

IMPLEMENTATION OF AN AUTHENTIC ASSESSMENT MODEL TO IMPROVE THE QUALITY OF LEARNING ASSESSMENT

Rini Agustiningsih¹ and Dwi Rahdiyanta²

¹Postgraduate School, Universitas Negeri Yogyakarta, Indonesia,

²Department of Mechanical Engineering Education, Universitas Negeri Yogyakarta, Indonesia

Email: rini.agustiningsih2016@student.uny.ac.id

ABSTRACT

An assessment instrument is one of the most important factors to guarantee objectivity, fairness, and accuracy in assessing the level of students' performances in the learning process. This study aims to produce an authentic assessment model in problem solving learning. The objectives of this study are to examine: (1) the feasibility of the developed authentic assessment model in the aspects of practicality, effectiveness, and feasibility of assessment; and (2) the accuracy of the assessment results from the authentic assessment model. This study was a research and development that adopts the Analyze, Design, Develop, Implement, and Evaluate (ADDIE) development model. The research subjects consisted of eight teachers and 44 students selected with a purposive sampling technique. This study was conducted in 4 Vocational High Schools in Yogyakarta. The research instrument validation was examined by expert judgment. The data were collected using questionnaires and observation sheets. Its data were analysed with descriptive statistical techniques, inter-rater reliability, inter-item correlation, and intraclass correlation. The results of the study show that (1) the developed authentic assessment model was feasible to be implemented since it fulfilled the requirements of practicality, effectiveness and implementation; and (2) the level of accuracy of the assessment results from the authentic assessment model has a very good criteria with κ mean of 0.978.

Keywords: authentic assessment, model development, problem solving

INTRODUCTION

Problem solving is very crucial for the growth and development of students to be creative and innovative individuals who have competitiveness in a wider scope. Problem solving learning is one of the High Order Thinking Skills (HOTS) learning, which is popular as 21st century learning. Thinking and solving problems cannot be separated in the learning process. The ability to solve problems is the main key in understanding naturally while the ability to solve problems is obtained by direct practice and approaches used to understand problems [1].

Problem solving learning is a learning model that applies scientific approach. Therefore according to the 2013 curriculum, the problem-based learning model is one of

the learning models that must be mastered by teachers. As explained in the 2013 curriculum training materials, according to ministerial regulation of Education and Culture Minister No. 103 of 2014, learning with a scientific approach consists of five steps of learning activities namely observing, asking questions, gathering information or trying (experimenting), reasoning or associating, and communicating which can be continued with creating.

This problem-based learning method is often referred as experimental methods, reflective thinking methods, or scientific methods [2]. Maier stated that the problem solving method is a way of presenting lessons by encouraging students to search for and to solve a problem in achieving teaching goals [3]. Therefore the teacher is

required to be able to manage the class well and can use various teaching styles in accordance with the presented materials [4]. Mulyatiningsih explained that problem solving learning methods are very potential to train students to think creatively in dealing with various personal problems and group problems to be solved alone or together [5].

Based on the various opinions mentioned above, it can be concluded that problem solving learning is learning that uses real-life (authentic) open-ended problems to be solved by students to develop thinking skills, problem solving skills, social skills for independent learning and building or acquiring new knowledge.

Ideal learning must be able to provide opportunities for students to build their own knowledge through interaction with parents, siblings, teachers, friends, and the environment. Learning activities must be in accordance with the principles that include: (1) fun, (2) oriented to students' development, (3) oriented to students' needs, (4) active learning, (5) oriented to the development of character values, (6) oriented to life skills development, (7) conducive environment, (8) oriented to democratic learning, and (9) the use of media and learning resources.

Assessment instrument is one of the most important factors to guarantee objectivity and fairness in assessing students' ability levels. Therefore, accurate instruments are required in each assessment activity. Payne [6] stated that the assessment instrument is considered to be good if it meets several criteria, namely (1) relevant in data collection in accordance with the purpose of the assessment, (2) there is balance of proportions of measurement multidimensionally, (3) it can be used efficiently, (4) objectivity in scoring, (5)

consistency of measurement, (6) describing honesty (not biased), (7) specifically measuring aspects being studied, (8) describing the level of difficulty addressed by the skills, knowledge, and abilities of assessed individuals, (9) can distinguish individual ability levels, and (10) do not measure individuals' speed. Whereas according to Johnson et al. assessment is a process carried out to measure the achievement of students' competencies on an ongoing basis in the learning process, to monitor progress and improve students' learning outcomes, and plan further learning activities [7].

Authentic assessment is a process of gathering significant information on students' learning outcomes in the aspects of attitudes, knowledge, and skills by the teacher by using various measurement techniques based on actual facts. It is further stated that authentic assessment is a form of assessment, in which students are asked to perform real-world assignments that demonstrate the application of the knowledge and skills they have learned [8]–[10]. Authentic assessment is a form of evaluation to describe the real knowledge and skills of the students as a result of learning activities. Authentic assessment can be conducted using several techniques, such as product assessment, project assessment, portfolio assessment, self-assessment, peer assessment, written examinations, journals, and observations [11].

Authentic assessment has strong relevance to scientific approaches in learning that are in accordance with the 2013 curriculum. Authentic assessments are able to describe improvement in student learning outcomes through a scientific approach, namely observing, asking, gathering information, associating, and communicating. Authentic assessment is

also be able to show students' abilities, provide feedback and direction for their growth and development [12]. Authentic assessment contributes to the development of self-concept and achievement motivation and attitudes, and students' self-confidence [13], [14]. Authentic assessment is able to develop student motivation and achievement in learning [15]. Furthermore, authentic assessment integrates continuous learning activities, and can enhance abilities and critical and creative thinking [16]. According to Gulikers et al., authentic assessment requires students to use and demonstrate competence, or a combination of knowledge, skills and attitudes that can be applied in real life [17]. In authentic assessments, students are asked to do assignments in real-world regulatory questions as applications of knowledge and skills acquired by Mueller [8]. Authentic assessment is focused on giving assignments to students as actualization of their knowledge and skills because they will be applied in the real world. This is in accordance with the opinion of Burton who stated that authentic assessment is directly related to real work practices [18]. According to Atac, there are several reasons for using authentic assessment in learning: (1) it can produce direct measurements, (2) it can capture the meaning of actual learning, and (3) it can integrate assessment and teaching [19]. Authentic assessment is not only seen from the results of the task (product), but also from the quality of the educational process that involves students to be more human [20].

Based on the findings in the field, the assessment instruments used to assess the development of students at present are not structured and detailed, they are not operational, not practical to use, and have

never been standardized to see their validity and reliability. As a result, it was found that the results of the assessment by two teachers were incorrect, or the results were different. Therefore it is very important to develop an authentic assessment model in problem solving learning so that the assessment of student learning outcomes carried out by teachers can be accurate. The authentic assessment model in problem solving learning must contain the form of problem solving activities, time allocation, observation sheet evaluation, description of problem solving abilities, and criteria in problem solving.

METHOD

This study is categorized as research and development that adopts the Analyze, Design, Develop, Implement, and Evaluate (ADDIE) development model [21]. Authentic assessment developed in problem solving learning consists of components of students' abilities, namely: (1) understanding problems, (2) planning solutions, (3) implementing solutions, and (4) evaluating solutions. The research subjects consisted of 8 teachers and 44 students selected with purposive sampling technique. It was conducted in four Vocational Schools in Yogyakarta. Experts in early childhood learning and assessment are invited to provide expert judgement for the validation.

Data collection techniques used questionnaires and observation sheets. Questionnaires were used for data collection in the aspects of practicality and effectiveness while the observation sheet is used for data collection in the aspects of the implementation of the assessment process. The outline of the research instrument is listed in Table 1.

Table 1. Instrument Grid for Assessing Practicality, Effectiveness, and Implementation of the Assessment Process

No.	Aspect	Indicator
A.	Practicality of the Authentic Assessment Model	
1.	Application	<ol style="list-style-type: none"> 1. Ease in using the valuation model 2. The simplicity of the assessment model format 3. Clarity of stages of problem solving 4. Clarity of authentic assessment characteristics 5. Clarity of score and criteria in data tabulation
2.	Language	<ol style="list-style-type: none"> 1. Standard and communicative language 2. Use of terms that are easy to understand
B.	Effectiveness of the Authentic Assessment Model	
1.	Stages of Problem Solving	<ol style="list-style-type: none"> 1. Demands for problem solving steps 2. 2. Conformity with Basic Competencies (KD) 3. 3. Conformity with problem solving cases / activities
2.	Characteristics of Authentic Assessment	<ol style="list-style-type: none"> 1. Representation of authentic assessment characteristics with cases / problems 2. Conformity of authentic assessment characteristics 3. Representation of authentic assessment characteristics to describe students' ability to solve cases / activities
3.	Case for Problem Solving	<ol style="list-style-type: none"> 1. Learning activities are displayed in the form of cases / problems 2. Students understand learning activities from the cases / problems given 3. Completion of cases / activities using problem solving approaches 4. Cases / problems that are displayed foster student curiosity 5. Suitability of cases / problems with everyday life 6. Suitability of cases / problems with Basic Competence (KD)
C.	Implementation of the Assessment Process Using the Authentic Assessment Model	
1.	Preparation	<ol style="list-style-type: none"> 1. Preparing tools, materials, and places for learning activities 2. Conditioning students to be ready to start problem solving activities 3. Prepare authentic assessment models
2.	Core activities	<ol style="list-style-type: none"> 1. Determine the sequence of problem solving learning 2. Determine the stages of problem solving learning 3. Organize the activities of student activities in problem solving learning in accordance with the topic taken. 4. Describe the ability of problem solving steps in each student in problem solving learning
3.	Closing	<ol style="list-style-type: none"> 1. Make the problem solving ability criteria for each student in accordance with the Guidebook. 2. Closing problem solving learning activities

The data analysis used to determine the feasibility of the authentic assessment model developed is the analysis of reliability inter-rater, inter-item correlation, and intraclass correlation. Whereas to determine the level of practicality, the effectiveness and implementation of the assessment process used criteria adapted from Azwar [22]. These are presented in the following Table 2.

Table 2. Criteria for Practicality, Effectiveness, and the Implementation of Authentic Assessment Model

No	Interval	Criteria
1	$X > X_i + 1.5 S_{bi}$	Very Good
2	$X_i + S_{bi} < X \leq X_i + 1.5 S_{bi}$	Good
3	$X_i - 0.5 S_{bi} < X \leq X_i + S_{bi}$	Moderate
4	$X_i - 1.5 S_{bi} < X \leq X_i - 0.5 S_{bi}$	Poor
5	$X \leq X_i - 1.5 S_{bi}$	Very Poor

X = Actual total score

X_i = ideal score = $1/2$ (ideal maximum score + ideal minimum score)

S_{bi} = Standard deviation of ideal score = $1/6$ (ideal maximum score - ideal minimum score)

RESULTS AND DISCUSSION

To find out the feasibility of authentic assessment models developed in terms of practicality, effectiveness, and the implementation of authentic assessment processes, a model implementation was carried out. Implementation of authentic assessment models is carried out after obtaining validation from the assessment expert. The implementation of authentic

assessment models was carried out in four classes with a total of 8 people and 44 students.

The following is a description of teacher assessment data on authentic assessment models that are developed both in terms of practicality, effectiveness, and implementation of the assessment process. Data from teacher assessment results on the practical aspects of the authentic assessment model can be seen in Table 3.

Table 3. Results of Practical Assessment of the Authentic Assessment Model

Resp.	Assessment Item Number							Score	Criteria
	1	2	3	4	5	6	7		
Teacher 1	4	4	4	4	3	3	3	25	Very Good
Teacher 2	3	4	4	3	3	3	3	23	Very Good
Teacher 3	4	4	4	4	4	3	3	26	Very Good
Teacher 4	3	3	4	3	3	3	4	23	Very Good
Teacher 5	3	4	3	3	4	3	3	23	Very Good
Teacher 6	3	3	4	3	4	2	3	22	Very Good
Teacher 7	3	3	3	3	3	3	3	21	Very Good
Teacher 8	4	3	3	3	3	4	3	23	Very Good
Total Score of Research Result (ΣX)									116
Average Total Score of Research Results (X)									29
Ideal Total Score (ΣXi) = 8 x 7 x 4									224
Ideal Average (Xi)									17.5
Standard Deviation Ideal (Sbi)									3.5

Table 3 shows that all of the 8 teachers gave very good criteria. This means that in the aspect of practicality, the teachers considered the developed authentic assessment model as very good. Data description of the authentic assessment instrument model can be seen in Table 4. Based on the teacher assessment data, it can be seen that there were five teachers who gave very good ratings, and three teachers gave good ratings thus it means that in general the teachers stated that the effectiveness of the authentic assessment model developed is good.

Data from teacher assessment results in the aspects of the implementation of the assessment process from the authentic

assessment model developed is presented in Table 5.

Based on the data from the teacher's assessment, it can be seen that there were five teachers who gave good ratings, and three teachers gave quite good ratings thus it means that in general the teachers state that the level of implementation in the assessment process using this authentic developed evaluation model is already good. Data from the assessment of students' ability in problem solving learning using authentic assessment models can be seen in Table 6. It shows that students who have the ability with good and very good criteria are 81.6%. Figure 5 shows the rod frequency diagram of student ability.

Table 4. Summary of the Authentic Assessment Instrument Model

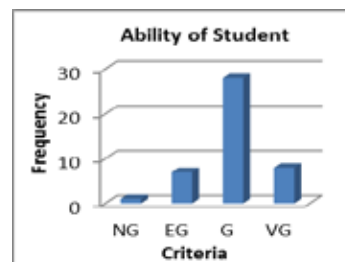
Resp.	Assessment Item Number												Score	Criteria
	1	2	3	4	5	6	7	8	9	10	11	12		
Teacher 1	3	3	3	4	3	3	3	3	3	3	3	4	38	Good
Teacher 2	4	3	3	4	4	3	4	4	4	3	4	4	44	Very Good
Teacher 3	4	3	3	3	3	3	3	3	3	3	3	3	37	Good
Teacher 4	4	3	3	3	3	3	3	3	3	2	3	3	36	Good
Teacher 5	3	4	4	3	4	3	3	4	4	3	3	3	41	Very Good
Teacher 6	3	3	3	4	3	3	3	4	3	4	3	3	39	Very Good
Teacher 7	4	3	4	4	3	4	2	3	3	3	3	4	40	Very Good
Teacher 8	4	4	4	4	3	3	2	4	3	4	2	3	40	Very Good
Total Score of Research Result (ΣX)													315	
Average Total Score of Research Results (X)													39.5	
Ideal Total Score (ΣXi) = 8 x 7 x 4													384	
Ideal Average (Xi)													30	
Standard Deviation Ideal (Sbi)													6.00	

Table 5. Summary of the Results of Evaluation of the Implementation of the Model

Resp.	Assessment Item Number									Score	Criteria	
	1	2	3	4	5	6	7	8	9			
Teacher 1	3	4	3	3	2	2	4	3	3	25	Moderate	
Teacher 2	3	4	4	3	2	4	3	3	3	29	Good	
Teacher 3	3	3	3	3	2	2	4	3	3	25	Moderate	
Teacher 4	3	3	3	3	2	3	3	3	4	26	Moderate	
Teacher 5	3	3	3	3	3	4	4	3	4	31	Good	
Teacher 6	3	2	3	4	3	4	3	3	3	28	Good	
Teacher 7	3	2	3	4	3	3	3	4	3	27	Good	
Teacher 8	3	3	3	4	2	2	3	3	3	27	Good	
Total Score of Research Result (ΣX)											218	
Average Total Score of Research Results (X)											27.25	
Ideal Total Score (ΣXi) = 8 x 7 x 4											288	
Ideal Average (Xi)											40.5	
Standard Deviation Ideal (Sbi)											3.5	

Table 6. Frequency of Student Ability in Learning Solution to Problem

Criteria	Freq	Percentage	Valid Percentage	Cum. Percentage
1	1	2.3	2.3	2.3
2	7	15.9	15.9	18.2
3	28	63.6	63.6	81.8
4	8	18.2	18.2	100.0
Total	44	100.0	100.0	



Where VG, G, EG and NG are Very Good, Good, Moderate and Not Good respectively.

Figure 5. Rod Frequency Diagram of Student Ability

The reliability of the authentic assessment model developed was analyzed based on the Cornbach's Alpha (α) value. The results of the reliability analysis of the authentic assessment model can be seen in Table 7.

Table 7. Reliability Analysis of the Authentic Assessment Model

Cronbach's Alpha	Cronbach's Alpha Based on Standardized Items	N of Items
.978	.990	8

Based on the reliability statistics table, it is known that the Cornbach's Alpha value of (α) = 0.978. This shows that the instruments developed in the authentic valuation model have a 'very reliable' level of reliability. This is in accordance with the criteria for the

reliability level of Cornbach's Alpha according to Hair et al. [23] presented in Table 7.

Table 7. Level of Reliability

Score Cornbach's Alpha	Reliability level
0.0 – 0.20	Very Poor
> 0.20 – 0.40	Poor
> 0.40 – 0.60	Moderate
> 0.60 – 0.80	Good
> 0.80 – 1.00	Very Good

Analysis of the level of agreement between teachers in the assessment of the model developed, using inter-item correlation and intraclass correlation coefficient. The results of the inter-item correlation analysis to determine the level of agreement among teachers can be seen in Table 8.

Table 8. Inter-Item Correlation Matrix on implementing the Assessment Model Kappa Coefficient (κ)

	Teacher 1	Teacher 2	Teacher 3	Teacher 4	Teacher 5	Teacher 6	Teacher 7	Teacher 8
Teacher 1	1.000	.866	1.000	1.000	.866	.866	.866	.866
Teacher 2	.866	1.000	.866	.866	1.000	1.000	1.000	1.000
Teacher 3	1.000	.866	1.000	1.000	.866	.866	.866	.866
Teacher 4	1.000	.866	1.000	1.000	.866	.866	.866	.866
Teacher 5	.866	1.000	.866	.866	1.000	1.000	1.000	1.000
Teacher 6	.866	1.000	.866	.866	1.000	1.000	1.000	1.000
Teacher 7	.866	1.000	.866	.866	1.000	1.000	1.000	1.000
Teacher 8	.866	1.000	.866	.866	1.000	1.000	1.000	1.000

Based on the Inter-Item Correlation Matrix table on the implementation of the valuation model, it can be seen that the value of the Kappa (κ) coefficient between all teachers is > 0.81. Thus, it means that among these teachers has a very strong level of agreement on the authentic assessment model that has been developed. This is in

accordance with the agreement level criteria or the Kappa coefficient value (κ) adapted from Landis et al. [24] in Table 9. To find out the level of agreement both individually and between teachers, an intraclass correlation coefficient analysis was carried out. The results of interclass correlation coefficient analysis can be seen in Table 10.

Table 9. Level of Agreement Strengths Kappa Coefficient Value

Kappa Value	(Strength of agreement)
< 0.20	(Poor)
0.21 – 0.40	(Fair)
0.41 – 0.60	(Moderate)
0.61 – 0.80	(Good)
0.81 – 1.00	(Very Good)

Tabel 10. Intraclass Correlation Coefficient

	Intraclass Correlation ^b	95% Confidence Interval		F Test with True Value 0			
		Lower Bound	Upper Bound	Value	df1	df2	Sig
Single Measures	.847 ^a	.510	.996	45.267	2	14	.000
Average Measures	.978 ^c	.893	.999	45.267	2	14	.000

Two-way mixed effects model where people effects are random and measures effects are fixed.

a. The estimator is the same, whether the interaction effect is present or not.

b. Type C intraclass correlation coefficients using a consistency definition. The between-measure variance is excluded from the denominator variance.

c. This estimate is computed assuming the interaction effect is absent, because it is not estimable otherwise.

Individual Kappa (κ) Coefficient Value Calculation of the assessment of practicality, effectiveness, and implementation by each teacher on the authentic assessment model in developing problem solving learning has a Kappa (κ) Coefficient of 0.847. It can be said that each individual has a very strong agreement level. The calculation of the practicality, effectiveness, and implementation between teachers on the authentic assessment model in developing problem solving learning has a Kappa (κ) Coefficient of 0.978, it can be said that all teachers have a very strong agreement level.

Individual Kappa (κ) Coefficient Value Calculation of the assessment of practicality, effectiveness, and implementation by each teacher on the authentic assessment model in developing problem solving learning has a Kappa (κ) Coefficient of 0.847, it can be said that each individual has a very strong agreement level.

The calculation of the practicality, effectiveness, and implementation between teachers on the authentic assessment model in developing problem solving learning has a Kappa (κ) Coefficient of 0.978, it can be said that all teachers have a 'very strong' agreement level.

Individual Kappa (κ) Coefficient Value Assessment of practicality,

effectiveness, and implementation of each teacher on the authentic assessment model in developing a problem solving learning has a Kappa (κ) Coefficient = 0.847, it can be said that each individual has a very strong agreement level.

The calculation of the practicality, effectiveness, and implementation of teachers on the authentic assessment model in developing a problem solving learning has a Kappa (κ) Coefficient = 0.978, it can be said that all teachers have a very strong agreement level agreement level.

The development of authentic assessment models in problem solving learning refers to the learning of High Order Thinking Skill (HOTS), one of which is problem solving learning. Problem solving learning has many advantages for students, namely, sharpening thinking skills [25], improving skills in solving social problems [26] [27] [28], improving math skills [29], [31], improve interpersonal skills [32], [33], and have an impact on student learning achievement [34].

The authentic assessment model is often used by teachers for learning mathematics with various media [35]. Policy assessments in learning and curriculum are divided into three, namely assessment approaches to development, assessment approaches to development and learning,

and assessment approaches to learning [36]. In this study using an assessment approach to learning.

The product quality of a development is based on 3 criteria. The three criteria according to Nieveen [37] are validity, practically, effectiveness. A product of development is called valid if it reflects the soul of knowledge (state-of-the-art knowledge), or referred to as content validity. Meanwhile the product components must be consistent with each other (construct validity). Furthermore, a product is categorized to be practical if the teacher and other experts state that the product can be used easily by the teacher and students as users. Then a product is categorized to be effective if it gives results in accordance with the objectives set by the developer.

Nieveen [37] stated that a product of development is called valid if the components of the material are based on state-of-the-art knowledge (content validity) and all components should be consistently linked to each other (construct validity). If the product meets these requirements, hence, it is considered to be valid. This means that product material must be based on the area of knowledge (content validity) and all components must be consistently related to each other (construct validity). If the product meets these requirements, then it is valid. The development results are said to fulfill the content validity if the components of the material developed are correct and in accordance with the reference taken. While construct validity is related to consistency and the relationship between one component and the other components. Furthermore, in this study valid criteria for product development are based on expert opinion with the following criteria: (1) the results of expert assessment show that the product development along with its supporting

aspects have strong theoretical aspects and; (2) the results of the expert assessment show that the development products and their supporting components are consistently interrelated and have a valid assessment level.

Practicality in Nieveen's opinion, teachers (and other experts) consider the materials to be useful and that it is easy for teachers and students to use the materials that are largely compatible with the developers' intentions, it means that the teacher (and other experts) declare that the development product can be used and easily for the teacher and students in a way that is largely compatible with the developer's intention [37]. For this reason, the practical aspects are related to 2 things, namely whether the user states that the product developed is easy to use. Then in real terms the effectiveness of learning using the developed tools includes good categories. For this purpose, the teacher's assessment sheet is used, and the observation sheet is the implementation of authentic assessment in problem solving learning.

Effective learning tools are learning tools which can deliver students to achieve learning goals in a predetermined time. Nieveen stated that a product of development is called effective when students appreciate the learning program and that desired learning takes place, meaning that students value learning and learning programs that are planned to be implemented according to the learning target [37]. Effective learning tools provide a rich learning experience and optimum students' learning outcomes. Learning devices in this case in the form of authentic assessment instruments developed are considered effective if the teacher can provide an assessment of student development quickly and precisely.

The development of authentic assessment models in problem solving learning refers to the learning of High Order Thinking Skill (HOTS), one of which is problem solving learning. The authentic assessment model developed consists of problem solving activities, time allocation, assessment observation sheets, description of problem solving abilities, problem solving ability criteria with all the tools to assess students' development.

Based on the results of reliability statistics calculations, the inter-item correlation matrix, and the inter-item correlation matrix for authentic assessment models in problem solving learning, it can be seen that this authentic assessment model in problem solving learning developed, from the aspect of practicality has very good criteria, from the aspect of effectiveness has very good criteria, and from the aspect of implementation, has good criteria. The level of reliability of this authentic assessment model is very good. Thus it means that the authentic assessment model developed has very strong consistency, stability, and stability to assess problem solving learning activities for students.

Based on data analysis, it can be seen that the level of agreement between teachers regarding the practicality, effectiveness, and implementation of this authentic assessment model has a very strong level of agreement. This means that the teachers have a very strong agreement that the authentic assessment model developed is very practical, very effective and this authentic assessment model can be easily implemented to assess students' abilities in problem solving learning. The development of authentic assessment models in problem solving learning refers to learning from High Order Thinking Skills (HOTS), one of which is problem solving learning. The authentic

assessment model developed consists of problem solving activities, time allocation, observation sheets assessment, description of problem solving abilities, problem solving abilities and all the tools to assess students' development.

The results of reliability statistics calculations showed that the inter-item correlation matrix, and the inter-item correlation matrix for authentic assessment models in problem solving learning, it can be seen that this authentic assessment model in problem solving learning developed, has a very good criterion, aspect of effectiveness has very good criteria, and from the aspect of implementation, has good criteria. The level of reliability of this authentic assessment model is very good. Thus, it means that the authentic assessment model has been developed very strong consistency, stability, and stability to assess problem solving learning activities for students.

It was revealed that the level of agreement between teachers regarding the practicality, effectiveness, and implementation of this authentic assessment model has a very strong level of agreement. This means that the teacher has a very strong agreement that the authentic assessment model developed is very practical, very effective and this authentic assessment model can be easily implemented to assess students' abilities in problem solving learning.

CONCLUSION

The authentic assessment model developed is feasible because it meets the requirements for practicality, effectiveness and implementation. From the aspect of practicality having very practical criteria, the effectiveness aspect has very effective criteria, and from the implementation aspect

it has good criteria and with a very reliable level of reliability (Cronbach's Alpha coefficient value (α) = 0.978). The level of teacher agreement in giving an assessment of the aspects of practicality, effectiveness, and implementation of authentic assessment models in problem solving learning is very good (individual Kappa (κ) coefficient = 0.847; and the average Kappa (κ) value = 0.978). Thus, the authentic assessment model developed is very suitable to be used as an instrument to assess the results of students' abilities carried out by an assessment team consisting of two or more teachers. The results of the ability assessment of 44 students in problem solving learning were: 8 students (18.2%) got Very Good criteria, 28 students (63.6%) with Good criteria, 7 students (15.9%) with Moderate criteria, and 1 student (2.3%) with Not Good criteria. The level of accuracy of the results of the assessment of practical learning skills carried out by the assessment team consisting of 8 teachers is very good criteria (κ mean = 0.978).

REFERENCES

- [1] S. Wang and H. Wang, "Teaching and Learning Higher-Order Thinking," *Int. J. Arts Sci.*, vol. 7, pp. 179–187, 2014.
- [2] N. Sudirman, *Ilmu Pendidikan*. Bandung: Remaja Rosdakarya.
- [3] D. Meier, *The Accelerated Learning Hand Book*. Bandung: Kaifa, 2005.
- [4] E. Supriyadi, S. Soenarto, F. Surwi, and E. Prianto, "Evaluating the Assessment System of Basic Courses in the Department of Electrical Engineering," *J. Pendidik. Teknol. dan Kejuru.*, vol. 23, no. 3, pp. 235–240, Apr. 2017.
- [5] E. Mulyatiningsih, *Riset Terapan Bidang Pendidikan dan Teknik*. Yogyakarta: UNY Press, 2012.
- [6] D. A. Payne, *Applied educational assessment, 2nd ed.* Canada: Wadsworth/Thomson Learning, 2003.
- [7] A. D. Johnson, R. S., Mims-Cox, J. S., & Nichols, *Developing Portfolios in Education*. United States: SAGE Publications, Inc., 2006.
- [8] J. Mueller, "The Authentic Assessment Toolbox: Enhancing Student Learning through on Line Faculty Development," *J. Online Learn. Teach.*, vol. 1, no. 1, pp. 13–35, 2005.
- [9] N. M. Ratminingsih, A. A. I. N. Marhaeni, and L. P. D. Vigayanti, "Self-Assessment: The Effect on Students' Independence and Writing Competence," *Int. J. Instr.*, vol. 11, no. 3, pp. 277–290, Jul. 2018.
- [10] A. A. I. N. Marhaeni and L. P. Artini, "Asesmen Autentik dan Pendidikan Bermakna: Implementasi Kurikulum 2013," *J. Pendidik. Indones.*, vol. 4, no. 1, pp. 499–511, 2015.
- [11] R. Olfos and H. Zulantay, "Reliability and Validity of Authentic Assessment in a Web Based Course," *J. Educ. Technol. Soc.*, vol. 10, no. 4, pp. 156–173, 2007.
- [12] V. J. Janesick, *Authentic Assessment Primer*. New York: Peter Lang Publishing, Inc., 2013.
- [13] J. K. Mintah, "Authentic Assessment in Physical Education: Prevalence of Use and Perceived Impact on Students' Self-Concept, Motivation, and Skill Achievement," *Meas. Phys. Educ. Exerc. Sci.*, vol. 7, no. 3, pp. 161–174, Sep. 2003.
- [14] A. E. Finch, "Authentic assessment: Implications for EFL performance testing in Korea," *J. Second. Educ. Res.*, vol. 49, pp. 89–122, 2002.
- [15] C. S. Fook and .G.K. Sidhu, "Authentic Assessment and Pedagogical Strategies in Higher Education," *J. Soc. Sci.*, vol. 6, no. 2,

- pp. 153–161, 2010.
- [16] Y. Pantiwati, “Authentic Assessment for Improving Cognitive Skill, Critical-Creative Thinking and Meta-Cognitive Awareness,” *J. Educ. Pract.*, vol. 4, no. 14, pp. 1–9, 2013.
- [17] J. Gulikers, T. Bastiaens, and P. Kirschner, “Authentic Assessment, Student and Teacher Perceptions: The Practical Value of the Five Dimensional-Framework,” *J. Vocat. Educ. Train.*, vol. 58, no. 3, pp. 337–357, Sep. 2006.
- [18] K. Burton, “A Framework for Determining the Authenticity of Assessment Tasks: Applied to an Example in La,” *J. Learn. Des.*, vol. 4, no. 2, pp. 20–28, Feb. 2011.
- [19] B. A. Ataç, “Foreign Language Teachers’ Attitude toward Authentic Assessment in Language Teaching,” *J. Lang. Linguist. Stud.*, vol. 8, no. 2, pp. 7–19, Oct. 2012.
- [20] T. T. Vu and G. Dall’Alba, “Authentic Assessment for Student Learning: An ontological conceptualisation,” *Educ. Philos. Theory*, vol. 46, no. 7, pp. 778–791, Jun. 2014.
- [21] R. M. Branch, *Instructional Design: The ADDIE Approach*. Boston, MA: Springer US, 2009.
- [22] S. Azwar, *Metode Penelitian*. Yogyakarta: Pustaka Pelajar, 2010.
- [23] J. Hair, *Multivariate Data Analysis, 7th ed.* New Jersey: Perason Education Inc., 2010.
- [24] J. R. Landis and G. G. Koch, “The Measurement of Observer Agreement for Categorical Data,” *Biometrics*, vol. 33, no. 1, pp. 159–74, Mar. 1977.
- [25] L. R. Gloeckler, J. M. Cassell, and A. J. Malkus, “Teacher Practices with Toddlers During Social Problem Solving Opportunities,” *Early Child Dev. Care*, vol. 184, no. 5, pp. 749–765, May 2014.
- [26] M. Lopez, J., Whittington, Susie, “Higher-order Thinking in a College Course: A Case Study,” *North Am. Coll. Teach. Agric.*, no. December 2001, pp. 22–29, 2001.
- [27] M. A. Shiakalli and K. Zacharos, “The Contribution of External Representations in Pre-School Mathematical Problem Solving,” *Int. J. Early Years Educ.*, vol. 20, no. 4, pp. 315–331, Dec. 2012.
- [28] T. D. Taylor, “Identifying Students for Intervention in Math Problem Solving: An Evaluation of Fluency-Based Word,” University of California, 2014.
- [29] S. Yoleri, “Investigation of Relationship between the Skills to Solve Interpersonal Problems and Concept Development of Preschool Children,” *Egit. ve Bilim*, vol. 39, pp. 82–90, 2014.
- [30] C. D. Rogers, “Poverty Analysis Using Alternative Poverty Measures and Geospatial Technology in the Buffalo-Niagara Falls, New York, Metropolitan Statistical Area: 1970-2000,” *URISA J.*, vol. 26, no. 2, pp. 51–59, 2014.
- [31] S. M. Lim, S. Rodger, and T. Brown, “Assessment of Learning-Related Skills and Interpersonal Skills Constructs within Early Childhood Environments in Singapore,” *Infant Child Dev.*, vol. 45, pp. 251–259, 2017.
- [32] S. J. Bagnato, “Developmental Diagnostic Reports: Reliable and Effective Alternatives to Guide Individualized Intervention,” *J. Spec. Educ.*, vol. 15, no. 1, pp. 65–76, Apr. 1981.
- [33] A. Karatas, I. & Baki, “The Effect of Learning Environments Based on Problem Solving on Students’ Achievements of Problem Solving,” *Int. Electron. J. Elem. Educ.*, vol. 5, pp. 249–268, 2013.
- [34] A. Herrington, J. Herrington, L. Sparrow, and R. Oliver, “Learning to Teach and Assess Mathematics Using Multimedia: A Teacher

- Development Project,” *J. Math. Teach. Educ.*, vol. 1, no. 1, pp. 89–112, 1998.
- [35] A. Pyle and C. DeLuca, “Assessment in the Kindergarten Classroom: An Empirical Study of Teachers’ Assessment Approaches,” *Early Child. Educ. J.*, vol. 41, pp. 373–380, 2013.
- [36] M. Ebbeck, G. L. C. Teo, C. Tan, and M. Goh, “A Study on Assessing Developmental Learning Outcomes in Toddlers,” *Early Child. Educ. J.*, no. 42, pp. 115–123, 2014.
- [37] N. Nieveen, *Prototyping to Reach Product Quality*. London: Kluwer Academic Publisher.