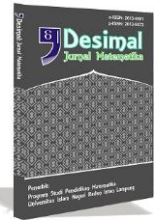




Contents lists available at DJM

## DESIMAL: JURNAL MATEMATIKA

p-ISSN: 2613-9073 (print), e-ISSN: 2613-9081 (online), DOI 10.24042/djm  
<http://ejournal.radenintan.ac.id/index.php/desimal/index>



# Development of a numeracy literacy ability test instrument based on high order thinking skills (hots) in junior high school mathematics learning

Fathrul Arriah\*, Sri Sulfiati Romba

Universitas Muhammadiyah Makassar, Indonesia

## ARTICLE INFO

### Article History

Received : 06-03-2022

Revised : 08-08-2022

Accepted : 07-12-2022

Published : 30-04-2023

### Keywords:

HOTS; Test Instrument; Numerical Literacy; Mathematical Literacy.

\*Correspondence: E-mail:  
[fatrulariah@unismuh.ac.id](mailto:fatrulariah@unismuh.ac.id)

Doi:  
[10.24042/djm.v6i1.11523](https://doi.org/10.24042/djm.v6i1.11523)

## ABSTRACT

*The aim of this study was to develop a numeracy literacy test instrument based on the High Order Thinking Skills (HOTS) for class VIII junior high school students. The type of research used is research and development (R&D) with the ADDIE model. The ADDIE model consists of five stages of development, namely: Analysis, Design, Development, Implementation, and Evaluation. A total of 15 questions have been prepared covering the material for straight line equations, functions, coordinate systems, linear equations of two variables and number patterns in odd semesters. The items were validated by three mathematicians, and the calculations were analyzed using the Aiken index. Furthermore, the questions that were declared valid were tested on 34 students in two junior high schools in Makassar City using convenience sampling as a sample selection technique. The experimental data are dichotomous. The data were analyzed using KR-20 to produce a reliability coefficient. This study resulted in a HOTS-based numeric literacy test instrument in the form of multiple choice with a total of 15 valid and reliable questions. 15 items have high category validity, while the reliability coefficient of the item is 0.91, or high category.*

<http://ejournal.radenintan.ac.id/index.php/desimal/index>

## INTRODUCTION

The demands for 21<sup>st</sup> century skills are increasing along with the complexity of the problems faced by humans. These skills are then developed in students to prepare human resources who are rich in skills and ready to face the progress of the times (Hendrowati & Sunanto, 2021). After the school curriculum echoed the

integration of 4C skills (Critical Thinking, Creativity, Collaboration, and Communication) in learning (Zubaidah, 2018), now numeracy literacy is present as a new breakthrough for Indonesian students (Fiangga, M. Amin, Khabibah, Ekawati, & Rinda Prihartiwi, 2019) through the Minimum Competency Assessment (AKM) in the Merdeka Learning program, which was first

introduced by the Minister of Education and Culture of Indonesia (Biro Komunikasi dan Layanan Masyarakat Kementerian Pendidikan dan Kebudayaan Indonesia, 2019).

Numerical literacy is a basic skill for students in calculating and interpreting real objects such as graphs and tables or quantitative information in everyday life (Rahmawati, Candra Pradana, Rinaldi, Syazali, & History, 2021; Sayekti, Sukestiyarno, Wardono, & Dwijanto, 2021; Setyawati, Aisyah, & Kusaeri, 2021). In short, these skills are used to solve math problems in real life, so they are essential for students. Numerical literacy is also the result of responses to PISA (Programme for International Student Assessment) scores in recent years. This international assessment reviews the educational progress of countries participating in PISA (Pratama & Retnawati, 2018) through improving students' reading literacy, mathematical literacy, and science literacy skills (Fiangga et al., 2019; Subroto, Suhadi, & Julyana, 2019). As recorded in the 2018 PISA results released by the OECD, Indonesia is ranked 72 out of 78 countries in mathematical literacy (Kementrian Pendidikan dan Kebudayaan, 2018). A figure that describes the ability of Indonesian students' mathematical literacy is still far behind the countries that are members of PISA.

There is a relationship between PISA and the numeracy literacy skills of students who have not been able to answer PISA questions due to their low mastery of numeracy literacy (Setyawati et al., 2021). Recent studies suggest that one of the factors causing low numeracy literacy is that Indonesian students are not used to solving problems that require an understanding of numeracy literacy (Kaka, Ate, & Making, 2021). On the other hand, they have difficulty solving questions that require HOTS (High Order Thinking Skills) or have not been able to reach the HOTS level (Hadi & Wijaya,

2020). This fact has become a stimulus for Indonesian education to further improve and focus attention on increasing students' numeracy literacy and HOTS.

HOTS also plays an important role in mathematics and as a support for numeracy literacy skills. HOTS helps students gain a complete and in-depth understanding of mathematical concepts (Bakry & Bakar, 2015). Even HOTS is able to improve students' cognitive and affective aspects through conceptual exploration and solving mathematical problems in real life. In addition, they do not just remember or restate but are required to develop ideas so that they are trained to think critically, creatively, and logically (Hadi & Wijaya, 2020; Hamdi, Suganda, & Hayati, 2018).

A number of skills-related problems can be overcome through assessment (Abdullah et al., 2016). Teachers are expected to be able to develop tests related to numeracy literacy skills (Pramono, 2021). Tests are considered a stimulus to build divergent ways of thinking (Dinnullah, Noni, & Sumadji, 2019; Kusaeri, Sadieda, Indayati, & Faizien, 2018). The items also train students to hone skills according to the demands of the 21<sup>st</sup> century (Hamdi et al., 2018). However, Nursakiah, Arriah, & Dharma (2022), and R S Pramono (2021) mention that class teachers have difficulty distinguishing numeracy literacy from learning mathematics. Mahendra, Parmithi, Hermawan, Juwana, & Gunartha (2020) also stated that on average, teachers do not understand the process of preparing HOTS questions, and many teachers still use test items that tend to test memorization rather than higher-order thinking skills. Even though students are needed in the future, namely HOTS (Hamdi et al., 2018). Thus, the HOTS-based numeracy literacy test is still very relevant for development because the test is an integral part of the mathematics

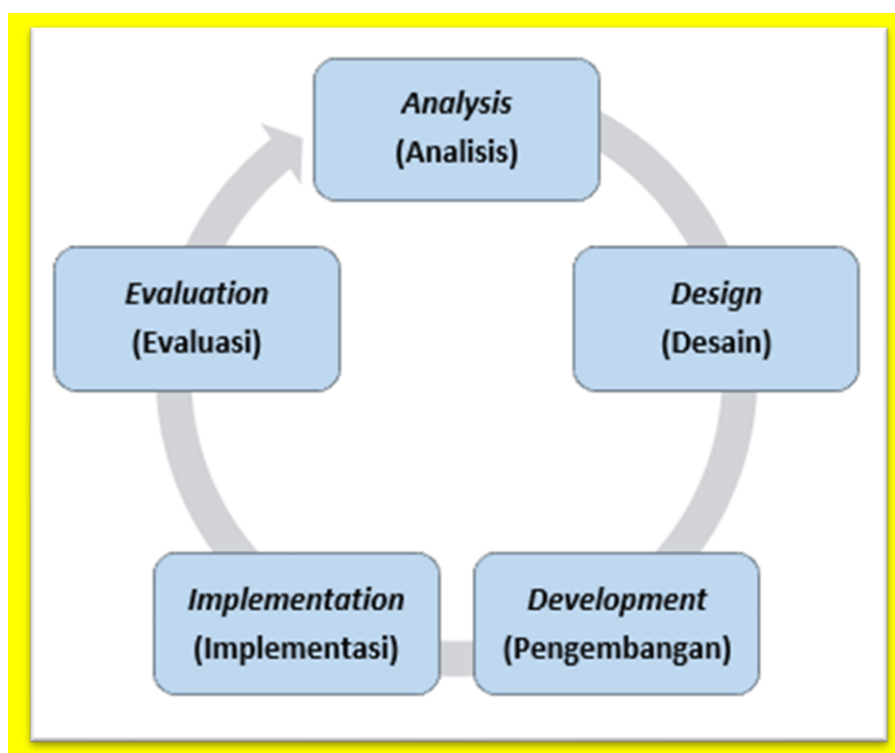
learning process and measures the achievement of learning objectives.

According to various previous studies, studies that develop instruments to measure HOTS-based numeracy literacy skills in students are very rare. Previously, Liswati, Sakinah, & Yuniarti (2021) had developed a numeracy literacy-based assessment instrument for high school students that was not HOTS-oriented. Therefore, the purpose of this study focuses on developing HOTS-based mathematical numeration assessment instruments for junior high school (SMP) students. This instrument is expected to be a practice test for students to develop numeracy and HOTS literacy skills as well

as a tool to measure these abilities. So the main characteristics of this instrument must meet the rules of good test development and have good validity and reliability (Azwar, 2012). The value of validity and reliability is one aspect of assessing the accuracy of an instrument to measure its ability.

## METHOD

This research was conducted using Research and Development (R&D) through the stages of Analysis, Design, Development, Implementation, and Evaluation, usually called the 5 ADDIE steps.



**Figure 1.** Stages of Instrument Development Research

Tegeh, Jampel, & Pudjawan (2015) write that the ADDIE stages are relevant to research that departs from learning theory, so they were chosen in the development of HOTS-based numeracy literacy instruments. The analysis stage begins with analyzing the needs in the field regarding the need to develop a

HOTS-based numeracy literacy test instrument. Furthermore, the design stage is carried out, namely making test grids that are relevant to learning mathematics in junior high schools and covering the syllabus and basic mathematics competencies of grade VIII junior high schools with the concepts of numeracy

literacy and High Order Thinking Skills (HOTS).

Development stage, starting with writing 15 multiple-choice items with correct answers given a score of 1 and wrong answers given a score of 0. Then, the items were reviewed and validated by three mathematicians. Scores from experts were analyzed using the Aiken technique; the research team used the limit value of Aiken mentioned by Retnawati (2016), namely 0.4 to 1, or the medium to high category.

In the implementation stage, questions that had been declared feasible by the validator were tested on 34 class VIII students at two SMP Negeri Makassar City. The fifteen questions were used as a data collection instrument to obtain responses from students. Sample selection used a convenience sampling technique, namely the freedom of the researcher to choose a sample based on sample availability and compatibility of sample characteristics based on the principle of convenience (Solehudin & Widodo, 2021). After the data is collected, it enters the evaluation stage, namely the analysis of test items to review the reliability of the test instrument developed using the KR-20 (Kuder Richardson) with the help of Microsoft Office Excel. The minimum value of the KR-20 reliability coefficient that is tolerated in this study is more than 0.70 (Fraenkel, Wallen, & Hyun, 2012). The KR-20 formula is attached as follows:

$$r_i = \left(\frac{n}{n-1}\right) \left(\frac{s^2 - \sum pq}{s^2}\right) \quad (1)$$

Remarks:

$r_i$  = overall test reliability

$p$  = the proportion of subjects who answered the item correctly

$q$  = the proportion of subjects who answered the item incorrectly

$\sum pq$  = the number of products multiplied between  $p$  and  $q$

$n$  = the number of items

$s^2$  = standard deviation of the test (standard deviation is the root of the variance).

## RESULTS AND DISCUSSION

The results of the development in this study are a HOTS-based numeracy literacy test instrument for mathematics learning that is equipped with scoring guidelines.

The initial stage in developing this test instrument was the analysis stage, which involved conducting a literature study related to the need for numeracy literacy instruments in mathematics. The development of the test was suggested by Pramono (2021) because this ability has only been echoed in the last 2 years. Of course, there is a need for a test in an assessment related to numeracy literacy skills. Next, the determination of test objectives, test forms, and the study of learning materials and KD in the 2013 curriculum for class VIII SMP Semester I was carried out (Table 1).

**Table 1.** Materials and Basic Competencies Developed

Learning Materials	Basic Competencies
Straight line equation	3.4: Determine the equation of a straight line and its graph.
Function	3.5: Presenting functions in various forms of relations, ordered pairs, function formulas, tables, graphs, and diagrams.
Coordinate system	3.1 0: Using Cartesian coordinates to explain the position of objects relative to a certain reference.
The linear equation of two variables	4.1: Create and solve mathematical models of real problems related to linear equations of two variables.
Number pattern	4.3: Use patterns and generalizations to solve real problems.

After the analysis phase, the design phase begins with the design of a prototype HOTS-based numeracy literacy test instrument. The test instrument is in the form of multiple-choice questions with five options. According to Mahendra et al. (2020) and Hamdi et al. (2018), multiple-

choice tests are relevant for the HOTS test. Furthermore, the grid of test items is arranged based on the basic competencies attached in Table 1. The grid of HOTS-based numeracy literacy test instruments can be seen in Table 2.

**Table 2.** HOTS-Based Numerical Literacy Test Instrument Lattice

Dimensions	Indicator	No.
Solving problems related to patterns in number sequences and HOTS-based object configuration sequences.	Determine the $n^{\text{th}}$ term of a geometric series.	1
	Calculates the sum of the first $n$ terms of an arithmetic sequence.	2
	Apply the rules of number patterns, sequences, and series to solving various real-life problems.	3
Solving problems related to the position of points in the HOTS-based Cartesian coordinate plane.	Calculates the area of a region on the Cartesian coordinate plane.	4
	Presenting learning outcomes about Cartesian coordinates.	5
	Solving problems about the Cartesian coordinate plane in the context of everyday life.	6
Solve problems related to relations and functions by using various representations based on HOTS.	Define a function based on a table or graph.	7
	Solve relational and functional problems.	8, 9
Solve contextual problems related to linear functions as straight-line equations based on HOTS.	Determine the equation of a straight line based on a graph.	10
	Solve straight line equations in the context of everyday life.	11, 12
Solving problems related to the HOTS-based two-variable linear equation system.	Solve problems related to linear equations with two variables.	13
	Solve the problem of linear equations with two variables in the context of everyday life.	14, 15

After preparing the grid, the development stage is carried out by writing the items. A total of 15 multiple-choice questions have been developed; each correct answer has a score of 1 and a wrong score of 0. In Table 3, question number 11 discusses straight-line equations.

**Table 3.** Developed Item Number 11

No.	Question Items	Score
<b>11</b>	Makassar City's population growth every year is constant or fixed. In 2008, there were 650,000 people, and in 2013, there were 900,000 people. What will the city's population be in 2021? a. 1,250,000 <b>b. 1,300,000</b> c. 1,350,000 d. 1,400,000 e. 1,450,000	<b>1</b>

**Answer: b**

Furthermore, the 15 items were assessed by three mathematicians. Three

validators are the minimum number and meet the instrument validation criteria by expert judgment (Retnawati, 2016). Heale & Twycross (2015) also wrote that ideally the instrument should be assessed by 3–5 experts. This stage is called content validation, which is testing how well the developed instrument is and whether it can measure certain concepts that are the purpose of measurement. This stage should not be missed in the development of the instrument because the aim is to ascertain whether the instrument will measure variables accurately, in this case, numeracy literacy (Bajpai & Bajpai, 2014).

Specifically, the validator will review all items for readability, clarity, and completeness and arrive at some level of agreement on which items should be included in the final instrument. The validator assesses aspects of the use of language rules in mathematical sentences on questions, aspects of material constructs, and aspects of alignment of basic competencies, indicators, and items with student characteristics.

The following results of content validity using the Aiken technique are presented in Table 4.

**Table 4.** Aiken Scores from HOTS-based Numeracy Literacy Test Items

No.	Question Items	Description
1	0.89	Valid
2	0.89	Valid
3	0.89	Valid
4	0.89	Valid
5	0.89	Valid
6	0.89	Valid
7	0.89	Valid
8	0.89	Valid
9	0.89	Valid
10	0.89	Valid
11	0.89	Valid
12	0.89	Valid
13	0.89	Valid
14	0.78	Valid
15	0.89	Valid

More than four decades ago, Aiken (1980) wrote that the value of content validity was between 0 and 1. If it is close to 1, it is in the high category, and vice

versa, if it is close to 0, it is in the low category. However, Retnawati (2016) categorizes it into 3 validity categories 0.4 (low), 0.4–0.8 (moderate), and > 0.8 (high). Meanwhile, Saukiyah, Sunardi, & Dinawati (2017) interpret the validity of an instrument as high if the validity coefficient is greater than 0.6. High validity indicates that the instrument can carry out its measuring function properly and produce measurement results that are relevant to the measurement objectives (Ndiung & Jediut, 2020).

Widodo, Ibrahim, Hidayat, Maarif, & Sulistyowati (2021) stated that an Aiken index score (V-Aiken) of more than 0.5 in a math test is good and can be used to measure students' mathematical abilities, but if it is less than 0.5, then revisions are made to the questions being developed. Referring to the sources previously described, 15 items are categorized in the high category. Empirically, the tests developed were valid, and no items were omitted because they had significant items (Taherdoost, 2016). Thus, the test is able to measure numeracy literacy skills precisely.

The test validates HOTS-based numeracy literacy skills, so it is able to encourage students to think deeply about teaching materials, especially in straight-line equations, functions, coordinate systems, two-variable linear equations, and number patterns. This is supported by the study of Barnett & Francis (2012), which found that a test that is valid and contains elements of higher-order thinking skills, or HOTS, will basically encourage test takers to think deeply.

The validated test items were then assembled into a test and tested on 34 Makassar City Middle School students in two schools (implementation stage). 34 students are sufficient as test-takers in the trial of an instrument that measures literacy (Gradini, Firmansyah B, & Edy Saputra, 2021). Work on the questions

using a paper test with a processing time of 90 minutes.



**Figure 2.** Implementation of Instrument Trials

The final stage in this study is evaluation. A number of data points from instrument trials were analyzed using the KR-20 (1) formula, assisted by Microsoft Office Excel, to obtain reliability values. In addition to validity, empirical testing must also involve reliability testing to show the extent to which a person's test score is stable and free from measurement errors (Thompson, 2013). In other words, an instrument shows the extent to which it is without bias and the measurement is consistent across time and across various items (Bajpai & Bajpai, 2014). Because the instrument answer scores are in the form of dichotomous data (0 and 1), Heale & Twycross (2015) suggest using KR-20 to determine reliability and strength (instrument reliability).

The reliability of the instrument is indicated by the reliability coefficient (Heale & Twycross, 2015). This study uses KR-20 to determine reliability in the range of 0.00–1.00. A value close to 1.00 means that the factor under investigation can be measured. The reliability coefficient is in the acceptable reliability score range of 0.7 and higher (Heale & Twycross, 2015). This is in accordance with Ferita & Retnawati (2016), who state that if the

value has reached 0.7, the instrument developed is reliable.

The following is the result of the calculation using the KR-20 formula:

$$r_{11} = \left(\frac{n}{n-1}\right)\left(\frac{S^2 - \sum pq}{s^2}\right) \Leftrightarrow r_{11}$$

$$r_{11} = \left(\frac{15}{15-1}\right)\left(\frac{18.52 - 2.87}{18.52}\right)$$

$$r_{11} = \left(\frac{15}{14}\right)\left(\frac{15.65}{18.52}\right)$$

$$r_{11} = (1.07)(0.85) \Leftrightarrow r_{11} = \mathbf{0.91}$$

Based on the results of the analysis, the reliability coefficient ( $r_i$ ) is 0.91 or more than 0.7 or in the very high category because it is more than 0.8 (Kereh, Liliyasi, Tjiang, & Subandar, 2015; Mohamad, Sulaiman, Sern, & Salleh, 2015). The higher the reliability coefficient number, the smaller the possibility of errors occurring (Ferita & Retnawati, 2016), or the higher the accuracy of the test in measuring the ability of the test takers (Ratnaningsih, Isfarudi, & Soleiman, 2013).

The reliability coefficient value of 0.91 can be used to make decisions about test takers based on the results of testing (Istiyono, Mardapi, & Suparno, 2014). This high-reliability score was obtained because it fulfilled content validity by expert judgment or because the items developed were derived from the ability indicator to be measured (Istiyono et al., 2014), namely HOTS-based numeracy literacy skills. Istiyono et al. (2014) and Ndiung & Jediut (2020) mentioned again that another cause was because the subjects being measured had taken the test seriously, resulting in a high correlation. A strong correlation indicates high reliability, while a weak correlation indicates that the instrument may not be reliable. Basically, a test instrument is reliable if the test taker's score has a high correlation with the total score (Ferita & Retnawati, 2016).



Thus, the items are good and appropriate to use without revision. Based on the stages of the ADDIE development carried out and producing 15 items that have fulfilled empirical testing (valid and reliable), it can be interpreted that the HOTS-based numeracy literacy assessment instrument in the form of a multiple choice test instrument can measure numeracy literacy skills in mathematics learning well. The HOTS-based numeracy literacy test that has been developed can be used by educators to measure the numeracy skills of class VIII students in junior high school mathematics learning covering straight-line equations, functions, coordinate systems, two-variable linear equations, and number patterns. This test is valid and reliable. Tests that fulfill these two things can measure students' abilities (Sa'idah, Yulistianti, & Megawati, 2018; Vincent & Shanmugam, 2020) and improve understanding and ability (Aziza, 2019; Hamdi et al., 2018) in this case the numeracy literacy skills and HOTS of junior high school students.

### CONCLUSIONS AND SUGGESTIONS

This development study produced 15 items in the form of multiple choice questions with five answer options that met good test standards, namely validity and reliability. The questions cover straight-line equations, functions, coordinate systems, linear equations of two variables, and number patterns. The HOTS-based numeracy literacy test instrument can be used by students to practice increasing numeracy and HOTS literacy skills and by teachers as a tool to measure the numeracy literacy of junior high school students.

Based on the development carried out, the suggestion for this study for future studies is the need to develop numeracy literacy tests on different materials and levels. As a result of this development study, the High Order Thinking Skills

(HOTS)-based numeracy test can be used by teachers to improve and measure the math literacy and HOTS abilities of junior high school students. In addition, it is recommended that students use it to practice math literacy skills and HOTS.

### REFERENCES

- Abdullah, A. H., Mokhtar, M., Abd Halim, N. D., Ali, D. F., Mohd Tahir, L., & Abdul Kohar, U. H. (2016). Mathematics teachers' level of knowledge and practice on the implementation of higher-order thinking skills (hots). *EURASIA Journal of Mathematics, Science and Technology Education*, 13(1).  
<https://doi.org/10.12973/eurasia.2017.00601a>
- Aiken, L. R. (1980). Content validity and reliability of single items or questionnaires. *Educational and Psychological Measurement*, 40(4), 955-959.  
<https://doi.org/10.1177/001316448004000419>
- Aziza, M. (2019). Kemampuan pemecahan masalah siswa dalam menyelesaikan soal tertutup dan terbuka pada pokok bahasan lingkaran. *Pythagoras: Jurnal Pendidikan Matematika*, 14(2), 126-138.  
<https://doi.org/10.21831/pg.v14i2.26563>
- Azwar, S. (2012). Reliabilitas dan validitas. *Yogyakarta: Pustaka Pelajar*.
- Bajpai, R., & Bajpai, S. (2014). Goodness of measurement: Reliability and validity. *International Journal of Medical Science and Public Health*, 3(2), 112.  
<https://doi.org/10.5455/ijmsph.2013.191120133>
- Bakry, B., & Bakar, M. N. Bin. (2015). Process of thinking among junior high school student in solving hots question. *International Journal of Evaluation and Research in Education (IJERE)*, 4(3), 138.



- <https://doi.org/10.11591/ijere.v4i3.4504>
- Barnett, J. E., & Francis, A. L. (2012). Educational psychology: An international journal of experimental using higher order thinking questions to foster critical thinking: A classroom study. *An International Journal of Experimental Educational Psychology*, 32(2).
- Biro Komunikasi dan Layanan Masyarakat Kementerian Pendidikan dan Kebudayaan Indonesia. (2019). *Mendikbud tetapkan empat pokok kebijakan pendidikan "merdeka belajar" | gerakan literasi nasional*. Retrieved from <https://gln.kemdikbud.go.id/glnsite/mendikbud-tetapkan-empat-pokok-kebijakan-pendidikan-merdeka-belajar/>
- Dinnullah, R. N. I., Noni, E., & Sumadji, S. (2019). Analisis kesalahan siswa pada penyelesaian soal cerita berdasarkan tahapan newman. *Jurnal Tadris Matematika*, 2(2). <https://doi.org/10.21274/jtm.2019.2.2.175-184>
- Ferita, R. A., & Retnawati, H. (2016). Pengembangan perangkat penilaian autentik untuk pembelajaran matematika di kelas vii semester 1 development authentic assessment kit for mathematics learning in grade vii semester 1. *PYTHAGORAS: Jurnal Pendidikan Matematika*, 11(1), 69. <https://doi.org/10.21831/pg.v11i1.9672>
- Fiangga, S., M. Amin, S., Khabibah, S., Ekawati, R., & Rinda Prihartiwi, N. (2019). Penulisan soal literasi numerasi bagi guru sd di kabupaten ponorogo. *Jurnal Anugerah*, 1(1), 9-18. <https://doi.org/10.31629/anugerah.v1i1.1631>
- Fraenkel, J. R., Wallen, N. E., & Hyun, H. H. (2012). How to design and evaluate research in education. *McGrawHill*.
- Gradini, E., Firmansyah B, & Edy Saputra. (2021). Mendesain tes literasi matematis menggunakan soal pisa-like konteks kultur lokal. *Jurnal Ilmiah Pendidikan Matematika Al Qalasadi*, 5(1), 29-43. <https://doi.org/10.32505/qalasadi.v5i1.2945>
- Hadi, S., & Wijaya, S. (2020). Junior high school students' ability to complete mathematical hots questions. *JURNAL ILMIAH GLOBAL EDUCATION*, 91-93. <https://doi.org/10.55681/jige.v1i2.42>
- Hamdi, S., Suganda, I. A., & Hayati, N. (2018). Developing higher-order thinking skill (hots) test instrument using Lombok local cultures as contexts for junior secondary school mathematics. *Research and Evaluation in Education*, 4(2), 126-135. <https://doi.org/10.21831/reid.v4i2.22089>
- Heale, R., & Twycross, A. (2015). Validity and reliability in quantitative studies. *Evidence Based Nursing*, 18(3), 66-67. <https://doi.org/10.1136/eb-2015-102129>
- Hendrowati, T. Y., & Sunanto, R. F. (2021). Numeracy and literacy skill in elementary school students: The utilization of kampus mengajar perintis program. *Desimal: Jurnal Matematika*, 4(3). <https://doi.org/10.24042/djm.v4i3.10462>
- Istiyono, E., Mardapi, D., & Suparno, S. (2014). Pengembangan tes kemampuan berpikir tingkat tinggi fisika (pysthots) peserta didik sma. *Jurnal Penelitian Dan Evaluasi Pendidikan*, 18(1), 1-12. <https://doi.org/10.21831/pep.v18i1.2120>
- Kaka, A. E. L., Ate, D., & Making, S. R. M. (2021). Analisis kemampuan literasi numerasi siswa smp n.1 kota

- tambolaka. *Jurnal Penelitian Pendidikan Matematika Sumba*, 3(2).  
Kementrian Pendidikan dan Kebudayaan. (2018). *Pendidikan di indonesia belajar dari hasil pisa 2018*. Retrieved from <http://repositori.kemdikbud.go.id/16742/1/Laporan Nasional PISA 2018 Indonesia.pdf>
- Kereh, C. T., Liliarsari, Tjiang, P. C., & Subandar, J. (2015). Validitas dan reliabilitas instrumen tes matematika dasar yang berkaitan dengan pendahuluan fisika inti. *Jurnal Inovasi Dan Pembelajaran Fisika*, 2(1).
- Kusaeri, Sadieda, L. U., Indayati, T., & Faizien, M. I. (2018). Developing an assessment instrument of higher order thinking skills in mathematics with in islamic context. *Journal of Physics: Conference Series*, 1097, 012151. <https://doi.org/10.1088/1742-6596/1097/1/012151>
- Liswati, T. W., Sakinah, N. G. A. P., & Yuniarti, Y. S. (2021). *Pengembangan instrumen penilaian berbasis literasi numerasi*. Jakarta: Direktorat Sekolah Menengah Atas.
- Mahendra, I. W. E., Parmithi, N. N., Hermawan, E., Juwana, D. P., & Gunartha, I. W. (2020). Teachers' formative assessment: Accessing students' high order thinking skills (hots). *International Journal of Innovation, Creativity and Change*, 12(12).
- Mohamad, M. M., Sulaiman, N. L., Sern, L. C., & Salleh, K. M. (2015). Measuring the validity and reliability of research instruments. *Procedia - Social and Behavioral Sciences*, 204, 164-171. <https://doi.org/10.1016/j.sbspro.2015.08.129>
- Ndiung, S., & Jediut, M. (2020). Pengembangan instrumen tes hasil belajar matematika peserta didik sekolah dasar berorientasi pada berpikir tingkat tinggi. *Premiere Educandum : Jurnal Pendidikan Dasar Dan Pembelajaran*, 10(1), 94. <https://doi.org/10.25273/pe.v10i1.6274>
- Nursakiah, N., Arriah, F., & Dharma, S. (2022). Developing mathematical literacy test with context of bugis-makassar local wisdom for junior high school students. *Jurnal Elemen*, 8(1), 16-28. <https://doi.org/10.29408/jel.v8i1.4049>
- Pramono, R. S. (2021). The feasibility study of moodle-based e-module to increase the numeration literacy of elementary school students. *DIDAKTIKA: Jurnal Pendidikan Sekolah Dasar*, (December).
- Pratama, G. S., & Retnawati, H. (2018). Urgency of higher order thinking skills (hots) content analysis in mathematics textbook. *Journal of Physics: Conference Series*, 1097, 012147. <https://doi.org/10.1088/1742-6596/1097/1/012147>
- Rahmawati, Y., Candra Pradana, K., Rinaldi, A., Syazali, M., & History, A. (2021). Curiosity and creative characters: The impact on students' numerical ability. *Desimal: Jurnal Matematika*, 4(2).
- Ratnaningsih, D. J., Isfarudi, I., & Soleiman, N. (2013). Analisis butir soal pilihan ganda ujian akhir semester mahasiswa di universitas terbuka dengan pendekatan teori tes klasik. *Jurnal Pendidikan Terbuka Dan Jarak Jauh*, 12(2).
- Retnawati, H. (2016). Analisis kuantitatif instrumen penelitian. In *Parama Publishing*.
- Sa'idah, N., Yulistianti, H. D., & Megawati, E. (2018). Analisis instrumen tes higher order thinking matematika smp. *Jurnal Pendidikan Matematika*, 13(1). <https://doi.org/10.22342/jpm.13.1.6619.41-54>

- Saukiyah, S., Sunardi, & Dinawati, T. (2017). Pengembangan soal literasi matematika berbasis budaya etnik madura untuk siswa smp/mts. *KadikmA*, 8(2).
- Sayekti, I., Sukestiyarno, Y., Wardono, & Dwijanto. (2021). Perception and understanding of madrasah tsanawiyah teachers on numerical literacy in mathematics learning. *Journal of Physics: Conference Series*, 1918(4), 042029. <https://doi.org/10.1088/1742-6596/1918/4/042029>
- Setyawati, M., Aisyah, H., & Kusaeri, K. (2021). Developing the problem of numeration literacy with islamic context. *MaPan*, 9(2), 347. <https://doi.org/10.24252/mapan.2021v9n2a11>
- Solehudin, Z., & Widodo, T. (2021). Influence of social media marketing activities on customer response with mediation of brand equity in e-wallet. *EProceedings of Management*, 8(2).
- Subroto, P. W., Suhadi, E. Y., & Julyana, R. (2019). Efektifitas model pembelajaran problem posing ditinjau dari kemampuan literasi matematis siswa smp pada materi aljabar. *Jurnal Tadris Matematika*, 2(2). <https://doi.org/10.21274/jtm.2019.2.2.185-194>
- Taherdoost, H. (2016). Validity and reliability of the research instrument; How to test the validation of a questionnaire/survey in a research. *SSRN Electronic Journal*. <https://doi.org/10.2139/ssrn.3205040>
- Tegeh, I. M., Jampel, I. N., & Pudjawan, K. (2015). Pengembangan buku ajar model penelitian pengembangan dengan model addie. *Jurnal Dimensi Pendidikan Dan Pembelajaran*, 3(1).
- Thompson, N. A. (2013). *Reliability & validity*. Assessment Systems.
- Vincent, W., & Shanmugam, S. K. S. (2020). The role of classical test theory to determine the quality of classroom teaching test items. *Pedagogia : Jurnal Pendidikan*, 9(1), 5–34. <https://doi.org/10.21070/pedagogia.v9i1.123>
- Widodo, S. A., Ibrahim, I., Hidayat, W., Maarif, S., & Sulistyowati, F. (2021). Development of mathematical problem solving tests on geometry for junior high school students. *Jurnal Elemen*, 7(1), 221–231. <https://doi.org/10.29408/jel.v7i1.2973>
- Zubaidah, S. (2018). Mengenal 4c: Learning and innovation skills untuk menghadapi era revolusi industri 4.0. *2nd Science Education National Conference*, (September).

