Check for updates

Developing mathematics problems using local wisdom context of Maluku to improve students' numeracy

Fahruh Juhaevah [*]

Mathematics Education Department, IAIN Ambon, Maluku, Indonesia

[*] Correspondence: fahruh.juhaevah@iainambon.ac.id
© The Author 2022

Abstract

Students' difficulty in solving numeracy problems is a problem that teachers must pay attention. Therefore, it needs media to help students improve their numeracy ability. The purpose of this study is to develop mathematics problems based on PISA (Programme for International Student Assessment) content and use local wisdom context in Maluku that obtained validity, practically, and effectively. Mathematics problems developed are expected to improve the numeration ability of junior high school students in Maluku. The study was research and development based on the Plomp model. The subject consists of 40 students (15 males and 25 females). The result showed that mathematics problems met validity criteria. The validity was tested using the Rasch model that consists of suitable, reliability, difficulty, and differentiation items. Then, it was validated based on experts in content, league, and media. All of them met validity criteria. Mathematics problems had met practically based on the user experience questionnaire (UEQ) that consisted of interest, clarity, efficiency, accuracy, stimulation, novelty, and prospered based on the system usability scale (SUS). Mathematics problems improve the student's numeration ability, as shown based on N-Gain 0.36. Thus, it effectively improves the numeracy ability of students in Maluku.

Keywords: local wisdom; mathematics problem; numeracy

Introduction

Indonesia is one of the countries that participates every three years in PISA activities, starting from 2000 until the last PISA assessment in 2018. Based on participation data in 2018, Indonesia was ranked 70th for science skills, while literacy and math skills were ranked 72 out of 79 countries (Schleicher, 2019). The government has made various efforts to improve the ranking of Indonesian students on PISA assessments; one of the programs carried out by the government is the National Literacy Movement with Numeracy Literacy content. In addition, the Minister of Education has abolished the National Examination, which was turned into a Minimum Competency Assessment and character survey. The problems presented at the Minimum Competency Assessment consist of two types, namely literacy, and numeracy.

Literacy and numeracy are very important because they can activate the ability to analyze, build arguments, interpret, and communicate ideas or ideas effectively in a problem-solving process (Govorova et al., 2020; Stacey & Turner, 2015; Tai & Lin, 2015). The problems or problems presented at Minimum Competency Assessment have similarities with PISA problems (Nusantara et al., 2020).

The cause of the low numeracy of students in Indonesia is that students are still not used to working on PISA standard problems (Charmila et al., 2016; Dasaprawira, 2019; Nusantara & Putri, 2020). In addition, the problems presented in the PISA problem have a general context because the problem presented is the same problem and then translated into the language of each country (Stacey & Turner, 2015), whereas students will be easier to solve problems in the context of the problem presented has been experienced or seen directly by students (Jonassen, 2010; Van den Heuvel-Panhuizen & Drijvers, 2020).

One of the contexts that can bring students closer to everyday life is the context of culture and local wisdom (Hasanah, 2017; Retnowati & UN, 2010; Supiyati & Hanum, 2019). In addition to bringing students closer to daily life problems, local culture is also very appropriate as a medium in instilling character education (Samo, 2018; Suradi, 2018). For this reason, it is essential to introduce children related to culture through learning mathematics in school, especially problems related to numeracy ability.

National Literacy Movement activities in 2021 get obstacles, especially implementation in the field because of the extraordinary events of the COVID-19 pandemic that forces all forms of learning activities to be diverted using online media based on the Ministry of Education Letter No. 4 of 2020. The results of initial observations in Ambon City showed that students' numeracy ability is still very low. From 40 students who were given random numeracy questions, only four students were able to answer the problem appropriately, and the rest had difficulty either understanding the problem, finding the right problem-solving strategy, or misusing the principle/concept in solving the given problem. Most students find it difficult to work on numeracy problems because they do not understand the context of the given problem. In addition, some research shows that students' numeracy ability is still very low based on students' inability to understand problems, transform problems, and use appropriate concepts in solving problems (Aliifah, 2020; Ekowati et al., 2019; Puspita & Ahda, 2020).

The context of local wisdom is chosen based on the that students will easily solve a given problem if they understand the context of the problem presented (Mungmachon, 2012; Uge et al., 2019). In addition, the problem with PISA consists of a social context that contains the local cultural content of a society (Johar, 2012; Stacey, 2015). The local cultural context of Maluku is very diverse and numerous (Andaya, 2018; Latinis & Stark, 2005), thus allowing it to be used as a context for problems to be developed. Students are considered to need to have numeracy ability so that students can solve their life problems in the future; of course, the problem is not far from their lives (Tan, 2020)

Using mathematics problem-based PISA content and the local wisdom context of Maluku, students can practice working on existing numeracy problems and being used as media for evaluating numeracy ability following the function used as a media of practice to solve problems and evaluation (Amtmann et al., 2020).

Several previous studies that used to develop PISA content mathematics problems, including using change and relationship content (Hasibuan & Fauzi, 2019; Nusantara et al., 2020), using shape and space content (Efriani & Putri, 2019; Nasution et al., 2019), using quantity content (Murtiyasa et al., 2018; Nusantara, 2021), and uses uncertainty and data content (Nizar & Putri, 2018; Yansen & Putri, 2019). Of the few existing studies, no one has simultaneously developed a mathematical problem using four PISA contents. In addition, some studies develop mathematical problems using the context of local wisdom, including using the context of local Javanese wisdom (Machromah & Sari, 2021; Mulyatna et al., 2021), using the context of Jambi (Charmila et al., 2016), Palembang (Mayari, 2020), Malang (Pratiwi et al., 2020), Riau (Zulfah & Insani, 2020), and Bangka Belitung (Putra & Vebrian, 2019). Based on the research, there has been no development of mathematical problems that use Maluku context, so researcher are interested in developing mathematical problems using four PISA contents using Maluku context, namely *Pela*, *Gandong*, *Famili*, and *Sasi*. In addition, the development of mathematical problems is presented using Moodle-assisted websites.

Based on the above description, the researcher is interested in developing mathematics problems that accommodate the content of PISA and the context of Maluku content which includes validity, practicality, and effectiveness (Nuraeni et al., 2021; Rosli et al., 2020; van den Akker et al., 2012). So that, students are used to solving problems that require numeracy ability. The ultimate goal of this media development is to improve the students' numeracy in Maluku.

Methods

This study used students at SMPN 14 Ambon and Muhammadiyah Ambon Junior High School in the development of Mathematics problems with details of 40 students (15 males and 25 females). The instruments used in this study are test question sheets, UEQ questionnaires, SUSquestionnaires, validation sheets, and test questions based on developed mathematics problems. The test question sheet consists of ten minimum competency assessment questions in 2020, including number material, algebra, geometry, and statistics. The UEQ questionnaire was adopted from Santoso et al. (2016), which contains 26 items with seven preferred scales

that represented 1) attractiveness (fun, kind, exhilarating, comfortable, attractive, user friendly), 2) clarity (understandable, easy to learn, simple, clear), 3) efficient (fast, efficient, practical, organized), 4) accuracy (predictable, supportive, safe, as expected) 5) stimulus (useful, interesting, not boring, motivating), 6) novelty (creative, inventive, unorthodox, innovative). The SUS questionnaire consists of ten statements with five answer scales. SUS questionnaires are used to assess media systems developed related to the overall use of media that consist of reuse, complex, easy, hard to use, stable, inconsistency, easy to learn, confused, no problem, need time. The mathematics problem validation sheet consists of three instruments consisting of content (suitability of the problem with the curriculum, completeness of representation, suitability of PISA problems and content, proportionality of the problem, and level of difficulty of the problem), language (choice of words, easy-to-understand sentences, ambiguity, sentence structure, default words, proper writing, scientific terms, effective sentences, and informative), and media (layout, background, color, typeface, navigator, consistency, archival, completeness of information, image suitability). The validity of content is related to the conformity of content with the material taught in schools, language validity is related to the readability of the bank of the problem, and the validity of the media is related to the appearance of the problem. The final test question used questions that have been developed in Mathematics problems. It consists of 16 questions, consisting of four PISA contents that consist of change and relationship, shape and space, quantity, uncertainty, and data; and four contexts of local wisdom of Maluku: *Pela*, *Gandong*, *Famili*, and *Sasi*.

The test question sheet is used as an initial description of the numeracy ability of students used at the preliminary research stage. At the development or prototyping stage, the researcher creates and develops problem items based on PISA content: change and relationship, shape and space, quantity, and uncertainty and data in the context of Maluku local wisdom related to *Pela, Gandong, Famili*, and *Sasi*. *Pela* is a form of a fraternal bond of friendship that connects the indigenous peoples of two different countries (Wakano, 2019). In addition, *Gandong* is a fraternal relationship or friendship on the basis of a genological relationship (Malatuny & Ritiauw, 2018). *Famili* is the existence of fraternal relations based on the clan (Wakano, 2019). *Sasi* is a form of management and utilization of all existing sources, especially natural resources traditionally carried out for generations in Maluku (Lonthor, 2018). After the development is tested, the problem point is based on the analysis of the Rasch Model, which measures the suitability of the problem item with the model, the different power, and the difficulty level of the problem item. The criteria of the Rasch Model that analyzed used Winsteps 5.7.1.0 version Boone et al. (2013) used are Table 1 as follows.

Table 1. Criteria for suitability of questions

Criteria	Interval	Description
Mean Square (MNSQ)	0.5 < MNSQ < 1.5	Accept
Z-Standard (ZSTD)	-2.0 < ZSTD < +2.0	Accept
Pt Measure Corr.	0.4 < pt measure corr < 0.85	Accept

The measure used to determine the problem's difficulty level is to look at the item: measure using the Winsteps app. Measure values have a range between -3 to 3. However, some experts also limit the lower and upper limits of the possible values to between the range of -2 to 2, if they are outside that range can be expressed as extreme or deviant values (Alagumalai et al., 2005; Bond & Fox, 2013; Dutt et al., 2019). Sumintono and Widhiarso (2015) guide assessing the item into four categories: the following Table 2.

Table 2. Difficulty level of problem

Interval	Category
$me < -1$	Very easy
$-1 < me \leq 0$	Easy
$0 < me \leq 1$	Hard
$me > 1$	Very hard

At the assessment stage, construct validation is carried out to measure the validity of the bank based on expert assessment of content/ material, language, and media. The criteria for its validity test decision based on Widoyoko (2009) are presented in the following Table 3.

Table 3. Validity Criteria

Interval	Validity	Description
$4.2 < V \leq 5$	Highly valid	No need revision
$3.4 < V \leq 4.2$	Valid	need revision
$2.6 < V \leq 3.4$	Enough	partial revision
$1.8 < V \leq 2.6$	Invalid	full revision
$0 \leq V \leq 1.8$	Highly invalid	full revision

The practicality and feasibility of the bank is measured using the UEQ (Schrepp et al., 2017), SUS (Martins et al., 2015). Here, in Table 4 is the EUQ decision criteria, and in Table 5 is the SUS decision criteria.

Table 4. Pratically criteria

Interval	Category
$1.75 \leq n \leq 2.50$	Highly practical
$1.60 < n \leq 1.75$	Practical
$1.20 < n \leq 1.60$	Average
$0.70 < n \leq 1.20$	Impractical
$-1.00 \leq n \leq 0.70$	Highly impractical

Table 5. SUS eligibility category

Interval	Category
$SUS \geq 70$	Proper
$SUS < 70$	Improper

To measure the effectiveness of mathematics problems, a comparison of initial test results (pretest) that uses Minimum Competency Assessment problems in 2020 and final test results (posttest) using the problem in mathematics problems. It is then analyzed using descriptive statistics with N-Gain parameters and inferential statistics with the Wilcoxon signed-rank test. The criteria for N-Gain (Nissen et al., 2018) are as follows in Table 6.

Table 6. N-Gain category

Interval	Category
$0.7 < g \leq 1$	High
$0.3 < g \leq 0.7$	Average
$0 < g \leq 0.3$	Low

Results

Mathematics problems are the result of the development that has been done. It before first use through the manufacturing and development process. It that has been developed can be utilized by students to improve numeracy ability. The study was developed based on three stages based on the following Plomp development model.

Preliminary research

At this stage, problem analysis is carried out to identify problems experienced by students related to numeracy ability. For this reason, the initial test is carried out using the Minimum Competency Assessment problem that has been used in 2020. Based on the results of tests given to students, information is obtained that students have low student numeracy ability. It is marked by an average student outcome of 38. For this reason, product development is needed in the website form that consists of mathematics problems that can improve students' numeracy ability.

Development or prototyping

At this stage, the creation and development of problem items and bank problems is carried out. The development of the problem item is based on the PISA content and context of local wisdom Maluku. The distribution of the mathematics problems made are as follows.

Table 7. Distribution of PISA content and local wisdom context of Maluku

Content	Context				Total
	Pela	*Gandong*	*Famili*	*Sasi*	
Change and Relationship	3	2	2	2	9
Shape and Space	2	3	2	2	9
Quantity	2	2	2	2	8
Uncertainty and Data	2	2	2	3	9
Total	**9**	**9**	**8**	**9**	**35**

Assessment

After the problem item is made, the process of analysis of the item is carried out using the Rasch model. The results of the analysis are as follows.

Suitability model (Fit)

Based on the results of the model conformity analysis using the Winstep application, information was obtained from 35 points of the problem developed. There were 33 qualified questions received related to the suitability of the problem using two test indicators, namely

MNSQ and ZSTD. The MNSQ test indicator is used to determine if the problem to be used can measure a student's numeracy ability. In contrast, the ZSTD test indicator is used to measure whether the grades obtained by students are likely rational. The MNSQ value in the item analysis conducted is in the range of 0.85 to 1.25. The ZSTD value in the item analysis conducted is in the range of -2.4 to 2.2. The ZSTD value on two problem items (number 17 and number 32) is not accepted because it has a value outside the specified range. Suitability model used to justify validity item based on Rasch model.

Different power

To determine the consistency of the problem in determining students who can answer correctly and students who cannot answer correctly. Pt. Measure Correlation is used as an indicator of different power tests. Based on the data analysis obtained information, three points of questions do not meet. It is because of the value of Pt. Measure Correlation is negative. The problem that does not meet is the numbers 5, 17, and 32. Problem items that do not meet the different power criteria are omitted and not used.

Difficulty level

The problem of difficulty level, measure orders are used. Based on the analysis of the difficulty level of the item obtained information that the problem item is very difficult is the problem item number 21, 18, 26, and 30 because the value of me > 1. In contrast, problem item number 20, 11, 15, 19, 31, 17, 23, 24, and 25 belongs to the category of difficult problems because they are at the interval of $0 < me \leq 1$. Easy problems are found in the item of the number 4, 10, 22, 5, 9, 14, 28, 35, 1, 29, 3, 6, 8, 12, 13, 2, 7, 16, 27, 32, and 33 because they are at an interval of $-1 < me \leq 0$. While the number of very easy problems is only found in problem number 34 because it is at the interval me < -1.

Reliability

The reliability test is carried out to find out the problem item can show consistency in describing the numeracy ability of students. Based on the reliability test results, the obtained information is that the value of the reliability of the item is 0.74, and the reality of the person is 0.69 with the value of Alpha Cronbach of 0.76. Thus, the problem item has fulfilled the element of inconsistency (Schrepp, 2020).

Based on the results of the analysis of the problem details, 32 questions are used to be further entered into the website, and three others are omitted. After all the problem items are entered into the website, an expert validation test is conducted based on content, language, and media. Expert validation results are presented in the following Table 8.
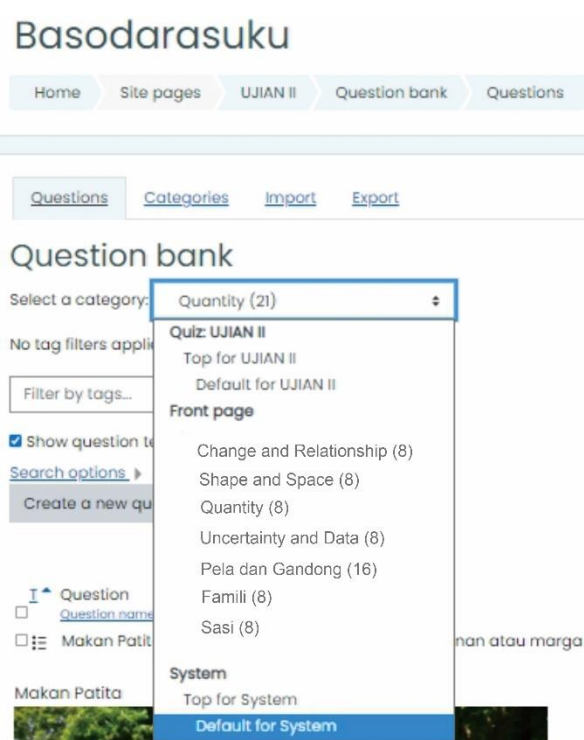
Table 8. Expert validation results

Expert	Mean		Revised	Criteria
	First	Second		
Content	3.8	4	Proportional Content	Valid
Language	3.83	4.11	Effectiveness of sentences	Valid
Media	3.9	4	Completeness of Information	Valid

Content validity

Here in Figure 1 and Figure 2 are the results of validation and revision based on expert suggestion:
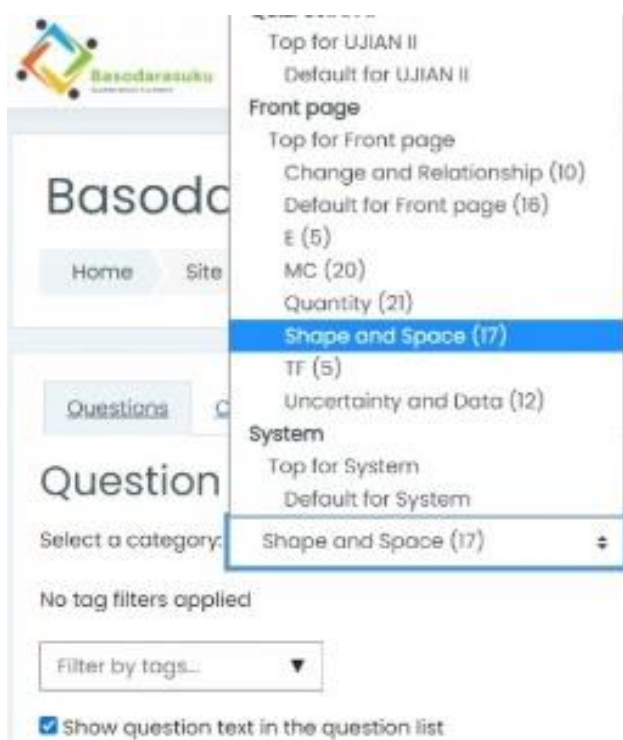
Figure 1. Before revised   Figure 2. After revised

The focus for revisions made by content experts is proportionality between the issues in each content and the context so that the content and context have the same contribution.

Language validity

Here are the results of validation and revision based on the linguists.

> *Residents of Rohomoni Country, Haruku Island, Central Maluku Regency, Maluku, follow the tradition of Maasiri Rumah Sigit or the change of roof of Uli Hatuhaha Old Mosque, in Rohomoni Village. The tradition of changing the roof of the mosque built in the 16th century AD is done once a year as a symbol of purification and self-cleansing to welcome the month of Ramadan.*

The mosque has three sections or levels of different sizes. The bottom two sections form four trapezoidal sides, and the tops are five-shaped quadrilateral without a base. The roof of the mosque is made of dried sago leaves. The longest side of the largest roof trapezoid is 24 m and the parallel side is 12 m long. If the ratio of the largest roof area and roof area in the second part is 3:1, then the number of roof sheets that are 2 m² wide is as much ... to cover the second part of the roof.

The focus of the linguist's input is the effective use of sentences. The last sentence on the problem indicates an ineffective sentence because it has the potential to give rise to a double meaning (ambiguous). Should be *If the ratio of the largest roof area and the roof area in the second part is 3:1, then how much roof is needed to cover the roof of the first and second parts, if the roof area used each sheet is 2 m² ...*

Media validity

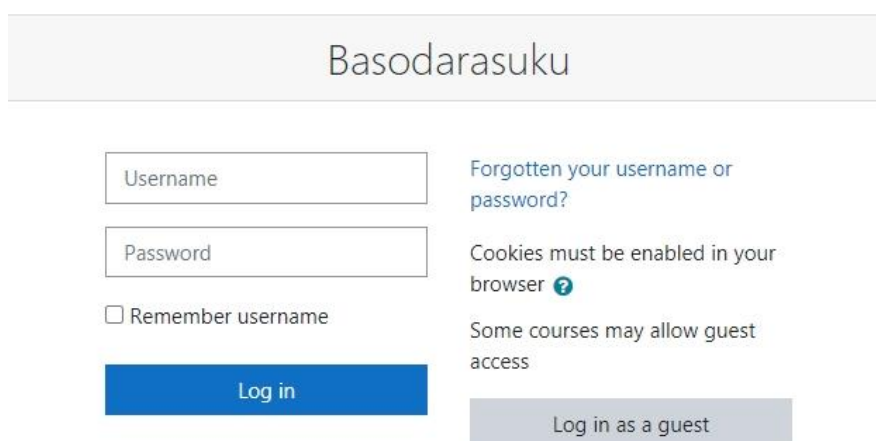Here in Figure 3 and Figure 4 are the results of validation and revision based on expert media suggestion:
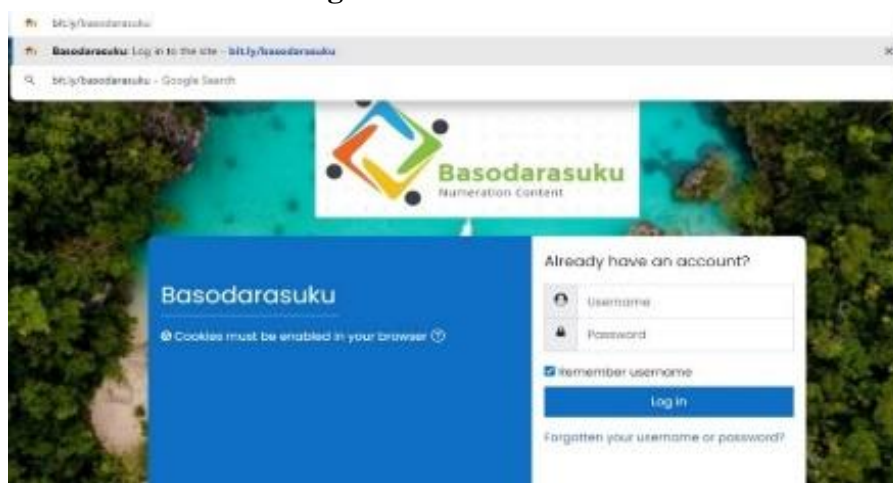

Figure 3. Before revised


Figure 4. After revised

The focus of revision of media experts is the front page of the website. The front page view should be made more interesting and show information related to the naming of the website.

To find out the practicality and feasibility of Mathematics problems, students who have used Mathematics problems are given the opportunity to fill out the usage questionnaire based on UEQ and SUS. Here in Table 10 is the results of UEQ data analysis.

Table 10. Data analysis results of UEQ

Component	Mean	Critea
Attraction	1.36	Enough
Clearity	1.51	Enough
Efficiency	1.42	Enough
Accuracy	1.65	Practical
Stimulation	1.78	Highly Practical
Novelty	1.65	Practical

Based on table 10 above, all components meet the criteria of practicality. So that students can use Mathematics problems. After knowing the practicalities, students can fill out the SUS questionnaire to determine the feasibility of Mathematics problems. Based on the results of the SUS questionnaire analysis, the average value of 75 or is on the criteria is proper. Here in Figure 5 is presented a picture of proper acceptance.
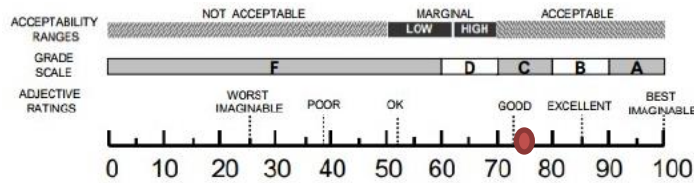
Figure 5. Acceptance diagram based on proper

To measure the effectiveness of mathematics problems, tests were conducted on students using the problems contained in Mathematics problems. Based on the test results obtained, the average student test result of 60.5 and N-Gain of 0.39, or if included in the category, the value can be declared to be in the moderate category. Thus, there is a difference in the numeracy of students who use minimum competency assessment problems and those who use mathematics problems. So it can be stated that Mathematics problems can improve students' numeracy. To find out the difference is statistically significant, the t-test with nonparametric criteria is obtained to obtain the following results in Table 11.

Table 11. t-test Wilcoxon

	W	df	p
Pre - Post	0.000		< .001

Based on the Wilcoxon t-test obtained information that the value of significantly p is smaller than $\alpha = 0.05$. So, it can be stated that there is a significant difference between the numeracy ability of students before and after using mathematics problems.

Discussion

Preliminary Research

The development stage begins with identifying problems, one way to provide diagnostic tests to students. It is in line with the Huda et al. (2020); Kusumaningsih et al.(2020), which states that before the development of media, first conducted initial tests so that the student's abilities could be known. The study results showed that the numeracy ability of students is still low; this is in line with Megawati and Sutarto (2021); Saefurohman et al. (2021). Thus, The media is needed that can facilitate students in improving their numeracy ability.

Development

The development of mathematical problems is based on PISA content and the context of local wisdom of Maluku, based on the results of research showing that there are 35 problems developed proportionally. Problem development refers to the context of local wisdom that can increase students' attractiveness in solving problems because they can easily understand the context of a given problem. (Hasanah, 2017; Hasibuan & Fauzi, 2019).

Assessment

The development of the mathematics problem has fulfilled the Rasch model consisting of reliability, conformity of the model, difficulty level, and other power. It must meet the criteria that have been determined so that the problem item used is valid. Some studies show that the problem item that has been tried and validated first is better than the problem item that is not validated first (Alavi et al., 2020; Erfan et al., 2020; Lu et al., 2021). Invalid items are negligible and unfit for use (Petersen et al., 2020; Robitzsch, 2021). In addition to item validation, the mathematics problems website has gone through an expert validation process for the website, including content, language, and media. Validation results show that the content is following the issue of PISA standards, has been following the rules of good Indonesian, and the media has been following the purposes of media creation. Thus, the item on the website has been validated to meet the element of validity and is worth to use.

The item that must be developed has fulfilled the validity of content with proportionally diverse issues. It is in line with Barrada et al. (2008), Koneputugodage et al. (2019), Widana (2014), who stated that a good item accommodates the diversity of types of questions presented and it can represent all materials proportionally. In addition, the developed item has met the validity of language characterized by the use of effective sentences in presenting problems. It is in line with Fadhilah (2021), Van Anh (2021), Winarti et al. (2021), which states that sentences on the problem must be effective and easily understood by students so as not to cause ambiguity. The item banks must also show an attractive display to increase students' attractiveness in using the item banks. It is in line with Mutmainah (2017), Nurpala and Munawaroh (2020) which states that the item must be presented attractively so that students are enthusiastic in working on the problems presented.

The website of mathematics problems on the results of development shows practicality based on indicators of attractiveness, clarity, efficiency, accuracy, stimulation, and novelty in

line with the opinion of Schrepp et al. (2017), which states that the development media is worth using when it has met the six indicators with a value above 1.20. In addition, the item banks are also worth using because the value of SUS is above 0.75 (Lewis & Sauro, 2018), so the mathematics problems are proper to use. Media development results are said to be effective if there are differences before and after using media (Chanda et al., 2019; Widodo, 2018). It supports the mathematics problems based on website development research on this issue because it improves students' ability to solve numeracy problems. It is characterized by increasing students' N-Gain grades before and after being given a problem using PISA content and local wisdom context with average category. It indicates that there is an increase in the numeracy of students. The average category simply reveals the difference before and after using the website.

The novelty in the research is in the products produced in the form of a collection of problems that contain problems related to the PISA content of the context of local wisdom of Maluku. The item's content and context are presented systematically, making it easier for students to choose and work on the type of problem they want. In addition, Mathematics problems can integrate student knowledge in the context of Maluku and numeracy issues so that students can solve numeracy problems through mathematics problems which have implications for improving students' numeracy ability in solving contextual problems.

Conclusion

Mathematics problems meet the elements of validity, practicality, and effectiveness. The validity consists of questions and item banks. The problem item's validity and reliability are based on the model's adjustable test with 32 valid question points with an Alpha Crombach value of 0.76 or can be said to be reliable. In addition, the validity of mathematics problems website based on experts shows valid criteria in content, language, and media. Mathematics problems that use the website effectively improve the numeracy ability of Maluku students because the N-Gain of students is 0.39 or is in the average category. In addition, the increase was accompanied by significant differences in students' numeracy abilities before and after using mathematics problems based on the website. The research is limited to the PISA content and socio-cultural context domains.

Conflicts of Interest

The author declares that no conflict of interest regarding the publication of this manuscript. In addition, the ethical issues, including plagiarism, misconduct, data fabrication and/or falsification, double publication and/or submission, and redundancies have been completely by the author.

References

Alagumalai, S., Curtis, D. D., & Hungi, N. (2005). *Applied Rasch measurement: A book of exemplars*. Springer.

Alavi, K., Isa, K., & Palpanadan, S. T. (2020). Application of Rasch model on resilience in higher education: An examination of validity and reliability of Malaysian Academician Happiness Index (MAHI). *International Journal of Higher Education*, *9*(4), 261–271. https://doi.org/10.5430/ijhe.v9n4p261

Aliifah, S. (2020). *Kajian terhadap kesalahan siswa dalam menyelesaikan masalah literasi numerasi [A study of student errors in solving numeracy literacy problems]*. Universitas Pendidikan Indonesia.

Amtmann, D., Bamer, A. M., Alschuler, K. N., Bocell, F. D., Ehde, D. M., Jensen, M. P., Johnson, K., Nery-Hurwit, M. B., Salem, R., & Silverman, A. (2020). Development of a resilience item bank and short forms. *Rehabilitation Psychology*, *65*(2), 145-157. https://doi.org/10.1037/rep0000312

Andaya, L. Y. (2018). Cultural state formation in eastern Indonesia. In *Southeast Asia in the early modern era* (pp. 23–41). Cornell University Press. https://doi.org/10.7591/9781501732171-005

Barrada, J. R., Olea, J., Ponsoda, V., & Abad, F. J. (2008). Incorporating randomness in the fisher information for improving item-exposure control in CATs. *British Journal of Mathematical and Statistical Psychology*, *61*(2), 493–513. https://doi.org/10.1348/000711007X230937

Bond, T. G., & Fox, C. M. (2013). *Applying the Rasch model: Fundamental measurement in the human sciences*. Psychology Press.

Boone, W. J., Staver, J. R., & Yale, M. S. (2013). *Rasch analysis in the human sciences*. Springer. https://doi.org/10.1007/978-94-007-6857-4

Chanda, K., Chanplin, P., & Charoenpoom, S. (2019). A development of creative instruction model through digital media. *International Journal of Educational and Pedagogical Sciences*, *13*(6), 785–788.

Charmila, N., Zulkardi, & Darmawijoyo. (2016). Pengembangan soal matematika model PISA menggunakan konteks Jambi [Development of PISA model math problems using the Jambi context]. *Jurnal Penelitian dan Evaluasi Pendidikan*, *20*(2), 198–207. https://doi.org/10.21831/pep.v20i2.7444

Dasaprawira, M. N. (2019). Developing mathematics questions of PISA type using Bangka context. *Journal on Mathematics Education*, *10*(2), 303–314. https://doi.org/10.22342/jme.10.2.5366.303-314

Dutt, A., Tan, M., Alagumalai, S., & Nair, R. (2019). Development and validation of the ability in behavior assessment and interventions for teachers using Delphi technique and Rasch analysis. *Journal of Autism and Developmental Disorders*, *49*(5), 1976–1987. https://doi.org/10.1007/s10803-019-03887-4

Efriani, A., & Putri, R. I. I. (2019). Sailing context in PISA-like mathematics problems. *Journal on Mathematics Education*, *10*(2), 265–276. https://doi.org/10.22342/jme.10.2.5245.265-276

Ekowati, D. W., Astuti, Y. P., Utami, I. W. P., Mukhlishina, I., & Suwandayani, B. I. (2019). Literasi numerasi di SD Muhammadiyah. *ELSE (Elementary School Education Journal): Jurnal Pendidikan dan Pembelajaran Sekolah Dasar*, *3*(1), 93–103. https://doi.org/10.30651/else.v3i1.2541

Erfan, M., Maulyda, M. A., Ermiana, I., Hidayati, V. R., & Widodo, A. (2020). Validity and reliability of cognitive tests study and development of elementary curriculum using Rasch model. *Psychology, Evaluation, and Technology in Educational Research*, *3*(1), 26–33.

https://doi.org/10.33292/petier.v3i1.51

Fadhilah, T. F. (2021). *The profile of students' grammatical ability and error shown in their writing of reading test questions*. Universitas Negeri Padang.

Govorova, E., Benítez, I., & Muñiz, J. (2020). How schools affect student well-being: a cross-cultural approach in 35 OECD countries. *Frontiers in Psychology*, *11*, Article 431. https://doi.org/10.3389/fpsyg.2020.00431

Hasanah, H. (2017). Efektivitas soal-soal matematika tipe PISA menggunakan konteks budaya Sumatera utara untuk mendeskripsikan kemampuan penalaran dan komunikasi matematis siswa SMP kota Medan [The effectiveness of PISA type math problems using the cultural context of North Sumatra to describe the mathematical reasoning and communication skills of Medan City Middle School students]. *AXIOM: Jurnal Pendidikan dan Matematika*, *6*(1), 1-12.

Hasibuan, S. A., & Fauzi, K. M. S. M. A. (2019). Development of PISA mathematical problem model on the content of change and relationship to measure students mathematical problem-solving ability. *International Electronic Journal of Mathematics Education*, *15*(2), Article em0570. https://doi.org/10.29333/iejme/6274

Huda, A., Azhar, N., Almasri, A., Hartanto, S., & Anshari, K. (2020). Practicality and effectiveness test of graphic design learning media based on android. *International Journal of Interactive Mobile Technologies, 14*(4), 192-203. https://doi.org/10.3991/ijim.v14i04.12737

Johar, R. (2012). Domain soal PISA untuk literasi matematika [PISA question domain for mathematical literacy]. *Jurnal Peluang*, *1*(1), 30-41. http://jurnal.unsyiah.ac.id/peluang/article/view/1296

Jonassen, D. H. (2010). *Learning to solve problems: A handbook for designing problem-solving learning environments*. Routledge. https://doi.org/10.4324/9780203847527

Koneputugodage, C. H., Healy, R., Lamont, S., Mallett, I., Brown, M., Walters, M., Attanayake, U., Zhang, L., Dean, R. T., & Hunter, A. (2019). Computer assisted composition in continuous time. *ArXiv Preprint ArXiv:1909.05030*.

Kusumaningsih, W., Buchori, A., & Cahyono, H. G. (2020). Uno stacko based on realistic mathematics: A developing learning media of trigonometry. *Journal of Physics: Conference Series*, *1663*(1), Article 12044. https://doi.org/10.1088/1742-6596/1663/1/012044

Latinis, D. K., & Stark, K. (2005). Cave use variability in central Maluku, Eastern Indonesia. *Asian Perspectives*, *44*(1), 119–136. https://doi.org/10.1353/asi.2005.0009

Lewis, J. R., & Sauro, J. (2018). Item benchmarks for the system usability scale. *Journal of Usability Studies*, *13*(3), 158-167. https://uxpajournal.org/item-benchmarks-system-usability-scale-sus/

Lonthor, A. (2018). Islam kultural dan interprestasi normatif pemikiran hukum islam di Maluku [Cultural Islam and normative interpretation of Islamic legal thought in Maluku]. *FIKRATUNA: Jurnal Penelitian Sosial Keagamaan*, *7*(2), 52-69. https://jurnal.iainambon.ac.id/index.php/FT/article/view/344

Lu, Z., Vincent, J. I., & MacDermid, J. C. (2021). Evaluation of the structural validity of the work instability scale using the Rasch model. *Archives of Rehabilitation Research and Clinical Translation*, *3*(1), Article 100103. https://doi.org/10.1016/j.arrct.2021.100103

Machromah, I. U., & Sari, C. K. (2021). Batik context as PISA-like problem to assess students' mathematical literacy. *Journal of Physics: Conference Series*, *1776*(1), Article 12013. https://doi.org/10.1088/1742-6596/1776/1/012013

Malatuny, Y. G., & Ritiauw, S. P. (2018). Eksistensi pela gandong sebagai civic culture dalam menjaga harmonisasi masyarakat di Maluku [The existence of pela gandong as a civic culture in maintaining the harmonization of society in Maluku]. *Sosio Didaktika: Social*

Science Education Journal, *5*(2), 35–46.

Martins, A. I., Rosa, A. F., Queirós, A., Silva, A., & Rocha, N. P. (2015). European Portuguese validation of the system usability scale (SUS). *Procedia Computer Science*, *67*, 293–300. https://doi.org/10.1016/j.procs.2015.09.273

Mayari, R. P. (2020). *Pengembangan soal matematika model pisa menggunakan konteks Palembang [Development of pisa model math problems using the Palembang context]*. Universitas Islam Negeri (UIN) Raden Fatah Palembang.

Megawati, L. A., & Sutarto, H. (2021). Analysis numeracy literacy skills in terms of standardized math problem on a minimum competency assessment. *Unnes Journal of Mathematics Education*, *10*(2), 155-165. https://journal.unnes.ac.id/sju/index.php/ujme/article/view/49540

Mulyatna, F., Imswatama, A., & Rahmawati, N. D. (2021). Design Ethnic-Math HOTS: Mathematics higher order thinking skill questions based on culture and local wisdom. *Malikussaleh Journal of Mathematics Learning (MJML)*, *4*(1), 48–51. https://doi.org/10.29103/mjml.v4i1.3059

Mungmachon, M. R. (2012). Knowledge and local wisdom: Community treasure. *International Journal of Humanities and Social Science*, *2*(13), 174–181.

Murtiyasa, B., Rejeki, S., & Setyaningsih, R. (2018). PISA-like problems using Indonesian contexts. *Journal of Physics: Conference Series*, *1040*(1), Article 12032. https://doi.org/10.1088/1742-6596/1040/1/012032

Mutmainah, S. (2017). Studi kelayakan pengembangan aplikasi bank soal dalam rumah belajar [Feasibility study for developing a question bank application in a study house]. *Jurnal Teknodik*, *21*(1), 14-26. https://doi.org/10.32550/teknodik.v21i1.263

Nasution, R. S., Fauzi, K. M. A., & Syahputra, E. (2019). Developing mathematics problem based on PISA level of space and shape content to measure student's mathematics problem solving ability. *American Journal of Educational Research*, *7*(10), 660–669. https://doi.org/10.12691/education-7-10-1

Nissen, J. M., Talbot, R. M., Thompson, A. N., & Van Dusen, B. (2018). Comparison of normalized gain and Cohen's d for analyzing gains on concept inventories. *Physical Review Physics Education Research*, *14*(1), Article 10115. https://doi.org/10.1103/PhysRevPhysEducRes.14.010115

Nizar, H., & Putri, R. I. I. (2018). Developing PISA-like mathematics problem using the 2018 Asian Games football and table tennis context. *Journal on Mathematics Education*, *9*(2), 183–194. https://doi.org/10.22342/jme.9.2.5246.183-194

Nuraeni, Z., Indaryanti, I., & Sukmaningthias, N. (2021). Pengembangan perangkat pembelajaran bercirikan CTL berbantuan GeoGebra menggunakan model flipped learning [Development of learning tools characterized by GeoGebra-assisted CTL using the flipped learning model]. *Jurnal Elemen*, *7*(1), 56–67. https://doi.org/10.29408/jel.v7i1.2723

Nurpala, D., & Munawaroh, M. (2020). Perancangan sistem aplikasi bank soal pada ujian online berbasis WEB (Studi kasus: SMA Negeri 1 Cibeber) [Design of a question bank application system for WEB-based online exams (Case study: SMA Negeri 1 Cibeber)]. *JOAIIA: Journal of Artificial Intelligence and Innovative Applications*, *1*(2), 51–55.

Nusantara, D. S., & Putri, R. I. I. (2020). Designing PISA-like mathematics problem in covid-19 pandemic (PISAComat). *Journal of Physics: Conference Series*, *1657*(1), Article 12057. https://doi.org/10.1088/1742-6596/1657/1/012057

Nusantara, D. S., Zulkardi, Z., & Putri, R. I. I. (2020). Designing PISA-like mathematics problem relating change and relationship using physical distancing context. *Journal of Physics: Conference Series*, *1663*(1), Article 12004. https://doi.org/10.1088/1742-6596/1663/1/012004

Nusantara, D. S. (2021). Designing PISA-like mathematics task using a COVID-19 context (PISAComat). *Journal on Mathematics Education*, *12*(2), 349–364. https://doi.org/10.22342/jme.12.2.13181.349-364

Petersen, L. A., Litteck, K., & Rohenroth, D. (2020). *NEPS technical report for mathematics: scaling results of starting cohort 3 for grade 12*. Bamberg: Leibniz Institute for Educational Trajectories.

Pratiwi, A. A., Effendi, M. M., & Ummah, S. K. (2020). Pengembangan instrumen evaluasi pembelajaran matematika tipe PISA berkarakteristik kebudayaan lokal [Development of PISA-type mathematics learning evaluation instruments with local cultural characteristics]. *Jurnal Pendidikan Matematika RAFA*, *6*(1), 39–53.

Puspita, S., & Ahda, Y. (2020). Validity of research based bioremediation text books. *International Conference on Biology, Sciences and Education (ICoBioSE 2019)*, 116–121. https://doi.org/10.2991/absr.k.200807.027

Putra, Y. Y., & Vebrian, R. (2019). *Literasi matematika (mathematical literacy) soal matematika model PISA menggunakan konteks Bangka Belitung [Mathematical literacy for PISA models using the Bangka Belitung context]*. Deepublish.

Retnowati, T. H., & UN, J. P. S. R. F. B. S. (2010). Membangun karakter siswa melalui pembelajaran batik di sekolah [Building student character through learning batik at school]. *Makalah Seminar Nasional dalam Rangka Dies-46 UNY*.

Robitzsch, A. (2021). *On the treatment of missing item responses in educational large-scale assessment data: The case of PISA 2018 mathematics*. https://doi.org/10.20944/preprints202110.0107.v1

Rosli, R., Abdullah, M., Siregar, N. C., Hamid, N. S. A., Abdullah, S., Beng, G. K., Halim, L., Daud, N. M., Bahari, S. A., & Majid, R. A. (2020). Student awareness of space science: Rasch model analysis for validity and reliability. *World Journal of Education*, *10*(3), 170–177. https://doi.org/10.5430/wje.v10n3p170

Saefurohman, S., Maryanti, R., Azizah, N. N., Al Husaeni, D. F., Wulandary, V., & Irawan, A. R. (2021). Efforts to increasing numeracy literacy of elementary school students through quiziz learning media. *ASEAN Journal of Science and Engineering Education*, *1*(3), 167–174.

Samo, D. D. (2018). Culture-based contextual learning to increase problem-solving ability of first year university student. *Journal on Mathematics Education*, *9*(1), 81–94. https://doi.org/10.22342/jme.9.1.4125.81-94

Santoso, H. B., Schrepp, M., Isal, R., Utomo, A. Y., & Priyogi, B. (2016). Measuring user experience of the student-centered e-learning environment. *Journal of Educators Online*, *13*(1), 58–79. https://doi.org/10.9743/JEO.2016.1.5

Schleicher, A. (2019). PISA 2018: Insights and interpretations. *OECD Publishing*.

Schrepp, M. (2020). On the usage of Cronbach's alpha to measure reliability of UX scales. *Journal of Usability Studies*, *15*(4), 247-248. https://uxpajournal.org/cronbachs-alpha-reliability-ux-scales/

Schrepp, M., Hinderks, A., & Thomaschewski, J. (2017). Construction of a benchmark for the user experience questionnaire (UEQ). *Int. J. Interact. Multim. Artif. Intell.*, *4*(4), 40–44. https://doi.org/10.9781/ijimai.2017.445

Stacey, K. (2015). The international assessment of mathematical literacy: PISA 2012 framework and items. *Selected Regular Lectures from the 12th International Congress on Mathematical Education*, 771–790. https://doi.org/10.1007/978-3-319-17187-6_43

Stacey, K., & Turner, R. (2015). Assessing mathematical literacy: The PISA experience. In *Assessing Mathematical Literacy: The PISA Experience*. https://doi.org/10.1007/978-3-319-10121-7

Sumintono, B., & Widhiarso, W. (2015). *Aplikasi pemodelan Rasch pada assessment*

pendidikan [Rasch modeling application in educational assessment]. Trim komunikata.

Supiyati, S., & Hanum, F. (2019). Ethnomathematics in Sasaknese architecture. *Journal on Mathematics Education*, *10*(1), 47–58. https://doi.org/10.22342/jme.10.1.5383.47-58

Suradi, A. (2018). Pendidikan berbasis multikultural dalam pelestarian kebudayaan lokal nusantara di era globalisasi [Multicultural-based education in the preservation of the local culture of the archipelago in the era of globalization]. *Wahana Akademika: Jurnal Studi Islam Dan Sosial*, *5*(1), 111–130. https://doi.org/10.21580/wa.v5i1.2566

Tai, W. C., & Lin, S. W. (2015). Relationship between problem-solving style and mathematical literacy. *Educational Research and Reviews*, *10*(11), 1480–1486. https://doi.org/10.5897/ERR2015.2266

Tan, C. Y. (2020). What PISA and ASPIRES studies tell us about the nuanced influence of cultural capital on student learning: Construct complexity, student outcomes and contexts. *British Educational Research Journal*, *46*(6), 1338–1356. https://doi.org/10.1002/berj.3635

Uge, S., Neolaka, A., & Yasin, M. (2019). Development of social studies learning model based on local wisdom in improving students' knowledge and social attitude. *International Journal of Instruction*, *12*(3), 375–388. https://doi.org/10.29333/iji.2019.12323a

Van Anh, H. H. (2021). Analisis kalimat efektif pada soal cerita bidang studi matematika di SMK [Effective sentence analysis on story problems in the field of mathematics at SMK]. *Anafora: Jurnal Penelitian Mahasiswa Pendidikan Bahasa dan Sastra Indonesia*, *1*(1), 1-11. https://publication.uniku.ac.id/index.php/anafora/article/view/5

van den Akker, J., Branch, R. M., Gustafson, K., Nieveen, N., & Plomp, T. (2012). *Design approaches and tools in education and training*. Springer Science & Business Media.

Van den Heuvel-Panhuizen, M., & Drijvers, P. (2020). Realistic mathematics education. *Encyclopedia of Mathematics Education*, 713–717. https://doi.org/10.1007/978-3-030-15789-0_170

Wakano, A. (2019). Nilai-nilai pendidikan multikultural dalam kearifan lokal masyarakat Maluku [The values of multicultural education in the local wisdom of the Maluku people]. *Al-Iltizam: Jurnal Pendidikan Agama Islam*, *4*(2), 26–43. https://doi.org/10.33477/alt.v4i2.1006

Widana, I. W. (2014). Pengembangan bank soal [Question bank development]. *Emasains*, *3*(2), 186–197.

Widodo, S. A. (2018). Selection of learning media mathematics for junior school students. *Turkish Online Journal of Educational Technology-TOJET*, *17*(1), 154–160.

Widoyoko, E. P. (2009). The evaluation of the learning program. *Student Library*.

Winarti, W., Hairida, H., & Lestari, I. (2021). Deskripsi kemampuan guru membuat soal berdasarkan pada kurikulum 2013 di sekolah menengah atas Kabupaten Landak [Description of the teacher's ability to make questions based on the 2013 curriculum in high school in Landak Regency]. *Jurnal Ilmiah Wahana Pendidikan*, *7*(2), 108–115.

Yansen, D., & Putri, R. I. I. (2019). Developing PISA-like mathematics problems on uncertainty and data using Asian Games football context. *Journal on Mathematics Education*, *10*(1), 37–46. https://doi.org/10.22342/jme.10.1.5249.37-46

Zulfah, Z., & Insani, S. U. (2020). Pengembangan soal matematika berbasis kearifan lokal dan daya tarik wisata Riau pada tahap preliminary research [Development of math problems based on local wisdom and Riau's tourist attraction in the preliminary research stage]. *Jurnal Cendekia: Jurnal Pendidikan Matematika*, *4*(2), 797–799. https://doi.org/10.31004/cendekia.v4i2.311