table 3.11

Criteria for Questionnaire

Percentage Value	Criteria
$0 \leq P < 25$	Less
$25 \leq P < 50$	Enough
$50 \leq P < 75$	Good
$75 \leq P < 100$	Very Good

e) Questionnaire Percentage

Table 3.12

Questionnaire Percentage

Percentage Value	Category
0%	Invalid
$25 \leq P < 50$	Enough
$50 \leq P < 75$	Good
$75 \leq P < 100$	Very Good

3.8 Instrument Analysis Result

In analyzing instruments, there is a process of validation concerning the respiratory system question, which has been assessed by experts, then tested to the students of junior high school grade 8 as many as 30 students by using the method of convenience sampling. The process of data processing is depicted as follows :

3.8.1 Student's Critical Thinking Test Item

Before this test item was used as an objective test in research, this test item was tested by experts in the field of biology and tested on 8th-grade junior high school students by taking a sample of 30 students online, who already learned about human respiratory system. The results of this test item were processed using ANATES to analyze the validity, reliability, level of difficulty, and discriminating power. The results of the test items are tabulated in Table 3.13

Table 3.13

Test Item Analysis

Number of Test Item	Validity	Level of Difficulty	Discriminating Power	Acceptance
1	0.365 (Low)	Easy	0.25 (Satisfactory)	Used
2	0.578 (Enough)	Easy	0.37 (Satisfactory)	Used
3	0.562 (Enough)	Easy	0.25 (Satisfactory)	Used
4	0.202 (Low)	Medium	0.25 (Poor)	Revised
5	0.343 (Enough)	Medium	0.50 (Good)	Revised
6	0.198 (Low)	Medium	0.12 (Poor)	Revised
7	0.334 (Low)	Medium	0.50 (Satisfactory)	Revised
8	0.000 (Low)	Difficult	0.50 (High)	Revised
9	0.223 (Low)	Medium	0.50 (Satisfactory)	Revised
10	0.391 (Low)	Very Difficult	0.87 (Good)	Revised
11	0.454	Very Easy	0.25	Used

Number of Test Item	Validity	Level of Difficulty	Discriminating Power	Acceptance
	(Low)		(Satisfactory)	
12	0.413	Medium	0.62	Used
	(Very Low)		(Satisfactory)	
13	0.396	Medium	0.37	Used
	(Enough)		(Satisfactory)	
14	0.416	Medium	0.37	Used
	(Enough)		(Satisfactory)	
15	0.653	Medium	0.50	Used
	(Enough)		(Good)	
16	0.454	Medium	0.50	Used
	(Enough)		(Good)	
17	0.103	Medium	0.50	Revised
	(Low)		(Good)	
18	0.216	Difficult	0.50	Revised
	(Enough)		(Good)	
19	0.584	Easy	0.00	Used
	(Very Low)		(Poor)	
20	0.377	Medium	0.25	Used
	(Very Low)		(Poor)	
21	0.206	Medium	-0.125	Revised
	(Very Low)		(Poor)	
22	0.377	Medium	0.375	Used
	(Enough)		(Satisfactory)	
23	0.206	Medium	0.75	Used
	(Enough)		(High)	
24	0.377	Medium	0.375	Used
	(Enough)		(Satisfactory)	
25	0.369	Medium	0.625	Used
	(Enough)		(Good)	
26	0.429	Easy	0.875	Used
	(Enough)		(High)	
27	0.385	Easy	0.50	Used
	(Enough)		(Good)	
28	0.647	Easy	0.25	Revised
	(High)		(Satisfactory)	

Based on Table 3.13, the results of ANATES analysis obtained data that there were no questions that were rejected; the result is 11 questions needed improvement. Then for the other 17 questions, it can be used as an instrument for the objective test on the aspects of critical thinking skills, For distribution about critical thinking indicators, not sequential due so that

students can be effective to answer the question correctly and effectively because of a matter that is distributed is varied and diverse. For value data, if the data showed the value of 0.5 then the data is valid, also with the reliability of the data the obtained data that the value of the validity of as much as 0.80 and a reliability of 0.72. This is in accordance with the assessment of expert judgment in biology topic, which says that "question can be used for validation", and other comments say "the level of questions is quite difficult and can train students ' critical thinking skills".

3.9 Research Procedure

Research and Development (RnD) methods aim to develop products and also test them. Research and development also aim to develop and validate products used in research (Borg & Gall, 1983). This research is also used to develop and validate the products used in the learning process, the research that is produced to produce a product through procedures, which is followed by the development process and ends with trials or evaluations, such as the following stages:

1) Research and information collecting

    a) Observation stage for a research field

    b) Literature Study about critical thinking and respiratory system

    c) Designing learning media using unity 2019 and research instrument

    d) Decide Software used to develop a game

    e) Validation of objective tests, rubrics and research instruments

2) Research Implementation Stage, include :

    a) Designing media by using Unity 2019 to create the game

    b) Testing media to experts on science, media, and school teachers

    c) Testing Product and inserting question based on critical thinking

indicators into games

3) Final Stage

   a) Revision media and prepared for taking data at school or online

   b) Design Revision and taking data or school or online

   c) Evaluation session, assessed rubric and students questionnaire

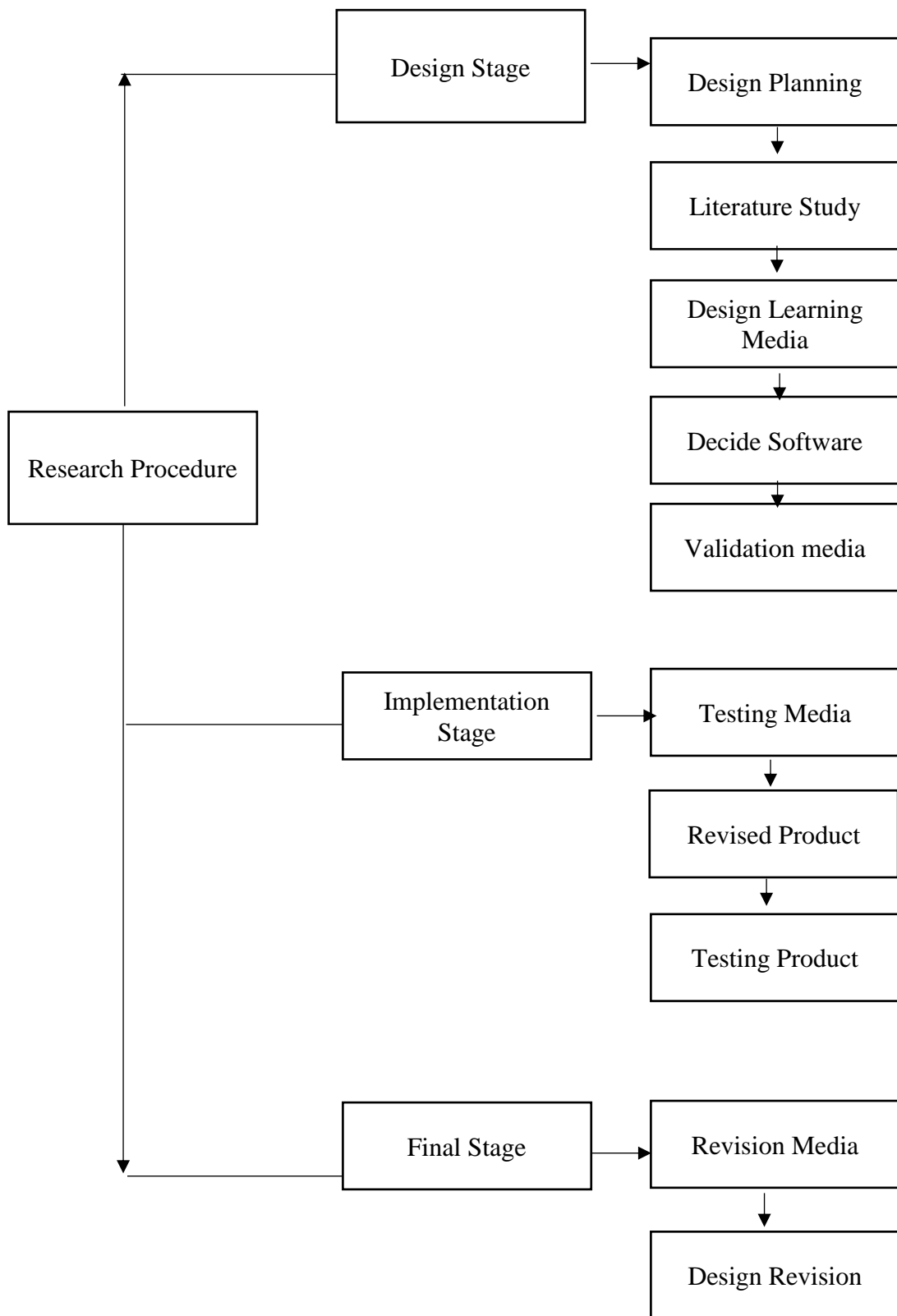   d) Judgement students questionnaire and rubric

The scheme of research procedure stages is shown in Figure 3.2

Figure 3.2 Research Procedure Design