
Development of Test Instruments in Substance Pressure to Measure Student's Numerical Literacy Ability

Rita Iva Fatmala✉, Ani Rusilowati, Budi Astuti
Postgraduate Universitas Negeri Semarang, Semarang, Indonesia

Article Info

Article history:
Submitted 22 November 2022
Revised 12 February 2023
Accepted 14 February 2023

Keywords:
Test Instruments,
Literacy, Numeracy, Physics,
Pressure

Abstract

The purpose of this study was to develop an instrument test to measure student's numeracy literacy skills on substance pressure and its application. The research was conducted using the Research and Development (R&D) development method through the ADDIE research design (Analysis, Design, Development, Implementation, and Evaluation) through several stages, namely curriculum analysis, content analysis, context analysis, and determining instrument framework, instrument development, pilot testing instrument and item analysis of the instrument. The trial was conducted on a sample determined by a random sampling technique. The item analysis of the instrument includes item validity test, item reliability test, item difficulty level, and item discriminatory power. The results of this study were the test instrument of numeracy literacy in the form of multiple-choice questions with 18 reliable questions with a reliability coefficient of 0.903. Based on the validity test, there are 5 questions with moderate validity category, 10 questions with high validity category, and 3 questions with very high validity category. Based on the results of the discriminatory power test and the level of difficulty of the questions, it was obtained that all questions had distinguishing power in the good category with the difficulty level of the questions in the easy, medium, and difficult category are 1, 13, and 4 questions, respectively. Therefore, it can be concluded that the numerical literacy test instrument on the substance pressure material and its application can be used as a test instrument to measure student's numeracy literacy skills.

INTRODUCTION

The literacy and numeracy abilities of students are the focus of the development of the Indonesian government in the field of education. This is because the numeracy literacy skills of students are in line with the Human Development Index (HDI) which is the reference for the quality of human resources in a country. If the numeracy literacy ability of students is in a good category, then the HDI of a country is also good, and vice versa.

Aisyah (2021) reveal that students must at least have mastery of six basic literacy skills, which include literacy, numeracy, scientific literacy, digital literacy, financial literacy, and civic cultural literacy to be able to face global challenges 21st century. In line with these challenges, several regulations have been developed by the government to support the strengthening of student's numeracy literacy skills. Then, this regulation was followed up with the implementation of the program, called Gerakan Sekolah Literasi (School Literacy Movement) as the implementation of Permendikbud No. 23 of 2015. Saputri & Diana (2021) states that the implementation of government support for strengthening the literacy and numeracy abilities of students can be seen in the development of the Minimum Capability Assessment (MCA), where literacy and numeracy abilities are one of the abilities measured in the MCA. The implementation of the MCA can be used as a reference for the literacy and numeracy ability profile of students on a national scale to be followed up by educators and students as an evaluation material for strengthening student's literacy and numeracy abilities for the education unit level.

So far, Trends in International Mathematics and Science Study (TIMSS) and Program for International Student Assessment (PISA) have become the reference for measuring the scientific literacy and numeracy abilities of students on an international scale. Based on TIMSS data, the level of scientific literacy and numeracy of Indonesian students always ranks in the bottom five in the world.

According to Fuadah *et al.* (2017), scientific literacy and numeracy skills are very important in determining the quality of education in a country as well as providing students with the ability to face changing times and technological developments. Fuadi *et al.* (2020) state that one of the causes of the low level of scientific literacy and numeracy of students is the instrument technique that does not fully accommodate the criteria for assessing scientific literacy and numeracy abilities. This is in line with Maulida & Hamama (2021) research, the test instruments used in schools generally only test cognitive aspects and mathematical calculations, without considering process and context aspects.

The development of scientific literacy and numeracy test instruments is one of the efforts to measure and improve student's literacy and numeracy skills in the science field. Sulistiawati (2015) states that one way to measure student's scientific literacy and numeracy skills is by using a test instrument. It is also supports to Martinah *et al.* (2022) research that the development of scientific literacy and numeracy test instruments used should include four categories of scientific literacy and numeracy, namely science as a body of knowledge, science as a way of thinking, science as a way of investigating, and the interaction between science, technology, and society.

According Dewi & Rochintaniawati (2016), science literacy and numeracy instruments can be adopted from the questions on PISA. However, Chasanah *et al.* (2022) stated that the development of test instruments to measure student's science literacy and numeracy skills tailored to the background of learners in Indonesia is still needed to determine students' level of understanding of science concepts and their application in daily life. Then, in line with Howard *et al.*, (2017) research state that the development of numeracy and literacy test can be arranged based on the interests of students, this is to anticipate the results of abilities that are not in accordance with the reality. Therefore, a test instrument is needed to measure the literacy and numeracy skills of learners, especially in science subjects that correspond to the background of learners. Its support to O'toole *et al.*, (2020) research who explain that science education a fertile background for the development of literacy and numeracy test because this subject could clarify discussions of literacy in other disciplines. Grotlüschen *et al.*, (2020) add that the literacy and numeracy test are needed because related to major aspects of life, they are recognizing and analyzing the use of methods of inquiry that lead to scientific knowledge and the ability to organize, analyze, and interpret quantitative data and scientific information. The purpose of this research is to develop a physics literacy and numeracy test instrument on substance pressure and its application that is valid, reliable, has a differential power and a good level of difficulty to measure the literacy and numeracy of learners in general.

METHODS

The research was conducted using the Research and Development (R&D) method through the ADDIE research design which includes analysis, design, development, implementation, and evaluation. In general, the stages of research carried out in developing numeracy literacy test instruments on the material pressure of substances and their application in everyday life can be seen in Figure 1.

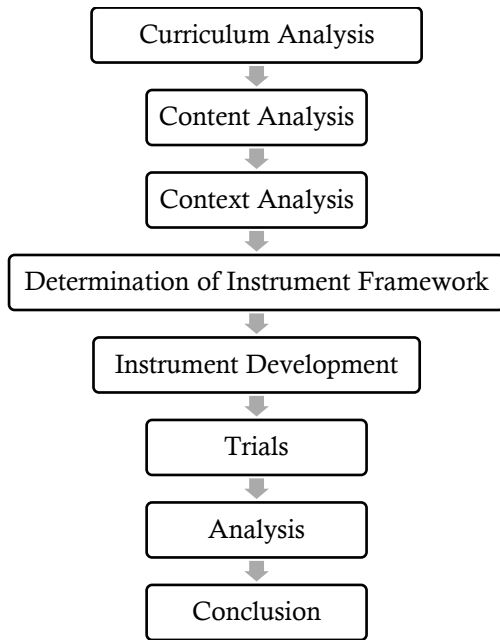


Figure 1. Stages of Development of Test Instruments

The examination of the test instrument was carried out on a sample determined by random sampling technique, namely the students of class VIII A, VIII B, VIII C, and class VIII H in MTs. Ma'ahid Kudus totaled 113 students. Data analysis carried out includes validity test, reliability test, level of difficulty, and discriminating power of questions. The instrument validity uses the product-moment correlation formula Arikunto (2015) as shown in equation (1).

$$r_{xy} = \frac{N \sum XY - (\sum X)(\sum Y)}{\sqrt{\{N \sum X^2 - (\sum X)^2\} \{N \sum Y^2 - (\sum Y)^2\}}} \quad (1)$$

where r_{xy} is the correlation coefficient of an item or item, N is the number of subjects, X is the score of an item or item and Y is the total score. Equation (1) above is used to measure the level of validity of the test questions per item with the question criteria said to be valid if the value of $r_{xy} > r_{Table}$. The r_{Table} data obtained from the auxiliary table is adjusted to the number of N , which is 113 with a significance level of 0.05, then the r_{Table} is 0.1848.

The determination of the category of the validity of the instrument developed refers to the classification of validity according to Guilford in Arikunto (2015), which is as follows. The validity category is very high if the r_{xy} is 0.81–1.00. High category if obtained r_{xy} 0.61–0.80. Sufficient category if obtained r_{xy} 0.41–0.60. Low category if obtained r_{xy} 0.21–0.40 and very low category if obtained r_{xy} 0.00–0.20.

The reliability of a test instrument produces a test that can still be said to be highly reliable. In this study, the formula for testing the reliability of the

instrument used is according to Arikunto (2015) as shown in equation (2) below.

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

where r_{11} is the reliability of the instrument, n is the number of instrument items, $\sum pq$ is the number of item variances, and S^2 is the total variance. The results of reliability calculations can be categorized using reliability criteria, namely if the reliability index $r_{11} > 0.70$ it is said that the developed instrument is reliable, but if $r_{11} < 0.70$ its can be said that the instrument is not reliable.

The level of difficulty is obtained by calculating according to equation (3) according to Arikunto (2015) as follows.

$$P = \frac{B}{JS} \quad (3)$$

where P is the difficulty index, B is the number of students who answered the questions correctly and JS is the total number of students who took the test. Good questions are questions that are not too easy or not too difficult. The difficulty index of 0.00–0.30 is a difficult question, for medium questions it has a difficulty index of 0.30–0.70, and for a difficulty index of 0.70–1.00 it is an easy question category.

Distinguishing power is obtained by the discrimination equation according to Arikunto (2015), which is shown in equation (4) as follows.

$$D = \frac{Ba}{Ja} - \frac{Bb}{Jb} \quad (4)$$

where D is the distinguishing power, Ba the number of correct answers in the upper group, Bb is the number of correct answers in the lower group, Ja is the number of students who did the correct test in the upper group, and Jb is the number of students who did the correct test in the lower group.

The criteria for discriminating power are determined as follows. The category of discriminatory power is very good if the results of the discrimination index are 0.70–1.00. Good category if the discrimination index is 0.40–0.69, sufficient category if the discrimination index is 0.20 – 0.39, and bad category if the discrimination index is 0.00 – 0.19, and if the discrimination index is less than 0, 00 then it is included in the bad category and the question should be discarded.

Septiani *et al.* (2019) stated that the literacy and numeracy skills of learners can be categorized with the average of the overall value of learners in working on the literacy and numeracy test instruments developed, where the category of literacy and numeracy skills is very good if obtained an average of 80–100, category literacy and numeracy ability are good for an average of 66–79, the literacy and numeracy ability category is adequate for an average of 56–65, our literacy and numeracy ability category for an average of 40–55, and the fail category for an average of 30– 39.

RESULTS AND DISCUSSION

The instrument was developed in the form of test questions in the form of multiple-choice questions in the form of 18 questions with four answer choices, this is in line with Patta & Muin (2021) research. The instrument was developed by containing aspects of literacy and numeracy that were adapted to the background of the students in

MTs. Ma'ahid is the sample of this study. The content of literacy and numeracy aspects according to students' backgrounds is obtained by interviewing students, teachers, and related stakeholders in the school environment. Then, the instrument was tested in April 2022 and the results were analyzed.

Validity analysis is used to determine the level of validity of a test item. The results of the test instrument validity can be seen in Table 1.

Table 1. Instrument Validity Test Results

Validity Index	Question Item	Amount	Percentage
>0,1848 (valid)	1; 2; 3; 4; 5; 6; 7; 8; 9; 10; 11; 12; 13; 14; 15; 16; 17; 18	18	100%
<0,1848 (not valid)	-	-	0%

The validity category obtained in the instrument test can be seen in Table 2 below.

Table 2. Category of Instrument Validity Test Results

Category	Question Item	Amount	Percentage
Very High	5, 10, 11	3	17%
High	1, 3, 6, 7, 8, 9, 12, 14, 17, 18	10	55%
Medium	2, 4, 13, 15, 16	5	28 %
Low	-	-	0%
Very Low	-	-	0%

Based on the results of the validity test of the instrument, it can be said that the instrument on the substance pressure material and its application developed is valid in measuring students' numeracy literacy abilities.

Reliability analysis is used to determine the level of consistency of the developed instrument. Based on the calculations obtained r_{11} of 0.903 or greater than the r_{Table} of 0.1848 so it can be stated that the test instrument developed is reliable. Based on the results of the reliability test of the items on the material pressure test instrument and its application,

it can be said that the test instrument is reliable and the results in measuring the scientific literacy and numeracy abilities of students can be trusted.

Based on the difficulty level test, it can be concluded that the question instrument is included in the good question category as evidenced by the large percentage of questions in the medium category, which is as much as 72%. This is supported by Arikunto (2015) who states that a good question is a question that is neither too easy nor too difficult. The measurement test data can be seen in Table 3.

Table 3. Category of Instrument Difficulty Level Test Results

Difficulty Level Indeks	Question Item	Amount	Percentage
Easy	7	1	6%
Medium	1, 2, 3, 4, 5, 8, 9, 10, 11, 12, 13, 14, 15	13	72%
Difficult	6. 16, 17, 18	4	22%

The discriminatory test is used to measure the ability of a question to distinguish between high-ability students and low-ability students. Based on the discriminatory power test, the different power value for each item is greater than equal to 0.5 so that it can be categorized that the material pressure

test instrument and its application are in the category of literacy and numeracy ability test instruments with good discriminating power.

Based on the test results with the developed test instrument, it was found that the literacy and numeracy abilities of students in MTs. Ma'ahid in

the various category. Testing the category of scientific literacy and numeracy levels obtained data as shown in Table 4.

Table 4. Category of Scientific Literacy Level Test Results

Science Numeracy Literacy Level Index	Amount	Percentage
Very Good	22	20%
Good	14	12%
Enough	32	28%
Not Enough	17	15%
Fail	28	25%

The OECD (2015) states that the factors causing the low literacy and numeracy of science in Indonesia include gender, economic and social. Nofiana & Julianto (2018) states that the low scientific literacy and numeracy skills of students are triggered by their unpreparedness of students to the times and problems in the surrounding environment. Therefore, there is a need for a series of actions to increase and strengthen the literacy and numeracy abilities of students, especially in the environment of MTs. Ma'ahid to later has quality graduates who are competent and able to compete in the global era in the 21st century.

CONCLUSION

Based on the validity test of the test instrument developed, 17% of the questions were in the very high valid category, 55% of the questions were in the high valid category and 28% were in the moderately valid category. Questions in the reliable category with a reliability value of 0.903. In addition, the questions are in the good category with the percentage of questions with a medium difficulty level of 72% and can categorize students with low and high abilities through a discriminatory power level of more than 0.5. Therefore, the test instrument developed on the material of substance pressure and its application in everyday life can be used as an instrument to measure students' numeracy literacy skills.

REFERENCES

- Aisyah, H. (2021). *Pengembangan Instrumen Penilaian Literasi Numerasi dengan Memasukkan Konteks Islam*. UIN Sunan Ampel Surabaya.
- Arikunto, S. (2015). *Prosedur Penelitian Suatu Pendekatan Praktik*. Rineka Cipta.
- Chasanah, N., Widodo, W., & Suprpto, N. (2022). Pengembangan Instrumen Asesmen Literasi Sains Untuk Mendeskripsikan Profil Peserta Didik. *PENDIPA Journal of Science Education*, 6(2), 474–483. <https://doi.org/10.33369/pendipa.6.2.474-483>
- Dewi, P. S., & Rochintaniawati, D. (2016). Kemampuan Proses Sains Siswa Melalui Pendekatan Saintifik Dalam Pembelajaran Ipa Terpadu Pada Tema Global Warming. *Edusains*, 8(1), 18–26. <https://doi.org/10.15408/es.v8i1.1564>
- Fuadah, H., Rusilowati, A., & Hartono. (2017). Pengembangan Alat Evaluasi Literasi Sains untuk Mengukur Kemampuan Literasi Sains Siswa Bertema Perpindahan Kalor dalam Kehidupan. *Lembaran Ilmu Kependidikan*, 46(2), 51–59.
- Fuadi, H., Robbia, A. Z., Jamaluddin, J., & Jufri, A. W. (2020). Analisis Faktor Penyebab Rendahnya Kemampuan Literasi Sains Peserta Didik. *Jurnal Ilmiah Profesi Pendidikan*, 5(2), 108–116. <https://doi.org/10.29303/jipp.v5i2.122>
- Grotlüschen, A., Desjardins, R., & Liu, H. (2020). Literacy and numeracy: Global and comparative perspectives. *International Review of Education*, 66(2–3), 127–137. <https://doi.org/10.1007/s11159-020-09854-x>
- Howard, S. J., Woodcock, S., Ehrich, J., & Bokosmaty, S. (2017). What are standardized literacy and numeracy tests testing? Evidence of the domain-general contributions to students' standardized educational test performance. *British Journal of Educational Psychology*, 87(1), 108–122. <https://doi.org/10.1111/bjep.12138>
- Martinah, A. A., Mubarak, V., Miarsyah, M., & Ristanto, R. H. (2022). Pengembangan Instrumen Tes Literasi Sains Berbasis Kontekstual pada Materi Pencemaran Lingkungan. *Bioedusia: Jurnal Pendidikan Biologi*, 6(2), 192–218. <https://doi.org/10.37058/bioed.v6i2.3251>
- Maulida, & Hamama, S. F. (2021). Pengembangan Instrumen Tes Tipe Pilihan Ganda Dalam Evaluasi Hasil Belajar Siswa Pada Konsep Sel Tingkat Sekolah Menengah Atas. *Jurnal Dedikasi Pendidikan*, 8848(1), 171–178.
- Nofiana, M., & Julianto, T. (2018). Upaya Peningkatan Literasi Sains Siswa Melalui Pembelajaran Berbasis Keunggulan Lokal. *Biosfer: Jurnal Tadris Biologi*, 9(1), 24. <https://doi.org/10.24042/biosf.v9i1.2876>
- O'toole, J. M., McKoy, K., Freestone, M., &

- Osborn, J. A. (2020). 'Scientific Literacy': An Exercise in Model Building. *Education Sciences*, 10(8), 1–16. <https://doi.org/10.3390/educsci10080204>
- OECD. (2015). *PISA RESULTS IN FOCUS - 2015*. Colombia University.
- Patta, R., & Muin, A. (2021). JIKAP PGSD : Jurnal Ilmiah Ilmu Kependidikan Kemampuan Literasi Numerasi Ditinjau Dari Gaya Kognitif Reflektif-Impulsif. *JIKAP PGSD: Jurnal Ilmiah Ilmu Kependidikan*, 212–217.
- Saputri, V., & Diana, H. A. (2021). Development of HOTS Problem-Based Test Instruments to Measure Level 4 Numeracy Capabilities Using Rasch Model. *Jurnal Ilmiah Mandala Education*, 7(4), 62–66. <https://doi.org/10.36312/jime.v7i4.2371>
- Septiani, D., Widiyawati, Y., & Nurwahidah, I. (2019). Pengembangan Instrumen Tes Literasi Sains Berbasis Pisa Pada Aspek Menjelaskan Fenomena Ilmiah Untuk Siswa Kelas Vii. *Science Education and Application Journal*, 1(2), 46. <https://doi.org/10.30736/seaj.v1i2.144>
- Sulistiawati. (2015). Analisa Pemahaman Literasi Sains Mahasiswa yang Mengambil Mata Kuliah IPA Terpadu menggunakan Contoh Soal PISA 2009. *Sainteks*, XII(1), 21–40.