Developing Critical Thinking Skills Assessment of Digestive System for Senior High Schools

Yenny Anwar*, Djunaidah Zen, Adinda Tiara

Department of Biology Education, Universitas Sriwijaya, Indonesia
**Corresponding author. Email: yenny_anwar@fkip.unsri.ac.id*

ABSTRACT
This study aims to produce valid, reliable and practical test instruments based on critical thinking skills. The method used in this study is the development research (DR) that consists of synthesizing theory and needs analysis, step design (variable construction, indicator development, questions arrangement, making the instrument, scoring) and evaluation. This study produced a valid questions in terms of content, construct, and language, with a final above 0,61. The reliability of the test was 0,88. The result of students questionnaire analysis in the phase of field trial gained an average of 3,72 The average of value obtained from the student questionnaire indicates that the Critical Thinking Skills questions of digestive system is easy to use (practical). Based on the data that has been analyzed, it shows good validity, practicality and reliability so that a question instrument based on critical thinking skills has been produced which is feasible and practical to use.

Keywords: *Biology, Practicality, Reliability, Validity*

1. INTRODUCTION

Biology is a scientific study that has a high level of difficulty. The object of Biology study is very broad and includes all living things. Some of the aspects studied in living things are anatomy, morphology, and physiology. Physiology is the study of mechanical, physical, and biochemical processes that support bodily functions. The complex biology material to study is physiology. One of the biology materials that demands an understanding of physiological aspects is the digestive system material [1].

Digestive System material demands learning competence at a high level of understanding. However, in reality currently students are more likely to memorize rather than understand, even though understanding is the basic initial capital for further mastery. This causes the results of students' science learning, especially in Indonesia, are still dominant at a low level. Evidence that the results of studying science in Indonesia are low can be seen from the 2012 PISA data which shows that the science ability score of students aged 15-16 years (SMA) in Indonesia is only 382, which is much lower than the international average score of science ability, which is 501 [2]. While PISA in the 2015 Science Competencies for Tomorrow's World, Indonesia was ranked 69 out of 76 countries [3].

One part of the vision of national education based on Law Number 20 of 2003 concerning the National Education System is to develop quality human beings who are capable and proactive in responding to the challenges of the ever-changing times. In Article 35 Paragraph 1, it is further explained that the quality of national education must be improved on a regular basis, one of which is to excel in competitions between nations in world civilization. Related to this, the low ranking of Indonesian students in scientific ability shows that Indonesian students have not been able to excel in global competition, which means that this vision has not been fully achieved.

The low learning outcomes of Indonesian students demand immediate improvement. Good learning should be able to explain how students should learn and think [4]. Permendikbud No 35/2018 also explains that the learning that has been carried out by educators is expected to be able to invite and train students to think to a higher level. Thus, students are not accustomed to solving high-order thinking problems is the reason for the low learning outcomes of Indonesian students. Therefore, there needs to be a change in learning and assessment, especially the questions developed. The questions developed are expected to encourage students to think critically. The ability to think critically is an internal factor that affects learning outcomes and is important to develop. The ability to think critically is related to the use of cognitive skills or strategies that increase the likelihood of getting the desired impact [5]. Someone who has critical thinking skills will be able to

think clearly and systematically, so that he is able to express ideas with good articulation [6]. This ability is also needed in logically analyzing structures in order to understand a text well.

Based on the results of interviews of researchers while carrying out the Development and Application of Learning Devices (P4) activities with SMA Negeri 5 Palembang educators and seeing the evaluation of the Learning Implementation Plan document, it was found that educators were still making questions with Low Level Thinking Ability. As a result, when students are faced with problems with a higher level of thinking, one of which is critical thinking, students will have difficulty solving them. In this regard, **[7]** research shows that students' critical thinking skills in building basic skills are still lacking. Another study regarding critical thinking skills of junior high school students were still low[8]. The low ability to think critically can have an adverse impact on further education. The ability to think critically and creatively must be possessed by all students at every level of education, therefore, critical thinking skills need to be trained [9]. Critical thinking can be taught and requires practice to have it [10]. One way to train students' critical thinking skills is through the development of questions based on critical thinking skills. Researchers tried to develop Biology questions based on critical thinking skills in the Digestive System material for high school, with K.D. 3.7, where the goal to be achieved is that students are able to analyze the relationship between the tissue structure of the organs in the digestive system and relate it to nutrients and bioprocesses so that they can explain the digestive process and functional disorders that may occur in the human digestive system through literature studies, observations, experiments , and simulations, with the hope that the questions that will be developed later can develop students' critical thinking skills.

This research aims to produce valid, reliable, and practical questions based on critical thinking skills on the digestive system material for high school.

2. METHOD

This research is a development research or Development Research which refers to the design flow of instrument development according to[11]. This development research aims to produce a product in the form of questions through the steps of developing an accountable instrument. The resulting product is a test instrument in the form of multiple choice questions with five answer choices totaling 40 questions that aim to measure critical thinking skills in the digestive system material for high school.

The population in this question development research was all students of XI IPA class at SMA Negeri 5 Palembang for the 2019/2020 school year. Determination of the research class sample in this study using the Random Sampling technique, which is a method of random sampling without paying attention to the strata in the population [12]. The class used in this research is confirmed to have received material on the Digestive System from the Biology subject educators. Determined the sample in this question development research is XI IPA 1 class and XI IPA 6 class, totaling 58 students.

This research was conducted by following the development steps of Djaali and Puji Mulyono's (2008) instrument with the following development steps:

2.1. Examining Theory and Analyzing Needs

The step of identifying and collecting theories related to instrument development in the form of questions, learning evaluation, and assessment. Needs analysis in the form of literature review through interviews with educators and analysis of learning tools in the form of critical thinking skills so as to determine research objectives.

2.2. Devising Questions Indicators

Devising question indicators can be done after examining theory and analyzing needs. This research on developing questions based on critical thinking skills uses indicators that think critically There are five indicators which are further grouped into twelve selected sub-indicators [13]. Then the question indicators are designed according to the basic competency 3.7 and the selected critical thinking indicators.

2.3. Creating Questions Grids

The grid is a description of the questions to be developed, in the form of a table of specifications which aims as a guide in writing questions. The question grid contains basic competencies, question indicators, and measured indicators.

2.4. Creating Questions Instruments and Scoring

Based on the grid that has been made, the assessment of the instrument refers to the material to be developed. The number of questions made was 40 multiple choice questions (several choices). Then determine the score of the items for which the correct answer is given a score of 1 and the wrong answer is given a score of 0, then interpret the value. The resulting instrument is an initial product called prototype 1 which is ready to be validated.

2.5. Doing Theoretical Validation

The theoretical validation is carried out by experts (validators). This research in developing critical thinking skills based on two experts as validators. The theoretical validation that is carried out is construct validation carried out by experts in the construct field. The construct validation indicators analyzed were the

clarity of the question instructions, the preparation of easy-to-understand question components, the readability of the tables and pictures in the questions, having homogeneous and logical answer choices. Then the content validation is carried out by experts in the content field. Content validation indicators are the extent and suitability of the material being asked for the expected learning objectives, correctness of concepts and definitions, suitability with facts in the environment so that they can be observed, distractors on multiple choice questions function well and have one correct answer key. Finally, language validation is carried out by linguists. The language validation indicator is in the questions using good and correct language according to EYD, the language used is communicative, the questions and answer choices do not provoke and offend a side, there are no sentences that have multiple interpretations, and do not use regional languages.

2.6. Doing Revision

Suggestions from experts are used to revise the development instruments made. Comments or suggestions from the validators are written on the validation sheet as material for revising the item development instrument. Based on the responses of the analyzed experts, then it is used as material for the revision of the product being developed. The results of this product revision are called prototypes 2. At this stage, the research product developed can be said to be valid if it has been declared valid by experts (validators) and is feasible to be tested on students.

2.7. Doing Empirical Validation

Empirical Validation is the final stage of developing a question instrument. Empirical validation was carried out by field trials involving 58 students of XI IPA class at SMA Negeri 5 Palembang. The results obtained will be analyzed the validity, reliability, degree of difficulty, distinguishing power, and distractor function using Anates V4. In addition, at this stage of field trials, a practicality test was also carried out on the question instruments being developed. At this stage, the research product developed is said to be practical if it has been declared practical by students through a questionnaire.

3. RESULT AND DISCUSSION

The results of the theory examine and needs analysis that have been carried out based on interviews with Biology educators at SMA Negeri 5 Palembang are that only a few students are able to work on questions based on critical thinking skills, the rest still find it difficult to understand the questions. Furthermore, the results of the analysis of questions in the lesson plan (RPP) show that the questions used are not based on critical thinking skills, however, they are still categorized as questions C1 to C6. This means that the critical thinking skills of students still need to be trained. This is supported by the opinion of Anwar (2020) which states that the factor that causes low critical thinking skills is the lack of training for students in answering questions that cause problems [9].

3.1. Devising Questions Indicators

Indicator design is carried out by designing item indicators according to basic competencies 3.7 and indicators of selected critical thinking skills. Indicators of critical thinking skills used are based on indicators of critical thinking skills there are five indicators which are further grouped into twelve sub-indicators [13].

3.2. Creating Questions Grids

The question grid is arranged based on the indicators that have been designed. The grid is used as a guide for writing questions. The questionnaire made by the researcher is derived from basic competencies 3.7.

3.3. Creating Questions Instruments and Scoring

The arrangement of the instrument begins with the creation of a grid of questions based on critical thinking skills. During the process of arranging the instrument the researcher communicated with the supervisor to design the initial prototype. The supervising lecturer comments on the initial prototype, namely in the previous draft the image included was not clear about the source so that the source information was corrected, in the previous draft some questions were still unclear in order to clarify the problem instructions, in the previous draft some tables and pictures still did not provide complete information so that equipped with information, in the previous draft some of the questions made did not match the indicators of critical thinking skills to be achieved so that the indicators were adjusted.

The supervisor's comments above serve as a guide for researchers to revise the initial prototype. The revised result of the initial prototype is called prototype 1. In addition to designing the grid, the researcher also determines the score for each item.

The supervisor's comments above serve as a guide for researchers. Before questions based on critical thinking skills enter the theoretical validation stage of the experts, the researcher also designs the front page, general instructions and answer sheets to complete a series of questions based on critical thinking skills. Prototype 1 was then consulted again to the supervisor, the supervisor stated that the prototype of critical thinking skills based questions was ready to go through the next evaluation stage, namely the theoretical validation stage. To revise the initial prototype. The revised result of the initial prototype is called prototype 1. In addition to designing the grid, the researcher also determines the score for each item.

3.4. Expert Judgment (Theoretical Validation)

The activity of the expert assessment stage, namely the validity test of prototype 1, was carried out by experts or validators who validated the design of questions based on critical thinking skills consisting of

content validators, construct validators, and language validators. The validator assesses the initial prototype by filling out the validation sheet. The results of theoretical validation are as follows:

3.4.1.Content Validation

The content validator validates questions based on critical thinking skills in terms of the suitability of the content of the question material. Comments from the content validator (content) prototype 1 that is, some question contents still do not match the indicators.

3.4.2.Construct Validation

The construct validator validates questions based on critical thinking skills in terms of the problem construct. Comments from the construct validator prototype 1 namely the placement of tables and proportional tables, pay attention again, make it sweeter, also pay attention to the relationship between the image and the contents of the question.

3.4.3.Languange Validation

Comments from the language validator prototype 1. that is, pay attention to the language pattern, adjust it to EYD.

3.5. Empirical Validation

Finding the value of empirical validity is done by examining the answers of students. Each correct answer is given a score of 1 and the wrong answer is given a score of 0. This calculation is then analyzed and the validity value is obtained by entering the students' answers into the Anates V4 program.
Table 1 shows the results of the empirical validation which includes the validity of item items based on critical thinking skills.

Table 1. Empirical Validation Measurement Results

Question Item	Interpretation		Correlation Coefficient (rpbi)
	Valid	Invalid	
1,2,3,4,5,6,7,8,9, 10,11,12,13,14,15,16,17,19,20,21,22,23,24,25,26, 27,28,,33,34,35, 37,39.	✔		0,300-0,500
5,18,29,30,31,32, 36, 38, 40		✔	0.100-0,000

Based on the results of the analysis of the validity of the questions, of the 40 items tested for their validity, 31 items were stated as valid items, while 9 items are invalid items.

The trial phase was also carried out to determine the students' assessment of the practicality of questions based on critical thinking skills by filling out a practicality questionnaire. The results of students' assessment of the practicality of questions based on critical thinking skills that have been developed can be seen in Table 2.

Table 2. Students' Assessment of Practical Questions Based on Critical Thinking Ability

No.	Indikator	The Value Of Students' Responses (N=58)
1	Instructions for questions	3.77
2	Sentences, tables, graphs, and pictures used in the questions	3.65
3	Characteristics of critical thinking	3.82
4	Languange	3.65
Final score		3.72
Category		Practical

Table 2 shows that the results of students' assessment of the practicality of questions based on critical thinking skills were 3.72 and questions based on critical thinking skills were in the practical category.

3.6. Reliability Test

The next step is to calculate the reliability value of questions based on critical thinking skills. The reliability test was carried out to see the stability of the questions based on the developed critical thinking skills. Students' answer sheets were analyzed and the value of the reliability coefficient was calculated. Table 3 below shows the results of the calculation of the reliability test of questions based on critical thinking skills.

Table 3. Results of the Calculation of Questions Reliability Test Based on Critical Thinking Ability

Product moment correlation coefficient	0.78
Reliability coefficient of the test	0.88

Table 3 that the reliability of critical thinking questions is 0.88, so it can be stated that the critical ability based questions have high reliability.

3.6.1. Item analysis

Item analysis was carried out after the test based on the critical thinking ability of prototype 2 was tested, its validity and reliability values were calculated. Each item will be analyzed the degree of difficulty, distinguishing power and function of the distractor. This evaluation is carried out to find out whether the items that make up the test can carry out its function properly or not.

3.6.2. Degree of Difficulty

The difficulty index calculation was carried out using the Anates V4 program. The results of the

calculation of the difficulty index are interpreted in three criteria, namely: questions with P <0.30 are considered difficult questions; questions with P 0.30 to 0.70 are classified as medium questions; and questions with P <0.70 are relatively easy questions. Based on the results of the item difficulty index analysis, there were 9 items or 22.5% included in difficult questions, 26 questions or 65% were moderate questions, and as many as 5 questions or 12.5% were included in easy questions. The following is the distribution of the results of the item analysis based on the degree of difficulty.

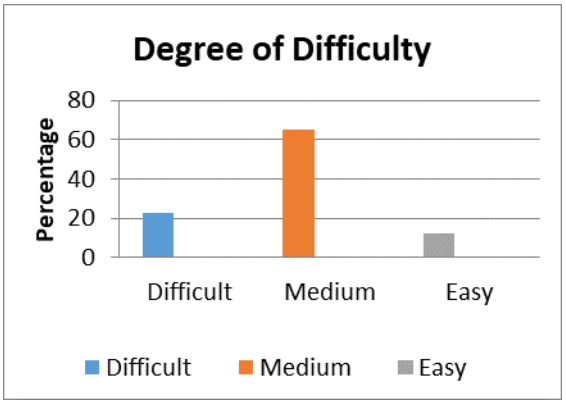

Figure 1. Degree of Difficulty

3.6.3. Discrimination Power

The results of the calculation of discrimination power with the Anates V4 program show that 18 items or 45% have good discrimination power, 10 questions or 25% have sufficient discrimination power, 5 items or 12.5% have excellent discrimination power. The following is the distribution of the items based on distinction.

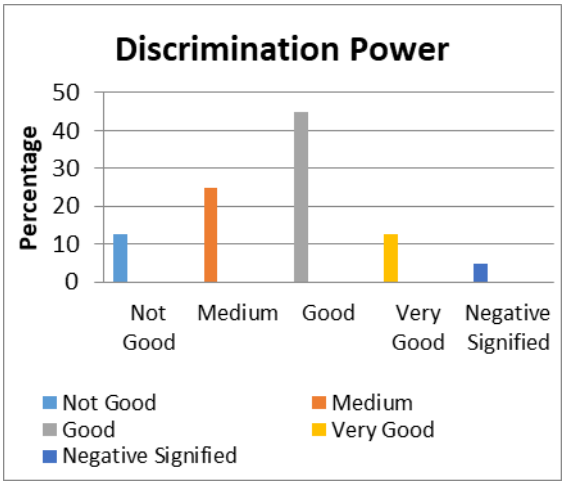

Figure 2. Discrimination Power

3.6.3. Distractor Function

Calculation of the effectiveness of the distractor (distractor function) on questions based on the critical thinking ability of the digestive system material using the help of the Anates V4 program.

The results of the analysis show that on average the answers of the deceivers have performed their function well. At least 5% of all test takers have chosen the cheating answers, from the 40 test items there are 29 questions that have good quality, namely questions number 4, 7, 9, 10, 11, 13, 14, 15, 16, 18 , 19, 20, 21, 22, 23, 24, 25, 26, 27, 29, 31, 32, 33, 34, 35, 36, 37, 38, 40.

Based on the results of the item analysis, then linked to the results of the validity and reliability test, the questions based on critical thinking skills developed have been classified into good quality test instruments. Questions based on critical thinking skills that have been tested for validity (theoretical and empirical), calculated the value of reliability, analyzed each item declared valid, reliable, and practical.

The development of the instrument begins with analyzing the importance of developing critical thinking skills based questions. This is supported by previous research on critical thinking skills. 21st century education requires the emergence of a superior generation who is able to think critically and creatively, the emphasis is expected to increase their abilities in the learning process that can foster and encourage critical thinking skills [9]. Emphasis The critical thinking ability of students can be trained through learning by providing critical thinking questions [14]. Furthermore, the indicator of critical thinking skills that will be used is an indicator of critical thinking skills according to Ennis (1985), namely there are five indicators which are further grouped into twelve sub-indicators [13]. The next stage, the indicators that have been determined are then set out in the question grid. The question grid based on critical thinking skills was developed in the form of a multiple choice test of 40 items with five answer choices.

The critical thinking skill-based question grid that has been created is called the initial prototype. The initial draft was then consulted with the supervisor and got suggestions used to revise the initial draft before entering the expert validation stage. The results of the revision of the initial prototype are called prototype 1. Prototype 1 is then tested for its theoretical validity by experts, including content experts, construct experts, and linguists. The conclusion obtained from the expert is that the question instrument is valid with several previous revisions. After passing the expert validation stage and being revised, the instrument is called the prototype 2. Prototype 2 is then tested on 58 students of class XI IPA 1 and XI IPA 6 at SMA Negeri 5 Palembang. The trial at this stage was carried out to determine the item's validity, reliability, practicality, degree of difficulty, differentiation and function of the distractor on the test item instrument based on critical thinking skills developed by the researcher.

A total of 31 questions based on critical thinking skills or 77.5% were declared valid and the remaining 9 questions based on critical thinking skills or 22.5% were

declared invalid. Valid items are due to a relationship between the alignment of the score on the item items and the total score. A valid question means that the item has been able to carry out its function, which is to measure what should be measured. Then for invalid questions it can be caused by various factors. This is in line with there are three factors that affect the validity of the test results, namely the instrument factor used for the test, administrative and scoring factors, and factors from students' answers [15]. Factors originating from the test, namely: 1) unclear clues; 2) Use of difficult vocabulary and sentence structures; 3) Ambiguity, namely the possibility of multiple interpretations in understanding and solving test questions; 4) excessive emphasis on certain levels, so it is too easy to predict the tendency of the answers to questions; 5) the quality of the test items is not coherent, and 6) the answer patterns are easy to predict.

Based on the researchers' observations during the research process, the validity of the items based on the developed critical thinking skills tended to be influenced by the use of difficult vocabulary and sentence structures and questions that were too long. Students who act as test takers have difficulty understanding sentences. This makes them need more time to solve problems based on critical thinking skills. The ambiguity factor also influences the validity of the test items. Valid items can be stored in the question bank. On the other hand, invalid items need to be corrected but must go through the theoretical and empirical validation stages again. The results of empirical validation show that the instrument based on critical thinking skills developed by researchers is categorized as a valid instrument in line with the results of expert validation that has been done previously.

In addition to validity, high test reliability is one of the requirements for a learning outcome test that can be said to be a good test. Reliability is a consistency or similarity in the results of measuring objects that are carried out many times at different times. Analysis of the items in terms of reliability is carried out to measure the consistency of the test. A test is said to be reliable if the test gives the same results when given to the same object at different times. The reliability value of the instrument based on critical thinking skills was 0.88, meaning that the test questions based on critical thinking skills developed by the researcher had high test reliability. Because the value obtained is above 0.70, so the question instrument based on critical thinking skills developed by the researcher can be said to be a reliable test instrument and can measure what is to be measured in accordance with the objectives of developing this critical thinking ability-based question. After obtaining the data on the validity of the questions, then the calculation of the achievement of the critical thinking ability indicators was carried out, namely the analysis of the critical indicators that were answered correctly by students. The indicator that most students answered correctly was the third indicator, namely concluding that the sub-indicator made and determined the results of the consideration with the percentage chosen by students 55.6%, this shows that the ability of students to make and determine the results of the consideration is good.

The instrument based on critical thinking skills that has been validated is then tested for practicality using a questionnaire to obtain practicality value. During the research process in the trial phase, the researcher succeeded in proving that the product of the questions based on critical thinking skills developed was considered practical by students. The practicality value obtained was 3.72. Assessment of product practicality is obtained by exploring students' opinions through practicality questionnaires. Students said that the questions presented were good and easy to understand, but there were still some sentences in long questions and there were some terms they didn't understand so they were difficult to understand.

The number of item questions in this research was 40 items and after being validated it became 31 items because there were 9 invalid items. If viewed from the level of difficulty, the items are found that are difficult for students to understand. Some of the sentences in the questions were difficult to understand and too long so that the research subjects gave answers that did not match the purpose of the questions. The items that are invalid are questions number 5, 18, 29, 30, 31, 32, 36, 38 and 40. Judging from the results of the analysis of the causes of the nine questions being declared invalid, it can be influenced by the level of difficulty and the distribution pattern of the answers. of the item. Problem number five is in the moderate category, but the uneven distribution of students' answers makes this question categorized as invalid. After analyzing other causes numbers 5, 18, 29, 30, 31, 32, 36, 38 and 40 are invalid are sentences that are difficult for students to understand so that they find it difficult to determine answers.

After calculating the validity and reliability and practicality of the questions, the last thing to do is to analyze the test items based on critical thinking skills. This analysis is carried out to determine whether the items that build critical thinking skills-based questions are able to carry out their functions properly. So that the resulting test device is useful as a measure of good quality critical thinking skills. The analysis carried out includes three things, namely (1) the degree of difficulty of the item, (2) the distinguishing power of the item, and (3) the distractor function.

The quality or not of the items based on critical thinking skills can first be seen from the degree of difficulty or level of difficulty that each item has. There are 9 questions or 22.5% including difficult questions, 26 questions or 65% are medium questions, and 5 questions or 12.5% are easy questions. Item that is difficult will make students not interested in trying again because it is beyond the ability of students. Easy item questions will not stimulate the ability of students to answer questions. An item is considered good if it has

a difficulty level between 0.30-0.70 [16]. So it can be concluded that the question instrument based on critical thinking skills includes questions that have a good level of difficulty because some items or 65% have a moderate level of difficulty. The follow-up that can be done after analyzing the difficulty level of the items is as follows [17], items that have a level of difficulty in the good category (moderate difficulty level), should be stored in the question bank so that they can be issued again at a later time. will come. There are 3 possible follow-up items, namely: the item is discarded and will not be issued again in the future learning outcome test. Re-examined the factors that cause the items concerned are difficult to answer by students. Improvements can be made by simplifying the sentences so that they do not cause multiple interpretations. Furthermore, these items can be issued again in future learning outcomes tests. The items are retained to be used again in very strict tests, in the sense that most of the students will not pass the selection test.

Furthermore, an analysis of the distinguishing power of the items was carried out. The discrimination power of an item is the ability of a test item to be able to distinguish between high-skilled students and low-ability students. The discrimination power of items is very important. One of the bases held for compiling the items of the test is the assumption that the abilities between one student and another are different. The test items must also be able to reflect the differences in abilities that exist among students. The results of the analysis of items based on critical thinking skills show that out of 40 items, there are 18 items or 45% have good discrimination power, 10 questions or 25% have sufficient discrimination power, 5 items or 12.5% have discrimination power very good question. The results showed that the instrument based on critical thinking skills developed by the researcher had good quality discrimination power so that it could be directly entered into the question bank.

Finally, after analyzing the degree of difficulty and discrimination power of the items, the next analysis is the confounding factor. The main purpose of installing a cheat on each item is so that from the many students who take the test some are interested or stimulated to choose it as the correct answer. The more students who are fooled, it can be stated that the distractor has been able to carry out its function properly. The results of the analysis show that on average the swindler's answers have performed their function well because they have been chosen by at least 5% of all test takers.

According to researchers' observations, there is a relationship between the validity and reliability value of the questions with the practicality of the question products. This study obtained a practicality value of 3.72. This value shows that the product that is made has good readability, is easy to understand, and is in accordance with the characteristics of critical thinking skills because most of the statements in the practicality questionnaire are considered good by students. Product questions that have good practicality will increase the significance value of each item so that the product moment correlation is high. If the product moment correlation is high, then the question is valid. The number of valid questions causes a high item reliability value.

A good measuring tool must be valid, reliable and practical. In addition to validity and reliability, a test instrument must also be analyzed for its constituent items. This evaluation is carried out to find out whether the items that make up the test can carry out its function properly or not. So that, in the future, the tests that are arranged can really carry out its function as a measuring device that has good quality. The measuring tool that has been developed can be used as a formative assessment used as a self-assessment instrument. with self-assessment students can reflect on their abilities. with self-reflection students' abilities will develop better [18].

The result is that the questions based on critical thinking skills are developed, seen from the degree of difficulty of the items, the distinguishing power and the distracting factors that have been functioning properly. The developed questions to measure the critical thinking skills of high school students and have good question reliability and can be used to measure students' critical thinking skills [14]. Based on this, the researchers concluded that the questions based on critical thinking skills developed were categorized good because they were valid, reliable and practical

4. CONCLUSION

Questions based on critical thinking skills produced for the material on the structure and function of tissue cells in the digestive system of class XI SMA are categorized as valid, reliable and practical. The question is categorized as valid because it has been validated by the validator and declared fit for use by students. This study resulted in 31 valid item questions. The questions are categorized as reliable because they have a high reliability rate, and are in the practical category because a practicality test has been carried out by the participants by filling out a practicality questionnaire and most students state questions based on critical thinking skills for practical digestive system material. This research produces a product in the form of a Question Bank based on Critical Thinking Ability that is feasible to use.

ACKNOWLEDGMENTS

This research was financially supported by a research grant (PNBP Universitas Sriwijaya 2020) from the Directorate research and technology, Ministry of National Education Republic of Indonesia.

REFERENCES

[1] Michael, Joel., what make physiology hard for student to learn? result of faculty survey, *Advances in Physiology Education* vol. 319, 2007, pp. 31-40.

[2] OECD, PISA 2012 Results: What *students Know and Can Do – Student Performance in Mathematics, Reading and Scienc*e, Vol.I, Revised edition, February 2014). PISA: OECD Publishing. DOI:http://dx.doi.org/10.1787/9789264201118-en.

[3] OECD, Program for International Students Assessment (PISA) 2015 Results in Focus,Science Competencies for Tommorow's World. PISA: OECD Publishing.

[4] Fzryah, Nurul, Pengaruh model pembelajaran terhadap hasil belajar ilmu pengetahuan alam sekolah dasar kota depok, *Jurnal Pendidikan Dasar*, vol. 6, 2015, pp. 49.

[5] D. F. Halpern, Thought and Knowledge: an introductional to critical thinking, Taylor & Francis, 2014.

[6] Sani, R. A, Pembelajaran berbasis hots (higher order thinking skills), Tsmart, 2019.

[7] S. Pradana, Pengembangan tes kemampuan berpikir kritis pada materi optik geometri untuk mahasiswa, *Jurnal Penelitian dan Evaluasi Pendidikan*, vol. 21, 2017, pp. 51.

[8] Lilis Nuryanti, Analisis kemampuan berpikir kritis peserta didik smp, *Jurnal Pendidikan*, vol. 3, (2018), pp. 155-158.

[9] Y. Anwar, S.Permata, Ermayanti, Measuring biology educations students' critical thinking skill, *Journal of Physics:* Conf. Series 1480, 2020. DOI:https://doi.org/10.1088/1742-6596/1480/1/012068

[10] L, Yuliati, Efektivitas bahan ajar ipa terpadu terhadap kemampuan berpikir tingkat tinggi peserta didik smp, *Jurnal Pendidikan Fisika Indonesia*, vol. 9, 2013, pp. 55-57

[11] Prof. DR. H. Dr. Pudji Muljono Djaali, Pengukuran dalam Bidang Pendidikan, Grasindo, 2008.

[12] Sugiyono, Metode penelitian (pendekatan kuantitatif, kualitatif, dan r&d), Alfabeta, 2016.

[13] R. H. Ennis, Goal critical thinking curriculum. Virginia: Association for Supervision and Curriculum Development (ASCD), 1985.

[14] Ipin Aripin, Pengembangan soal-soal pilihan ganda untuk mengukur kemampuan berpikir kritis siswa pada konsep sistem regulasi manusia untuk jenjang sma, *Jurnal Manggifera edu*, 2018, vol. 3, pp. 13-25.

[15] Zainal Arifin, Evaluasi pembelajaran, PT Remaja Rosdakarya. 2016.

[16] Suharsimi Arikunto, Dasar-dasar evaluasi pendidikan (edisi 2), Bumi Aksara, 2018.

[17] Anas, Sudijono, Pengantar Evaluasi Pendidikan, Rajawali Press, 2016.

[18] Y. Anwar., Enhancing the prospective biology teachers' Pedagogical Content Knowledge (PCK) through a peer coaching based model. *J of Physics*: Conf. Series, 1022 (2018), 012059. DOI: https://doi.org/10.1088/1742-6596/1022/1/012059