

Development of Authentic Assessment Instruments to Measure Students' Critical Thinking Skills on Human Digestive System Material

Elfira Yulia Sasahan¹, Masra Latjompoh², Lilan Dama³, Elya Nusantari⁴,
Margaretha Solang⁵

^{1,2,3,4,5}Biology Education, Postgraduate, State University of Gorontalo, Gorontalo, Indonesia

Corresponding Author: Elfira Yulia Sasahan

DOI: <https://doi.org/10.52403/ijrr.20250653>

ABSTRACT

This research is a development study aimed at producing an authentic assessment instrument on the human digestive system material that is valid, practical, and effective for measuring students' critical thinking skills. The development of the assessment instrument uses the Borg & Gall development model, with the resulting products including attitude assessment instruments, skill assessment instruments, student worksheets (LKPD), and critical thinking test instruments. The study was tested in Grade VIII at SMP Negeri Dayow and SMP Negeri Dumagin. Data collection instruments included validation sheets, observation sheets, student response questionnaires, and critical thinking skill test instruments. The results of the study show that: 1) The attitude assessment instrument, skill assessment instrument, critical thinking test, and LKPD were rated as highly valid by experts; 2) The practicality of the instruments is demonstrated by the implementation of learning by both teachers and students being in the "very good" category, as well as very positive student responses, indicating that the instruments are practical and ready to be used in the classroom; 3) In terms of effectiveness, the average scores of students' critical thinking skills in the limited and large-scale trials reached 62.50% and

64.09%, respectively, which fall into the high category, proving that the instruments are effective in measuring critical thinking ability. Write abstract/summary of your research paper here.

Keywords: Assessment Instruments, Authentic Assessment, Critical Thinking Skills

INTRODUCTION

Education in Indonesia faces significant challenges in developing students' critical thinking skills (Bachtiar, 2022). According to Zebua et al. (2024), critical thinking skills are essential for helping students not only memorize facts but also analyze, evaluate, and draw logical conclusions based on the available information. This aligns with the demands of the curriculum, which emphasizes competency-based learning and the development of 21st-century skills, including critical thinking, which every student must possess (Hanipah, 2023).

One of the essential topics in Natural Science (IPA) education is the human digestive system. This topic not only involves biological facts but also requires students to think critically, such as analyzing the relationships among organs within the digestive system and evaluating the impact of dietary patterns and disorders on digestive health (Andini & Qomariyah, 2022). Research by Sirajudin et al. (2022)

revealed that students' understanding of the human digestive system varied, with 14% demonstrating good conceptual understanding, 45% experiencing misconceptions, 12% making errors, and 29% failing to understand the concept. A study by Mu'arikha and Qomariyah (2020) further indicated that students experienced misconceptions in various subtopics of the digestive system, with varying percentages: the digestion process (65.50%), diseases or disorders of the digestive system (62.22%), digestive chemicals (58.75%), food substance testing (58.33%), food nutrition (57.22%), and the structure and function of digestive organs (43.75%). These data show a high prevalence of misconceptions and low levels of understanding, highlighting the urgent need to reinforce this material in classroom instruction.

Many existing assessment instruments remain at a low level and primarily focus on testing memory or factual understanding, thus failing to comprehensively reflect students' critical thinking skills (Hadi et al., 2022). This underscores the urgency of developing authentic assessment instruments that can evaluate critical thinking in a more contextual and applicable manner. According to Fityana et al. (2017), authentic assessments prioritize the measurement of skills relevant to students'

real-life experiences and provide a deeper understanding of their level of comprehension and abilities.

The use of authentic assessment instruments is crucial for measuring critical thinking skills, as they are expected to provide a more accurate depiction of students' ability to identify problems, analyze data, and draw evidence-based conclusions. Properly designed authentic assessments not only measure students' learning outcomes but also the thinking processes they undergo—an essential aspect of 21st-century education (Sumaryatun & Nugroho, 2016).

According to Ennis (as cited in Ardiyanti & Nuroso, 2021), critical thinking is the ability to think logically and reflectively in analyzing and evaluating information in a structured manner. Therefore, to measure critical thinking skills, assessment instruments must include aspects such as the ability to understand arguments, evaluate evidence, and draw valid conclusions. These skills are essential for comprehensively and applicatively understanding the topic of the human digestive system. Various theories have proposed indicators of critical thinking. One such theory, by Ennis (as cited in Affandy et al., 2019), categorizes critical thinking activities into five key indicators. The following are critical thinking indicators as proposed by Ennis:

Table 1. Indicators and Sub-Indicators of Critical Thinking Skills

No	Indicator	Sub-Indicator
1	Providing elementary clarification	Focusing on questions, analyzing questions, as well as posing and answering questions about an explanation or inquiry.
2	Building basic support	Considering the credibility of information sources, and observing and evaluating the results of observations.
3	Drawing inferences	Making deductions and evaluating their outcomes, making inductions and evaluating their outcomes, and making and evaluating value judgments.
4	Providing further explanation	Identifying terms and definitions of reasoning and dimensions, and identifying assumptions.
5	Developing strategies and tactics	Determining actions and engaging in interaction with others.

(Source: Affandy et al., 2019)

Research conducted by Hasan et al. (2020) highlights the importance of developing authentic assessment instruments to measure critical thinking skills, emphasizing their role in training and habituating students to

think critically so that they become accustomed to solving problems. Similarly, a study by Putri et al. (2019) reinforces the significance of developing assessment instruments for critical thinking skills as an

effort to foster and cultivate students' critical thinking abilities.

Therefore, the development of authentic assessment instruments aimed at measuring critical thinking skills in the topic of the human digestive system is expected to habituate students to think critically, construct their own knowledge, and relate it to everyday life.

Based on the above background, this study aims to develop a valid, practical, and effective authentic assessment instrument for measuring students' critical thinking skills in the topic of the human digestive system. It is hoped that the findings of this study will contribute to improving the quality of science education and serve as a reference for teachers in designing assessments grounded in critical thinking skills.

MATERIALS & METHODS

This study employed the Borg and Gall development model, which was modified into nine steps: needs analysis, data collection, product design, design validation, design revision, limited trial, product revision, large-scale trial, and final revision. The research was conducted during the 2024/2025 academic year at SMP Negeri Dayow and SMP Negeri Dumagin, located in South Bolaang Mongondow Regency. The instruments used in the study included validation sheets, learning implementation observation sheets, student activity observation sheets, student response questionnaires, and critical thinking essay tests.

The data analysis technique encompassed three main aspects: validity, practicality, and effectiveness. The analysis of validity and practicality was carried out using a Likert Scale, as presented in Table 2.

Table 2. Likert Scale Criteria

Score	Category
4	Very Good
3	Good
2	Fair
1	Poor

(Source: Sugiyono, 2013)

STATISTICAL ANALYSIS

The obtained score results were then used to calculate the percentage of validity using the following formula:

$$\text{Validity Percentage} = (\text{Total Score}) / (\text{Maximum Score}) \times 100\%$$

The results of the expert review sheets were used to determine the feasibility of the assessment tools. The obtained validity percentage is presented in the following table. Furthermore, the average score was converted into qualitative values based on the Likert scale using the following assessment criteria:

Table 3. Validation Criteria

Score Interval	Category
81% – 100%	Very Valid
60% – 80%	Valid
41% – 61%	Fairly Valid
21% – 40%	Invalid
0% – 20%	Very Invalid

(Source: Sugandi et al., 2021)

The authentic assessment product developed was deemed feasible for testing if the validation level from the validators' assessments was at least in the "Valid" category.

Practicality Analysis

Analysis of Learning Implementation

Teacher classroom management was assessed using a checklist and a 1–4 rating scale based on Likert criteria. The percentage of classroom management by the teacher was calculated using the following formula:

$$\text{Percentage} = (\text{Total Score}) / (\text{Maximum Score}) \times 100\%$$

Table 4. Conversion of Teacher Activity Percentage in Learning Process

Percentage (%)	Category
81% – 100%	Very Good
60% – 80%	Good
41% – 61%	Fair
21% – 40%	Poor
0% – 20%	Very Poor

(Source: Dayelma et al., 2019)

Analysis of Student Activity

Student activity during authentic assessment was observed comprehensively. The data from these observations were analyzed and described. The percentage was calculated as:

$$\text{Percentage} = (\text{Total Score}) / (\text{Maximum Score}) \times 100\% \text{ (Sugandi et al., 2021)}$$

Student activity was categorized as practical if at least 80% of the activities planned in the Lesson Plan (RPP) were successfully carried out by the students.

Table 5. Student Activity Assessment Criteria

Percentage	Category
0%–20%	Not Practical
21%–40%	Less Practical
41%–60%	Fairly Practical
61%–80%	Practical
81%–100%	Very Practical

(Source: Irsalina & Dwiningsih, 2018)

Analysis of Student Responses

Student responses were analyzed through questionnaires distributed during the trial. The scoring was based on the Likert scale. The percentage of student responses was calculated using the formula:

$$\text{Percentage of Student Responses} = (\text{Total Score}) / (\text{Maximum Score}) \times 100\% \text{ (Source: Sugandi et al., 2021)}$$

The questionnaire data were then analyzed descriptively and interpreted using the criteria shown in Table 6.

Table 6. Criteria for Student Response Questionnaire

Percentage	Category
81% – 100%	Very Good
60% – 80%	Good
41% – 61%	Fair
21% – 40%	Poor
0% – 20%	Very Poor

(Source: Dayelma et al., 2019)

Effectiveness Analysis

Critical Thinking Skills Test Instrument

The results of the critical thinking skills test were analyzed to determine the effectiveness of the developed assessment

product. The steps of the analysis were as follows:

- 1) Assigning scores to each student response based on a predetermined rubric.
- 2) Calculating the total score for each indicator for every student.
- 3) Determining the percentage of critical thinking skills for each student using the formula:
Percentage of Critical Thinking Skills = (Student's Total Score per Indicator) / (Maximum Score per Indicator) × 100%

The percentage (P) of students who gave accurate and complete answers for each indicator was then converted using the following criteria:

Table 7. Critical Thinking Criteria

Percentage	Category
80 < P ≤ 100	Very High
60 < P ≤ 80	High
40 < P ≤ 60	Moderate
20 < P ≤ 40	Low
0 < P ≤ 20	Very Low

(Source: Solikhin & Fauziah, 2021)

RESULTS

This development research was conducted to assess three main aspects. First, to evaluate the validity of the developed products, including the authentic assessment tools (for attitude and skills), student worksheets (LKPD), and critical thinking test instruments (for cognitive assessment). Second, to assess the practicality of the instruments based on observations of teacher activities and student engagement. Third, to evaluate the effectiveness of the instruments based on student response questionnaires and critical thinking test results.

Based on the validation results of the attitude and skills assessment instruments, the critical thinking test instruments, and the LKPD, all components were categorized as valid, although several suggestions and corrections from the validators required revisions. The initial revisions were made

based on the feedback received during the validation phase, prior to the limited-scale trial implementation. The attitude and skills assessment instruments, the critical thinking test instruments, and the LKPD were revised in response to the feedback and suggestions provided.

Product Trial Phase

The practicality level of the developed assessment instruments was determined through observation sheets evaluating the implementation of the learning process (teacher activities), student activities, and student response observation sheets.

Analysis of Observation Results on Learning Implementation

Based on the observation results during the limited-scale trial, the scores related to the implementation of learning by the teacher varied across each meeting. The observer-assigned scores indicated the following: the first meeting obtained a score of 90 with a percentage of 93.75%; the second meeting scored 82 with a percentage of 93.18%; the third meeting scored 62 with a percentage of 86.11%; the fourth meeting scored 63 with a percentage of 87.50%; and the fifth meeting scored 53 with a percentage of 88.33%. These observer scores were used to determine the category of learning implementation. The assessment results are presented in Figure 1.

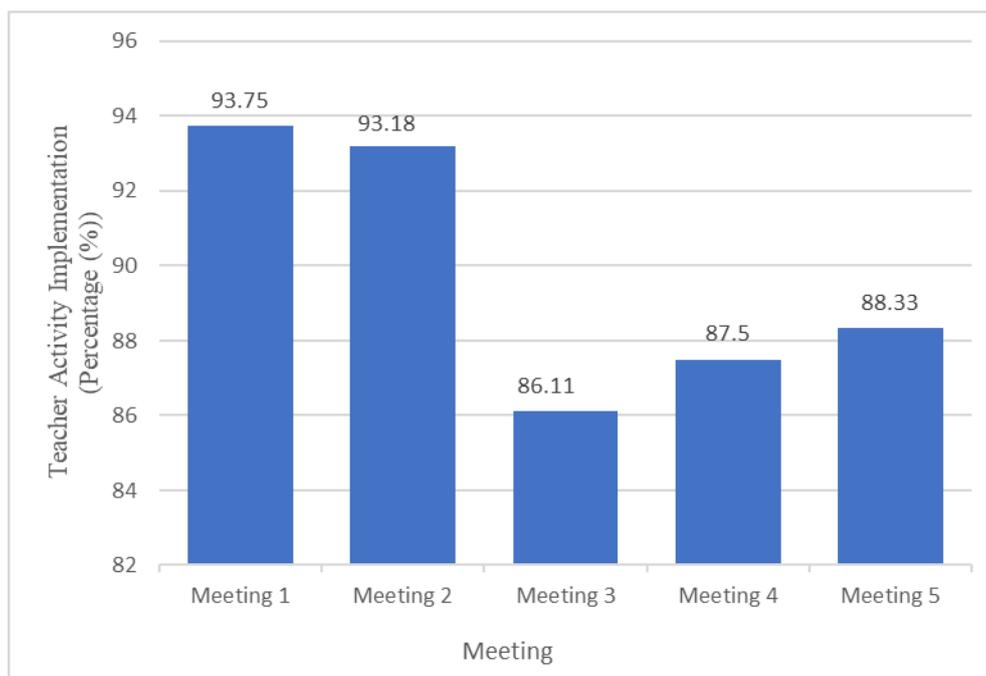


Figure 1. Assessment Results of Teacher Activity Implementation During the Learning Process in the Limited-Scale Trial

Based on the assessment by two observers of the implementation of teacher activities during the learning process, it was found that the teaching activities fell within the "very good" criteria, with percentage ranges between 81%–100% during the large-scale

trial. These results are presented in Figure 2. These findings indicate that the implementation of teacher activities based on the instructional design in the classroom was carried out very effectively.

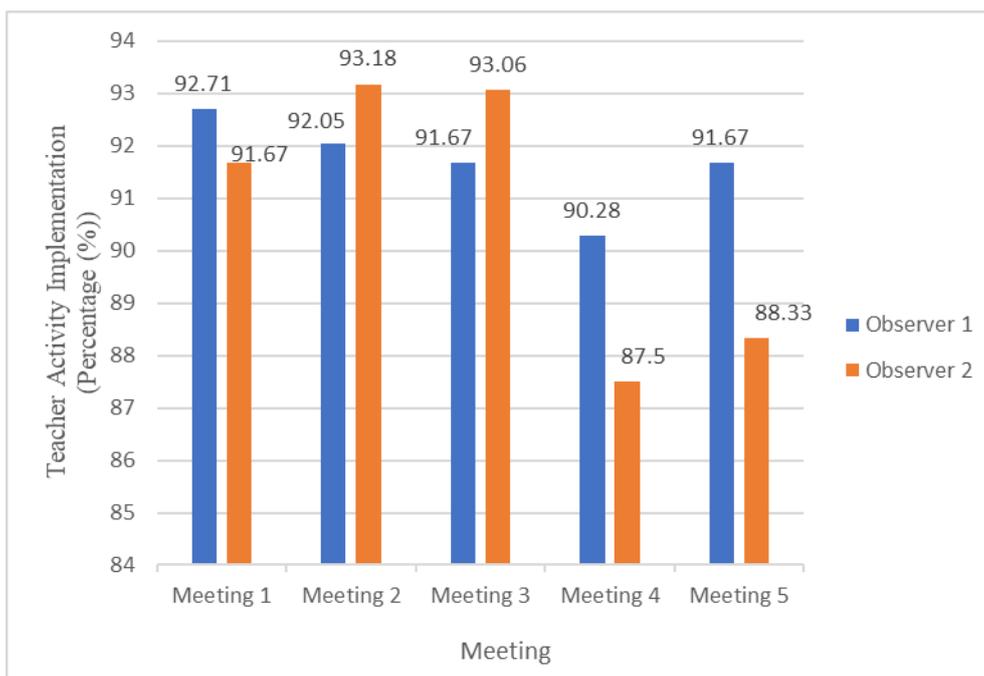


Figure 2. Data Analysis of Teacher Activity Implementation During the Learning Process in the Large-Scale Trial

Analysis of Student Activity Assessment Results During the Learning Process

Student activity was assessed by observers during the limited-scale trial involving Grade VIII students of SMP Negeri Dayow. The goal was to evaluate the practicality of the developed learning tools. Overall, the assessment results showed varying total scores across the five meetings. In the first meeting, a score of 56 was obtained with a

percentage of 87.50%; the second meeting scored 53 with a percentage of 88.33%; the third meeting scored 60 with 88.24%; the fourth and fifth meetings both scored 52 with a percentage of 86.67%. Based on the 81%–100% percentage range, the practicality of student activities is categorized as very practical. These results are illustrated in Figure 3.

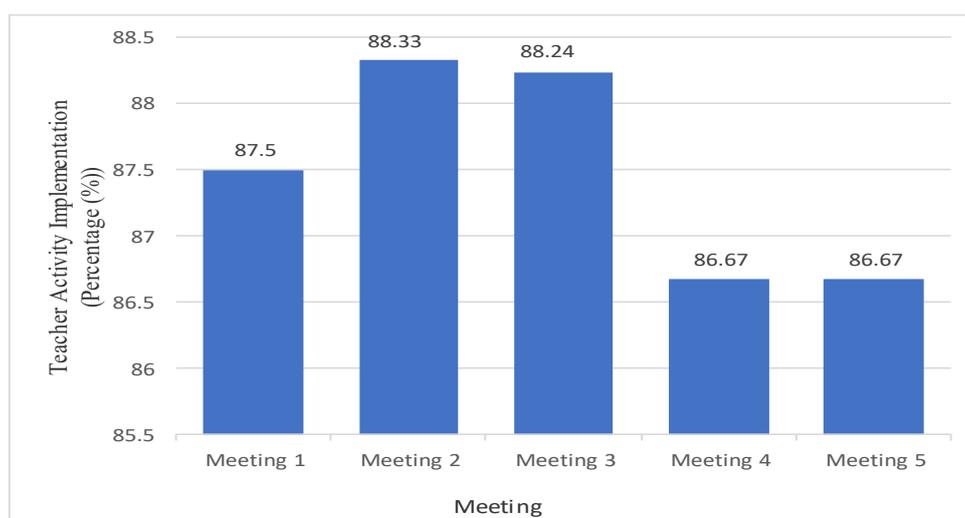


Figure 3. Student Activity Assessment Results During the Learning Process in the Limited-Scale Trial

Student activity assessment during the large-scale trial involved Grade VIII students of SMP Negeri Dumagin, aiming to evaluate the practicality of the learning tools.

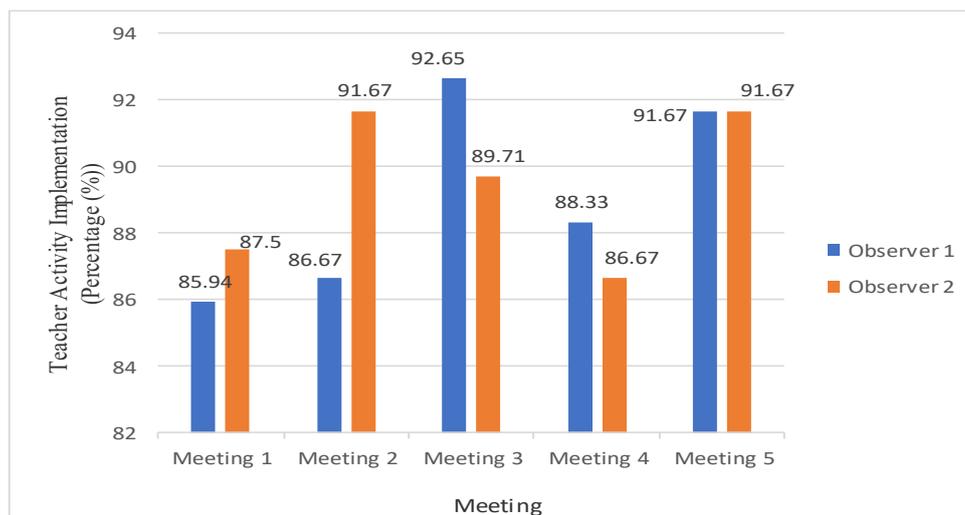


Figure 4. Student Activity Assessment Results During the Learning Process in the Large-Scale Trial

Overall, based on the practicality criteria assessed by two science teachers, the scores varied in each meeting. The first observer recorded: 55 (85.94%) for the first meeting, 52 (86.67%) for the second, 63 (92.65%) for the third, 53 (88.33%) for the fourth, and 55 (91.67%) for the fifth. Meanwhile, the second observer recorded: 56 (87.50%) for the first meeting, 55 (91.67%) for the second, 61 (89.71%) for the third, 52 (86.67%) for the fourth, and 55 (91.67%) for the fifth. Based on these results, the practicality criteria for all five meetings

ranged between 81%–100%, categorized as very practical. The complete results are shown in Figure 4.

Analysis of Student Response Questionnaire Results

The student response questionnaire was administered to obtain data on student responses during the learning process. It consisted of six questions answered by students based on their individual perceptions. The results of this assessment are shown in Figure 5.

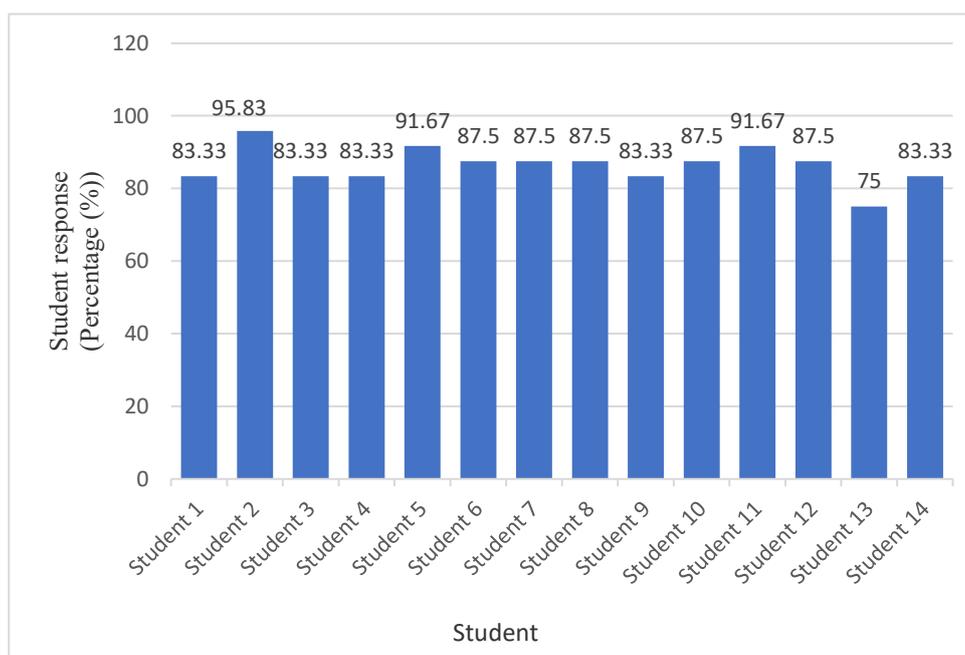


Figure 5. Results of Student Response Questionnaire in the Limited Trial

The student response data from the large-scale trial, as shown in Figure 6, indicate a percentage range of 81%–100%, falling under the *very good* category. Based on the analysis of the learning implementation data

(teacher activities), student activity, and student responses—which all show positive reactions to the learning process—it can be concluded that the developed learning tools are *highly practical* for classroom use.

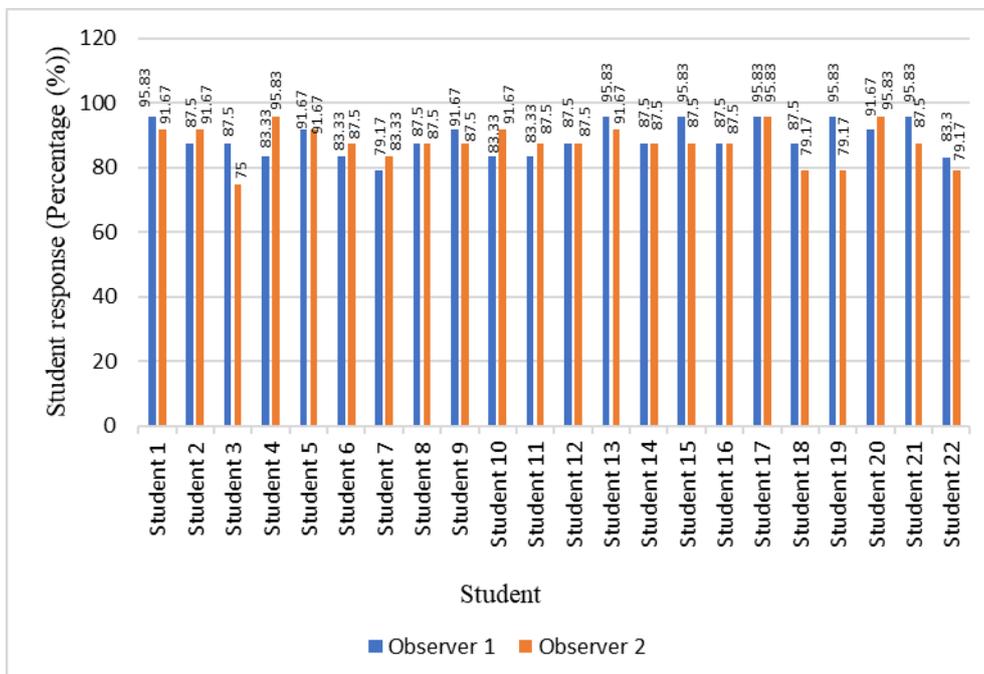


Figure 6. Student Response Questionnaire Results in the Large-Scale Trial

Analysis of the Effectiveness of Authentic Assessment Instruments to Measure Students' Critical Thinking Skills on the Digestive System Topic

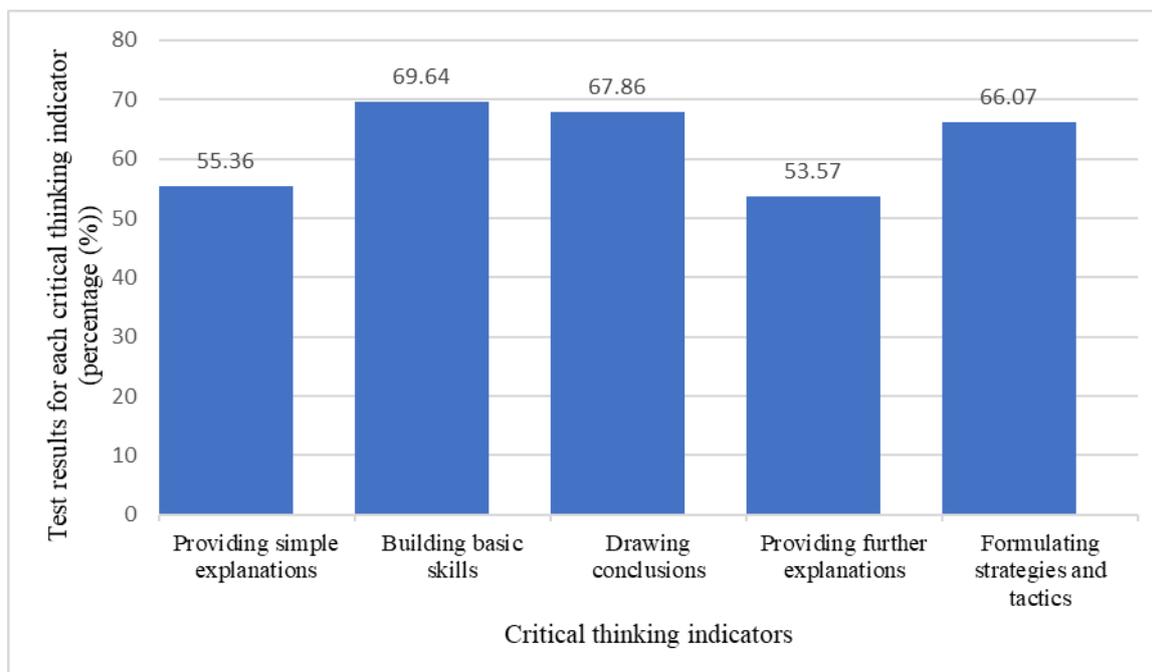


Figure 7. Results of Each Critical Thinking Indicator in the Limited Trial

Based on the data from the critical thinking skills test conducted during the limited trial phase, the achievement of each critical thinking indicator is as follows: elementary clarification reached 55.36%, categorized as moderate; basic support reached 69.64%,

categorized as high; inference reached 67.86%, categorized as high; advanced clarification reached 53.57%, categorized as moderate; and strategy and tactics reached 66.07%, categorized as high.

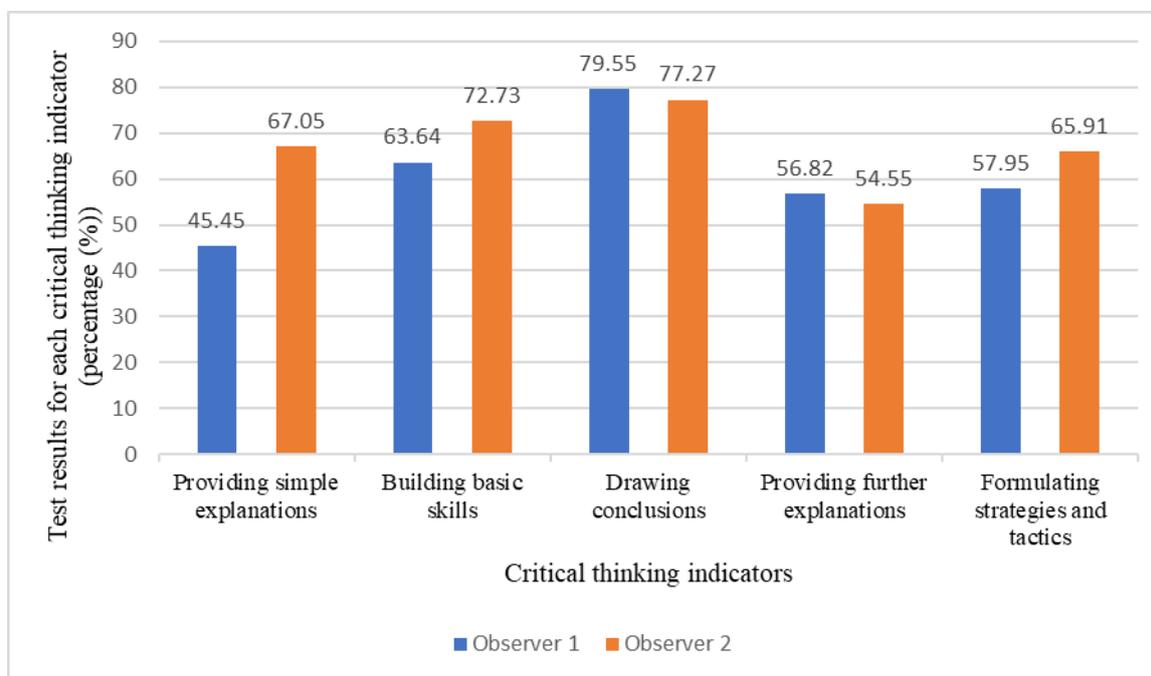


Figure 8. Results of Each Critical Thinking Indicator in the Large-Scale Trial

Effectiveness data were obtained from the analysis of a critical thinking skills test instrument consisting of five items based on Ennis's critical thinking indicators. This instrument was administered to 44 eighth-grade students at SMP Negeri Dumagin, divided into two classes, during the large-scale trial.

The test results for Class VIII A showed the following indicator achievements: elementary clarification at 45.45%, basic support at 63.64%, inference at 79.55%, advanced clarification at 56.82%, and strategy and tactics at 57.95%.

Meanwhile, the results for Class VIII B were: elementary clarification at 67.05%, basic support at 72.73%, inference at 77.27%, advanced clarification at 54.55%, and strategy and tactics at 65.91%. The results for each indicator of critical thinking skills are presented in Figure 8.

DISCUSSION

Based on the results of the research analysis, this section discusses the research questions and objectives related to the validity, practicality, and effectiveness of the authentic assessment tools developed to measure students' critical thinking skills in the digestive system topic. Assessment results from three validators indicated that all instructional tools were categorized as highly valid, although some suggestions for improvement were provided for the authentic assessment instruments. The attitude assessment instrument received an average percentage of 91.67%, categorized as very valid. The skills assessment instrument obtained an average percentage of 92.36%, also categorized as very valid. The critical thinking skills test achieved an average score of 91.67%, indicating a very valid classification. Meanwhile, the student activity sheet obtained an average score of 94.44%, also falling within the very valid

category. The criteria for validity were determined based on the 81–100% percentage range according to the Likert scale qualitative conversion (Sugiyono, 2013). These findings are supported by Yudha (2019), who developed an authentic assessment instrument to measure students' critical thinking skills and concluded that the instrument was appropriate for use. Similarly, Mukhlisa (2023) stated that a good test instrument must undergo validity testing prior to implementation.

Instructional tools are considered practical when they meet three main indicators: the implementation of the learning process (teacher activities), student activities, and students' responses, all of which should fall under at least the "good" category. This aligns with the view of Alik et al. (2023), who assert that the practicality of instructional tools can be observed from the implementation of learning activities, student engagement, and student responses. These findings are further reinforced by Hapsari and Gularso (2024), whose research demonstrated that instruments based on critical thinking skills were categorized as highly practical for use in classroom instruction.

The implementation score for teacher activities during the learning process, as assessed by a science teacher observer during the limited trial, yielded an overall average of 89.78% across five sessions. In the broader trial, evaluated by two observers, the average was 91.21%. Both results are categorized as very good. Observations comprehensively covered the introduction, core activities, and conclusion stages. These results indicate that the learning implementation was carried out optimally and in accordance with the prepared lesson plans. This finding supports Makhrus et al. (2018), who emphasized the importance of alignment between lesson planning and classroom implementation, including teacher engagement during the introductory, core, and concluding phases which reflect reflection and content reinforcement.

Regarding student activities, the average implementation score was 87.47% in the limited trial and 89.25% in the wider trial, both falling within the very practical category. These results suggest that the implementation of the instructional tools was highly effective overall. The findings indicate that the learning process was conducted effectively, in accordance with the lesson plans, and had a positive impact on students' active engagement. This success also demonstrates that the instructional tools were implementable and contributed to achieving learning objectives optimally. This is consistent with the findings of Fatmasuci (2017), who stated that the practicality of instructional tools is reflected in positive teacher evaluations and high levels of student engagement during instruction.

Based on the data analysis, the average student response in the limited trial was 86.31%, and in the broader trial, it was 88.45%, both categorized as very good. This indicates that students gave positive responses to the developed assessment tools in terms of content clarity, engaging activities, and the ability to train critical thinking skills. The high response scores suggest that the tools met student needs and supported the development of critical thinking skills.

The effectiveness of the instructional tools in this study was determined through the results of a critical thinking test developed based on Ennis' indicators. The data showed that the average overall critical thinking test score was 62.50% in the limited trial and 64.09% in the broader trial, both of which fall into the high category. This indicates that students' critical thinking skills were well-developed. When examined using Ennis' critical thinking indicators, students demonstrated relatively strong performance in several areas.

For the elementary clarification aspect, students in Class VIII A scored 45.45%, categorized as moderate. This aligns with Rahmawati et al. (2016), who also found moderate scores in this aspect. Arum (2014)

in Ramadhani et al. (2021) explained that low scores in elementary clarification stem from students' unfamiliarity with focusing questions and analyzing problems before answering. Meanwhile, students in Class VIII B scored 67.05%, which falls under the high category, indicating that they were able to understand and identify problems and generate relevant questions. These findings are consistent with Setyawan and Patonah (2022), who reported high scores in elementary clarification. However, Amarila et al. (2021) still found moderate results in this aspect, attributed to the transitional cognitive development stage of junior high school students, who are still progressing toward higher-order thinking skills.

In the basic support aspect, students achieved scores of 63.64% and 72.73%, both categorized as high. These results indicate that students could provide reasons and simple evidence to support their opinions or answers. This aligns with the findings of Cahyani et al. (2022), who also reported high scores in this indicator. Fitriani et al. (2018) emphasized that basic support is closely related to problem-based learning, which trains students to engage in scientific inquiry, evaluate source credibility, conduct investigations, and assess findings.

For the inference aspect, students scored 79.55% and 77.27%, both categorized as high. This indicates that students were able to draw conclusions from the given information and logically connect concepts. These findings are consistent with the studies of Purwanti (2023) and Nurdiyanti (2025), which both reported high levels of inference ability among students.

In the advanced clarification aspect, students scored 56.82% and 54.55%, both falling into the moderate category. This shows that while students were beginning to develop skills in evaluating arguments and critically comparing different opinions, some weaknesses remained and required further reinforcement. These results are consistent with Suharyani and Siswanto (2022) and Safitri and Hidayanti (2024),

who also found that students' advanced clarification skills were in the moderate category. Ardiansyah (2020) in Solihah et al. (2023) identified three factors influencing critical thinking skills: (1) parental education level, (2) social environment, and (3) peer support.

For the strategies and tactics aspect, students in Class A scored 57.95%, categorized as moderate. This aligns with the findings of Rahmawati et al. (2016) and Wijayanti and Siswanto (2020), who also reported moderate performance in this aspect. In contrast, students in Class B scored 65.91%, which falls into the high category. This suggests that most students in this group were able to determine appropriate steps for solving problems. These results are in line with the findings of Cahyani et al. (2022), who also found high scores in this aspect. According to Ramadhani et al. (2021), strong strategy and tactics skills enable students to manage, adjust, modify, or improve their thinking processes and determine the most appropriate actions when facing problems. Conversely, low scores in this area are often due to students' lack of awareness regarding real-world issues around them, resulting in underdeveloped strategic and tactical thinking skills.

CONCLUSION

Based on the results of the research and development of authentic assessment tools to measure students' critical thinking skills in the human digestive system topic, three main conclusions can be drawn. First, in terms of validity, the instruments—comprising attitude assessment, skills assessment, critical thinking tests, and student worksheets (LKPD)—were rated as highly valid by experts, demonstrating appropriateness in content, construct, and language. Second, regarding practicality, the implementation of learning by both teachers and students fell into the "very good" category, and students' responses were also highly positive, indicating that the tools are practical and ready for classroom use. Third,

in terms of effectiveness, the average scores of students' critical thinking skills during the limited and broader-scale trials were 62.50% and 64.09%, respectively, both of which fall into the high category, proving that the assessment tools are effective in measuring critical thinking abilities.

The following recommendations are proposed: (1) the developed product should be disseminated to teachers to encourage widespread use and ongoing refinement; (2) similar assessment tools should be developed for all science topics in order to comprehensively enhance students' critical thinking skills; and (3) future development of authentic assessment tools should ensure the fulfillment of the criteria of validity, practicality, and effectiveness to be considered suitable for instructional use.

Declaration by Authors

Acknowledgement: None

Source of Funding: None

Conflict of Interest: No conflicts of interest declared.

REFERENCES

1. Affandy, H., Aminah, N. S., & Supriyanto, S. Analisis keterampilan berpikir kritis siswa pada materi fluida dinamis di SMA Batik 2 Surakarta. *Jurnal Materi dan Pembelajaran Fisika*. 2019; 9(1):25-33.
2. Alik, I. P., Paramata, D. D., & Supartin, S. Analisis kepraktisan perangkat pembelajaran model discovery learning berbantuan media ispring suite pada materi fluida statis. *Jurnal Pendidikan MIPA*. 2023; 13(1):46-53.
3. Amarila, S., Raula, B. S., & Saptono, S. Analisis Kemampuan Berpikir Kritis Siswa Pada Pembelajaran IPA Terpadu Tema Lingkungan. *Improvement: Jurnal Ilmiah Untuk Peningkatan Mutu Manajemen Pendidikan*. 2021; 8.
4. Andini, A. R., & Qomariyah, N. Validasi E-Book tipe Flipbook materi sistem pencernaan manusia berbasis PBL untuk melatih kemampuan berpikir kritis siswa SMA. *Berkala Ilmiah Pendidikan Biologi (BioEdu)*. 2022; 11(2):330-340.
5. Ardiyanti, F., & Nuroso, H. Analisis tingkat keterampilan berpikir kritis siswa kelas XI MIPA dalam pembelajaran Fisika. *Karst: Jurnal Pendidikan Fisika dan Terapannya*. 2021; 4(1):21-26.
6. Bachtiar. Tantangan dan Strategi Penerapan Berpikir Kritis pada Pembelajaran Online: Kajian Pustaka. *JP2SD (Jurnal Pemikiran dan Pengembangan Sekolah Dasar)*. 2022; 10(2):145-159
7. Cahyani, P. A. D., Setiono, S., & Juhanda, A. Analisis Kemampuan Berpikir Kritis Siswa Sma Melalui Model Inquiry Based Learning Berbantuan Media Wayang Sukuraga Pada Materi Sistem Indra:(Analysis of Critical Thinking Ability of High School Students Through an Inquiry-Based Learning Model Assisted by Wayang Sukuraga Media on The Material). *BIODIK*. 2022; 8(4):116-124.
8. Dayelma, Y., Octarya, Z., & Refelita, F. Hubungan literasi sains dengan keterampilan berpikir kritis siswa pada materi ikatan kimia. *JEDCHEM (Journal Education and Chemistry)*. 2019; 1(2):72-78.
9. Fatmasuci, F. W. Pengembangan perangkat pembelajaran berbasis masalah berorientasi pada kemampuan komunikasi dan prestasi belajar matematika siswa SMP. *Jurnal Riset Pendidikan Matematika*. 2017; 4(1):32-42.
10. Fitriani, R. Peningkatan kemampuan berpikir kritis peserta didik SMP kelas VII melalui pembelajaran berbasis masalah. *Bioedusia: Jurnal Pendidikan Biologi*. 2018; 3(1):8-14.
11. Fityana, I. N., Sarwanto, S., & Sugiyarto, S. Pengembangan Instrumen Penilaian Autentik Pada Pembelajaran IPA Berbasis Proyek Untuk Siswa SMP/MTs Kelas VII. *Jurnal Materi dan Pembelajaran Fisika*. 2014; 7(2):23-27.
12. Hadi, M. S., Zarnuji, A., Ikhwanuddin, M., & Asy'ari, H. Efektivitas Instrumen Penilaian Pendidikan Agama Islam Untuk Meningkatkan Kemampuan Berpikir Kritis Dan Kreatif Peserta Didik Di Smk Muhammadiyah Mlati. *Berkala Ilmiah Pendidikan*. 2022; 2(2):57-64.

13. Hanipah, S. Analisis kurikulum merdeka belajar dalam memfasilitasi pembelajaran abad ke-21 pada siswa menengah atas. *Jurnal Bintang Pendidikan Indonesia*. 2023; 1(2):264-275.
14. Hapsari, Y. H. P. & Gularso, D. Pengembangan Kepraktisan Instrumen Penilaian Berbasis Keterampilan Berpikir Kritis Siswa Kelas IV SD Negeri Wirosaban. *Jurnal Pendidikan Matematika*. 2024; 8(2).
15. Hasan, S. W., Auliah, A., & Herawati, N. Pengembangan Instrumen Penilaian Kemampuan Berpikir Kritis Siswa SMA. *Chemistry Education Review*. 2020; 3(2):185-1939.
16. Irsalina, A., & Dwiningsih, K. Analisis kepraktisan pengembangan lembar kegiatan peserta didik (LKPD) berorientasi blended learning pada materi asam basa. *Jurnal Kimia dan Pendidikan Kimia*. 2018; 3(3):171-182.
17. Makhrus, M., Harjono, A., Syukur, A., Bahri, S., & Muntari, M. Analisis Rencana Pelaksanaan Pembelajaran (RPP) terhadap Kesiapan Guru sebagai "Role Model" Keterampilan abad 21 pada Pembelajaran IPA SMP. *Jurnal Penelitian Pendidikan IPA*. 2019; 5(1).
18. Mu'arikha, M. A., & Qomariyah, N. Analisis Miskonsepsi Materi Sistem Pencernaan Dengan Menggunakan Three-Tier Test Pada Siswa Kelas XI SMA. *Berkala Ilmiah Pendidikan Biologi (BioEdu)*. 2020; 9(2):199-206.
19. Mukhlisa, N. Validitas Tes. *Jurnal Pendidikan dan Pembelajaran Sekolah Dasar*. 2023; 2(1).
20. Purwanti, E. Analisis Keterampilan Berpikir Kritis Siswa SMP pada Materi Sistem Kehidupan Tumbuhan. *Bioscientist: Jurnal Ilmiah Biologi*. 2023; 11(1):864-873.
21. Putri, O. D., Nevrita, N., & Hindrasti, N. E. K. Pengembangan instrumen penilaian keterampilan berpikir kritis siswa SMA pada materi sistem pencernaan. *BIOEDUKASI: Jurnal Pendidikan Biologi*. 2019; 10(1):14-27.
22. Rahmawati, I., Hidayat, A., & Rahayu, S. Analisis keterampilan berpikir kritis siswa SMP pada materi gaya dan penerapannya. *Pros. Semnas Pend. IPA Pascasarjana UM*. 2016; 1(1):1112-1119.
23. Ramadhani, P., Mustafid, F.Z.A & Karenina, A. Profil Keterampilan Berpikir Kritis Siswa dalam Pembelajaran Sistem Gerak pada Manusia di Madrasah Tsanawiyah di Ponorogo. *Proceeding of Integrative Science Education Seminar*. 2021; 1:340-344
24. Safitri, L & Hidayati, N. Profil Tingkat Kemampuan Berpikir Kritis Siswa pada Materi Pencemaran Lingkungan Kelas VII SMPN 3 Lirik Tahun Ajar 2023/2024. *Journal of Social Humanities and Education*. 2024; 3(3).
25. Setyawan, J & Patonah, S. Profil Kemampuan Berpikir Kritis Siswa Pada Materi Senyawa Hidrokarbon Melalui Pendekatan Steam Di Kelas XI IPA SMA Negeri 2 Mranggen. *Jurnal Inovasi Pembelajaran di Sekolah*. 2022; 3(2):166-180.
26. Sirajudin, H. M., Hadi, P. W., Wahyuni, A. E., Rosidi, I & Yasir, M. Analisis Miskonsepsi Siswa Pada Materi Sistem Pencernaan Menggunakan Instrumen Four-Tier Test. *Jurnal Natural Science Educational Research*. 2022; 5(2).
27. Solihah, B., Hadi, W.P., Qomaria, N., Tamam, B & Rakhmawan, A. Identifikasi Kemampuan Berpikir Kritis Siswa Pada Materi Pencemaran Lingkungan. *Jurnal Natural Science Educational Research*. 2023; 6(1).
28. Sugandi, A. I., Sofyan, D & Maesaroh, S. Pengembangan Perangkat Pembelajaran Menggunakan Deduktif Induktif Berbantuan Geogebra Dalam Meningkatkan Kemampuan Berpikir Kreatif Siswa Pada Masa Pandemi. *Jurnal Pembelajaran Matematika Inovatif*. 2021; 4(1).
29. Sugiyono. *Metode Penelitian Kuantitatif, Kualitatif dan R&D*. ALFABETA. 2013.
30. Suharyani, L. A & Siswanto, J. Analisis Kemampuan Berpikir Kritis Siswa Kelas X SMA Untuk Pengembangan Modul Ajar Konsep Perubahan Lingkungan Melalui Penerapan Education For Sustainable Development (ESD). *Journal of Comprehensive Science*. 2022; 1(5).

31. Solikhin, M & Fauziah, A. N. M. Analisis Kemampuan Berpikir Kritis Siswa SMP Pada Pelajaran Ipa Saat Pembelajaran Daring Selama Pandemi Covid-19. *Jurnal Pendidikan Sains*. 2021; 9(2):188-192.
32. Sumaryatun, R. A & Nugroho, S. E. Pengembangan Instrumen Penilaian Autentik Kurikulum 2013 Berbasis Literasi Sains Pada Materi Bioteknologi. *Journal of Primary Education*. 2016; 5(1).
33. Wijayanti, R. & Siswanto, J. Profil Kemampuan Berpikir Kritis Siswa SMA pada Materi Sumber-sumber Energi. *Jurnal Penelitian Pembelajaran Fisika*. 2020; 11(1):109-113
34. Yudha, R. P. Pengembangan Instrumen Asesmen Otentik Tes Keterampilan Berpikir Kritis Pada Pembelajaran Matematika. *Jurnal Matematika Dan Pendidikan Matematika*. 2019; 4(1):9-20.
35. Zebua, J. Y., Zega, Y & Telaumbanua, Y. N. Analisis Kemampuan Berpikir Kritis Siswa dalam Menyelesaikan Soal Matematika. *Jurnal Kependidikan*. 2024; 13(1).

How to cite this article: Elfira Yulia Sasahan, Masra Latjompoh, Lilan Dama, Elya Nusantari, Margaretha Solang. Development of authentic assessment instruments to measure students' critical thinking skills on human digestive system material. *International Journal of Research and Review*. 2025; 12(6): 457-470. DOI: <https://doi.org/10.52403/ijrr.20250653>
