# Hots-Based Test Instrument Design on Geometry Transformation With Culture Context

1st Budi Halomoan Siregar[1], 2nd Hasratuddin Siregar[2], 3rd Asmin Panjaitan[3], 4th Mulyono[4], 5th Kairuddin[5], 6th Edi Syahputra[6], 7th Dwi Novita Sari[7], 8th Yasifati Hia[8].

{{budihalomoan@unimed.ac.id, siregarhasratuddin@yahool.com, asminpanjaitan@gmail.com}

Mathematics Education Department, Faculty of Mathematics and Natural Sciences, Universitas Negeri Medan, Medan, 20221, Indonesia, Universitas Negeri Medan, Indonesia[1, 2, 3]

**Abstract.** The purpose of this research is to design a test instrument based on high-order thinking skills on the topic of geometric transformations that can be used to measure this ability. The research was conducted based on the Plomp model which consisted of three phases: (1) preliminary research, (2) prototyping, and (3) assessment. The subjects of this study were students of class IX Alhijrah 2 in Deliserdang Regency. The results showed that the HOTS-based test instrument on the topic of geometric transformation met the Nieveen quality standard, which consisted of valid, practical and effective criteria. Five validators stated that the instrument was valid in terms of language, content, and construct, with each category being very valid. This intervention product was also practical, this means that the test instrument was easy to use by students and teachers. In addition, the test instrument was effectively used to measure students' high-level mathematical thinking abilities.

**Keywords:** Design Research, Instrument Test, HOTS, Geometric Transformation

## 1 Introduction

Science and technology have developed dramatically and are accompanied by complex problems in all areas of people's lives. The massive use of information technology in every aspect of life is one of the characteristics of this progress. Thus, the ability to think critically, creatively, collaborate to solve complex problems is an important skill for students to be able to survive facing complex life problems [1;2].

In learning mathematics, the development of higher order thinking skills is very important for students, because these skills are required to solve unusual problems [3,4]. Higher-order thinking skills are the abilities to analyse, evaluate, and create [3,4], to process thoughts to always be creative in solving problems [7]. With this ability, a person will be able to build and develop new ideas to see opportunities and solve complex problems [8]. There are several characteristics of higher order thinking skills, such as: (1) non-algorithmic, (2) complex, (3) finding many alternative solutions, (4) multiple interpretations, (5) full of meaning and (6)

AISTEEL 2023, September 19, Medan, Indonesia

DOI 10.4108/eai.19-9-2023.2340428

deep impressions [9]. Higher-order thinking skills are critical (analysing and evaluating) and creative (formulating, planning, and producing) thinking skills [10;11]. To enhance the competencies, several steps are required: (1) providing real problems, (2) conducting an investigative process, (3) and a discussion process [12].

There are several differences in defining higher order thinking skills, it means this terminology has many definitions [13;14]. Some experts state that the ability to think is not only at the level of remembering facts, but requires more complex thinking processes to occur in order to find solutions to complex problems [6;8]. Conversely, some experts argue that higher order thinking skills can only be known based on some characteristics of the problem that can be overcome, such as: non-algorithmic, complex, more than one solution, multiple interpretations, multiple criteria, and is independent in thought processes. Where all these characteristics can be known through learning activities that involve complex thinking processes. In addition, many experts argue that higher-order thinking skills include: (1) problem solving [11], (2) critical thinking skills [15;16] and creative thinking [10;11;17;18], (3) decision making [17], (4) think logically, reflectively, and metacognitively [11]. Where this competency is always used as learning targets in mathematics learning.

Referring to Anderson & Krathwohl's revised Bloom's taxonomy that the learning objectives are divided into two points of view: (1) cognitive processes and (2) knowledge, thus the higher order thinking skills in Bloom's taxonomy need to be adjusted [5]. When associated with the dimensions of cognitive processes, higher-order thinking skills consist of analysing, evaluating, and creating. Furthermore, from a knowledge point of view, higher-order thinking abilities include conceptual knowledge, procedural knowledge, and metacognitive knowledge [19].

With the complexity of the definition of higher order thinking skills, it is necessary to formulate the concept. This is necessary to develop a HOTS-based test instrument, then the test items developed are in accordance with their objectives, namely to gather information related to the level of students' thinking abilities.

The definition of higher order thinking skills in this study considers two perspectives, namely cognitive and knowledge aspects. Where the ability to think at a high level is an intersection of the dimensions of analysing (C4), evaluating (C5), and creating (C6) towards the dimensions of conceptual, procedural, and metacognitive knowledge [20]. Where, operational verbs can be used to determine and formulate high-level thinking indicators easily. Moreover, operational verbs denote certain processes that represent the dimensions of cognitive processes (analyse, evaluate, create) [5]. Conversely, aspects of knowledge (conceptual, procedural, metacognitive) represent nouns that are useful as objects of the processes being carried out. As a result, these two parts (verbs and nouns) can make it easier to formulate indicators of higher-order thinking skills.

The process of developing learning device like test instruments needs to consider the local cultural context, whether in the form of artefacts, social interaction, economy, health, value, etc. [21;22;23]. Culture is that complex whole which includes knowledge, belief, art, morals, law, customs and other capabilities acquired by man as a member of society [24]. Culture includes all forms of technology including social, ideological, religious, as well as art and objects, all of which are social heritage [25].

The cultural context needs to be integrated in the development of learning teaching materials, such as media, modul, HOTS based test instrument [21;26], context includes the internal elements of a text and all aspects that externally surround the text. Meanwhile, Kristanto and Setiawan stated that context refers to (1) elements of the physical or social environment related to certain utterances, and (2) shared knowledge between the speaker and those who hear them, so that they can understand what the speaker means [27]. Based on these two perspectives, it can be said that context is an utterance or sentence that is intended to understand the meaning in the relevant context. By involving the right context, the questions will be more interesting for students and easier to understand.

Geometry transformation is one of the mathematical topics related to reflection, rotation, dilation and translation. Where this topic is abstract and has different characteristics from other materials. The learning objectives on this topic are that students are able to: (1) discover the properties of an object that undergoes a process of translation, reflection, dilation and rotation, (2) rediscover translation, reflection, dilation and rotation formulas through solving complex problems. Thus, it is necessary to integrate cultural aspects and high-level thinking principles to design the test instruments, so that the test instruments become more attractive and can reduce students' cognitive load. In addition, this instrument can be used to evaluate the extent to which students' thinking skills in solving problems at the HOTS level. Therefore, it is necessary to develop a HOTS-based test instrument on the topic of geometric transformation based on high-order thinking principles and cultural context.

## 2 Method

This research is a research design with the type of development study with the Plomp model which consists of three phases: (1) preliminary research, (2) prototyping phase, (3) assessment [28;29]. The type of formative evaluation applied to this research is formative Tessmer which consists of four processes: (1) self-evaluation or screening, (2) expert review, (3) one-to-one evaluation, (4) small group or micro evaluation, and then continue to field test or tryout as a summative evaluation [30;31;32]. This step is described in the figure 1.

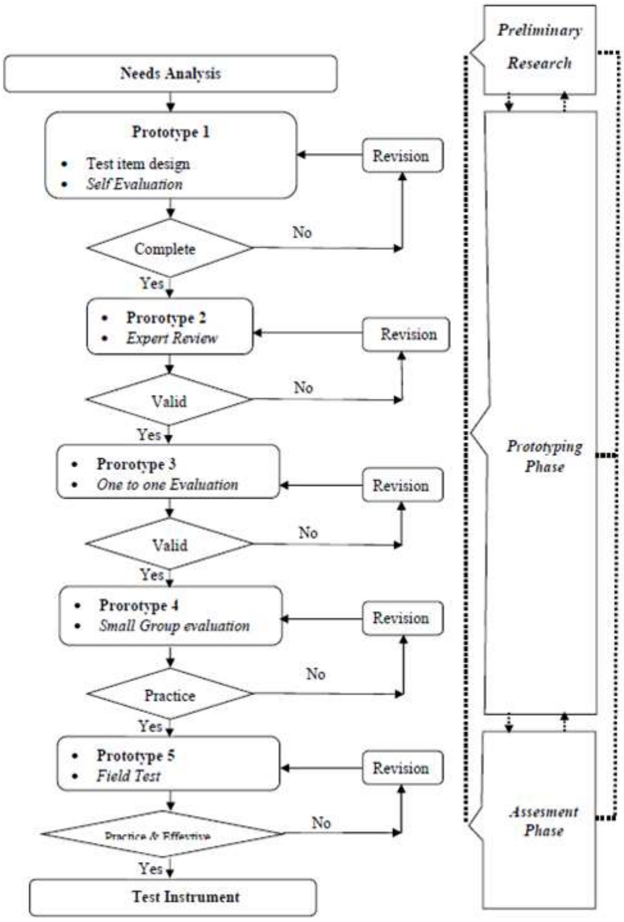


**Fig 1.** Research Flow (developed from Plomp [28])

The preparation and design stages are two parts of the preliminary stage. At this stage the researchers analyzed aspects of the curriculum, literature, needs and concepts required to develop the test instrument [33;34]. Then, at the design stage, an overview of the HOTS-based test instrument was develoved which includes basic competencies, achievement indicators, and the desired level of thinking.

The prototyping process was the second step in developing a HOTS-based test instrument. At this stage, a formative evaluation was carried out consisting of self-assessment, expert review, one-on-one evaluation, and small group trials. In the self-evaluation phase, the researchers conducted his own assessment of the results of the test instrument design. Where the researchers conducted his own evaluation of the content, construction, and language for each item, as a result of this evaluation was prototype 1. Furthermore, prototype I was evaluated by five experts, there were 2 evaluation experts and 3 mathematics education experts, after revising it, then the result was called prototype 2. Then the next phase was one-to-one evaluation, where the test instrument was used on three students with high, medium and low category in competence. At this phase, the researchers collected all information related to: (1) how students use the instrument test, (2) the readability of the text of each test item, (3) the difficulty of understanding the context, etc. All information obtained was analyzed and revised, then it became prototype 3. The next evaluation was a limited trial of 12 students, consisting of 4 students with high abilities, 4 students with moderate abilities, and 4 students with low abilities. At the end of the trial, the teacher and students filled out a practical questionnaire. This aims to find out how useful and easy to use the test instrument is (Nieveen, N. & Folmer, E., 2013). After the test instrument was declared to meet the practical requirements, the test instrument was declared expected practicality, then prototype 4 was obtained.

The third phase of this research was the assessment process. In this phase, field trials were carried out, where the test instrument was tested on the target group to determine the level of actual practicality and actual effectiveness. Field test was conducted in class IX Alhijrah 2 which consisted of 30 students.

The quality standard of a product in this study uses Nieveen's definition. He argued that the quality standards of interventions can be viewed from three aspects: (1) validity, this aspect is divided into two: (a) content validity (relevance/content validity); and (b) construct validity (consistency/construct validity); (2) practicality; and (3) effectiveness [35]. The validation instrument consists of several statements related to content, construct, and language. The score on this validation sheet uses a Likert scale of 5: very valid (5), valid (4), quite valid (3), and not valid (2) and very invalid (1). Then the Aiken validity test was used to state the level of validity by formula 1.

$$V_i = \frac{\sum s}{(n(c-1))}, \quad (1)$$

where s=r-lo, $V_i$ : Aiken's validation index, r: the score given by the validator, lo: the lowest validation rating, c: the highest validation rating, and n: the number of validators. After obtaining the average score, then the results are consulted in table 1 [36].

**Table 1.** Validity Criteria

| Average Score Range | Category |
|---|---|
| $0.80 \le V_i \le 1.00$ | Very high |
| $0.60 \le V_i \le 0.80$ | High |
| $0.40 \le V_i \le 0.60$ | Enough |
| $0.20 \le V_i \le 0.40$ | Low |
| $0.00 \le V_i \le 0.20$ | Very low |

Furthermore, the researchers used a practicality questionnaire to measure the usefulness and ease of use of the intervention product [37]. Expected practicality data by teacher and students were obtained during the small group evaluation phase (micro-evaluation), and actual practicality was obtained after field trial. The practicality questionnaire is the basis for measuring the extent to which the test instrument is "usable and easy to use". The practicality questionnaire was developed using a Likert scale of 5: strongly disagree (1), disagree (2), Neutral (3) agree (4) and strongly agree (5) [38;39]. To determine the average score of the total score on practicality of the instrument, the formula 2 was applied.

$$Prac = \frac{\sum_i^n p_i}{n} \quad (2)$$

Where: Prac= Practicality Data; n= the number of teachers/ students; $\sum_i^n p_i$ = average score of practicality.

After finding a practicality score, then the results were described from quantitative to qualitative data with reference to the practicality criteria of Table 2.

**Table 2**. Practicality Criteria

| Score Range | Category |
|---|---|
| $Prak > 4{,}2$ | Very Practical |
| $3{,}4 < Prak \leq 4{,}2$ | Practical |
| $2{,}6 < Prak \leq 3{,}4$ | Neutral (Fairly Practical) |
| $1{,}8 < Prak \leq 2{,}6$ | Impractical |
| $Prak \leq 1{,}8$ | Very Impractical |

To find out how effective the test instrument is, it is necessary to get information related to the response of teacher and students after using the instrument. This instrument aims to find out whether they feel interested, happy, comfortable, etc. after using the instrument test [35]. The questionnaire aims to determine students' reactions or feelings towards the intervention product after using it [35]. In this study, to make sense of the data, the Guttman scale was used with the options "Yes" and "No". Where the "Yes" option indicates a positive answer criterion with a score of 1 and the "No" option for a negative answer criterion with a score of 0 [40]. Student responses were then analyzed with the formula 3.

$$\text{PRS} = \frac{N(A)}{\text{N(B)}}\ x\ 100\% \tag{3}$$

Where:
PRS : Percentage of students who gave a positive response
N($A$) : The proportion of students who gave a positive response
N($B$) : Number of students (respondents)

Student responses were declared positive if the number of respondents who responded positively was greater than or equal to 80% of the number of respondents present [41].

In addition, the scores obtained by students after working on HOTS-based questions during field trials were considered as a measure of the effectiveness of the test instrument. Then the score was interpreted into five groups of higher order thinking skills, which were determined by the International Center for the Assessment of Higher Order Thinking (ICAT).

**Table 3.** Criteria for Higher Order Thinking Ability (*International Centre for the Assessment of Higher Order Thinking (ICAT)*).

| Score | Category |
|---|---|
| 81 – 100 | Very Good |
| 61 – 80 | Good |
| 41 – 60 | Neutral (Fairly) |
| 21 – 40 | Not Good |
| 0 – 20 | Very not Good |

# 3 Finding and Discussion

## 3.1 Preliminary Research

At this stage, the process of preparation and design of the test instrument is carried out. Before designing the questions, an analysis process was carried out on the curriculum, learning achievement targets, students and needs. In this phase, basic competencies, indicators, and cognitive levels are formulated as a reference for designing test items.

**Table 4.** Basic competences, indicators, cognitive level, and form of questions

| Basic Comp. | Test Item Indicator | Levels Cognitive | Quest. Type | Quest. Number |
|---|---|---|---|---|
| a. Reinvent the concepts of geometric transformations (translations, reflections, rotations and dilations) and be able to use them to solve complex problems.<br>b. Solving contextual problems related to geometric transformations (translations, reflections, rotations, and dilations). | Presented contextual problems related to distance, students are able to find (describe) the final position of an object after experiencing a translation (shift). | Analyze (C4) | essay | 1a |
| | Presented contextual problems related to distance, students can find and prove the shortest distance of an object that translates from its initial position to another position. | Evaluate (C5) | essay | 1b, 4 |
| | Presented contextual problems related to reflection, students are able to find (describe) the position of an object to its image | Analyze (C4) | essay | 2a |
| | Presented contextual problems related to reflection, students are able to find or prove the distance of an object to its image. | Evaluate (C5) | essay | 2b |
| | Presented contextual problems related to reflection, students are able to find or prove the minimum length of cable used to connect two houses and an electric pole. | Analyze (C4) | essay | 3 |
| | Presented contextual problems related to dilation, students are able to find and prove the size of an object after experiencing the dilatation process. | Analyze (C4) | essay | 5a |
| | | Evaluate (C5) | essay | 5b |
| | Presented contextual problems related to rotation, students are able to prove the position of the object after rotation occurs. | Evaluate (C5) | essay | 6 |

At this stage, the researchers designed HOTS-based test items on geometry transformation material based on the formulation of basic competencies, question indicators, cognitive levels, and context, as described in the table 4. At this stage, the initial design of the test instrument based on students' high-order thinking skills on the topic of geometric transformation was developed, its result then called the design product.

### 3.2 Prototyping Phase

In order to obtain high quality intervention products, a multilevel evaluation process was implemented consisting of: self-evaluation, expert reviews, one to one evaluation, and small group which is described as follows.

### A. Self-Evaluation

Formative evaluation implemented on design products was self-evaluation, where this evaluation carried out by the researchers by examining the consistency (construct validity) of the HOTS-based test instrument. Where, this evaluation sheet contains: content validity, appearance, language appropriateness, and graphics.

**Table 5.** Assessment aspects of the self-evaluation instrument

| Aspect | Information |
|---|---|
| content | • Researchers have synchronized each item with learning goals, learning outcomes of geometry transformation materials, and learning indicators.<br>• Researchers have considered the duration of time required to complete each item of questions.<br>• The researchers have chosen the right cultural context for each item |
| Display and language | Researchers have matched:<br>• Display and format for each question item.<br>• The effectiveness of the sentences used.<br>• Using of punctuation marks for each item. |
| Graphics | Researchers have determined:<br>• Accuracy in the use of fonts.<br>• Conformity of color selection<br>• Integration of image illustrations<br>• Alignment of each part |

After the self-evaluation stage, then prototype 1 is obtained.

### B. Expert Reviews

In this phase, the test instrument was validated by five experts, the first expert was a professor in the field of evaluation, the second validator was a doctor in the field of evaluation, and the other 3 were doctors in the field of mathematics education. This validation aims to obtain their review regarding the quality of the test instrument developed based on HOTS principles integrated with the cultural context. The results of the expert reviews were used by the researchers as a basis for improving each test item. The type of expert review instrument is in the form of a checklist. The aspects evaluated on the instrument are content, construct and language.

In this study, the validation process applied the Aiken Validation Test (AVT). Where the researchers divided the Eiken validation index for each aspect, such as: content, construct and language. The results of the validity are shown as in table 6.

**Table 6.** Instrument validation test results

| Validation Aspect | Vi (Average Aiken Validation Index) | Category |
|---|---|---|
| Content | 0.817 | Very high |
| Construct | 0.814 | Very high |
| Language | 0.758 | High |

After revising each item based on the review of the five validators, the HOTS-based test instrument became prototype 2.

**C. One to One Evaluation**

In the One-to-One Evaluation phase, a closed interview technique was applied to 3 students. In order for the evaluation process to run well, the researchers prepared several important things to find out the necessary information. The outline of the questions and statements used were related to: (1) attractiveness and cultural context (2) readability, and (3) ease of use, and (4) time. The outcome of the revision at this stage resulted prototype 3. The results of the interviews are described as in table 7.

**Table 7.** Interview results at the one-to-one evaluation stage

| Aspect | Summary of interview results |
|---|---|
| Attraction and cultural context | The three students stated that the cultural context used in the items could stimulate them to work on the questions. In addition, they admitted that the questions were easier to imagine. The pictures used really help them to understand the problem. |
| Readability | In general, students can understand the questions well, except for item number 1, due to language ambiguity. |
| Ease to use | The students admitted that each item was very challenging to work on, because completing each item required a high-order thinking process. |
| Duration | Students admitted that the duration of the time allotted to work on the items was not enough, so it was necessary to add approximately 1 minute for each item. |

**D. Small Group**

At this stage, an evaluation was implemented using a small group evaluation technique (micro-evaluation) aimed at: (1) measuring the actual practicality of prototype 3, and (2) knowing whether students are able to understand the instructions contained in each test item. This evaluation was attended by 12 students who were selected by purposive sampling technique, meaning that the 12 students were selected based on certain characteristics, namely the value of mathematics in report book, they consist of: 4 students with high ability category, 4 students with moderate ability category, and 4 others with low ability category. In order for the evaluation process to run systematically, the researchers used practical questionnaires and interviews. As for the practical results in small group trial are shown in table 8.

**Table 8.** Practical results in small group trials

| Respondent | Practicality average score (p) | Category |
|---|---|---|
| A teacher | 3.579 | Practical |
| 12 Students | 3.481 | Practical |

In addition to the data in table 8, the researchers also interviewed 6 students, where they stated they could understand instructions for each item except for question number 2, they were confused about identifying the information needed to solve the problem. Then the researchers revised based on the information and suggestions, then resulted prototype 4.

### 3.3 Assessment Phase

After formative evaluation, then a summative evaluation was implemented by conducting experiments in the classroom (Field Test). This evaluation aims to: (1) determine the actual practicality of the test instrument, (2) measure the effectiveness of the intervention. To obtain this data, four types of instruments were used: (1) actual practicality questionnaires, (2) HOTS-based test instruments, and (3) student response questionnaires to measure effectiveness. Then the results of practicality in field test are shown in table 9

**Table 9.** Practical results in field test

| Respondents | Practicality average score (p) | Category |
|---|---|---|
| teacher | 3.627 | Practical |
| 27 Students | 3.995 | Practical |

The results of student responses to the higher-order thinking-based test instruments are described in Table 10.

**Table 10**. Results of student responses to the instrument test

| Category | The number of students | Percentage |
|---|---|---|
| Positive Response | 24 | 88.89% |
| Negative Response | 3 | 11.11% |
| Not attending class | 3 | - |
| Total | 30 | 100% |

After the learning process in class, students were then asked to work on the HOTS-based test instrument (prototype 4). The test results are described in Table 11.

**Table 11.** Test results in field test

| Score Range | The number of students | Category | Information |
|---|---|---|---|
| 81-100 | 4 | Very Good | |
| 61-80 | 20 | Good | |
| 41-60 | 3 | Fairly | |
| 21-40 | 0 | Not Good | |
| 0-20 | 2 | Very not Good | Not attending class |

## 3 Conclusion

The results of the data analysis show that the HOTS-based test instrument on the topic of transformation has met the high-quality standards. Where the product was declared very valid by the five validators. Then the intervention product was stated to be included in the very feasible category, meaning that the product was useful and easy to use by teacher and students to measure the students' higher-level mathematical thinking abilities. In addition, teacher and students responded positively to the intervention product. In addition, teacher and students responded positively to the intervention product. Therefore, the researchers suggest mathematics teachers to use this test instrument to measure students' high-level mathematical thinking skills on the topic of geometric transformations. Then the researchers suggested continuing this research by developing a culture-integrated HOTS-based test instrument by considering student learning modalities.

**Acknowledgments.** I would like to thank my research colleagues, teachers, experts and validators who have collaborated well so that this research can be completed properly.

## References


[1] Siregar, B. H., Kairuddin, Siregar, N., Fudholi, A. (2019). Enhancing the Prospective Teachers' Higher Order Thinking Skills in Solving Pedagogical Problems. International Journal of Innovative Technology and Exploring Engineering (IJITEE). ISSN: 2278-3075, Volume-8 Issue-12, October 2019.

[2] Siregar B. H., Kairuddin, Mansyur, A. (2021) Developing Interactive Electronic Book Based on TPACK to Increase Creative Thinking Skill. Al- Ishlah: Jurnal Pendidikan, December 2021, vol. 13 (3), Pages 2831-2841.

[3] Siregar B. H., Kairuddin, Mansyur, A. Siregar, N. (2021). Development of Digital Book in Enhancing Students' Higher-Order Thinking Skill. Journal of Physics: Conference Series, 1819 012046.

[4] Anselmus, A., Risalah, D., & Sandie, S. Pengembangan Instrumen Tes HOTS dalam Mengukur Kemampuan Berpikir Analitis Siswa pada Materi Program Linear di Kelas XI SMA Sungai Kehidupan. JURING (Journal for Research in Mathematics Learning), 4(4), 371-378.

[5] Anderson, L.W., & Krathwohl, D.R., (2001). A Taxonomy for Learning, Teaching, and Assessing: A Revision of Bloom‟s Taxonom y of Educational Objectives. New York: Addison Wesley Longman, In

[6] Thomas, A. & Thorne, G. (2014). How to Increase Higher Order Thinking. Metarie, LA: Center for Development and Learning.

[7] Brookhart, S. M. (2010). How to Assess Higher Order Thinking Skill in Your Classroom. Alexandria, VA: ASDC.

[8] Lewis, A., & Smith, D., (1993) Defining Higher Order Thinking. Theory into Practice, 32 (3), 131-137. https://doi.org/10.1080/00405849309543588

[9] Resnick, L. B. (1987). Education and Learning to Think. Washington, D.C: National Academy Press.

[10] Conklin, W. (2012). Higher Order Thinking to Develop 21st Century Learners. California, CA: Shell Education Publishing.

[11] King, F.J., Goodson, L., & Rohani, F. (2010). Higher Order Thinking Skills: Definition, Teaching Strategies, Assessment. Florida: A Publication of the Educational Services Program, Now Known as the Center for Advancement of Learning and Assessment, Florida.

[12] Miri, B., David, B.C., & Uri, Z., (2007). Purposely Teaching for The Promotion of Higher Order Thinking Skills: A Case of Critical Thinking. Research in science Education.

[13] Barahal, S. L. (2008). Thinking About Thinking. Phi Delta Kappan, 90(4), 298–302

[14] Ariyana, Y., Pudjiastuti, A., & Bestary, R. (2018). Buku Pegangan Pembelajaran Berorientasi pada Keterampilan Berpikir Tingkat Tinggi: Program Peningkatan Kompetensi Pembelajaran Berbasis Zonasi.

[15] Tsujioka, B. T., & Yeo, A. J. (Eds.). (2015). Handbook of Research on Advancing Critical Thinking in Higher Education. IGI Global.

[16] Paul, R., & Elder, L. (2008). Critical Thinking: The Nature of Critical and Creative Thought. Journal of Developmental Education, 32(2), 2-3.

[17] Presseisen, B. B., (1985). Thinking Skill: Meaning and Model. Dalam Costa, A.L. (Eds.), Developing Mind: A Resoursce Book for Teaching Thinking (pp.43-48). Alexandria, VA: ASCD. Resnick, L. B. (1987). Educational and Learning to Think. Washington, DC: National Academy Press.

[18] Krulik, S. and Rudnick, J.A. (1999). Innovative Task to improve Critical and Creative Thinking Skills. In I. Stiff (Ed.). Developing Mathematics Reasoning in Grade K-12. Reston: National Council of Teachers of Mathematics.
[19] Nilson, L. B. (2016). Teaching at Its Best: A Research-Based Resource for College Instructors (4th ed.). Jossey-Bass
[20] Setiawati, W., Asmira, O., Ariyana, Y., Bestary, R., Pudjiastuti, A., (2018). Buku Penilaian Berorientasi HOTS: Program Peningkatan Kompetensi Pembelajaran Berbasis Zonasi.
[21] Au, K. H. (2018). Culturally Responsive Instruction: What is it and Does it Matter? International Journal of Multicultural Education, 20(2), 1-15.
[22] Ahn, S. (2022). Enhancing Cultural Relevance in Mathematics Education: A Case Study in Elementary Schools. Journal of Cultural Studies in Education, 15(1), 35-48.
[23] Banks, J. A. (2018). Cultural Diversity and Education: Foundations, Curriculum, and Teaching (7th ed.). Routledge.
[24] Taylor, Edward B. (1871). Primitive Culture: Researches into the Development of Mythology, Philosophy, Religion, Art, and Custom. Volume 1. London: John Murray.
[25] Mellvie, J., & Stern, J., B., (1961). General Anthropology. Barners & Noble, 1952.
[26] Nieto, S. (2017). Reinventing Multicultural Education: From “Culture” to “Cultures” and from “Cultural Competence” to “Cultural Humility.” In J. A. Banks (Ed.), Cultural Diversity and Education: Foundations, Curriculum, and Teaching (6th ed., pp. 16-35). Routledge
[27] Kristanto, P. D., & Setiawan, P. G. F. (2020). Pengembangan soal hots (higher order thinking skills) terkait dengan konteks pedesaan. In PRISMA, Prosiding Seminar Nasional Matematika (Vol. 3, pp. 370-376).
[28] Plomp, T. (1997). Educational Design: Introduction. From Tjeerd Plomp (eds). Educational & Training System Design: Introduction. Design of Education and Training (in Dutch). Utrecht (the Netherlands): Lemma. Netherland: Faculty of Educational Science and Technology, University of Twente.
[29] Plomp, T. (2010). Educational Design Research: An Introduction. In T. Plomp & N. Nieveen (Eds), Educational Design Research-Part A: An Introduction (pp.10-51). Enschede, Netherlands: Netherlands Institute for Curriculum Development (SLO).
[30] Plomp, T & Wolde, J.V.D. (1992). The General Model for Systematic Problem Solving. From Tjeerd Plomp (Eds.). Design of Educational and Training (in Dutch). Utrecht (the Netherlands): Lemma. Netherland: Faculty of Educational Science and Technology, University of Twente. Enschede the Netherlands.
[31] Tessmer, M. (1993). Planning and Conducting Formative Evaluations. London: Kogan Page.
[32] Plomp, T., Nieven, N., Gustafon, K., Branch, R.M., dan van den Akker, J (eds). (1999). Design Approach and Tools in Education and Training. London: Kluwer Academic Publisher.
[33] Barab, S. & Squire, K. (2004). Design-Based Research: Putting A Stake in The Ground. Journal of the Learning Sciences, 13(1), 1-14.
[34] Bakker, A. (2018). Design Research in Education: A practical Guide for Early Career Researchers. New York, NY: Routledge.
[35] Nieveen, N. & Folmer, E. (2013). Formative Evaluation in Educational Desain Research. In Plomp, T. & Nieveen, N. (Eds) Educational Design Research, Part A: An Introduction. New York: Routledge.
[36] Arikunto, S., (2018). Dasar-Dasar Evaluasi Pendidikan (Edisi ke-3), Jakarta: Bumi Aksara
[37] Akker, J., Gravemeijer, K., McKenney S., & Nieveen. (2006). Educational Design Research. London: Routledge Taylor and Francis Group.

[38] Creswell, J. W., & Creswell, J. D. (2018). Research Design: Qualitative, Quantitative, And Mixed Methods Approaches. Los Angeles, California: Sage Publications.
[39] Sugiyono (2022). Metode Penelitian & Pengembangan (Research and Development). Bandung: PT. Alfabeta.
[40] Sugiyono. (2017). Metode Penelitian Kuantitatif, Kualitatif, dan R&D. Bandung: Alfabeta, CV.
[41] Siswono, T. Y. E. (2019). Paradigma Penelitian Pendidikan: Pengembangan Teori dan Aplikasi Pendidikan Matematika. Bandung: PT. Remaja Rosdakarya.