JPPIPA 10(8) (2024)

Jurnal Penelitian Pendidikan IPA

Journal of Research in Science Education

http://jppipa.unram.ac.id/index.php/jppipa/index

Fostering Superior Characters: Development of Innovative Instruments for Critical Reasoning and Independent Character in the Realm of Science Topic

Esty Setyo Utaminingsih[1*], Ellianawati[2], Muhammad Habib Ramadhani[3], Maria Ayu Puspita[4], Sri Sumartiningsih[5], Muhamad Subhi Apriantoro[6], Aldi Ihsandi[7]

[1] Department of Madrasah Ibtida'yah Teacher Education, Sekolah Tinggi Agama Islam Muhammadiyah Blora, Blora, Indonesia.
[2] Department of Physics, Faculty of Mathematics and Natural Science, Universitas Negeri Semarang, Semarang, Indonesia.
[3] Department of Elementary School Teacher Education, Faculty of Teacher Training and Education, Universitas Bengkulu, Indonesia.
[4] Postgraduate School, Universitas Negeri Semarang, Semarang, Indonesia.
[5] Department of Sports Science, Faculty of Sports Science, Universitas Negeri Semarang, Semarang, Indonesia.
[6] Department of Sharia Economic Law Study Program, Faculty of Islamic Religion, Universitas Muhammadiyah Surakarta, Indonesia.
[7] Sekolah Tinggi Agama Islam Negeri Bengkalis, Bengkalis, Indonesia.

Received: January 24, 2024
Revised: June 14, 2024
Accepted: August 25, 2024
Published: August 31, 2024

Corresponding Author:
Esty Setyo Utaminingsih
estyutami@staimuhblora.ac.id

DOI: 10.29303/jppipa.v10i8.7053

Abstract: The purpose of this study was to develop a critical reasoning and independent character instrument on the topic of human blood circulation. The method used was research and development (R&D) with a 4D model (Define, Design, Develop, and Disseminate). The results showed that the validity value of the content of critical reasoning instruments obtained a final value of 0.97, while independent instruments obtained a value of 1, which was perfect. Judging from empirical validity, vital and independent reasoning instruments were declared valid because they had a value above 0.355 from 31 students tested. Critical and independent reasoning instruments were declared reliable. The Alpha Cronbach value >0.70 had good distinguishing power because the value of the Corrected Item-Total Correlation did not exceed the Alpha Cronbach value. Developing assessment instruments for a critical and independent character in human blood circulation for fifth-grade elementary students was successful, demonstrating high validity and reliability, and held the significant potential to enhance science education through technology integration.

Keywords: Critical reasoning; Human Circulation; Independent; Instruments

Introduction

In a dynamic and future-oriented educational environment (Gulson & Webb, 2023), the ability to reason critically and independently is a fundamental aspect that learners must develop (Fazryn et al., 2023). It is also conveyed in Regulation of the Minister of Education and Culture Number 22 of 2020 to create superior Human Resources (HR) with character. The regulation is outlined through the establishment of 6 (six) character dimensions in the Pancasila Student Profile that must be developed among current students (Muqit et al., 2023), of which two are critical reasoning and independent (Ministry of Education and Culture, 2020).

The ability to reason critically is one of the critical competencies that must be developed in students (Hayati & Setiawan, 2022). The ability to reason critically and think analytically, critically (Slam, 2021) and logically form a robust intellectual foundation at every

How to Cite:
Utaminingsih, E. S., Ellianawati, Ramadhani, M. H., Puspita, M. A., Sumartiningsih, S., Apriantoro, M. S., & Ihsandi, A. (2024). Fostering Superior Characters: Development of Innovative Instruments for Critical Reasoning and Independent Character in the Realm of Science Topic. *Jurnal Penelitian Pendidikan IPA*, *10*(8), 4650–5661. https://doi.org/10.29303/jppipa.v10i8.7053

stage of education (Jafari & Keykha, 2023; Jokhan et al., 2022). This character must be strengthened at the primary school education level (Utaminingsih, Puspita, et al., 2023), considering that students in this period are developing their fundamental cognitive foundations (Ernawati & Rahmawati, 2022). They engage in a highly dynamic and critical learning process to develop their thinking (Nursalam & Suardi, 2022). Learners begin to explore the world of knowledge in a more systematic way (Utaminingsih, Ellianawati, et al., 2023).

The formation of critical reasoning character in students is essential (Utaminingsih, Ihsandi, et al., 2023), especially in science education in today's information age (Rumtini et al., 2022). Science learning presents complex and exciting content to explore (Suminar, 2022). In the context of science education in elementary schools, the topic of human blood circulation provides a rich opportunity to hone these skills (Windfall et al., 2022). The human circulatory system is an important topic that provides a rich platform for developing critical reasoning abilities (Siswati, 2022). However, educators' most prominent challenge is evaluating students' ability to use critical reasoning to understand and apply this concept.

The urgency of developing an instrument to evaluate critical reasoning and independence in elementary school students, particularly fifth graders, on human blood circulation must be considered within the context of education aimed at building superior and globally competitive human resources. The Pancasila Student Profile dimensions, which consist of six key elements, emphasize critical reasoning and independence as foundational attributes that students must possess. Critical reasoning is essential for cultivating a generation capable of evaluating information, solving problems, and making decisions grounded in thorough analysis in an era increasingly shaped by global complexity. Meanwhile, independence equips students with the ability to manage their learning, take initiative, and develop strong self-discipline skills that are indispensable in today's competitive global landscape.

The research to be developed seeks to address these challenges by providing an empirically validated instrument that can measure and foster the enhancement of these two traits among fifth-grade students, specifically focusing on the human circulatory system. The importance of developing this instrument is supported by empirical findings from interviews with educators and educational observers, which reveal that students' critical reasoning skills are currently low. Furthermore, students' independence is also identified as being suboptimal, raising concerns about their ability to thrive in an ever-evolving world. Additionally, the instruments used to evaluate students' character are mainly limited to student worksheets, which teachers have deemed ineffective, indicating the need for innovative methods to assess these traits. Therefore, developing more innovative and comprehensive evaluation instruments is urgently needed (Azizah et al., 2020).

Existing evaluation instruments, such as student worksheets, often focus on factual knowledge (Bahri et al., 2021). However, it must be more supportive of developing critical reasoning skills. This raises questions about how evaluation instruments can be designed to test learners' understanding (Pratiwiningtyas et al., 2017), mainly on human blood circulation, and evaluate the extent to which they can apply critical and independent reasoning in the learning process.

Independent character building in students is also one of the main goals, especially in science learning (Kusumawati et al., 2023). Independence in learning is one of the critical skills that must be developed early, especially in science learning, which demands exploration, initiative, and discovery (Idawati et al., 2023). The topic of human blood circulation in primary school offers a valuable opportunity to develop independence in the learning process (Lestari et al., 2023). So, to determine the extent of students' independent character and whether it is by the elements contained in the Pancasila student profile, it is necessary to evaluate (Asiyah et al., 2023). However, existing evaluation instruments often need to be more effective in measuring and supporting the development of student's independent character in understanding complex science concepts (Diptera et al., 2022). Therefore, there is a need to develop more innovative and comprehensive evaluation methods that can assess the extent to which students can learn and apply knowledge independently.

Innovative evaluation methods can also be done using technology (Agusta et al., 2022) by collaborating evaluation instruments with digital products (Hamid et al., 2020). This makes the evaluation process more innovative and exciting (Calamlam, 2021). Something that looks interesting will encourage student motivation to understand better the material taught to be able to master the learning topic (Syahroni et al., 2016; Yanuarti et al., 2022), which later can be evaluated with questions that have been developed with the use of technology (Istuningsih et al., 2018). These alternative learning evaluation methods also address the digital era's challenges, which require education practitioners to utilize technology in the learning process (Utaminingsih, Raharjo, et al., 2023).

Research on the development of evaluation instruments has been developed by Pomalato et al. (2021). Unlike the research, this study will develop more innovative problems by utilizing technology to make it

look more attractive. Meanwhile, development research has yet to use yet to use technology.

The study aims to develop a new instrument to evaluate the critical and independent reasoning character of grade V elementary school students, especially in learning human circulatory topics. Through the development of this evaluative instrument, the aim is not only to provide an accurate tool for assessing these two traits but also to offer clear guidance to educators on how to more effectively support student development, significantly contributing to the practice of science education in elementary schools. Educators can better identify and develop learners' critical and independent reasoning abilities with more innovation. As such, this research has significant potential to contribute substantially to shaping Indonesia's human resources into globally competitive individuals aligned with the national educational objectives embodied in the Pancasila Student Profile.

Method

The methodology of this research was Research and Development (R&D). Model of Research methodology using 4D models, including Define, Design, Development, and Disseminate (Wardani et al., 2019). The chart of the 4D model is described in Figure 1.

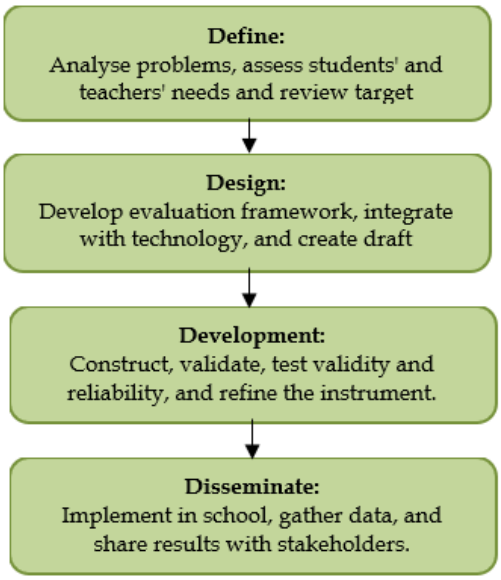

Figure 1. 4D models of research technology

The first stage is Define or definition. The stage in this step is to review information related to the problems found, student needs, and student characteristics through needs questionnaires and interviews with teachers and students. The problem is that class V Subtheme 1 material on human blood circulation is difficult for students to master. In addition, there is no inspirational evaluation instrument related to the topic to measure learning achievements, especially in critical and independent reasoning. The next step is to conduct a literature review to analyze the solutions to the problems found.

The second stage is Design. Researchers design a predetermined solution, namely by developing evaluation instruments that collaborate with technology on the topic of human blood circulation. This evaluation instrument will later be able to be displayed on Android devices. This step begins with making an initial design of instrument development, starting from the grid of questions to be developed. The grid of science literacy questions consists of 5 description questions, while for independent character questions, there are four questions, three multiple-choice questions, and one description question. The next step is to collect the software for use, such as relevant references, and check the completeness of Ms. Power Point, which has iSpring Suite 10 Features and Programs *Statistics SPSS 25*.

The next step is to Develop. This stage makes the evaluation instrument as designed. The results of developing critical and independent reasoning character instruments are presented in Table 1.

Table 1. Instrument Brief Grid

Ability Aspect	Competency Achievement Indicators	Number of Questions
Critical Reasoning	Acquire and process information and ideas	1 3
	Analyze and evaluate reasoning	2 4
	Reflect and evaluate his thoughts	5
Self-sufficient	Acquire and process information and ideas.	1 2
	Analyze and evaluate reasoning	3 4

Research instrument trials are needed to be valid and can be used to measure variables. Test instrument trials to measure validity, reliability, differentiating power, and difficulty. Validity tests consist of two types, namely logical validity and empirical validity. Logical validity refers to established theories and provisions, while empirical validity refers to testing instruments. Empirical validity is obtained by arranging instruments based on provisions and logical validity and must be proven by testing the product or instrument (Bashooir & Supahar, 2018).

The logical validity test includes two things, namely content validity and construct validity. Content validity refers to the suitability of the content in the assessment instrument to the material or concept to be assessed. Testing the validity of the content is carried out by experts (expert judgment) (Sugiyono, 2015).

Specification tables are prepared for context experts to measure their adequacy and compatibility with test items (Zulyusri et al., 2017). The calculation of content validity uses Aiken's V equation to calculate the Content-Validity Coefficient (CVI) based on the results of several validators' assessments of the Item. The content's validity is analyzed by comparing the developed instrument with the instrument grid.

The results of this expert validation will be used as a reference for revising the instruments that have been developed. The results of expert validation are qualitative and quantitative data. Qualitative data is in the form of criticism and suggestions, while quantitative data is in the form of an assessment of instruments that have been developed. Scoring guidelines on validation sheets using the Likert Scale by Sugiyono (2015), which consists of points 1 to 4. The Likert Scale scoring categories are presented in Table 2.

Table 2. Categories Likert Scale Scoring (Sugiyono, 2015)

Answer Categories	Score
Excellent	4
Good	3
Bad	2
Very not good	1

Aiken formulated Aiken's V formula for calculating the content-validity coefficient based on the assessment of a panel of experts of n people on an item in terms of the degree to which it represents the measured construct. The evaluation is carried out by giving a checklist mark from the category "No Appropriate" with a score of "1" to "Very Appropriate" with a score of "5.". Aiken's V formula is as follows (Aiken, 1985).

$$V = \frac{\Sigma S}{[n(c-1)]} \quad (1)$$

Information:

S = r – lo
r = number given by the appraiser
lo = lowest validity assessment number
n = number of appraisers
c = highest validity assessment number

According to Aiken (1985), Determining each Item's validity can be analyzed by looking at the value of the validity coefficient V contained in the Validity Coefficient Value Table presented in Table 3.

Kappa statistic or interrater reliability is a measure used to test agreement between two people (raters) on categorical variables. This technique is used by some researchers, such as Wynd et al. (2003), who use CVI and kappa Multi-rater to validate the scale content it develops. They argue that kappa statistics are an essential supplement, not a CVI substitute. Kappa provides information about the extent of the deal beyond the possibilities (Polit & Beck, 2006). If there are more than two raters, the Kappa multi-rater technique can be used. Kappa Cohen's statistical measure of interrater reliability generally ranges from 0 to 1.0, where a large number means better reliability; a value close to or less than zero indicates that the deal was caused by chance alone (Polit et al., 2007). Landis et al. (1977) provide assessment guidelines for Kappa statistics presented in Table 4.

Table 3. Value of Validity Coefficient (Aiken, 1985)

Raters	Number of Twigs Categories							
	2		3		4		5	
	V	P	V	P	V	P	V	P
2							1	0.04
3							1	0.01
3			1.00	0.04	1	0.02	0.92	0.03
4					1	0.01	0.92	0.03
4			1.00	0.01	0.92	0.02	0.88	0.02
5			1.00	0.01	0.93	0.01	0.90	0.01
5	1.00	0.03	0.90	0.03	0.87	0.02	0.80	0.04

Table 4. Kappa Statistic Assessment Guidelines (Landis & Koch, 1977)

Kappa	Interpretation
< 0	Poor Agreement
0.0 – 0.20	Slight Agreement
0.21 – 0.40	Fair Agreement
0.41 – 0.60	Moderate Agreement
0.61 - 0.80	Substantial Agreement
0.81 – 1.00	Almost perfect Agreement

The next validity test is empirical validity. Empirical validity tests were carried out on the questions developed on 31 students. The validity test of the question items is carried out using the SPSS 25 Program. Validity testing correlates each indicator item's score with the total score. The significance level used is 0.05. The test criteria are: H_0 is accepted if r is calculated > r table, which means (the measuring instrument used is valid or valid). H_0 is rejected if the statistical r ≤ r table means that the measuring instrument used is invalid or invalid). Interpret the score of each Item with the r table. The interpretation of r table values is presented in Table 5.

Reliability analysis was performed with the SPSS 25 Programs. The reliability obtained from the results of data analysis is reviewed using the value of Alpha Cronbach. Test instruments are reliable when the Alpha Cronbrach coefficient value≥ 0.70 or at least meets the interpretation criteria of fixed/good reliability (Payadnya & Jayantika, 2018). Furthermore, reliability levels are classified by the interpretation criteria in Table 6.

Table 5. Distribution of r Values Significance Table 5% (Janna & Herianto, 2021)

N	The Level of Significance (5%)
21	0.433
22	0.432
23	0.413
24	0.404
25	0.396
26	0.388
27	0.381
28	0.374
29	0.367
30	0.361
31	0.355
32	0.349

Table 6. Reliability Interpretation (Rofiyadi & Handayani, 2021)

Correlation Coefficient	Correlation	Reliability Interpretation
0.90 ≤ r ≤ 1.00	Very high	Very fixed/ very good
0.70 ≤ r ≤ 0.90	Tall	Fixed / good
0.40 ≤ r ≤ 0.70	Keep	Quite fixed/ good enough
0.20 ≤ r ≤ 0.40	Low	Not fixed/ bad
R < 0.20	Very low	Very impermanent/ very bad

The next step is to analyze the differentiating power. The perceived difference is still sufficient for a question if it is equal to or greater than 0.30 or the value of Cronbach's Alpha coefficient if Item Deleted does not exceed the value of Cronbach's Alpha coefficient (Payadnya & Jayantika, 2018). Differentiating power interpretations are classified in Table 7.

Table 7. Interpretation of Discriminating Power (Nani, 2021)

Interval	Interpretation
0.00 – 0.19	Poor
0.20 – 0.39	Satisfactory
0.40 – 0.69	Good
0.70 – 1.00	Excellent

In addition to the discriminating power, the question items must also be analyzed for the difficulty level. The difficulty level of the question item shows the possible value of the number of respondents who can answer the question item correctly. The ideal question has moderate difficulty (Susanto et al., 2015). The difficulty level of the question items can be calculated using the Formula 2.

$$IK = \frac{\bar{X}}{SMI} \quad (2)$$

Information:
IK : Question Item Difficulty Index
X : the average answer score of each question item
SMI : Ideal maximum score (Max score)

The difficulty index of a question item is interpreted in categories, as shown in Table 8.

Table 8. Interpretation of Difficulty Levels

Value	Interpretation of the Difficulty Index
IK = 0.00	Too difficult
0.01 ≤ IK ≤ 0.30	Difficult
0.31 ≤ IK ≤ 0.70	Keep
0.71 ≤ IK ≤ 0.99	Easy
IK = 1.00	Too easy

The final stage of the study is Disseminate. After the instrument is deemed feasible based on validity tests by experts and empirical validity test results, the instrument will be disseminated in elementary schools where pre-research is carried out.

Result and Discussion

Results of Question Development

The development of evaluation instruments starts with drafting the question items. The question design that has been developed is outlined in a grid of questions that represent the learning objectives to be achieved. The instrument grids of critical and independent characters are presented in Table 9.

Table 9. Evaluation Instrument Grille

Ability Aspect	Competency Achievement Indicators	Question Indicator	Question No.
Critical Reasoning	Acquire and process information and ideas	Presenting information related to diseases caused by smoking and heart health, then students are asked to give ideas	1 3
	Analyze and evaluate reasoning	Analyze disorders of blood vessels Present data on cases of heart disease and ask students to analyze and reason from the data presented.	2 4
	Reflect and evaluate his thoughts	Create a mind map of the circulatory system	5
Independent Character	Acquire and process information and ideas.	Presenting cases about diseases in the blood vessels, then learners are asked to give attitude statements	1 2
	Analyze and evaluate reasoning	Analyze disorders of blood vessels	3

Ability Aspect	Competency Achievement Indicators	Question Indicator	Question No.
		Mention activities that support the health of circulatory people	4

Critical Reasoning Instrument

Instrument Quality Based on Content Validity

The design of the science literacy assessment instrument and the character of Pancasila students that had been prepared were then validated by experts, namely four validators, consisting of one lecturer and three grade V elementary school educators who were experts in their field. The developed assessment instruments are analyzed based on the assessment results of validators using validation sheets. The results of expert validation of critical reasoning question items are presented in Table 10.

Table 10. Results of Validity of Critical Reasoning Character Question

Question Point	V value	CVI Aiken	Category KS Validity Status
1	0.94	Valid	Almost Perfect
2	1	Valid	Almost Perfect
3	0.94	Valid	Almost Perfect
4	1	Valid	Almost Perfect
5	1	Valid	Almost Perfect
Final V Value	0.98	Valid	Almost Perfect

Based on Table 10, the results of expert validation of critical reasoning instruments obtained a final V value of 0.98. The results are then reviewed from the value of the Aiken validity coefficient with the number of raters, as many as 4 (four) experts, and the questionnaire scale used. There are 5 (five) scales, with $p < 0.05$. Based on the validity level analysis results with Aiken's equation, the V value of 0.98 is included in the valid category. In addition, judging from the validity status of Kappa Statistics, it is also stated to be almost perfect. It can be concluded that the critical reasoning question item is declared valid.

Instrument Quality Based on Empirical Validity

Question Point Validation

Questions that have been developed, in addition to being validated by validators, are also tested for empirical validity. The questions were tested on 31 students, and then, to find out whether each question item was valid or not, it was analyzed using the SPSS Program. The result of the validity of the Item of critical reasoning character is presented in Table 1.

The results of the analysis of the question items in Table 10 show that the five question items developed to assess the character of critical reasoning, each question item is declared "valid." Based on the distribution of the r value of the table with a significance level of 5%, each question item shows an r value above 0.36, so it can be concluded that all critical reasoning character measurement items are declared valid.

Table 11. Results of Validity of Critical Reasoning Character Question Items

Number Question Point	Value	Criterion
1	0.61	Valid
2	0.42	Valid
3	0.60	Valid
4	0.83	Valid
5	0.51	Valid

Question Item Reliability

Reliability tests are performed on instruments of critical reasoning character. The analysis was conducted using the Statistical SPSS 25 program. The learners' critical reasoning character test instrument is reliable if the correlation coefficient value≥ 0.70 or at least meets the interpretation criteria of fixed reliability (Payadnya & Jayantika, 2018). The trial was conducted on thirty-one SD Supriyadi 01 Semarang grade V students. The results of the reliability test analysis obtained a reliability value of 0.74 with a fixed interpretation.

The results of the analysis showed that the developed instrument met the correlation coefficient value of > 0.70 and that the critical reasoning character assessment instrument was declared "reliable" because the Cronbach Alpha value obtained was included in the "fixed/good" category. Cronbach's Alpha value indicates the interaction between respondents and the Item.

Differentiating Power

The differentiating power test is performed on the critical reasoning character instrument. The purpose of this test is to determine the level of ability of question items in finding different powers, namely to determine whether the question items that have been developed can distinguish high-achieving groups (upper group) from low-achieving groups (lower group) among test participants. The difference that is considered to still exist for a question item is when it is equal to or greater than 0.30 (Payadnya & Jayantika, 2018). The results of the difference power test analysis of the developed science literacy question items are presented in Table 12.

Based on the results of the differentiation power analysis on the critical character instrument, it was found that all question items, namely question numbers 1 to 5, had a "strong" or "satisfactory" interpretation value. This interpretation is shown from the value of Corrected Item-Total Correlation, which is more than

0.3, and the value of Cronbach's Alpha Item Deleted is not more than 0.75. This shows that the question of the character of critical reasoning has the value of interpreting the difference power as "strong" or "satisfactory."

Table 12. Results of Instrument Differentiation Analysis of Critical Reasoning Character

Grain Question	Corrected Item-Total Correlation	Cronbach's Alpha Item Deleted	Interpretation
1	0.56	0.72	Powerful/ satisfying
2	0.39	0.75	Powerful/ satisfying
3	0.58	0.71	Powerful/ satisfying
4	0.61	0.68	Powerful/ satisfying
5	0.54	0.72	Powerful/ satisfying

The Difficulty Level of the Question Items

The difficulty level of the question item shows the possible value of the number of respondents who can answer the question item correctly. Difficulty level analysis is intended to find out whether the problem is classified as easy or difficult (Solichin, 2017). A good question is easy enough and easy (Payadnya & Jayantika, 2018). The critical reasoning character instrument totals five questions in the form of description questions tested for difficulty using Microsoft Excel. The results of the analysis of the difficulty level of the critical reasoning character instrument are presented in Table 13.

Table 13. Results of Difficulty Level Analysis Critical Reasoning Instrument

Question Type	Question Criteria	Percentage (%)	Question Number
Essay	Easy	0	-
	Keep	20	1
	Difficult	80	2, 3, 4, and 5

Table 13 shows that critical reasoning question items have "moderate" and "difficult" difficulty levels. Although only one question has a "medium" difficulty level and four other questions are classified as "difficult," none of these questions were eliminated because they still met the two criteria for a good question: medium and difficult. In addition, four questions that have a "difficult" difficulty level are used to measure the achievement of critical reasoning elements so that the questions are retained.

Independent Character Instruments

The character of Pancasila students tested for validity based on expert validation is independent. The developed test instruments are analyzed based on the assessment results of validators using validation sheets. The results of expert validation of independent character instruments are presented in Table 14.

Table 14. Results of Validity of Independent Character Question Items

Question Point	V value	CVI Aiken	Category KS Validity Status
1	1	Valid	Perfect
2	1	Valid	Perfect
3	1	Valid	Perfect
4	1	Valid	Perfect
Final V Value	1	Valid	Perfect

Based on Table 14, the results of expert validation of the independent character assessment instrument obtained a final V value of 1 with valid/perfect criteria. The results were then reviewed both from the value of Aiken's validity coefficient with the number of raters as many as 4 (four) experts and the questionnaire scale used were 5 (five) scales, with $p < 0.05$. Based on the validity level analysis results with Aiken's equation, the value of V "1" is included in the valid category.

Empirical Validity

In question number four, the empirical Validity of Independent Character, the instrument of self-governing character, is not carried out independently. Validity tests are only carried out based on the validity of the content or only carried out by validators. This is because number 4 in independent characters is the sole or only problem with the description question type. If the analysis is carried out through SPSS or with formulas calculated in MS Excel, it will not get the appropriate results because one Item cannot be analyzed.

Question Item Validity

Validity tests are also carried out on critical reasoning question items, and validity tests are carried out on independent character assessment instruments. Each question item on an independent character is measured to determine the validity of each question item. The results of the analysis of independent character instruments are presented in Table 15.

Table 15. Results of Validity of Independent Character Question Items

Number Question Point	Value	Criterion
1	0.44	Valid
2	0.44	Valid
3	0.61	Valid

Table 15 shows that five questions on the assessment instrument were developed to measure independent character, tested on thirty-one students, and each question item was declared "valid." Based on the distribution of table r values with a significance level

of 5%, each question item shows an r value above 0.35, so it can be concluded that all independent character measurement items are declared valid.

Reliability of Independent Question Items

The results of the reliability test analysis of independent character instruments obtained a reliability value of 0.70 with a fixed / reasonable interpretation, so it was declared "reliable" because the Alpha Cronbach value obtained was included in the "fixed/good" category. Cronbach's Alpha value indicates the interaction between respondents and the Item.

The results of the analysis showed that the developed instrument met the correlation coefficient value of > 0.70 and that the independent character instrument was declared "reliable" because the Alpha Cronbach value obtained was included in the "fixed/good" category. Cronbach's Alpha value indicates the interaction between respondents and the Item.

The Discriminating Power of Independent Character

Power difference analysis is also performed on independent character measurement instruments. There are five question points developed. The results of the independent character differentiating power analysis are presented in Table 16.

Table 16. Results of Differentiating Power Analysis of Independent Character Instruments

Grain Question	Corrected Item-Total Correlation	Cronbach's Alpha Item Deleted	Interpretation
1	0.284	0.715	Powerful/satisfying
2	0.246	0.724	Powerful/satisfying
3	0.422	0.672	Powerful/satisfying

Table 16 shows that the difference power analysis on independent character instruments in all question items, namely question numbers 1 to 5, has a solid or satisfactory interpretation value. This is indicated by a value of n not more than 0.704. It can be concluded that the independent character instruments that have been developed have a "strong" or "satisfactory" differentiation interpretation value.

Analysis of the Difficulty Level of the Question Items

Analysis of the difficulty level of the next question item is done using an independent character measurement instrument. The instrument developed consists of four questions: three multiple-choice questions and one description question. The results of the difficulty level analysis of independent character question items can be seen in Table 17.

Table 17. Results of Difficulty Level Analysis of Independent Character Instruments

Question Criteria	Percentage (%)	Question Number
Easy	50	1 and 2
Keep	25	3
Difficult	25	4

Based on Table 17, the question items of independent character instruments have a good spread. The analysis results of the two instruments can be used as a reference for choosing the ideal problem. However, the selection needs to consider the results of validity, reliability, and differentiation based on validity tests, which consist of validity tests and reliability tests (Nani, 2021), which are eliminated. Referring to these results, the two instruments can be used to evaluate the character of Pancasila students to reason critically and independently.

Quiz Display in STEAM-Based E-Modules

Instruments declared valid, attractive, and suitable for use are combined with STEAM-based e-modules. The instrument will then evaluate circulatory topics assessed from critical and independent reasoning. The display of instruments packaged as quizzes is named "Sequizi," an acronym for Core Quiz Set. This is done to make the evaluation process more exciting and less stressful. The quiz display in the E-Module is presented in Figure 2.

Figure 2. Quiz view

The presentation of instruments transformed into interesting quizzes is also designed to resemble playing games. There is a speech on the second page of the quiz, followed by filling in the identity; then there are instructions for using the quiz. Learners' answers, both true and false, will be notified directly. The correct answer will be given a congratulatory sentence while encouraging sentences to be given to study more diligently for answers that have yet to be determined. After completing the process, students can see their scores immediately. The teacher also gets a report from the student's grades sent by the system via email. The Quiz display is presented in Figure 3.

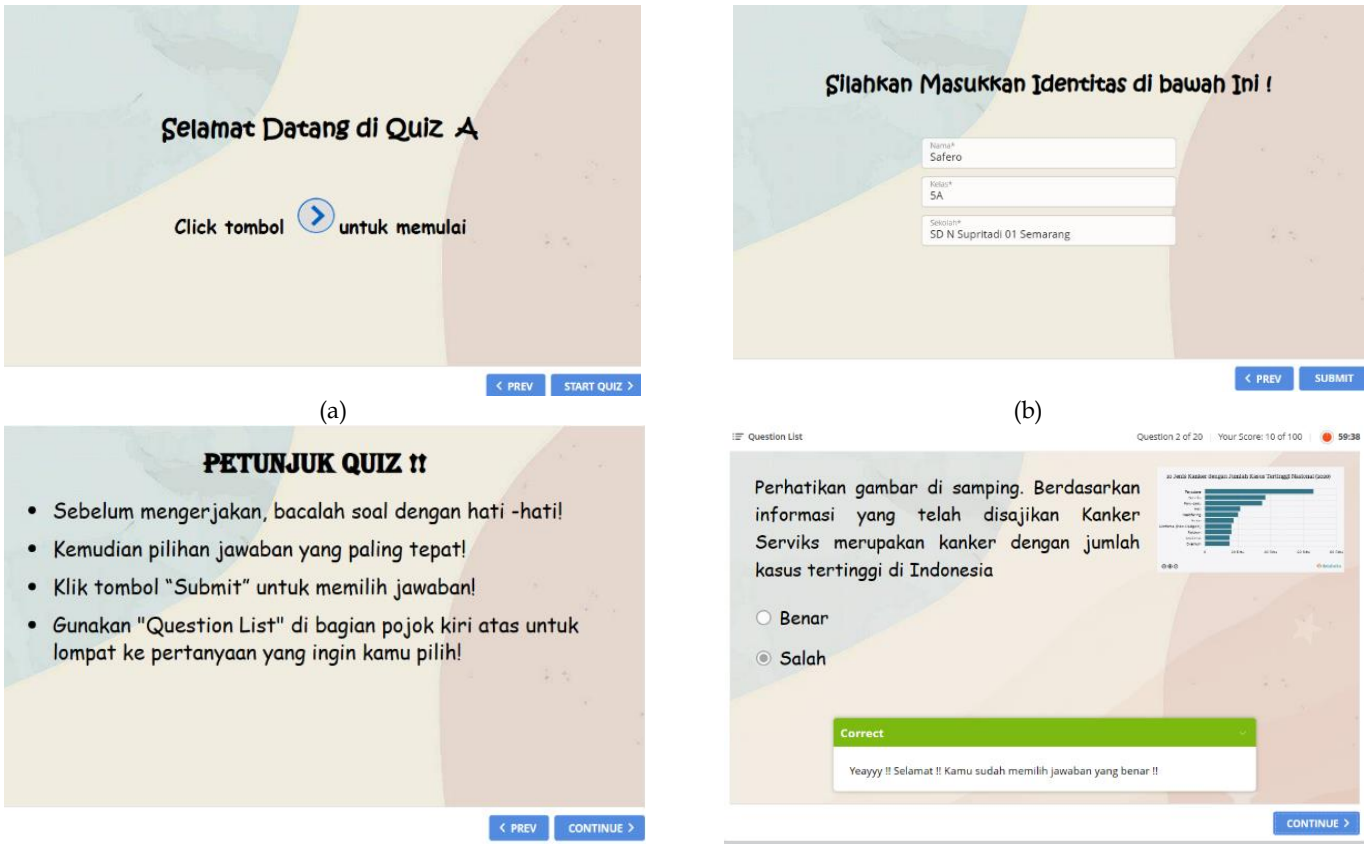

Figure 3. Quiz Overview: (a) First display; (b) User login; (c) Instructions for working on questions; and (d) Example questions

Based on these results, according to learners, the developed device has good feasibility and is innovative and feasible. The study's results align with the findings of Nurhasanah et al. (2020). The developed instrument has a feasibility percentage of 90.26%, which is in the "very viable" category. Accraf et al. (2019) also supported these findings; the validation results showed a percentage of 85%, so the product developed was "valid."

Evaluation instruments declared valid, reliable, and have good distinguishing power, with moderate difficulty, are disseminated to pre-research schools. In addition, it was disseminated to an elementary school research group at the Sekolah Tinggi Agama Islam Muhammadiyah Blora.

Conclusion

Based on the research results, the developed instrument was suitable for evaluating fifth-grade students' critical reasoning and independent character, particularly on human blood circulation. The instrument demonstrated high content and strong empirical validity, making it reliable. Its discriminative power was satisfactory, with difficulty levels effectively supporting measuring students' abilities. The use of technology in digital module-based evaluation further enhanced the relevance and innovation of this instrument within the context of modern education. The findings indicated that the instrument held significant potential for application in science education at the elementary school level. This aligns with the national education goals to shape globally competitive students through the Pancasila Student Profile. Further research is recommended to expand the instrument's testing to a larger population across different subject matters. Further exploration into technology-based instrument development is needed to support adaptive and personalized learning.

Acknowledgments

During the process of completing this research, the researcher received much support, both moral and material, from various parties. Therefore, on this occasion, the author would like to thank Mrs. Ellianawati as supervisor for their motivation, guidance, direction, and instructions while completing this article. The author is grateful for the solid teamwork that allowed the writing of this article to be completed.

Author Contributions

The principal author, Esty Setyo Utaminingsih, and the second author, Mrs. Ellianawati, contributed to the conceptualization of the research and designing methodology; the third author, Mr. Muhammad Habib Ramadhani, helped the first author

analyse, investigate, and curate the data. The fourth author, Mrs. Maria Ayu Puspita, prepared and made software; the fifth author, Mrs. Sri Sumartiningsih, was a validator and designing the methodology. The sixth author, Mr. Muhamad Subhi Apriantoro, and the first author, Esty Setyo Utaminingsih, are writing the original draft preparation, writing the review, and editing. The last author, Mr. Aldi Ihsandi, contributed to the instrument grid, project administration, and funding acquisition. All authors have read and agreed to the published version of the manuscript.

Funding
This research received no external funding.

Conflicts of Interest
The authors declare no conflict of interest.

References

Accraf, L. B. R., Suryati, S., & Khery, Y. (2019). Pengembangan E-modul Interaktif Berbasis Android dan Nature of Science Pada Materi Ikatan Kimia dan Gaya Antar Molekul Untuk Menumbuhkan Literasi Sains Siswa. *Hydrogen: Jurnal Kependidikan Kimia*, *6*(2), 133. https://doi.org/10.33394/hjkk.v6i2.1607

Agusta, F. R., Kartono, K., & Zulaeha, I. (2022). Creative Thinking Skill in Solving Mathematics Story Questions for Eight Graders. *Journal of Educational Research and Evaluation*, *11*(1), 9–14. https://doi.org/10.15294/jere.v11i1.55422

Aiken, L. R. (1985). Three coefficients for analyzing the reliability and validity of ratings, Educational and Psychological Measurument. *Journal Articles; Reports - Research; Numerical/Quantitative Data*, *45*(1), 131–142. https://doi.org/10.1177/0013164485451012

Asiyah, S., Sudana, I. M., & Soedjono, S. (2023). Manajemen Pembelajaran Berbasis Projek Penguatan Profil Pancasila Dimensi Mandiri dan Bergotong Royong di Sekolah Dasar. *Jurnal Inovasi Pembelajaran Di Sekolah*, *4*(2), 690–700. https://doi.org/10.51874/jips.v4i2.156

Azizah, W. A., Sarwi, S., & Ellianawati, E. (2020). Implementing Project -Based Learning Model (PjBL) Using STREAM-Based Approach in Elementary Schools. *Journal of Primary Education*, *9*(3), 238–247. https://doi.org/10.15294/jpe.v9i3.39950

Bahri, S., Simbolon, M., & Alhad, K. (2021). Development of Assessment Instrument of Scientific Literacy Ability for Students at Musamus University. *Advances in Social Science, Education and Humanities Research Proceedings of the International Joined Conference on Social Science (ICSS 2021)*, *603*(Icss), 395–398. https://doi.org/10.2991/assehr.k.211130.072

Bashooir, K., & Supahar. (2018). Validitas dan reliabilitas instrumen asesmen kinerja literasi sains pelajaran Fisika berbasis STEM. *Jurnal Penelitian dan Evaluasi Pendidikan*, *22*(2), 168–181. https://doi.org/10.21831/pep.v22i2.20270

Calamlam, J. M. M. (2021). The Development of 21st-Century e-Learning Module Assessment Tool. *Journal of Educational Technology Systems*, *49*(3), 289–309. https://doi.org/10.1177/0047239520953792

Diputera, A. M., Damanik, S. H., & Wahyuni, V. (2022). Evaluasi Kebijakan Pendidikan Karakter Profil Pelajar Pancasila dalam Kurikulum Prototipe untuk Pendidikan Anak Usia Dini. *Jurnal Bunga Rampai Usia Emas*, *8*(1), 1. https://doi.org/10.24114/jbrue.v8i1.32650

Ernawati, Y., & Rahmawati, F. P. (2022). Analisis Profil Pelajar Pancasila Elemen Bernalar Kritis dalam Modul Belajar Siswa Literasi dan Numerasi Jenjang Sekolah Dasar. *Jurnal Basicedu*, *6*(4), 6132–6144. https://doi.org/10.31004/basicedu.v6i4.3181

Fazryn, M., Adiansha, A. A., Syarifudin, Mariamah, & Diana, N. (2023). Implementasi Model Problem Based Learning Berbasis Aplikasi Geogebra untuk Meningkatkan Kemampuan Bernalar Kritis Matematika pada Siswa Sekolah Dasar. *JagiMIPA: Jurnal Pendidikan Matematika dan IPA*, *3*(1), 42–51. Retrieved from http://jurnal.bimaberilmu.com/index.php/jagomipa/article/view/279

Gulson, K. N., & Webb, P. T. (2023). Steering the mind share: technology companies, policy and Artificial Intelligence research in universities. *Discourse*, *44*(2), 195–207. https://doi.org/10.1080/01596306.2021.1981828

Hamid, M. A., Yuliawati, L., & Aribowo, D. (2020). Feasibility of electromechanical basic work e-module as a new learning media for vocational students. *Journal of Education and Learning (EduLearn)*, *14*(2), 199–211. https://doi.org/10.11591/edulearn.v14i2.15923

Hayati, N., & Setiawan, D. (2022). Dampak Rendahnya Kemampuan Berbahasa dan Bernalar terhadap Kemampuan Berpikir Kritis Siswa Sekolah Dasar. *Jurnal Basicedu*, *6*(5), 8517–8528. https://doi.org/10.31004/basicedu.v6i5.3650

Idawati, Prayitno, H. J., Harsono, & Sutopo, A. (2023). Pembelajaran Bahasa Indonesia dengan Media Audio Visual: Pembudayaan Dimensi Mandiri. *Jurnal Didactique*, *4*(2), 83–92. Retrieved from https://ejournal.univ-tridinanti.ac.id/index.php/Didactique/article/download/151/94

Istuningsih, W., Baedhowi, & Sangka, K. B. (2018). The Use Of Electrinic Modules For Learning Effectiveness. *Ijere*, *03*(03), 75–85. Retrieved from https://dergipark.org.tr/en/pub/ijere/issue/36244/449313#article-authors-list

Jafari, F., & Keykha, A. (2023). Identifying the opportunities and challenges of artificial intelligence in higher education: a qualitative study. *Journal of Applied Research in Higher Education*. https://doi.org/10.1108/JARHE-09-2023-0426

Janna, N. M., & Herianto. (2021). Konsep Uji Validitas Dan Reliabilitas Dengan Menggunakan SPSS. *Jurnal Darul Dakwah Wal-Irsyad (DDI)*, *18210047*, 1–12. https://doi.org/10.31219/osf.io%2Fv9j52

Jokhan, A., Chand, A. A., Singh, V., & Mamun, K. A. (2022). Increased Digital Resource Consumption in Higher Educational Institutions and the Artificial Intelligence Role in Informing Decisions Related to Student Performance. *Sustainability (Switzerland)*, *14*(4). https://doi.org/10.3390/su14042377

Kemdikbud. (2020). Peraturan Menteri Pendidikan Dan Kebudayaan Republik Indonesia Nomor 22 Tahun 2020 Tentang Rencana Strategis Kementerian Pendidikan Dan Kebudayaan Tahun 2020-2024. In *Kementerian Pendidikan dan Kebudayaan*. Retrieved from https://shorturl.asia/d93Y4

Kusumawati, E. T., Abdul, Mulkhan, M., & Sari, Z. (2023). Pemaknaan Dimensi Profil Pelajar Pancasila Dalam Praksis Pendidikan Kh Ahmad Dahlan. *Al Qalam: Jurnal Ilmiah Keagamaan Dan Kemasyarakatan*, *17*(3), 2023. Retrieved from https://jurnal.stiq-amuntai.ac.id/index.php/al-qalam

Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, *33*(1), 159–174. Retrieved from https://www.jstor.org/stable/2529310

Lestari, H. A., Anggraeni Siwi, D., & Prasetyo, K. (2023). Implementasi Penguatan Profil Pelajar Pancasila Dimensi Nilai Karakter Mandiri Dalam Menumbuhkan Budaya Akademik Siswa Kelas IV Sekolah Dasar Negeri 2 Bolopleret. *Jurnal Pendidikan Dasar Flobamorata*, *4*(3), 723–731. https://doi.org/10.51494/jpdf.v4i3.1063

Muqit, A., Mukarromah, M., & Amrullah, A. (2023). The Formation of Pancasila Student Profile in Student Independent Dimensions Through Peer Tutor in Pai Learning. *Edupedia : Jurnal Studi Pendidikan Dan Pedagogi Islam*, *8*(1), 89–98. https://doi.org/10.35316/edupedia.v8i1.3450

Nani, F. (2021). Analisis Tingkat Kesukaran, Daya Pembeda, dan Efektivitas Pengecoh Soal Pelatihan Kewaspadaan Kegawatdaruratan Maternal Dan Neonatal. *Paedagoria : Jurnal Kajian, Penelitian Dan Pengembangan Kependidikan*, *12*(2), 199–205. https://doi.org/10.31764/paedagoria.v12i2.4956

Nurhasnah, & Sari, L. A. (2020). E-Modul Fisika Berbasis Contextual Teaching And Learning Menggunakan Aplikasi Kvisoft Flipbook Maker Untuk Meningkatkan Literasi Sains Peserta Didik SMA/MA Kelas XI. *Natural Science: Jurnal Penelitian Bidang IPA Dan Pendidikan IPA*, *6*(1), 29–40. Retrieved from https://ejournal.uinib.ac.id/jurnal/index.php/naturalscience/article/view/1554

Nursalam, N., & Suardi, S. (2022). Penguatan Karakter Bernalar Kritis Berbasis Integratif Moral untuk Siswa Sekolah Dasar dalam Program Kampus Mengajar di Indonesia. *Jurnal Pendidikan: Teori, Penelitian, Dan Pengembangan*, *7*(8), 335–342. https://doi.org/10.17977/jptpp.v7i8.15416

Payadnya, I. P. A. A., & Jayantika, I. G. A. N. T. (2018). *Panduan Penelitian Eksperimen Beserta Analisis Statistik Dengan SPSS*. Sleman: Deepublish.

Polit, D. F., & Beck, C. T. (2006). The Content Validity Index: Are You Sure You Know What's Being Reported? Critique and Recommendations. *ResearchinNursing&Health*, *29*, 489–497. https://doi.org/10.1002/nur

Polit, D. F., Beck, C. T., & Owen, S. V. (2007). Focus on Research Methods Handling Missing Data in Self-Report Measures. *ResearchinNursing&Health*, *30*, 459–467. https://doi.org/10.1002/nur

Pomalato, S. W. D., Ismail, R., Amin Otoni Harefa, Imawan, O. R., Ningsi, B. A., & Wulandar, D. (2021). Instrument test development of mathematics skill on elementary school. *Mathematics and Statistics*, *9*(2), 106–111. https://doi.org/10.13189/ms.2021.090204

Pratiwiningtyas, B. N., Susilaningsih, E., & Sudana, I. M. (2017). Pengembangan Instrumen Penilaian Kognitif untuk Mengukur Literasi Membaca Bahasa Indonesia Berbasis Model Pirls pada Siswa Kelas IV SD. *Journal of Educational Research and Evaluation*, *6*(1), 1–9. Retrieved from http://journal.unnes.ac.id/sju/index.php/jere

Rejeki, A. S., Purnamasari, I., & Sutono, A. (2022). Keefektifan Media Build Geometry Dalam Meningkatkan Kemampuan Bernalar Kritis Warga Belajar Pendidikan Kesetaraan Paket A. *Refleksi Edukatika : Jurnal Ilmiah Kependidikan*, *13*(1), 74–78. https://doi.org/10.24176/re.v13i1.7147

Rofiyadi, Y. A., & Handayani, S. L. (2021). Pengembangan Aplikasi E-Modul Interaktif Berbasis Android Materi Sistem Peredaran Darah Manusia Kelas V Sekolah Dasar. *JPDI (Jurnal Pendidikan Dasar Indonesia)*, *6*(2), 54. https://doi.org/10.26737/jpdi.v6i2.2575

Rumtini, Kasimin, Arent, E., & Jalil, A. (2022). Analisis

Penilaian Afektif Kemampuan Bernalar Kritis Ditinjau Dari Jenis Kelamin Peserta Didik di SMAN 5 Yogyakarta. *Wacana Akademika: Majalah Ilmiah Kependidikan*, *6*(September), 115–120. Retrieved from https://garuda.kemdikbud.go.id/documents/detail/3149752

Siswati, S. (2022). Merdeka Belajar: Menciptakan Siswa Bernalar Kritis, Kreatif Dan Mandiri. *GUAU: Jurnal Pendidikan Profesi Guru*, *2*, 289–296. Retrieved from http://studentjournal.iaincurup.ac.id/index.php/guau/article/view/594

Slam, Z. (2021). Pembelajaran Make A Macth Online untuk Meningkatkan Kemampuan Bernalar Kritis Mahasiswa dalam Pendidikan Pancasila. *Elementar : Jurnal Pendidikan Dasar*, *1*(1), 124–140. https://doi.org/10.15408/elementar.v1i1.20891

Solichin, M. (2017). Analisis Daya Beda Soal, Taraf Kesukaran, Validitas Butir Tes, Interpretasi Hasil Tes dan Validitas Ramalan dalam Evaluasi Pendidikan. *Dirasat: Jurnal Manajemen & Pendidikan Islam*, *2*(2), 192–213. Retrieved from www.depdiknas.go.id/evaluasi-proses-

Sugiyono. (2015). *Metode Penelitian Kuantitatif, Kualitatif dan R&D*. Bandung: Alfabeta.

Suminar, D. Y. (2022). Penerapan video interaktif alur merrdeka untuk meningkatkan kemampuan bernalar kritis di SMAN 10 Pontianak. *Jurnal Pembelajaran Prospekti*, *7*(1), 34–39. Retrieved from https://jurnal.untan.ac.id/index.php/lp3m/article/view/55057/75676593326

Susanto, H., Rinaldi, A., & Novalia. (2015). Analisis Validitas Reabilitas Tingkat Kesukaran dan Daya Beda pada Butir Soal Ujian Akhir Semester Ganjil Mata Pelajaran Matematika. *The Journal of the Japan Society for Respiratory Endoscopy*, *37*(3), 343. http://doi.org/10.24042/ajpm.v6i2.50

Syahroni, M. W., Dewi, N. R., & Kasmui. (2016). The effect of using digimon (Science digital module) with scientific approach at the visualization of students' independence and learning results. *Jurnal Pendidikan IPA Indonesia*, *5*(1), 116–122. https://doi.org/10.15294/jpii.v5i1.5800

Utaminingsih, E. S., Ellianawati, E., Widiarti, N., Sumartiningsih, S., & Puspita, M. A. (2023). A Systematic Review : Digital Literacy for Strenghthening Character in Facing the Era of Society 5.0. *Research and Development Journal Of Education*, *9*(2), 638–647. http://doi.org/10.30998/rdje.v9i2.15173

Utaminingsih, E. S., Ihsandi, A., & Mutiarawati, I. S. (2023). Pancasila as Integration Philosophy of Education and National Character. *Jurnal Ilmiah Profesi Pendidikan*, *8*(4), 2443–2449. https://doi.org/10.29303/jipp.v8i4.1752

Utaminingsih, E. S., Puspita, M. A., Ihsandi, A., Intania, B. Y., & Tri, A. (2023). A Systematic Literature Review : The Role of Character-Based Digital Literacy in 21st Century Learning in Elementary Schools. *Jurnal Penelitian Pendidikan IPA*, *9*(10), 829–840. https://doi.org/10.29303/jppipa.v9i10.4858

Utaminingsih, E. S., Raharjo, T. J., & Ellianawati. (2023). Development of an E-module Based on STEAM on the Topic of Human Blood Circulation. *Jurnal Penelitian Pendidikan IPA*, *9*(7), 5333–5340. https://doi.org/10.29303/jppipa.v9i7.3719

Wardani, D. L., Degeng, I. N. S., & Cholid, A. (2019). Developing Interactive Multimedia Model 4D for Teaching Natural Science Subject. *International Journal of Educational Research*, *7*(1), 63–72. Retrieved from https://shorturl.asia/uYk19

Wynd, C. A., Schmidt, B., & Schaefer, M. A. (2003). Two quantitative approaches for estimating content validity. *Western Journal of Nursing Research*, *25*(5), 508–518. https://doi.org/10.1177/0193945903252998

Yanuarti, R., Utari, I., & Harianti, D. (2022). Evaluation of E-Module Utilization as Self-directed Learning Materials in Teachers' Competency Improvement Program. *Jurnal TEKNODIK*, *26*(2), 101–114. Retrieved from https://jurnalteknodik.kemdikbud.go.id/index.php/jurnalteknodik/article/view/972

Zulyusri, Sumarmin, R., & Miswati. (2017). Pengembangan Soal Biologi Berbasis Literasi Sains untuk Siswa SMA Kelas X Semester 1. *Bioeducation Journal*, *1*(1), 88–94. Retrieved from http://ejournal.unp.ac.id/index.php/bioeducation/article/view/7158