Jurnal Pendidikan Dasar Nusantara Vol 11, No 2, January 2026
ISSN 2579-6461 (Online) ISSN 2460-6324 (Print)
DOI: https://doi.org/10.29407/jpdn.v11i2.27895

Development of Creative Thinking Instruments On The Material of Changes In The Form of Objects In Elementary Schools

Submitted:
December 19, 2025
Accepted:
January 12, 2026
Published:
January 31, 2026

Arum Dewi Rinjani [1], Cindy Putri Ranny Azhari [2],
Diana Madya Ratri [3], Eal Ednov Athallah [4], M. Anas Thohir [5*]
arum.dewi.2401516@students.um.ac.id[1]
cindy.putri.2401516@students.um.ac.id[2]
diana.madya.2401516@students.um.ac.id[3]
eal.ednov.2401516@students.um.ac.id[4]
anas.thohir.fip@um.ac.id[5]
Elementary School Teacher Education, Faculty of Education,
Universitas Negeri Malang[1,2,3,4,5]

Abstract: This study aims to develop an assessment instrument for creative thinking skills on the material of changes in the state of matter for elementary school students. The study used a research and development (R&D) method with the ADDIE model which includes the stages of analysis, design, development, implementation, and evaluation. The instrument was compiled based on four indicators of creative thinking, namely fluency, flexibility, originality, and elaboration. In the initial stage, 21 items were developed which were then validated by four validators using Aiken's V and tested for inter-rater reliability using the Intraclass Correlation Coefficient (ICC). The results of content validity showed that 10 items met the valid criteria with an Aiken's V value ≥ 0.92. The inter-rater reliability test showed that the ICC value for each indicator was in the reliable to very reliable category, with a value range between 0.846 and 0.901, which indicated good consistency in inter-rater assessment. The selected instrument was then empirically tested on 27 elementary school students. The results of the empirical validity test using Pearson Product Moment correlation showed that 9 out of 10 questions were declared valid with a rT value of > 0.381. Furthermore, the empirical reliability test using Cronbach's Alpha showed a reliability value in the sufficient category, with a coefficient ranging from 0.427 to 0.570. Based on these results, the creative thinking instrument developed was declared suitable for use as an evaluation tool for students' creative thinking abilities in science learning on the material of changes in the state of objects in elementary schools.

Keyword: Creative thinking, Instrument validity, Reliability, Changes in the form of objects

PRELIMINARY

In the 21st century era, the education system is not only required to equip students with content mastery skills, but also higher skills such as critical thinking skills, creativity, communication and collaboration (Astuti, 2024; Wahyuni et al., 2025; Maimun et al., 2024). The ability to think creatively is one of the important skills that

ISSN 2579-6461 (Online) ISSN 2460-6324 (Print)
DOI: https://doi.org/10.29407/jpdn.v11i2.27895

students must have, especially in the context of learning in the 21st century era. In this case, creative thinking skills are very important for students to master, because this skill allows students to generate new ideas, think more flexibly, and find innovative solutions when facing various challenges both in the learning process and in everyday life (Kastur et al., 2025; Irman et al., 2025). In Indonesia itself, the demand for mastering this kind of competency is in line with efforts to form student profiles that are adaptive and have character and abilities according to the demands of the times, which is realized through the 8 Dimensions of the Graduate Profile.

In the context of basic education, especially science (Natural Science) learning in class V, developing creativity through science learning is proven to have great potential. The ability to think creatively in science is not only needed so that students can memorize concepts, but allows students to more actively explore natural phenomena, ask questions, design predictions, and provide alternative explanations, thereby making the learning process more meaningful and in-depth (Gilvando et al., 2025; Krismanita & Qosyim, 2021; Cahyaningsih et al., 2023). That way, students' creativity will support a deeper and more meaningful learning process, where students not only know about "what" but also "why" and "how".

One of the science materials at elementary school level that is relevant for training creative thinking is the material on changes in the form of objects. Through this material, students can observe changes in the form of objects that occur in their daily lives (for example ice melts, water boils, steam condenses, camphor sublimates), then explain these phenomena based on what they have learned, as well as predict events, for example "what changes will occur in the form of an object if the temperature changes?", or "what are examples of changes in the form of objects that exist?" around me?" Apart from that, students can also design new appropriate examples, compare several different situations, and construct analogies or similes to understand changes in the form of objects. With the characteristics of material changing the shape of objects which involves observation, prediction and reasoning, this material can open up opportunities for students to observe physical changes and provide explanations based on experience, analogies or their own creations, so it is very suitable as a medium for measuring and developing creative thinking abilities.

ISSN 2579-6461 (Online) ISSN 2460-6324 (Print)
DOI: https://doi.org/10.29407/jpdn.v11i2.27895

However, based on the results of initial observations carried out in class V of UPT SDN 2 Karangtengah, it was found that the practice of assessing science learning is still dominated by the use of written questions in the form of multiple choice and short essays which emphasize mastery of concepts and memorization. The results of short questions and answers with class teachers show that teachers do not yet have assessment instruments that are specifically designed to measure students' creative thinking abilities, especially on the indicators of fluency, flexibility, originality, elaboration. Teachers tend to assess creativity subjectively based on student activity during learning, without being supported by a structured rubric or instrument. Apart from that, analysis of the results of student learning evaluations on material changing the shape of objects shows that most of the questions only require one correct answer and do not provide space for students to put forward various alternative answers, provide explanations from different points of view, or develop ideas in depth. This condition indicates that students' creative thinking abilities have not been optimally facilitated and measured through assessment instruments used in science learning in elementary schools.

Apart from that, the availability of instruments specifically designed to measure the creative thinking abilities of elementary school students, especially in science subjects, is still very limited. Most of the research on the development of creative thinking ability instruments is focused on the junior high school (SMP) level. For example, instrument development carried out by Trimawati et al. (2020) and Widayanti et al. (2024) which focuses on students' critical thinking abilities at the junior high school level. The instruments developed for elementary school level, as done by Syaifullah et al. (2024) and Harta et al. (2021), has not specifically measured students' creative thinking abilities and has not focused on subjects certain IPAs. This shows that there is a gap between the need for assessing creative thinking abilities at the elementary school level and the availability of instruments that are relevant, contextual and have been tested.

Based on these conditions, an evaluation instrument is needed that is valid, reliable and appropriate to the level of development of students, which is specifically designed to measure students' creative thinking abilities at elementary school level in science subjects, especially in the material on changes in the form of objects. It is hoped that this instrument can help teachers to assess students' creative thinking abilities more precisely and objectively, and can support the implementation of science learning which

ISSN 2579-6461 (Online) ISSN 2460-6324 (Print)
DOI: https://doi.org/10.29407/jpdn.v11i2.27895

not only prioritizes cognitive abilities but also develops creativity and high-level thinking abilities.

Based on the results of observations and analysis of the characteristics of class V students at UPT SDN 2 Karangtengah, it is known that students are at the stage of concrete operational cognitive development towards early formal operations. At this stage, students more easily understand learning concepts when presented in a real context and close to everyday life. However, the assessment instruments used by teachers still tend to measure the ability to remember and understand, so they do not fully describe the abilities related to remembering and understanding, so they do not yet have a picture of students' creative thinking abilities. In these conditions, it shows that there is a gap between student characteristics and the assessment instruments used, so it is necessary to develop assessment instruments that are appropriate to students' cognitive development and learning needs.

Therefore, based on various considerations based on the conditions above, an instrument was developed with the title "Development of a Creative Thinking Instrument on Material for Changes in the Form of Objects in Class V Elementary Schools", which is expected to provide an instrument for assessing creativity in science subjects, as well as making a contribution to teachers and schools in evaluating and developing students' creative thinking abilities systematically and more meaningfully.

METHOD

Research Procedures

This research refers to the ADDIE development model which is a generic framework in instructional system design. Allen (2006) revealed that the ADDIE model consists of five phases, namely: Analysis, Design, Development, Implementation, Evaluation.

1. **Analysis stage**

   At this stage the main foundation for determining "what" will be developed. According to Allen (2006), analysis aims to identify performance gaps and determine the material that needs to be studied. This research includes a needs analysis to identify the lack of specific assessment instruments; Audience analysis to examine

ISSN 2579-6461 (Online) ISSN 2460-6324 (Print)
DOI: https://doi.org/10.29407/jpdn.v11i2.27895

the characteristics of class V students at UPT SDN 2 Karangtengah so that the instruments created are relevant to the students' cognitive development.

2. **Design stage**

This stage focuses on planning "how" the content will be presented. According to Allen (2006), this phase includes the creation of blueprints and design specifications including: arranging grids based on four indicators (fluency, flexibility, originality, elaboration) combined with material changes in the shape of objects.

3. **Development Stage**

This phase is the process of realizing design specifications into a physical product. As stated by Allen (2006), this stage involves material production and validation to ensure its quality before distribution. This stage includes: creating items that will be used for tests and validation sheets; Validation by fellow students supervised by the lecturer to examine the appropriateness of content, language and construction; and measuring inter-rater reliability using the Interclass Correlation Coefficient (ICC) technique to ensure the objectivity of validator assessments.

4. **Implementation Stage**

Implementation is the stage for delivering the product to the actual audience. According to Allen (2006) this stage is the process of sending or implementing instructions to participants including: independent testing of participants research subjects, namely 27 class V UPT SDN Karangtengah students to produce empirical score data.

5. **Evaluation Stage**

This final stage aims to measure the effectiveness of the product that has been developed. Allen (2006) states that evaluation is carried out to ensure whether development objectives have been achieved. This research focuses on psychometric analysis of trial results including: Empirical validity (Pearson's Product Moment) and Empirical Reliability (Cronbach's Alpha).

Types of research

The method used in this research is a test with creative thinking indicators sourced from various experts such as López Martínez et al. (2004), Sahliawati & Nurlaelah (2020), Lu et al. (2025) , and Wechsler (2006) . Based on 4 journals used as references, it was concluded that there were 4 indicators that were in accordance with the creative

ISSN 2579-6461 (Online) ISSN 2460-6324 (Print)
DOI: https://doi.org/10.29407/jpdn.v11i2.27895

thinking instrument in the material on changes in the form of objects in grade 5 elementary school, namely: Fluency (fluency in thinking), Flexibility (flexibility in thinking), Originality (originality of ideas), Elaboration (development of ideas).

Table 1. Four Creativity Indicators and the items that represent them

No.	Indicator	Questions
1.	Fluency	6
2.	Flexibility	6
3.	Originality	6
4.	Elaboration	3

Ten questions were further developed by fellow students supervised by lecturers from the Department of Elementary School Teacher Education, State University of Malang. Validation is carried out to obtain an assessment covering four aspects, namely (1) Theoretical Truth, (2) Construction and correctness of content, (3) Appropriateness in presentation, and (4) Use of language.

Instrument Validity Test

The validity test in this research was carried out through validity per item to ensure that the items developed were in accordance with the competency to be measured, namely the ability to think creatively on material changing the shape of objects in grade 5 elementary school. Content validity is obtained through the assessment of fellow students who have been supervised by the lecturer. The validators assess each instrument item based on 4 aspects, namely: Theoretical correctness, Construction and content correctness, Presentation, and Language. The assessment uses a 1-4 Likert scale, with information

1 = very invalid 3 = valid

2 = invalid 4 = very valid

The data from student peer assessments was then analyzed using the Aiken's V formula to determine the level of agreement of the validators regarding the suitability of the items with the indicators as proposed by Aiken (1985). Aiken's V formula is:

with description:

s = r-lo (r = value from the validator, lo = the lowest value on the scale, namely 1)

n = number of validators,

c = number of assessment categories (c = 4)

Aiken's V value is then interpreted based on Valid criteria if it is more than 0.92 (Aiken, 1985)

ISSN 2579-6461 (Online) ISSN 2460-6324 (Print)
DOI: https://doi.org/10.29407/jpdn.v11i2.27895

Reliability Test

Reliability testing was carried out to determine the level of internal consistency of the creative thinking ability instrument used in this research. Reliability testing carried out using the Intraclass Correlation Coefficient (ICC) coefficient with the help of the SPSS program based on data from instrument testing results totaling 21 items.

Reliability interpretation criteria use the standards in the Koo and Li (2016) guidelines as follows:

$\alpha \geq 0{,}90$	: very reliable
$0{,}70 \leq \alpha < 0{,}90$	: reliable
$0{,}60 \leq \alpha < 0{,}70$	: quite reliable
$\alpha < 0{,}60$	: less reliable

Based on the results of the reliability analysis, analysis via the Intraclass Correlation Coefficient (ICC) shows an Average Measures value of 0.758 with a significance value of 0.000, which means that the consistency between items is declared stable and statistically significant. Thus, the instrument developed can be said to be feasible and reliable for use in primary research.

Empirical Validity Test

After the instrument has been declared valid through content validity and Aiken's V, the next stage is to carry out an empirical validity test. The empirical test was carried out using data from instrument trials given to grade 5 students at UPT SDN 2 KarangtengahEmpirical validity is calculated using Corrected Item Total Correlation. Each item score (Xi) is correlated with the total instrument score (Yi) using the Pearson Product Moment correlation formula (Arikunto, 2013):

$$r_{xy} = \frac{N\sum XY - (\sum X)(\sum Y)}{\sqrt{[N\sum X^2 - (\sum X)^2][N\sum Y^2 - (\sum Y)^2]}}$$

with description:

N = number of respondents

X = student's score on one item

Y = student's total score (minus the score of the items tested)

XY = product of X and Y

ISSN 2579-6461 (Online) ISSN 2460-6324 (Print)
DOI: https://doi.org/10.29407/jpdn.v11i2.27895

The rxy values obtained are then compared with the r-table at the significance level $\alpha = 0.05$ with degrees of freedom (df) = N-2.

Decision making criteria:

$r > 0.381 \rightarrow$ item is declared valid

$r < 0.381$ items are invalid

The empirical validity test was carried out by calculating the item-total correlation using the Pearson Product Moment formula. The calculated r value for each item is then compared with the r-table at a significance level of 0.05 with df = N - 2. The item is declared valid if the calculated r > r-table, and invalid if the calculated r < r-table (Arikunto, 2013).

Empirical Reliability Test

After the instrument is declared valid in content and empirically, the next step is to carry out a reliability test to determine the level of internal consistency of the instrument. Reliability shows the extent to which the instrument can provide stable and reliable results in measuring creative thinking abilities on material changes in the form of objects. The reliability test in this research used the Cronbach's Alpha method, because the instrument consists of several questions that measure the same construct. The criteria for interpreting Cronbach's Alpha values according to Azwar (2012) are as follows.

No	Interval	Kriteria
1.	< 0,200	Sangat rendah
2.	0,200 – 0,399	Rendah
3.	0,400 – 0,599	Cukup
4.	0,600 – 0,799	Tinggi
5.	0,800 – 1,000	Sangat Tinggi

An instrument is declared reliable if it has a Cronbach's Alpha (α) coefficient value > 0.60, which confirms sufficient and acceptable internal consistency, especially as the aim of a resesearch (Taber, 2018).

ISSN 2579-6461 (Online) ISSN 2460-6324 (Print)
DOI: https://doi.org/10.29407/jpdn.v11i2.27895

RESULTS

The results of student peer validation are presented in Table 2.

Table. 2 Instrument Validity

Indicator	Item	V	Description
Fluency	1	0,96	Valid
	2	0,96	Valid
	3	0,66	Invalid
	4	0,64	Invalid
	5	0,58	Invalid
	6	0,59	Invalid
Flexibility	7	0,67	Invalid
	8	0,65	Invalid
	9	0,96	Valid
	10	0,98	Valid
	11	0,74	Invalid
	12	0,98	Valid
Originality	13	0,97	Valid
	14	0,97	Valid
	15	0,75	Invalid
	16	0,75	Invalid
	17	0,99	Valid
	18	0,75	Invalid
Elaboration	19	0,97	Valid
	20	0,94	Valid
	21	0,69	Invalid

The content validity of this instrument was tested using the Aiken's V method which involved four validators, namely fellow students. Assessment is carried out on the suitability of each question item with the ability indicators being measured. The Aiken's V calculation is used to determine the extent to which the instrument items are considered relevant by the validators.

Table 3. Inter-Rater Reliability

Indicator	Reliability Value	Description
Fluency	0,901	Very Reliable
Flexibility	0,894	Reliabel
Originality	0,855	Reliabel
Elaboration	0,846	Reliabel

ISSN 2579-6461 (Online) ISSN 2460-6324 (Print)
DOI: https://doi.org/10.29407/jpdn.v11i2.27895

Based on the results of the inter-rater reliability test using the ICC coefficient for each indicator of creative thinking ability, reliability values were obtained which were in the high to very high category. The fluency indicator shows a reliability value of 0.901 which is in the very reliable category, so that the inter-rater assessment scores on this indicator are consistent and stable. The flexibility indicator obtained a reliability value of 0.894, while the originality and elaboration indicators each had a value of 0.855 and 0.846. These three values are in the reliable category, which means that the assessors gave relatively uniform assessments of the items in the indicator. Overall, these results indicate that the instrument has levels. Good inter-rater consistency, so it is suitable to be used to assess students' creative thinking abilities objectively and consistently.

Table 4. Empirical Test Validity Results

Indicator	Item	rit (n=27)	r tabel (n=27)	Criteria
Fluency	1	0,51	0,381	Valid
	2	0,54	0,381	Valid
	3	0,59	0,381	Valid
Flexibility	4	0,66	0,381	Valid
	5	0,61	0,381	Valid
	6	0,46	0,381	Valid
Originality	7	0,61	0,381	Valid
	8	0,62	0,381	Valid
Elaboration	9	0,50	0,381	Valid
	10	0,32	0,381	Invalid

The results of the empirical validity analysis show that of the 10 items developed, 9 items were declared valid, while 1 item was declared invalid. The criteria for determining validity use a comparison between the rit value and the r value of the product moment table at a significance level of 5% with a total of 27 students as respondents (r table = 0.381). Question item number 10 has a rit value of 0.32 which is smaller than the table r value (0.32 < 0.381), so it is declared invalid. This shows that these items have not been able to measure students' creative thinking abilities consistently and do not have an adequate relationship with the total score of the instrument. Thus, question number 10 needs to be revised before being used in the final instrument.

Meanwhile, the other nine items have rit values that are greater than the r table, which is in the range of 0.46 to 0.66. This shows that these items have good

ISSN 2579-6461 (Online) ISSN 2460-6324 (Print)
DOI: https://doi.org/10.29407/jpdn.v11i2.27895

discriminatory power and are able to differentiate between students who have high and low creative thinking abilities. Thus, the nine items were declared suitable for use in further measurements.

Table. 5 Empirical Reliability Test Results

Indicator	Reliability Value	Description
Fluency	0,565	Quite reliable
Flexibility	0,570	Quite reliable
Originality	0,445	Quite reliable
Elaboration	0,427	Quite reliable

Based on the results of the empirical reliability test using the Cronbach's Alpha coefficient on each indicator of the creative thinking ability of class V students at SDN Karangtengah 2, a reliability value was obtained which was in the sufficient category. The fluency indicator has a reliability value of 0.564, while the flexibility indicator has a value of 0.570. These two values are included in the quite reliable category, which shows that the items on this indicator have adequate internal consistency. Furthermore, the originality indicator gets a reliability value of 0.445 and the elaboration indicator gets a value of 0.427. These two indicators are also in the sufficient category, although the level of consistency between items is relatively lower than the other two indicators.

DISCUSSION

Developing instruments to measure students' creative thinking abilities is very important in education, especially at the elementary school level. In this context, the four indicators of creativity that are the focus are Fluency, Flexibility, Originality, and Elaboration. Each of these indicators has questions designed to measure specific aspects of students' creative thinking abilities. Research shows that the ability to think creatively contributes to developmentproblem-solving skills needed in various fields (Alfarisa et al., 2025). Thus, appropriate measurement of student creativity can provide valuable information for teachers in designing more effective learning strategies.

Fluency, which refers to students' ability to generate many ideas, is represented by two items. Fluency is one of the main components of creativity and can be measured by asking students to produce as many answers as possible in a certain amount of time. Therefore, instruments that measure fluency need to be designed carefully so that they

ISSN 2579-6461 (Online) ISSN 2460-6324 (Print)
DOI: https://doi.org/10.29407/jpdn.v11i2.27895

can accurately reflect students' abilities. Flexibility, on the other hand, refers to a student's ability to move from one idea to another easily. Three questions were designed to measure this aspect. Flexibility in thinking is the key to finding creative solutions in varied situations. Students who are able to demonstrate flexibility in their thinking are better able to adapt to the various challenges they face. Thus, instruments that measure flexibility must be able to capture variations in students' thinking and provide a clear picture of their abilities.

Originality, which is measured by three questions, shows the extent to which the ideas produced by students are unique and unusual. Originality gives birth to new, unique ideas and ideas, thinking of new ways and unusual combinations (Susilowati et al., 2021). In an educational context, developing students' ability to think originaly is very important to prepare them to face challenges in the real world. Therefore, the instrument designed must be able to evaluate the extent to which students can produce ideas that are not only creative but also innovative. Elaboration, measured by two items, refers to students' ability to develop their ideas with detail and depth. This is very important in the learning process, where students need to explain and support their ideas with strong arguments. Thus, instruments measuring elaboration should be designed to assess students' ability to provide in-depth and detailed explanations.

Instrument Validity

In this research, the validity of the instrument was measured using the Aiken's V method, which involved four validators from fellow students. Content validity is an important aspect in instrument development, as it ensures that the items used actually measure the intended abilities. Based on the validation results presented in Table 2, there are a number of items that are declared valid and invalid. For example, for the fluency indicator, items 1 and 2 have an Aiken's value of 0.96, which shows that these items are very relevant and able to measure students' creative thinking abilities. However, items 3 to 6 show lower values, below 0.60, so they are declared invalid. This indicates that the items do not have sufficient connection with the ability being measured, and need to be revised. According to (Suryadi et al., 2023), good content validity is the key to ensuring that instruments can be used effectively in measuring learning outcomes.

Furthermore, content validity can also be understood as the extent to which the items in an instrument reflect the theoretical construct that is intended to be measured. In

ISSN 2579-6461 (Online) ISSN 2460-6324 (Print)
DOI: https://doi.org/10.29407/jpdn.v11i2.27895

this context, the construct in question is the ability to think creatively which consists of fluency, flexibility, originality and elaboration. Therefore, it is important for researchers to carry out in-depth analysis of each item before the instrument is used in broader measurements (Henukh et al., 2025). Meanwhile, empirical validity analysis shows that of the 10 items developed, 9 questions were declared valid. This shows that the instrument developed has good discriminatory power, able to differentiate students with high and low creative thinking abilities. Determining validity using a comparison between the rit value and the r table value at a significance level of 5% shows that these items can be used for further measurement.

In addition, revision of invalid items is essential to improve the quality of the instrument. According to (Kuntoro Bambang & Wardani Naniek, 2020) a test is said to have high validity if the tool carries out the function of measuring accurately or providing measurement results that are in accordance with the purpose of the measurement. Meanwhile, reliability is the belief in producing a score that is steady and does not change. Therefore, researchers need to make improvements and retest items that do not meet the validity criteria.

Inter-Rater Reliability

Inter-rater reliability is an important aspect in the development of instruments that measure creative thinking abilities. The results of the inter-rater reliability test show that the reliability value for each indicator is in the high to very high category. For example, the fluency indicator obtained a reliability value of 0.901, which is included in the very reliable category. This shows that the inter-rater assessments on this indicator are consistent and stable. The high reliability of the fluency indicator shows that the items used to measure students' creative thinking abilities in terms of agility of thinking are reliable. On the other hand, the originality and elaboration indicators each have values of 0.855 and 0.846, which are also in the reliable category. This shows that the assessors provide relatively uniform assessments of the items on the indicator, which is important to ensure that the instrument can be used consistently in assessing students' creative thinking abilities (Firdaus et al., 2025).

However, even though the reliability values for these indicators are considered good, it is important to continue evaluating and revising items that have lower reliability values. This is to ensure that each indicator can provide accurate and consistent results in

ISSN 2579-6461 (Online) ISSN 2460-6324 (Print)
DOI: https://doi.org/10.29407/jpdn.v11i2.27895

measuring students' creative thinking abilities. According to (Akdemir-Beveridge et al., 2025), instruments that have high reliability will give researchers more confidence in making decisions based on the data obtained. Apart from that, the reliability analysis also shows that the flexibility indicator has a slightly lower value, namely 0.894, but is still in the reliable category. This suggests that although there is variation in inter-rater assessments. This instrument can still be used to measure students' creative thinking abilities effectively. Therefore, researchers need to pay attention to items that may require improvement to increase the overall reliability of the instrument (Pradipta & Lasmawan, 2020). Overall, the results of this inter-rater reliability test show that the instrument developed has a good level of inter-rater consistency. This is especially important in educational contexts, where objective and consistent assessments are necessary to accurately evaluate student abilities.

Empirical Test Validity Results

The results of the empirical validity analysis show that of the 10 items developed, nine of them were declared valid. This empirical validity was tested using a comparison between the rit value and the r table value at a significance level of 5% with a total of 27 students as respondents. For example, the first item in the fluency indicator has a rit value of 0.51, which is greater than the r table (0.381), so it is declared valid. Determining validity is very important to ensure that each item in the instrument can measure the construct in question accurately (Puji Cahyani & Sutrisno, 2018). However, there was one item that was declared invalid, namely item number 10, which had a rit value of 0.32. This value is smaller than the r table, thus indicating that this item has not been able to consistently measure students' creative thinking abilities. This indicates the need for revision and improvement of these items so that they can be used in subsequent measurements. According to (Santoso & Widiana, 2022), revising invalid items is an important step in developing a quality instrument.

Of the nine valid questions, the rit value ranged from 0.46 to 0.66. This range of values shows that these items have good discriminatory power and are able to differentiate students with high and low creative thinking abilities. This is important in the context of measurement, because a good instrument must be able to provide clear information regarding differences in abilities between students (Hanum et al., 2025). In instrument development, it is important to not only focus on content validity, but also on empirical

ISSN 2579-6461 (Online) ISSN 2460-6324 (Print)
DOI: https://doi.org/10.29407/jpdn.v11i2.27895

validity. Empirical validity provides a clearer picture of the effectiveness of items in measuring desired abilities. Thus, the results of this empirical validity analysis show that the instrument developed has good potential for use in assessing students' creative thinking abilities in elementary schools (Sarmiasih et al., 2020).

Overall, empirical validity testing is a crucial step in instrument development, and the results obtained indicate that this instrument is suitable for use in measuring students' creative thinking abilities. Researchers need to continue to evaluate and improve existing items to ensure that the instruments remain relevant and effective in educational contexts.

Empirical Reliability Test Results

Empirical reliability tests were carried out to measure the internal consistency of the instruments developed. The results of the empirical reliability test show that all indicators have reliability values that are in the sufficient category. For example, the fluency indicator has a reliability value of 0.564, while the flexibility indicator has a value of 0.570. These two values indicate that the items on this indicator have adequate internal consistency, although there is still room for improvement. Sufficient reliability of these indicators shows that the instrument can be used to measure students' creative thinking abilities with an adequate level of confidence. However, the originality and elaboration indicators each have a value of 0.445 and 0.427, which shows that even though they are in the sufficient category, the level of consistency between items is relatively lower compared to the other two indicators. This indicates the need for further evaluation to improve the quality of the items in this indicator (Firdaus et al., 2025).

It is important to note that good reliability is crucial in the educational context, where measurement results will be used to make decisions that have an impact on the teaching and learning process. An instrument that has high reliability will provide more accurate information about student abilities, so teachers can design more appropriate interventions (Kholid et al., 2022). In the context of instrument development, this sufficient reliability value also shows that even though the instrument can be used, there is still potential for improvement. Researchers are advised to carry out a more in-depth analysis of items that have low reliability values, so that the resulting instruments can be more effective in measuring students' creative thinking abilities (Pradipta & Lasmawan, 2020). Overall, the results of this empirical reliability test provide an illustration that the instrument developed has adequate internal consistency, but still requires evaluation and

ISSN 2579-6461 (Online) ISSN 2460-6324 (Print)
DOI: https://doi.org/10.29407/jpdn.v11i2.27895

improvement to achieve a higher level of reliability. Thus, this instrument can be used as an effective tool in assessing students' creative thinking abilities in elementary schools.

CONCLUSION

This article succeeded in developing an instrument to see how far the creative thinking abilities of fifth grade elementary school students are on the subject of changing the shape of objects. This instrument was prepared based on the four main indicators of creativity, namely fluency, flexibility, originality, and elaboration which were adapted from various references relevant to the science learning context in elementary schools. The content validity results using the Aiken's V calculation show that all instrument items are in the valid category, with a V value above 0.90. This confirms that the question items are in accordance with the indicators of creative ability that will be measured. The consistency of assessment between validators is in the good to very good category. The Intraclass Correlation Coefficient produces values between 0.846 to 0.901, which indicates that the instrument has good internal consistency and is reliable for use in measuring students' creative thinking abilities.

The results of the empirical validity test showed that 9 out of 10 questions were declared valid with a value greater than the r-table (0.381), on the other hand, one question item needed revision because it did not meet the criteria. However, the empirical reliability of the instrument using Cronbach's Alpha is in the "sufficient" category with a range of values 0.427 to 0.570. Overall, the instrument that has been developed is declared suitable for use as a tool for evaluating students' creative thinking abilities on material changing the shape of objects. It is hoped that this instrument can help educators carry out more comprehensive assessments through instruments that have been tested and test encouraging the application of science learning that does not only focus on cognitive, but supports the development of creativity and high-level thinking abilities in students.

REFERENCES

Akdemir-Beveridge, Zaghi, A., & Syharat, C. (2025). *Understanding and Evaluating Engineering Creativity: Development and Validation of the Engineering Creativity AssessmentTool(ECAT)*.

Alfarisa, F., Arisetyawan, A., Maulana, F. R., Lathifah, N. Y., Afriani, S., & Mashudi, M. (2025). Creative thinking in elementary education: Reliable assessment instruments for science and social studies. *Jurnal Penelitian Dan Evaluasi Pendidikan*, *29*(1), 13–25. https://doi.org/10.21831/pep.v29i1.79429

Alvarez Espinosa, M., Filgueiras, M. L., Saunders, A., & Suárez, J. (2025). Validation of the Management Model for International Cooperation in Renewable Energy Sources: Reliability Analysis with Cronbach′s Alpha. Preprints. https://doi.org/10.20944/preprints202510.1198.v1

Astuti, M. L. (2024). The Role of 6C Skills in 21st Century Learning of Elementary School Students. *DIDAKTIKA: Jurnal Pendidikan Sekolah Dasar*, *7*(2), 154–161. https://doi.org/https://doi.org/10.21831/didaktika.v7i2.80220

Cahyaningsih, R. D., Purwanto, A., & Khaerudin, K. (2023). Analysis of Students' Creative Thinking in Science Learning. *Prisma Sains : Jurnal Pengkajian Ilmu Dan Pembelajaran Matematika Dan IPA IKIP Mataram*, *11*(3), 719. https://doi.org/10.33394/j-ps.v11i3.7934

Firdaus, F., Wiyanto, W., Putra, N. M. D., & Isnaen, W. (2025). Design of instruments for scientific creative thinking skills and creative thinking digital skills: Rasch models and confirmatory factor analysis. *Eurasia Journal of Mathematics, Science and Technology Education*, *21*(5). https://doi.org/10.29333/ejmste/16310

Gilvando, G., Subhanadri, & Aprizan. (2025). Analisis Kemampuan Berpikir Kreatif Siswa dalam Pembelajaran IPAS pada Materi Sumber Energi. *Jurnal Penelitian Pendidikan*, *1*(1), 17–23. https://doi.org/10.58740/jpp.v1i1.525

Hanum, F., Muliono, T. P. R. Z., Safira, F., Wulandari, A. T., & Putri, M. D. (2025). Pengembangan Instrumen Tes Kemampuan Berpikir Kreatif Pada Materi Fluida Untuk Siswa SMA. *Relativitas: JurnalRiset Inovasi Pembelajaran Fisika*, *8*(2), 177–188. https://doi.org/10.29103/relativitas.v8i2.20505

Harta, J., Ida Bagus Putu Arnyana, & Ni Ketut Suarni. (2021). Pengembangan Instrumen Kemampuan Berpikir Kritis Dan Literasi Humanistik Pada Pembelajaran IPA Kelas V SD. *PENDASI: Jurnal Pendidikan Dasar Indonesia*, *5*(2), 270–280. https://doi.org/10.23887/jurnal_pendas.v5i2.394

Henukh, A., Setiawan*, A., Rusyati, L., Rusdiana, D., Kamin, Y., & Yuliatun, T. (2025). Content Validity and Inter-Rater Reliability in Assessing Environmental Literacy and Creative Thinking Skills of Students. *Jurnal Pendidikan Sains Indonesia*, *13*(4), 1310– 1321. https://doi.org/10.24815/jpsi.v13i4.49212

Irman, Endang Surahman, Dita Agustian, Diana Herawati, & Liah Badriah. (2025). Profil Kemampuan Berpikir Kreatif Peserta Didik dalam Pembelajaran IPA. *JURNAL PENDIDIKAN MIPA*, *15*(1), 60–67. https://doi.org/10.37630/jpm.v15i1.2318

Kastur, A., Hidayati, F., Kurniasari, K., Julianto, J., & Rahmawati, E. (2025). Menguatkan Keluwesan Dan Kebaruan Berpikir Kreatif Siswa Sekolah Dasar Melalui Project-Based Learning (PjBL) yang Kontekstual. *DIDAKTIKA : Jurnal*

Pemikiran Pendidikan, *31*(2), 277–292. https://doi.org/10.30587/didaktika.v31i2.10032

Kholid, Hidayat S., & Syafrizal. (2022). Jurnal Ilmu Pendidikan Pengembangan Instrumen Penilaian Kreativitas Berpikir Siswa Dalam Menulis Teks Berbahasa Inggris. *Eduscience: Jurnal Ilmu Pendidikan*. https://ejurnal.esaunggul.ac.id/index.php/EDU/article/download/4313/3027

Krismanita, R., & Qosyim, A. (2021). Analisis Kemampuan Berpikir Kreatif Pada Pembelajaran Ipa Berbasis Inkuiri Terbimbing. *PENSA E-JURNAL :PENDIDIKAN SAINS*, *9*(2), 156–164. https://doi.org/https://doi.org/10.26740/pensa.v9i2.37336

Kuntoro Bambang, & Wardani Naniek. (2020). Pengembangan Instrumen Penilaian Sikap Sosial Pembelajaran Tematik Kelas III SD. *Jurnal Ilmiah Wahana Pendidikan*. https://doi.org/https://doi.org/10.5281/zenodo.3752471

Maimun, Ibrahim, & Bahtiar. (2024). The Role of Elementary Teachers in Developing 21st Century Skills: Creativity, Collaboration, Communication, and Critical Thinking. *Macca:Science-Edu Journal*,*1*(2). https://doi.org/https://doi.org/10.51574/msej.v1i2.1677

Pradipta, I., & Lasmawan, I. (2020). Pengembangan Instrumen Kemampuan Berpikir Kreatif Dan Literasi Matematika Pada Materi Geometri Peserta Didik Kelas IV Sekolah Dasar (Vol. 10, Issue 1). https://repo.undiksha.ac.id/4438/

Puji Cahyani, V., & Sutrisno, H. (2018). Validation of Multiple Representation Instrument to Measure Student's Multiple Representation Skill. *American Journal of Educational Research*, *6*(8), 1198–1205. https://doi.org/10.12691/education-6-8-20

Santoso, A. M., & Widiana, H. S. (2022). Penyusunan Skala Learning Agility Dengan Aikenâ€TMs V Dan Uji Reliabilitas. *JURKAM: Jurnal Konseling Andi Matappa*, 14–22. https://doi.org/10.31100/jurkam.v6i1.1699

Saputra, I. P. A. A., Jampel, I. N., & Suwatra, I. I. W. (2021). Pengembangan Instrumen Penilaian Kompetensi Pengetahuan Ipa Siswa SD *Journal for Lesson and Learning Studies*, *4*(1). https://doi.org/10.23887/jlls.v4i1.29794

Sarmiasih, N., Putrayasa, I. B., & Marhaeni, A. (2020). Pengembangan Instrumen Keterampilan Berpikir Kreatifdan Hasil Belajar Ipa Siswa Kelas *V SD* (Vol. 10, Issue 2). https://ejournal-pasca.undiksha.ac.id/index.php/jurnal_ep/article/view/3513

Suryadi, T., Alfiya, F., Yusuf, M., Indah, R., Hidayat, T., & Kulsum, K. (2023). Content Validity For The Research Instrument Regarding Teaching Methods Of The Basic Principles Of Bioethics. *Jurnal Pendidikan Kedokteran Indonesia: The Indonesian Journal of Medical Education*, *12*(2), 186. https://doi.org/10.22146/jpki.77062

ISSN 2579-6461 (Online) ISSN 2460-6324 (Print)
DOI: https://doi.org/10.29407/jpdn.v11i2.27895

Susilowati, E., Dewantara, D., & Winarno, dan N. (2021). Pengaruh Blended Learning Terhadap Keterampilan Berpikir Kritis Dan Ketrampilan Berpikir Kreatif (Vol. 6).

Syaifullah, I. M., Rosidin, U., Rohman, F., & Perdana, R. (2024). Analisis tentang Pengembangan Instrumen Asesmen Five-Tier Berbasis Masalah Pada Pembelajaran IPA Fase B untuk Mengukur Kemampuan Berpikir Kritis dan Kreatif Peserta Didik Sekolah Dasar. *Jurnal Penelitian Pendidikan IPA*, *10*(11), 9189–9199. https://doi.org/10.29303/jppipa.v10i11.7450

Trimawati, K., Tjandrakirana, & Raharjo. (2020). Pengembangan Instrumen Penilaian Ipa Terpadu Dalam Pembelajaran Model Project Based Learning (Pjbl) Untuk Meningkatkan Kemampuan Berpikir Kritis dan kreatif siswa SMP. *QUANTUM: Jurnal Inovasi Pendidikan Sains*, *11*(1), 36–52.

Wahyuni, M. T., Rodhiah, R. T. A., & Fadillah, M. (2025). Strategi Peningkatan Keterampilan Abad ke-21 Siswa SD dengan Memanfaatkan Aliran Filsafat Rekonstruksionisme. *JEMARI (Jurnal Edukasi Madrasah Ibtidaiyah)*, *7*(2), 88–103. https://doi.org/10.30599/0cbvwb74

Widayanti, A., Ridlo, R., & Ahmad, N. (2024). Pengembangan Instrumen Penilaian Berbantuan Media Nearpod Untuk Mengukur Kemampuan Berpikir Kreatif Siswa Pada Pembelajaran IPA SMP. *Eduproxima: Jurnal Ilmiah Pendidikan IPA*, *6*(3), 963–975. http://jurnal.stkippgritulungagung.ac.id/index.php/eduproximaEDUPROXIMA6