

Critical Thinking Ability Test Instrument Design with Validity Test Pearson Correlation

 <https://doi.org/10.31004/jele.v10i1.642>

* Nursichati, Purwo Susongko, Tri Mulyono ^{1,2,3} 

^{1,2,3} Universitas Pancasakti Tegal

Corresponding Author: nursichati20@gmail.com

ABSTRACT

This study aims to develop a valid test instrument model to measure critical thinking skills as an achievement of the Pancasila learner profile at the end of phase D. In developing this instrument, researchers used a validity test approach with the Pearson correlation method to ensure that each item can accurately reflect the understanding and ability of learners to process information objectively, build links between various information, analyze, evaluate and conclude. Data obtained from instrument testing involves learner responses which are then analyzed. The results of the analysis showed that all items had significant correlation coefficients, with values above the specified threshold, so they were declared valid. The validity of this instrument is very important to ensure that the evaluation of learning outcomes is in accordance with the standards set by the Merdeka Curriculum, and supports the achievement of curriculum goals in improving the quality of P5 results in schools. This research is expected to make a positive contribution to the development of more effective and efficient educational assessments.

Keywords: Assessment, Curriculum, Critical Thinking, Pancasila Learner Profile, Phase D

Article History:

Received 10th December 2024

Accepted 17th January 2025

Published 01st February 2025



INTRODUCTION

The rapid development of technology in various sectors of society has become a major challenge in the 21st century. This development affects various aspects of life, including education. Related to the dynamics of global education in the 21st century, Hadiansah (2022) in his book said to address this, as an educational institution, schools must provide quality services, schools must develop character, schools must cultivate literacy habits and prepare basic 21st century skills. Along with the dynamics of education above, the performance of Indonesian schools is now faced with two fundamental problems, namely weakening character and low student literacy (Kemendikbud, 2018). These two issues are the starting point of 21st century skills.

Indonesia itself has shown its efforts in improving the quality and quality of education. This is indicated by the publication of the "Merdeka Curriculum" by the Minister of Education, Culture, Research and Technology of the Republic of Indonesia (Kemendikbud Ristek RI). The Merdeka Curriculum comes as an answer to the tight competition for human resources globally in the era of society 5.0 (Indarta, 2022). Apart from the need for students to have skills in utilizing technology, monotonous or one-way learning is one of the reasons why this independent curriculum is also present. One of the efforts to improve the quality of education in the independent curriculum is the Pancasila student profile strengthening program.

The profile of Pancasila students is one of the efforts to improve the quality of education in Indonesia which prioritizes character building (Rachmawati, 2022). In the current era of globalization technological progress, the role of value and character education is needed to provide a balance between technological development and human development (Faiz & Kurniawaty, 2022). The profile of Pancasila students contains competencies and characters contained in 6 dimensions, namely (1) faith, fear of God, and noble character; (2) global diversity; (3) mutual cooperation; (4) independence; (5) critical reasoning; and (6) creativity (Kemendikbud, 2021).

One of the elements in the Pancasila learner profile is critical reasoning. Students who think critically are able to process information objectively, build links between various information, analyze information, evaluate and conclude (Kemendikbud, 2021). To achieve critical reasoning skills, critical thinking is needed. According to Ennis in Zakiah (2019) critical thinking ability is a thinking process that focuses on making decisions according to what is believed or done. There are 4 indicators of critical thinking ability, namely (1) interpretation, (2) analysis, (3) evaluation, and (4) inference. The government provides three Independent Independent Curriculum options which basically provide the same flexibility as the emergency curriculum. The choice depends on the desire of each education unit to implement independent learning. Namely, the independent curriculum for independent learning, independent change, and independent sharing.

Good critical thinking skills are important competencies for every learner. Critical thinking is very important to use in facing challenges in the era of society 5.0. A critical thinker will be able to evaluate and analyze any new information he receives (Nuryanti, 2018). According to Afifah (2021), critical thinking skills are important for students to have in learning mathematics because they help students understand and solve math problems and improve cognitive aspects. According to Syafitri (2021) critical thinking skills are important because students are able to build the quality of thinking so as to build good learning. In line with the opinion of Arofah (2019) which states that by thinking critically, students will easily adjust to changes in the era of society 5.0 which is expected to be able to utilize technology properly and wisely.

In line with the era of globalization, science and technology are developing very fast and increasingly sophisticated, with an increasingly broad role, teachers who have character are needed. A nation whose people are not prepared will almost certainly fall by the enormity of natural changes and the rapid advancement of science and technology as the hallmark of globalization itself. Therefore, the quality of education must be improved. Schools as educational institutions are required to have the skills of creative thinking, critical thinking and problem solving, communication, and collaboration or commonly referred to as 4C, In the 2013 curriculum there are changes, especially in Permendikbud number 20 of 2016. These changes are about the skills that are needed by the nation's children. Therefore, it requires the involvement of all parties, especially schools, in preparing the nation's children to have a number of skills needed in life in the 21st century. To be able to play a meaningful role in the era of globalization in the 21st century, every citizen is required to have the ability to answer the demands of the times.

This requires the role of educators to develop both hard skills and soft skills in students in learning at school so that they can enter the world of work and are ready to compete with other countries. By developing 21st century skills in learning, it is expected that each individual has the skills to live in the 21st century with the various opportunities and challenges that will be faced in the era of technological and information advances. Some experts explain the importance of mastering various 21st century skills as a means of success in a century where the world is developing rapidly and dynamically.

Indicators of success are based more on the ability to communicate, share and use information to solve complex problems, adapt and innovate in response to new demands and changing circumstances, and extend the power of technology to create new knowledge. The inability of children to express their desires, feelings and actualize what is in them

makes the problems faced by children even greater. So children need an ability and skill to express the problems they face to others. All these skills can be possessed by learners if educators are able to develop lesson plans that contain activities that challenge learners to think critically in solving problems. Activities that encourage learners to work together and communicate should be visible in every lesson plan they create.

Based on the 2018 PISA results, Indonesia is ranked 7th lowest out of 78 countries which shows that students' mathematics skills are still very low, including critical thinking skills (Puslitjakdikbud, 2021). According to Pertiwi (2018), students' critical thinking skills are still low as seen from the number of students who fulfill each aspect of critical thinking skills below 50%. In line with Lestari's research (2021), it states that students' critical thinking skills are in the very poor category. Firmansyah (2022) stated that of the four subjects who were given critical thinking test questions, only one was able to fulfill the four indicators of mathematical critical thinking ability. According to Prasetyo (2022), the reason students cannot fulfill the indicators of critical thinking ability is because students are not used to HOTS-based problems and are always used to low-level problems.

On 11 December 2019, the Minister of Education and Culture of the Republic of Indonesia, Nadiem Anwar Makarim, introduced the "Merdeka Belajar" program. One of the policies implemented is the replacement of the National Exam (UN) with the Minimum Competency Assessment (AKM). This AKM aims to measure competence in reading and math literacy, logical-systematic thinking skills, and skills in selecting and processing information. In 2022 the Merdeka Belajar policy has been implemented in several schools. This policy is an effort to reform education in Indonesia so that the quality of human resources (HR) increases. According to the Ministry of Education and Culture (2020), Merdeka Belajar policy has the aim of providing quality education to all Indonesians by increasing participation in all levels of education, improving quality learning outcomes, and improving the quality of education equally both geographically and socio-economically.

Learners in their lives are often faced with various kinds of problems, and some of them still have difficulty in finding solutions and even choose to despair. This can happen because they do not have the ability to solve their problems by using their thinking skills. One of the thinking skills that every individual should have is critical thinking. One way to foster their critical thinking skills is of course through education, namely from the learning process. The learning process is an activity carried out by teachers and students with the aim of changing the situation from not good to better, from not knowing to knowing, from not being able to become able through active interaction to achieve learning goals. One of the learning objectives in question is the ability to think critically and cognitive domain learning outcomes in education.

Education is part of the national development process which has a very important role in improving the quality of human resources, because human resources are the driving force of national development. Human resources that have critical thinking skills play a role in facing the challenges of globalization. Critical thinking skills are important in education so they need to be developed in students. Students with critical thinking skills can examine and analyze the opinions expressed by others. Critical thinking skills that are directed through learning in schools are higher order thinking skills. One of the higher order thinking skills is critical thinking. Critical thinking is the ability of students to solve problems and make decisions from various aspects and points of view.

The application of strengthening the Pancasila Student Profile in the Merdeka curriculum is through project learning. The project learning in question is to enable students to be active in critical reasoning. One of the critical reasoning carried out is that the teacher gives project assignments to students and conducts reviews and is finally able to solve the problems faced. The process of measuring/assessing the ability to reason critically is not only obtained by observation alone, but assessment data is needed that can measure the ability of students' critical thinking levels so as to achieve the character profile of Pancasila Students. The expected indicators of critical reasoning skills are 1) obtaining and processing

information and ideas, 2) analyzing and evaluating reasoning and procedures, and 3) reflecting on thoughts and thought processes. These indicators will be developed in the preparation of critical reasoning character assessment instruments in students.

Research conducted by Anastasia Lia, et al. In 2023 on Strengthening the Profile of Pancasila Students Critical Reasoning Through Scientific Writing, based on the results of research on Strengthening the Profile of Pancasila Students Critical Reasoning Through Scientific Writing shows that strengthening the profile of Pancasila students through scientific writing can be an effective approach in building their critical reasoning skills and encouraging them to become smart, critical, and responsible citizens in facing the changes and challenges of the times.

Research conducted by Restu Ayu Gustianingrum, et al. in 2023 on analyzing the critical thinking skills of students in supporting the strengthening of the Pancasila student profile, based on the results of the study. Critical thinking skills are very important for students to have in realizing the profile of Pancasila students and facing the era of society 5.0. The results of the critical thinking skills test show that the critical thinking skills of SMAN 1 Kampar students in the interpretation indicator have a percentage of 18% with a category, the analysis indicator has a percentage of 23%, the evaluation indicator is 30% with a category and the inference indicator is 24% with a category. Overall, the critical thinking skills of students are classified as low, namely with an average percentage of 23.75%. In learning, teachers should familiarize students in solving critical thinking skills questions in order to support the project of strengthening the profile of Pancasila students. Therefore, further research is needed related to critical thinking skills and the development of critical thinking skills test questions.

Based on observations at SMP Negeri 2 Pagerbarang and several other schools, information was obtained that the junior high school has implemented the Pancasila Student Profile Strengthening Project (P5) in the project there is a critical thinking ability. This can be seen from the activities every week with a period of 2-3 weeks used for project learning, where the teacher provides material exposure in the form of PPT and learning videos which are then continued with students to carry out project learning according to the given theme. The monitoring carried out by the teacher through observation has not been able to show the level of critical reasoning of students in project learning to realize the achievement of the Pancasila Learner Profile. Project-based learning integrated into the Merdeka Curriculum aims to strengthen the Pancasila Learner Profile and develop skills, especially 21st century skills, but at the time of its implementation at the end of the phase no special assessment was carried out, especially to measure the success of critical thinking ability as an achievement of the Pancasila learner profile.

Research has shown that the Rasch method can provide more accurate and reliable assessment results compared to classical test theory, for example, research conducted by Susdelina, et al in 2018 which compared the use of classical test theory and Rasch method in assessing the understanding of mathematical concepts of MTs class IX students. The results of the research from the quality analysis of the pretest instrument measuring concept understanding through the classical test theory approach and the Rasch model there are differences in the assessment of the validity of the questions between the two approaches. The Rasch model shows a variety of difficulty levels, while classical test theory does not have good quality. This study shows differences in the evaluation of concept understanding measurement instruments between classical test theory and the Rasch model. The validity, reliability, difficulty level, and discriminating power of questions are important aspects that need to be considered in the development of evaluation instruments.

Based on observations made, by reading existing research journals, it was found that the preparation of assessment tests still uses a lot of classical test theory. This traditional approach, although practical, has potential weaknesses in producing accurate measurements of students' knowledge abilities. The shortcomings of classical test theory, such as its dependence on group homogeneity and its inability to consider the influence of the test

itself, encourage educational practitioners to look for more modern and appropriate alternatives. This solution is found in the Rasch method, an approach based on Item Response Theory (IRT) that offers several significant advantages.

The Rasch method is able to measure individual abilities independently, identify items that do not function properly, and is applied to various test contexts, including adaptive tests. Research has shown the effectiveness of the Rasch Method in improving the quality of mathematics learning assessments, such as the research of Susdenila, et al in 2018 and Novi Indriyani Kones & Raden Rosnawati in 2021. The paradigm shift from classical test theory to the Rasch Method signifies a significant step forward in improving the quality of mathematics education in Indonesia. The application of the Rasch Method is expected to result in a more objective, comprehensive and fair measurement of students' abilities, thus providing a more accurate picture of their potential and learning needs.

Based on the background of the problem above, a study was made to examine the preparation of a critical reasoning assessment instrument model as an achievement of the Pancasila Student Profile in phase D (SMP). Based on the background of the problem above, it can be obtained identification of problems that arise in this study is in developing a critical reasoning assessment instrument model which is expected to be used as a measuring tool to determine the achievement of the Pancasila Student profile, especially in the project learning process. In developing an assessment instrument model, it is necessary to pay attention to the development model to ensure the quality of the assessment instrument in supporting learning effectiveness, because development is basically a process that is linear with the learning process. One of the assessment instrument development designs that is often used is the ADDIE Model through 5 stages: Analysis, Design, Development, Implementation and Evaluation. The development process requires several times of expert team testing, individual research subjects, limited scale and wide scale (field) and revision to improve the final product so that although the development procedure is shortened, it includes a testing and revision process so that the product developed has met the criteria for a good product, empirically tested and there are no more mistakes.

The development of the Critical Thinking Ability model assessment as an achievement of the Pancasila Learner Profile in Phase D is focused on the use of the Research and Development (R&D) method with the ADDIE model which is only carried out up to the ADD stage. This restriction includes the Analysis, Design, and Development stages, without involving the Implementation and Evaluation stages. Thus, the development of the assessment model will only result in designs and prototypes that have not been tested or evaluated thoroughly in the field. This can limit the validity and reliability of the assessment model produced, as well as reduce the opportunity to obtain empirical input that can be used for further improvement.

The problem limitation in this study is testing the validity of the assessment instrument of the Critical Thinking Ability model as an achievement of the Pancasila Student Profile in Phase D which will be developed using the Rasch model modern test theory. This test instrument is designed to measure Critical Thinking Ability as an Outcome of Pancasila Learner Profile in junior high school. The Rasch model approach was chosen because of its ability to provide a more detailed and accurate analysis of test items and student abilities, as well as allowing the identification of the suitability of test questions to the level of student abilities. By using the Rasch model, it is expected that the test instrument developed can validly measure the achievement of the Pancasila Learner Profile in accordance with the independent curriculum standards.

Based on the problem boundaries above, the problem formulations in this study are how is the analysis of the needs of teachers and principals on the critical thinking ability assessment instrument as the achievement of the Pancasila learner profile in phase D of the independent curriculum?, how is the research instrument design of the critical thinking ability assessment model as the achievement of the Pancasila learner profile in phase D of the independent curriculum used?, how is the validity of the Critical Thinking Ability

assessment instrument as the achievement of the Pancasila learner profile in phase D developed by measuring the Rasch model approach?.

Based on the formulation of the problem above, the research objectives in this study are to analyze the needs of teachers, and school principals for the critical thinking ability assessment instrument model as the achievement of the Pancasila learner profile in phase D of the independent curriculum, develop a critical thinking ability assessment test instrument model as the achievement of the Pancasila learner profile in Phase D of the independent curriculum, Proving the validity of the critical thinking ability assessment instrument model as the achievement of the Pancasila learner profile in Phase D developed using the Rasch model.

METHOD

This is research and development, which is meant by development is the development of a product in the form of an instrument for assessing students' critical thinking skills as an achievement of the Pancasila learner profile in phase D. Product development activities involve multi-disciplines in terms of design, process, product quality, assurance, technological innovation, marketing and business management. In creating new products, R&D utilizes the latest developments in processing technology and raw materials to obtain quality products and safe for consumption.

Research and development serves to validate products and develop products. Validating a product means that the product already exists, and researchers only test the effectiveness or validity of the product. Developing products in a broad sense can be in the form of updating existing products (so that they become more practical, effective, and efficient) or creating new products (which have never existed before). Borg and Gall use the name Research and Development R&D which can be translated into research and development. Richey, and Kelin use the name Design and Development Research which can be translated into Design and Development Research. Thiagarajan uses the 4D Model which stands for Define, Design, Development and Dissemination, Dick and Carry uses the term ADDIE (Analysis, Design, Development, Implementation, Evaluation), and Development Research, which can be translated into development research.

The R&D development research used is from Dick and Carry, ADDIE (Analysis, Design, Development, Implementation, Evaluation) with the Rasch model approach. The product produced in this study is a mathematics learning achievement test question at the end of phase D of the independent curriculum. The object of this research is the mathematics learning achievement assessment instrument at the end of phase D. In this research, the development of mathematics learning achievement test questions at the end of phase D was developed by testing the validity of the content aspects, the validity of the psychometric aspects and the validity of the test construct.

This research was conducted at SMP Negeri 2 Pagerbarang, this location is considered relevant and strategic to achieve the research objectives. By conducting research at SMP Negeri 2 Pagerbarang, it is hoped that the data and findings obtained can make a significant contribution to the need to implement the critical thinking ability assessment as an achievement of the Pancasila learner profile in Phase D. The population in this study were all students in grade IX of SMP Negeri 2 Pagerbarang, while the sample was taken from students in grade IX of SMP Negeri 2 Pagerbarang totaling 240 students.

Sampling technique is a sampling technique. To determine the sample to be used in research, there are various sampling techniques (Sugiyono 2019). This R&D research uses saturated samples, which are sampling techniques when all populations are sampled. This saturated sample provides equal opportunities for each element (member) of the population to become a sample member. In this study, the sample was ninth grade students at SMP Negeri 2 Pagerbarang. There are several factors for researchers to take saturated samples, including: a. high validation of results, by using the entire population as a sample,

researchers can ensure that the research results are more valid because no element is missed in the analysis, b. saturated samples to avoid bias that might arise if only part of the population is taken as a sample, c. the development of the test instrument model uses all the data that is very important to produce an accurate and reliable model. This research focuses on designing and compiling critical thinking ability assessment instruments as an achievement of the Pancasila learner profile in phase D. The research aims to identify students' critical thinking skills in achieving the Pancasila learner profile in phase D, create a valid assessment instrument to measure critical thinking skills as an achievement of the Pancasila learner profile, test the assessment model instrument made on a group of students and analyze the results.

Research instruments refer to tools or means used by researchers to collect research data with the aim of facilitating and improving the quality of their work, such as accuracy, completeness and systematicity, so that data can be more easily processed. In this study, the instruments used include expert validation sheets and items tested on students to evaluate their ability to respond to questions. The data collected in this study by means of documentation, expert validation where in the form of reviews from experts related to the design of critical thinking skills questions as the achievement of the Pancasila learner profile in phase D made by researchers. In this case, expert validation can provide suggestions and input related to aspects of content, construct and language which are then made improvements by researchers on these suggestions and input, test devices in the form of test instruments in the form of a set of test questions tested on students. The test instrument is used to determine the achievement of students in mastering aspects of critical thinking skills as the achievement of the Pancasila learner profile in phase D. Data collection techniques are things or methods used by researchers in obtaining data related to the problems taken. This technique aims to ensure that the data collected is relevant, accurate, and can be used to answer research questions or achieve research objectives. Data collection, especially in schools, can be done in several ways, namely questionnaires or questionnaires, interviews, test questions and documentation. R&D (Research and Development) research on critical thinking ability assessment model instruments as an achievement of the Pancasila student profile in phase D uses the Rasch model as the main approach in data analysis. The Rasch model is a statistical measurement model used to analyze test or assessment data. Data analysis techniques in this study will most likely include the following steps:

Content Aspect Validity Test

Content validity on the critical thinking ability assessment model instrument as an achievement of the Pancasila student profile in phase D must have a domain that will be clearly measured and described in the grids and test instruments. Content validity test is one type of validity test that aims to assess the extent to which items in measuring instruments, such as tests or questionnaires, actually represent the domain or concept to be measured. In other words, the content validity test ensures that the content of the measuring instrument is relevant, appropriate, and adequate to measure what is intended.

Four experts will be involved in the research, each of whom is an expert in the field of P5 as seen from their functional positions, structural positions and the quality of their scientific publications.

Psychometric Validity Test

The psychometric validity test is a process to evaluate how well a measuring instrument (such as a questionnaire or psychological test) measures what should be measured. This validity test is important to ensure that the data obtained from the measuring instrument reflects the intended construction. Psychometric validity testing is a series of procedures to assess the extent to which a measuring instrument, such as a test or questionnaire, actually measures what it is intended to measure.

The validity of this psychometric aspect is to ensure that the test items meet the psychometric rules in the preparation of items which include material, construction and language aspects. For the process of assessing the validity of psychometric aspects, using 4

sources each from psychometricians and school teachers. The results obtained from the review of the instruments of the three resource persons will be categorized according to the evaluation result criteria in table 5.

Content Validity Test

Content validity indicates whether all test items or tasks that involve cognitive processes in answering them are really appropriate and representative of the construct field being measured. The content aspect of construct validity is at least related to three things, namely the suitability of content representation and technical quality. The extent to which test items or tasks are appropriate and representative of the construct field is generally carried out by experts. The technical quality of test items relates to issues about the level of reading accuracy, ambiguous language and the accuracy of the answer key. (Susongko Purwo, 2019)

Decision-making techniques in R&D Research with the Rasch model approach on the critical thinking ability assessment model instrument as an achievement of the Pancasila student profile in phase D can be used to make decisions related to a number of things as follows. When developing an assessment instrument, the results of the Rasch model analysis can be used to identify items that are not functioning properly and need to be revised or removed from the assessment instrument. This aims to increase the validity of the assessment instrument.

The results of the Rasch model analysis can be used to identify learners who have difficulty working on the questions. This information can be used to develop appropriate learning solutions to help these learners. Education policy, the results of the Rasch model analysis can be used to obtain information about the overall achievement of P5 Phase D critical reasoning of learners. This information can be used to develop more effective education policies to improve the quality of education.

FINDINGS AND DISCUSSION

This research uses the Research and Development (R&D) research method, so that the steps in the research procedure must be followed to produce a product in the form of a test instrument that is in accordance with the needs analysis. The needs analysis activity of the critical thinking ability test instrument as the achievement of the Pancasila student profile was carried out using interview techniques, literacy analysis by reading previous studies and curriculum analysis by studying the achievement of critical thinking skills in strengthening the Pancasila student profile project in the independent curriculum. The analysis conducted by researchers shows that the unavailability of a critical thinking ability test instrument as an achievement of the Pancasila student profile at the end of phase D of the independent curriculum is a significant deficiency in the education system. Previous studies have shown that the independent curriculum has very high goals in improving the quality of education, but its implementation still has several obstacles. One of these obstacles is the absence of an effective test instrument to measure the achievement of critical thinking skills in learning to strengthen the Pancasila learner profile of students at the end of phase D.

The learning process to find out whether the learning objectives are maximally achieved or not requires evaluation. Evaluation in the learning process requires a measuring tool in the form of assessment. Assessment or also known as assessment is an application and use of various ways and tools to obtain a series of information about the learning outcomes and competency achievements of students. Assessment can also be called a process that provides and determines value to various specific objects based on certain criteria. The benefit of learning assessment is to provide a complete explanation of learning targets that can describe things such as how the level of knowledge of learners, and the information needed about the knowledge of learners.

The development model that researchers use is the ADDIE development model, which is a development model consisting of Analysis, Design, Development, Implementation, and

Evaluation. The research and development that researchers do is research and development in level 1, namely not researching but testing in other words that research that does not make product designs through research, but only conducts internal product validity tests.

Table 1. Test Instrument Content Validity Assessment Results

No	Aspects assessed	Expert	
		I	II
1	Are the items in line with the learning outcome objectives listed in the independent curriculum?	Meet	Meet
2	Do the items measure all aspects of the learning outcome indicators?	Meet	Meet
3	Is the language used in the items clear and easy to understand?	Meet	Meet
4	Correct answer key	Meet	Meet

After the development of test instruments is carried out through several stages, starting from the formulation of indicators, the preparation of grids, to the making of questions. Furthermore, the test items that have been prepared are tested for validity. Based on the analysis of the content validity test that has been carried out, the results show that: (1) The items are in accordance with the learning outcome objectives listed in the independent curriculum, (2) The items measure all aspects of the learning outcome indicators, (3) The language used in the items is clear and easy to understand, and (4) The answer key is correct. Based on the results of the psychometric validity review, it can be stated that the critical thinking ability test items that have been prepared are feasible from the psychometric aspect. The results of the assessment of psychometric aspects from assessors 1, 2, 3 and 4 showed good results in terms of material, construction, and language used in the question items. The aspects assessed from the psychometric aspect include: (1) The subject matter must be formulated clearly and clearly, (2) The formulation of the subject matter and answer choices must be a necessary statement only, (3) The subject matter should not give clues towards the correct answer, (4) The subject matter does not contain double negative statements, (5) Answer choices should not contain the statement "All answer choices above are correct", or "All answer choices below are correct", (6) The problem written on the question is clear and easy to understand (7) Each question must contain language that is in accordance with Indonesian language rules, (8) Do not use local language, if the question will be used for other regions or nationally, (9) Each question must use communicative language, (10) Answer choices must not repeat words or phases that do not constitute a unified understanding.

No	Aspek yang dinilai	Expert	
		I	II
1	The subject matter should be formulated clearly and firmly	Meet	Meet

Critical Thinking Ability Test Instrument Design with Validity Test Pearson Correlation

2	The formulation of question items and answer choices should be necessary statements only.	Meet	Meet
3	The subject matter should not give clues to the correct answer	Meet	Meet
4	The subject matter should not contain double negative statements	Meet	Meet
5	Answer options should not contain the statements "All answer options above are correct", or "All answer options below are correct".	Meet	Meet
6	The problem written on the problem is clear and easy to understand	Meet	Meet
7	Each question must contain language that is in accordance with Indonesian language rules	Meet	Meet
8	Do not use local language, if the question will be used for other regions or nationally.	Meet	Meet
9	Each question must use communicative language	Meet	Meet
10	Answer choices should not repeat words or phases that do not constitute a unified understanding.	Meet	Meet

Table 2. Results of the Validity Assessment of Psychometric Aspects of the Test Instrument

The results of validation by experts who assess the content aspects and psychometric aspects obtained the results that the test instrument is suitable for testing. After the test instrument is complete, product trials are carried out to respondents. The results of the analysis of 240 respondents working on 20 items of question items using the person correlation validity test. To determine the level of validity of the construction of a research instrument, it can use the formula for calculating the Aiken's V validity coefficient:

$$V = \frac{\sum s}{[n(c-1)]}$$

$\sum s$ = the grand total of the difference between the score given by the expert and the lowest score of the importance rating.

$S = r - lo$

r = the number given by the assessor

lo = lowest validity rating number

n = number of experts involved

c = highest validity rating number

Table 3. Results of Reviewing the Validity of Psychometric Aspects of Critical Thinking Ability Instrument

Butir soal	Aspek Psikometri										Aiken'S V	Kategori
	1	2	3	4	5	6	7	8	9	10		
1	1	1	1	1	1	1	1	1	1	1	1	Valid
2	1	1	1	1	1	1	1	1	1	1	1	Valid
3	1	1	1	1	1	1	1	1	1	1	1	Valid
4	1	1	1	1	1	1	1	1	1	1	1	Valid
5	1	1	1	1	1	1	1	1	1	1	1	Valid
6	1	1	1	1	1	1	1	1	1	1	1	Valid
7	1	1	1	1	1	1	1	1	1	1	1	Valid
8	1	1	1	1	1	1	1	1	1	1	1	Valid
9	1	1	1	1	1	1	1	1	1	1	1	Valid
10	1	1	1	1	1	1	1	1	1	1	1	Valid
11	1	1	1	1	1	1	1	1	1	1	1	Valid
12	1	1	1	1	1	1	1	1	1	1	1	Valid
13	1	1	1	1	1	1	1	1	1	1	1	Valid
14	1	1	1	1	1	1	1	1	1	1	1	Valid
15	1	1	1	1	1	1	1	1	1	1	1	Valid
16	1	1	1	1	1	1	1	1	1	1	1	Valid
17	1	1	1	1	1	1	1	1	1	1	1	Valid
18	1	1	1	1	1	1	1	1	1	1	1	Valid
19	1	1	1	1	1	1	1	1	1	1	1	Valid
20	1	1	1	1	1	1	1	1	1	1	1	Valid

Source: processed data

The results of validation by 4 validators / judgment experts for each item can be concluded that the instrument for achieving critical reasoning skills at the end of phase D that has been made is feasible from the psychomotor aspect and can be followed up with empirical trials. The results of the analysis showed that all items had significant correlation coefficients, with values above the established threshold. This shows that the instrument developed is declared valid. The validity of this instrument is very important to ensure that the evaluation of learning outcomes is in accordance with the standards set by the Merdeka Curriculum. With a valid instrument, it is expected to provide an accurate picture of students' ability to solve math problems. The results of the Pearson correlation analysis showed that all items had a significant correlation with the total score of the instrument, confirming the construct validity of this instrument. This finding is in line with the research objective to develop an instrument that can accurately measure students' mathematics learning outcomes in accordance with the Merdeka Curriculum. The instrument development process involving in-depth study of curriculum materials, preparation of relevant items, and empirical testing of the instrument has resulted in a valid and reliable instrument for measuring students' critical thinking skills.

CONCLUSIONS

The results of psychometric validity involving 4 raters / experts confirmed that the critical thinking ability assessment instrument as an achievement of the Pancasila learner profile in phase D has been declared feasible in terms of construction and language based on the Aiken Validity formula. This shows that the instrument has successfully met the quality standards required in measuring learner knowledge effectively and reliably. This research is expected to make a positive contribution to the development of educational assessments that are more effective and efficient, and support the achievement of curriculum objectives. The construct validation of the critical thinking ability assessment model in phase D obtained the following results: a) At the 100% confidence level, the items fit the model, b) The level of difficulty of the items in the range between -0.4977 to 2.1681, 17 questions or 85% are included in the good category (based on the range of measure values from easy to difficult categories) in accordance with the required criteria, namely between -2 to +2, c) At the

99.17% confidence level, the test takers' abilities fit the model, d) The results of the model feasibility test, substantive validity are said to be valid and feasible to use. Goodness of Fit output results are as follows: Collapse Deviance test 0.012, Hosmer-Lemeshow 0.001, and Casewise Deviance 0.000. A p-value of less than 0.01 satisfies model goodness of fit tests such as Collapse Deviance and Casewise Deviance. In addition, the p-value for accuracy is 62.1%, sensitivity is 47.9%, and specificity is 0.736 which means all are close to the value of 1.0, e) The validity of the structural aspects of the unidimensional test explains 53.795% of the variation and the scree plot graph slopes down so that it can be confirmed that Ho is accepted and the test contains one dimension, f) Anderson LR invariance test p-value is 0, Ho is accepted so that the parameters are invariant, g) At the 100% level there are no items detected DIF.

ACKNOWLEDGEMENTS

The author would like to thank all those who have supported and participated in the completion of this thesis. Special gratitude goes to Dr. Taufiqulloh, M. Hum., Rector of Pancasakti University Tegal; Dr. Fajar Ari Sudewo, S.H., M.H., Director of the Postgraduate Program; and Dr. Suriswo, M.Pd., Head of the Master of Pedagogy Study Program. The author also sincerely appreciates Prof. Dr. Purwo Susongko, M.Pd., as the first supervisor, and Dr. Tri Mulyono, S.Pd., M.Pd.I., as the second supervisor, for their invaluable guidance, direction, and time throughout the thesis preparation process.

REFERENCES

- Afifah, S. N., & Kusuma, A. B. (2021). Pentingnya Kemampuan Self-Efficacy Matematis Serta Berpikir Kritis pada Pembelajaran Daring Matematika. *JURNAL MathEdu (Mathematic Education Journal)*, 4(2), 313-320.
- Arikunto, S., (2019), *Prosedur Penelitian: Suatu Pengantar Praktik*, Jakarta: Rineka Cipta.
- Arofah, L., & Nawantara, R. D. (2019). Pentingnya Critical Thinking Bagi Siswa dalam Menghadapi Society 5.0. Kediri: Universitas Nusantara PGRI Kediri.
- Alec Fisher. 2009. *Berpikir Kritis: Sebuah Pengantar*. Jakarta: Erlangga.
- Facione, P. A. (2015). *Critical Thinking : What It Is and Why It Counts*. In *Insight assessment* (Issue ISBN 13: 978-1-891557-07-1.). <https://www.insightassessment.com/CT-Resources/Teaching-For-and-AboutCritical-Thinking/Critical-Thinking-What-It-Is-and-Why-It-Counts/CriticalThinking-What-It-Is-and-Why-It-Counts-PDF>
- Faiz, A., & Kurniawaty, I. (2022). Urgensi Pendidikan Nilai di Era Globalisasi. *Jurnal Basicedu*, 6(3).
- Firmansyah, D. (2022). Analisis Kemampuan Berpikir Kritis Matematis Siswa Sma Pada Materi Barisan Aritmatika. *Jumlahku: Jurnal Matematika Ilmiah STKIP Muhammadiyah Kuningan*, 8(1), 30-44.
- Gustianingrum, R., Murni, A., & Maimunah, M. (2023). Analisis Kemampuan Berpikir Kritis Peserta Didik dalam Menunjang Penguatan Profil Pelajar Pancasila. *PRISMA, Prosiding Seminar Nasional Matematika*, 6, 465-470.
- Hadiansah, Deni. (2022). *Kurikulum Merdeka dan Paradigma Pembelajaran Baru*. Bandung: Penerbit Yrama Widya.
- Hernawan, Anggi. (2018). "Pengembangan Instrumen Asesmen Berpikir Kritis pada Pembelajaran Tematik Kelas IV Sekolah Dasar di Kecamatan Tumijajar". *Journal of Chemical Information and Modeling*, 7(1), 1-10.
- Indarta, Y., Jalinus, N., Abdullah, R., & Samala, A. D. (2021). 21st Century Skills : TVET dan Tantangan Abad 21. *Edukatif: Jurnal Ilmu Pendidikan*, 3(6), 4340-4348.
- Kemendikbud. (2018). *Rencana Strategis (Renstra) Kementerian Pendidikan dan Kebudayaan Tahun 2015-2019*. Jakarta: Kemendikbud RI. Tersedia di [https://Kementerian Pendidikan dan Kebudayaan » Republik Indonesia \(kemdikbud.go.id\)](https://Kementerian Pendidikan dan Kebudayaan » Republik Indonesia (kemdikbud.go.id))

- Kurniastuti, Rahmaniar, D. (2022). Implementasi Profil Pelajar Pancasila Sebagai Salah Satu Bentuk Pendidikan Karakter Pada Siswa SMP. *Seminar Nasional Sosial Sains, Pendidikan, Humaniora (SENASSDRA)*, 1, 287–293.
- Kurniawan, Dian. 2020. *Assessemnt Learning (AFL) dalam Pendidikan Matematika*. Yogyakarta: CV Budi Utama
- Kusaeri dan Suprananto. 2012. *Pengukuran dan Penilaian Pendidikan*. Yogyakarta: Graha Ilmu.
- Lestari, S. Z. D., & Roesdiana, L. (2021). Analisis Kemampuan Berpikir Kritis Matematis Siswa SMP Pada Materi Himpunan. *MAJU: Jurnal Ilmiah Pendidikan Matematika*, 8(1).
- Lia, A., Penguatan Profil Pelajar Pancasila Bernalar Kritis Melalui Karya Tulis Ilmiah. Vol 4, No 1 (2023).
- Messick, S. (1996). Standar validitas dan validitas standar dalam kinerja. *penilaian. Pengukuran Pendidikan: Isu dan Praktik*, 14 (4), 5–8. <https://doi.org/10.1111/j.1745-3992.1995.tb00881.x>
- Nugraheni, Rachmawati. (2022). *Projek Penguatan Profil Pelajar Pancasila dalam Implementasi Kurikulum Prototipe di Sekolah Penggerak Jenjang Sekolah Dasar*. *Jurnal Basicedu Volume 6 Nomor 3 Tahun 2022* Halaman 3613 – 3625.
- Nuryanti, L., Zubaidah, S., & Diantoro, M. (2018). Analisis Kemampuan Berpikir Kritis Siswa SMP. *Jurnal Pendidikan: Teori, Penelitian, dan Pengembangan*, 3(2), 155–158.
- Pertiwi, W. (2018). Analisis Kemampuan Berpikir Kritis Matematis Peserta Didik SMK Pada Materi Matriks. *Jurnal Pendidikan Tambusai*, 2(2), 821–831.
- Prasetyo, N. H., & Firmansyah, D. (2022). Analisis Kemampuan Berpikir Kritis Matematis Siswa Kelas VIII dalam Soal High Order Thinking Skill. *Jurnal Educatio FKIP UNMA*, 8(1), 271–279.
- Priyadi, R., Mustajab, A., Tatsar, M. Z., & Kusairi, S. (2018). Analisis kemampuan berpikir kritis siswa SMA kelas X MIPA dalam pembelajaran fisika. *JPFT (Jurnal Pendidikan Fisika Tadulako Online)*, 6 (1), 53–55.
- Puslitjakdkbud. 2021. *Meningkatkan Kemampuan Literasi Dasar Siswa Indonesia Berdasarkan Analisis Data PISA 2018*. Pusat Penelitian Kebijakan: Kemdikbud
- Rosmalah, Asriadi, & Shabir, A. (2022). Implementasi Profil Pelajar Pancasila Dimensi Bernalar Kritis dalam Proses Pembelajaran di Sekolah Dasar. *Seminar Nasional Hasil Penelitian*, 969–975. <https://ojs.unm.ac.id/semnaslemlit/article/download/39822/18859>
- Sugiyono. (2019). *Metode Penelitian Kuantitatif Kualitatif dan R&D*. Bandung: Alfabeta.
- Suharsimi Arikunto, *Dasar-dasar Evaluasi Pendidikan*, Jakarta: Bumi Aksara, 2013.
- Sumintono, B dan Widhiarso, W. (2013). *Aplikasi Model Rasch untuk Penelitian Ilmu-ilmu Sosial*. Cimahi: Trim Komunikata Publishing House.
- Susongko Purwo. (2019). *Aplikasi Model Rasch Dalam Pengukuran Pendidikan Berbasis Program R*. Badan Penerbit UPS Tegal.
- Syafitri, E., Armanto, D., & Rahmadani, E. (2021). Aksiologi Kemampuan Berpikir Kritis (Kajian Tentang Manfaat dari Kemampuan Berpikir Kritis). *Journal of Science and Social Research*, 4(3), 320–325.
- Uno, H.B. & Koni, S. (2014). *Assesment Pembelajaran*. Jakarta : Bumi Aksara.
- Widana, I. W., Sumandya, I. W., & Dini Prastanti, N. P. (2023). Implementasi Metode Star Untuk Meningkatkan Kemampuan Guru Mengembangkan Modul Projek Penguatan Profil Pelajar Pancasila. *Jurnal PkM Pengabdian Kepada Masyarakat*, 5(6), 696. <https://doi.org/10.30998/jurnalpkm.v5i6.15621>
- Yusuf, M. (2015). *Asesmen Dan Evaluasi Pendidikan*. Jakarta: Prenada Media Group.
- Zakiah, Linda & Lestari, Ika. (2019). *Berpikir Kritis Dalam Konteks Pembelajaran*. Bogor: Erzatama Karya Abadi