JERE 8 (1) (2019) 30 - 38

Journal of Educational Research and Evaluation

http://journal.unnes.ac.id/sju/index.php/jere

Development of Assessment Performance Materials Cylinder Students Grade IX SMP Adiwiyata-Based

Zelmy Adista Vembriliya[1]✉, Muhammad Khumaedi[2], Masrukan[2]

[1] SMP N 1 Dawe, Kudus, Indonesia
[2] Universitas Negeri Semarang, Indonesia

Article Info

Article history:
Received 3 November 2019
Approved 12 January 2019
Published 10 March 2019

Keywords:
Development, Assessment, Performance, Cylinder, Adiwiyata

Abstract

The assessment instrument for mathematics performance of Adiwiyata-based cylinder material is not standard, although there are instruments but not yet valid and reliable, and there is no clear rubric to assess learning outcomes. This study aims to develop assessment instruments for cylinder material performance based on adiwiyata class IX students who are valid and reliable. This research method uses Djemari Mardapi model development research with preliminary stages, development, and evaluation. The study of instruments by experts was analyzed using the Aiken's V formula, the reliability of the instrument was based on the agreement of experts with analysis using two way ANOVA and re-analyzed using the Hoyt formula. Construct validity was assessed using exploratory factor analysis, the reliability of the instrument based on field results was analyzed using Cronbach Alpha with the help of SPSS. The performance assessment of cylinder material developed consists of 10 items. The results of the evaluation of the measured content have a value of> 0.3, that the skill instrument developed is valid. The reliability results based on expert agreement show a value of 6 0.6, which means experts agree in assessing. Data analysis in the field test can be seen from the results of Kaiser Meyer Olkin (KMO) which shows KMO values> 0.5 so that the instrument items can be analyzed further. There are three factors formed from each assessment, each item in each factor has a value of loading factor> 0.3 while the three factors are preparation, implementation, and final results. The results of reliability in the field have a value of> 0.6 so that the instruments developed are consistent in carrying out the assessment. The results of this study can be used as a guideline for teachers in junior high schools in conducting a performance assessment of adiwiyata-based cylinder material.

✉Alamat korespondensi
Jl. Raya Kudus - Colo No.KM.11, Piji Wetan, Lau, Dawe,
Kabupaten Kudus, Jawa Tengah 59353, Indonesia
E-mail: adistavembriliya@gmail.com

p-ISSN 2252-6420
e-ISSN 2503-1732

INTRODUCTION

Assessment is very important and strategic in learning activities. Assessment or assessment is the activity of interpreting or interpreting the results of a measurement data based on criteria or standards and certain rules. Assessment function for students to identify the level of success of learning, while for teachers to identify successes in teaching. The assessment technique used is very dependent on the competencies assessed (Kunandar, 2015, p.69). One of the assessment techniques is a test. The test is a tool to collect information on the characteristics of an object. Tests are more suitable to be used to determine students' abilities in aspects of knowledge and skills, not suitable to be used to measure attitudes because attitudes cannot be interpreted into the category of right or wrong (Widoyoko, 2016, p.65). This is in line with the research (Rusilowati, 2017 pp. 19-20) that the test is divided into three, namely objective tests, description tests, and performance. Assessment of performance is based on the results of observations on performance, behavior, or student interaction in a particular task. To assess performance, instruments and assessment rubrics are needed. Hugh Burkhardt and Malcolm Swan stated that performance appraisal is an important part of learning in any field, whether playing sports or musical instruments or doing mathematics (Burkhardt & Swan, 2008, p.3).

Learners to easily learn mathematics by applying forms of building space, especially Cylinder, using items in the environment, or using used materials as a medium or tool to help the learning process, this is called the adiwiyata program. The Adiwiyata program is a program that aims to encourage and shape schools in Indonesia in maintaining environmental preservation and sustainable development for the interests of the present and future (Adiwiyata, 2012, p.3). The benefit for students is to foster awareness of the environment and the importance of maintaining the environment to stay good. SMP 1 Dawe is one of the schools that will now take part in the national Adiwiyata program.

Based on the results of interviews with researchers with mathematics teachers who participated in the National Learning Innovation (INOBEL), said the increase in student learning outcomes increased with learning to utilize the waste in the environment around the students. With students learning and finding a formula directly. Students become aware of and remember what they have learned. In assessing student learning outcomes still, subjectively, the good results get good grades. Because there is no clear assessment or instrument. So the teacher assesses directly the results of student work. The next problem that is often faced by mathematics teachers is that the assessment of performance lies in the validity and reliability of the used measurement tools (Yudha, Masrukan, & Djuniadi, 2014, p.64). The preparation of student performance tests is still very limited to the teacher's knowledge and understanding of simulation tests.

Based on the explanation above, the researcher seeks to develop assessment instruments for the performance of tube material adiwiyata. By utilizing tubular used goods can be used as an instrument to make it easier for students to understand the shape of the curved side space, and make it easier for the teacher to make judgments accompanied by the rubric.

METHOD

The method used in this research is development research. This study uses a development research design put forward by the development research model (Mardapi, 2016, p.132) where there are 10 steps. 10 steps in the development of Mardapi are (1) Determining to Instrument Specifications, (2) Writing Instruments, (3) Determining Instrument Scale, (4) Determining the Scoring System, (5) Analyzing Instruments, (6) Conducting Tests, (7) Analyzing Instruments, (8) Assemble instruments, (9) Carry out measurements, (10) interpret measurements.

Research procedure

The steps of each development do not have to use standard steps so that they can choose and determine the most appropriate steps for the researcher based on the conditions faced

(Dwiyogo, 2004, p.6). The development procedure in this study was modified into 3 important stages, namely the preliminary study phase, development, and evaluation. In the preliminary study, phase some steps must be taken including a preliminary study description of the analysis of findings, writing the instrument up to examining the instrument. Development studies include testing, analysis, revision, and evaluation stages consisting of product analysis and improvement processes.

RESULTS AND DISCUSSION

Assessment Instrument Performance of Cylinder Material

Performance assessment developed by researchers at 1 Dawe Middle School in the form of assessment rubrics to find the surface area of a cylinder by utilizing waste in the surrounding environment, while the instrument specifications are as follows:

a. Instrument Grid

The grid designed based on KD 3.7 and 4.7 based on the 2013 curriculum syllabus produces 3 performance assessments namely assessment finding the area of a tube, cone, and bo. It is based on adiwiyata which uses used goods.

b. Scoring scale

The rating scale used in the observation instrument and assessment rubric uses a rating scale model with the lowest score of 1 and the highest 4.

c. Accounting Technique

Scoring techniques used in Adiwiyata-based performance assessment are (score obtained) or (maximum score) × 100.

Based on the scoring technique, the score criteria in the performance assessment of the Adiwiyata-based Cylinder material are scores (0-30) with fewer criteria, scores (31-70) with sufficient criteria, scores (71-100) with good criteria.

Content Validity

Instruments that have been developed before field trials must go through the content validation stage first carried out by experts, which in the preparation of product design is an assessment of performance by following rational thinking from experts. The instrument validation was carried out by 3 experts, namely 1 expert in the field of mathematics, 1 expert in the field of measuring and developing instruments, and 1 mathematics teacher, experts were asked to provide assessments, input, opinions on grids, assessment rubrics, scoring techniques and instruments assessment in the form of observation sheets that have been compiled by the researcher. The three experts gave assessment scores on the validation sheet of the assessment instrument for adiwiyata-based tube material performance, the highest score with number 4 and the lowest score with number 1. After obtaining the assessment score from the next expert, it will be analyzed with Aiken's V. Formula. The work of adiwiyata-based tube material can be seen in Table 1.

Table 1. Results of Content Validity with Formula Aiken's V

No Butir	Ahli 1	Ahli 2	Ahli 3	∑S	V Index	Keterangan
1	4	4	3	11	0.889	Valid
2	4	4	3	11	0.889	Valid
3	4	3	4	11	0.889	Valid
4	4	4	4	11	1	Valid
5	4	4	4	12	1	Valid
6	4	4	4	12	1	Valid
7	3	4	4	11	0.889	Valid
8	4	4	3	11	0.889	Valid
9	4	4	3	11	0.889	Valid
10	4	4	4	12	1	Valid

Based on the results of the content validity test using the Aiken's V formula, it shows that all aspects assessed as having the Aiken coefficient (3 0.3) mean that the instrument performance assessment of Cylinder material is determining the surface area of the adiwiyata-based tube has good content validity. This is in accordance with the validity coefficient stated (Azwar, 2016, pp.147) that if the validity coefficient> 0.3 means that it can be said to be adequate (valid) otherwise if the coefficient of validity is <0.3 then it is declared inadequate (invalid).

Based on the study of experts analyzed using the Aiken's V formula where there were 10 criteria measured in the performance instrument, namely the performance assessment instrument found that the cylinder surface area was declared valid. This is in accordance with the validity coefficient stated

(Azwar, 2016, p. 147) that if the validity coefficient> 0.3 means that it can be said to be adequate (valid) otherwise if the coefficient of validity is <0.3 then it is declared inadequate (invalid). This is in line with the research carried out by Majid on the Development of Authentic Assessment Instruments Performance at Science Subjects at Jlamprang Elementary School and SDN Wonosari 03 Batang Regency. The results of the study show agreement between the rater mean that the instrument has a high enough quality of quality (Majid, Nur K., Tri Joko R., & Supriyadi, 2017). This is in line with the research conducted by Eris Fahmi Rahmawan that performance appraisal is worthy of being used as a form of assessment. (Rahmawan, Eris F, Sumaryanto, & Supriyadi, 2016).

Reliability

Calculating the level of agreement between the three experts using the inter-rater consistency reliability test analyzed using a different test through two-factor ANOVA (two way anava) and then proved through the Hoyt Formula analysis by calculating the reliability coefficient value, the results of calculation analysis are done through two-way test procedures ANOVA uses the SPSS 16.0 program shown in table 2.

Table 2. Reliability Calculation Results Using

Tests of Between-Subjects Effects

Dependent Variable:SkorPenilai

Source	Type III Sum of Squares	Df	Mean Square	F	Sig.
Corrected Model	5.867[a]	29	.202	.	.
Intercept	418.133	1	418.133	.	.
P	1.667	2	.833	.	.
Butir	.533	9	.059	.	.
P * Butir	3.667	18	.204	.	.
Error	.000	0	.		
Total	424.000	30			
Corrected Total	5.867	29			

Based on the results of the reliability test carried out through the two-way ANOVA test using SPSS 16.0, it shows that the variance between the items stated by the Appraisal Mean Square is MK_s = 0.833 while the error variance is expressed by the Mean Square * Item Grading ie (MK)_is = 0.204. Then it was recalculated through the reliability of Hoyt (1941) with an average of three rater people

$$r_{xx\prime} = 1 - \frac{MK_{is}}{MK_s}$$

$$r_{xx\prime} = 1 - \frac{0{,}204}{0{,}833}$$

$$r_{xx\prime} = 0{,}756$$

The results calculated using the Hoyt formula resulted in a reliability coefficient of 0.756 meaning that between the rater judging agreed on the suitability of the content and this also indicated that the score given by each rater was consistent. This is in line with the research conducted by

(Sujarwanto & Rosilawati, 2015: 785), that if the reliability coefficient value is> 0.6, then experts are consistent in assessing, which means that the three experts are consistent in assessing the instrument. This means that the three experts are consistent in assessing instruments. According to (Khumaedi, 2012, p.3) reliability is a coefficient that shows the extent to which an instrument and measuring device can be trusted.

The analysis carried out by the researcher showed that the performance assessment instrument found that the surface area of the adiwiyata-based tube that had been developed had been tested for validity and reliability based on expert agreement, that the instrument developed was appropriate and fulfilled the validity and reliability requirements so that it could be used for field testing.

Test Validity of Constructs in the Field

Research conducted at Dawe 1 Junior High School towards 112 IX grade students, on the implementation of the instrument test, the researcher asked for help from 2 teachers of mathematics studies to make observations and assess the performance of students' skills.

The results of the teacher's assessment of the students' skills were then recapitulated to obtain data in the field trials. Then the data is then analyzed using factor analysis with the Exploratory Factor Analysis (EFA) approach. The results of the exploratory factor analysis with the help of the SPSS 16.0 program, on each instrument, can be seen in the following explanation:

1. KMO and Bartlett's Test

The main requirement for analyzing factors to test construct validity is to see the value of KMO (Kaiser Meyer Olkin Measure of Sampling). KMO values vary from 0 to 1. If the value of KMO> 0.5, factor analysis can be done, but if the value of KMO is <0.5, factor analysis cannot be continued (Ghozali, 2016). The following are the results of testing the feasibility of instruments in the field which can be seen in Table 3.

Table 3. Value of KMO and Bartlett's Test Assessment of First Performance

Kaiser-Meyer-Olkin Measure of Sampling Adequacy.		,702
Bartlett's Test of Sphericity	Approx. Chi-Square	1.071E3
	Df	45
	Sig.	.000

Based on the results of the analysis show that the KMO and Bartlett's numbers are 0.702 with a significance of 0,000, seeing the numbers listed in table 3, the indicators and existing samples actually meet the criteria and can be further analyzed. The next step is to analyze the correlation between performance assessment items to find the tube area.

2. Anti Image Correlation Matrix

Anti Image Correlation aims to determine the correlation of each item in the instrument being analyzed. Each item in the instrument is said to be correlated if the correlation value of each item is more than 0.5. The results of the item correlation analysis in the first performance assessment can be seen in Table 4 below:

Table 4. Anti Image Correlation Results

Butir 1	Butir 2	Butir 3	Butir 4	Butir 5	Butir 6	Butir 7	Butir 8	Butir 9	Butir 10
.505[a]	-.800	-.055	.032	.090	-.112	-.007	.001	.103	-.115
-.800	.507[a]	-.031	-.004	.009	.036	.030	.050	-.019	.030
-.055	-.031	.756[a]	-.197	-.328	.103	-.862	-.081	.005	-.002
.032	-.004	-.197	.794[a]	.177	-.178	-.011	-.675	-.104	.048
.090	.009	-.328	.177	.784[a]	-.666	-.004	.024	-.085	.032
-.112	.036	.103	-.178	-.666	.749[a]	.024	.006	.052	.001
-.007	.030	-.862	-.011	-.004	.024	.781[a]	.031	.020	-.001
.001	.050	-.081	-.675	.024	.006	.031	.801[a]	.077	-.027
.103	-.019	.005	-.104	-.085	.052	.020	.077	.501[a]	-.949
-.115	.030	-.002	.048	.032	.001	-.001	-.027	-.949	.501[a]

Anti image correlation was obtained> 0.5, in the overall rotated component matrix the overall criteria measured in the performance assessment instrument found that the surface area of the adiwiyata-based tube had an Eigenvalue> 1 and Factor Loading> 0.3 with the theory put forward (Azwar, Saifuddin, 2016, p. 123) if the value of KMO> 0.5, Anti Image Corelation> 0.5, Eigenvalue ≥ 1 and Factor Loading 3 0.3 factor analysis can be carried out.

3. Total Variance Explained

Total Variance Explained shows the percentage of total diversity that is capable of being explained by the diversity of factors formed. To determine how many components and factors are used to explain the total diversity, it can be seen from the eigenvalues value, components with more than one eigenvalue are the components used. To see how many factors are formed in factor analysis can be seen in Table 5.

Table 5. Results of Analysis of Total Variance Explained factor analysis.

	Initial Eigenvalues			Extraction Sums of Squared Loadings			Rotation Sums of Squared Loadings		
Component	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	4.270	42.705	42.705	4.270	42.705	42.705	4.253	42.525	42.525
2	1.962	19.618	62.323	1.962	19.618	62.323	1.961	19.613	62.139
3	1.808	18.082	80.405	1.808	18.082	80.405	1.827	18.267	80.405
4	.898	8.984	89.389						
5	.490	4.904	94.293						
6	.191	1.914	96.207						
7	.169	1.692	97.899						
8	.129	1.293	99.192						
9	.047	.473	99.665						
10	.034	.335	100.000						

Based on Table 5 shows 3 components are formed and can represent indicators, 10 criteria are measured and then analyzed, it turns out to have a value of eigenvalues> 1 which means that the 10 criteria measured can be grouped into 3 factors. Factor 1 has a value of 4,270 and can explain the variance of 42,705, Factor 2 has a value of 1,962 and can explain the variance of 19,618, Factor 3 has a value of 1,808 and can explain the variance of 80,405. This is in accordance with the existing criteria, namely factors that have eigenvalues of more than 1 will be maintained and factors that have eigenvalues of less than 1 are not included in the model (Suranto, Muhyadi, 2014, p.33) In addition to looking at the total variance table explained by many factors can also be done by looking at the scree plot. The factors used to be able to explain the total diversity are seen from the large eigenvalues, components with eigenvalues> 1 are the components used. Overall the results of the Total of Variance can be described in the scree plot that can be seen below:

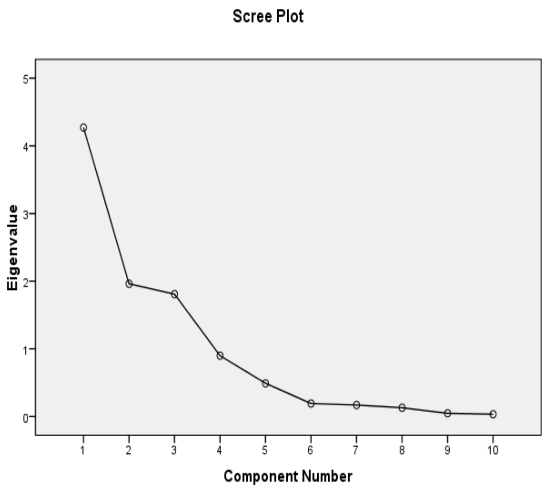

Figure 1. Scree Plot

Based on the Scree Plot above, it appears that 3 points are above the value of 1 and the other points are below the value 1. This illustrates that there is a component that has an eigenvalue above 1 means that there are only 3 factors formed.

4. Rotated Component Matrix

Rotated Component Matrix shows the magnitude of the correlation of each variable in the formed factor. A factor is formed if each instrument has a value of> 0.30. Grouping items into factors can be done by looking at the Rotation Component Matrix table. The following is the result of the Rotation Component Matrix in Table 6.

Table 6. Rotated Component Matrix

Rotated Component Matrix[a]

	Component		
	1	2	3
Equipment tools	.068	-.004	.948
Material supplies	-.012	-.017	.949
The accuracy of removing labels on used milk cans	.937	-.009	.086
Neatness in removing the label	.847	.024	-.042
Suitability in sticking to milk labels	.837	.051	.015
Accuracy in attaching milk labels	.703	-.060	.066
Conformity to draw circles	.915	-.023	.076
Accuracy in drawing circles	.787	-.079	-.079
Write the basic formula	-.015	.987	-.034
Concludes the broad formula	-.038	.986	.012

Based on Table 6 shows the results of the rotation factor, it can be seen that the grouping of indicators into factors and the magnitude of loading factors can be seen in the table above. It can be seen that the determination of indicator inputs to certain factors follows in the magnitude of the correlation between variables and factors, namely to those with a large correlation. The factors formed along with the items are presented in Table 7.

Table 7. Grouping each instrument into a factor

NO	Butir	Faktor yang terbentuk	Nilai korelasi faktor	Nama faktor
1	Equipment tools	1	0.948	Persiapan
2	Material supplies		0.945	
3	The accuracy of removing labels on used milk cans	2	0.937	Pelaksanaan
4	Neatness in removing the label		0.847	
5	Suitability in sticking to milk labels		0.837	
6	Accuracy in attaching milk labels		0.703	
7	Conformity to draw circles		0.915	
8	Accuracy in drawing circles		0.787	
9	Write the basic formula	3	0.987	Hasil akhir
10	Concludes the broad formula		0.986	

Based on table 7 According to the results of the grouping of 10 criteria measured into 3 factors formed and naming each factor. The components formed in factor 1 are named preparations, instrument criteria no. 1 and 2. Components formed in factor 2 are given the name of the manufacturing process, instrument criteria no. 3,4,5,6,7, and 8. Components formed in factor 3 given the final result, criteria no. 9 and 10.

Field Reliability Test

After knowing the results of construct validity on testing in the field, then the next step is to test the reliability of the instrument to determine the reliability coefficient of the instrument being developed. The reliability test of the assessment instrument for the first work was carried out using the Cronbach Alpha reliability test through the SPSS 16.0 program. The following are the reliability test results which can be seen in table 8 below:

Table 8. The reliability test

Cronbach'Alpha	N Of Items
.752	10

Based on the value of the coefficients generated in the field tests performed the reliability testing value is 0.752> 0.60 so that it can be said that the assessment instrument for the first work is consistently used in conducting the assessment. The is by following the results of the study of Widya Puji Astuti which shows that the performance assessment instrument is reliable and the results of trial 2 have a higher level of reliability than the results of trial 1 (Astuti, Widya P, Wibawanto, H & Khumaedi, M, 2015). This is in line with Ibnu Wachyudin's research with the title Research for Developing Performance Evaluation Instruments. The results of the PUK Instrument research developed proved to be valid, reliable and effective (Wachyudi, I, Sukestiyarno, & Waluya, B, 2015).

CONCLUSION

Based on the results of the research and discussion that have been conducted, it can be concluded that the assessment of valid and reliable tube material performance based on adiwiyata consisting of grids, the rating scale is an adiwiyata-based performance assessment rubric. The development of an assessment instrument for the performance of adiwiyata-based tube material has not been carried out by teachers in junior high schools. Therefore, further research and development are needed to make assessment instruments for the performance of adiwiyata-based cylinder material for other materials. Further research can be disseminated the results of the development of adiwiyata-based cylinder material performance assessment instruments on a broader scale.

ACKNOWLEDGEMENTS

Acknowledgments the researchers convey to those who have helped during the research process, including; 1) Prof. Kartono, M.Pd as expert validator in the field of mathematics, 2) Dr. Wahyu Lestari, M. Pd. as an expert validator in the field of education research and evaluation, 3) Dian Rosita, M.Pd as a mathematics teacher at SMP 1 Dawe Kudus, 4) SMP 1 Dawe Kudus as a place of research that has provided opportunities for researchers to conduct research activities.

REFERENCES

Adiwiyata, Tim Teknis. 2012. "Panduan Adiwiyata Sekolah Peduli Dan Berbudaya Lingkungan."

Astuti, Widya P., Wibawanto, Hari., and Muhammad & Khumaedi. 2015. "Pengembangan Instrumen Penilaian Unjuk Kerja Praktik Perawatan Kulit Wajah Berbasis Kompetensi Di Universitas Negeri Semaran." *Innovative Journal Of Curriculum and Educational Technology* 4 (1):13.

Azwar, Saifuddin. 2016. *Reliabilitas Dan Validitas*. Yogyakarta: Pustaka Pelajar.

Burkhardt, Hugh, and Malcolm Swan. 2008. "Designing Assessment of Performance in Mathematics," no. May:1–24.

Dwiyogo, W. D. 2004. "Konsep Penelitian Dan pengembangan.Makalah. Lokakarya Metodelogi Penelitian Jurusan Kepelatihan Olahraga Fakultas Ilmu Kedokteran." Semarang.

Khumaedi, Muhammad. 2012. "Reliabilitas Instrumen Penelitian Pendidikan." *Jurnal Pendidikan Teknik Mesin* 12 (1):29.

Kunandar. 2015. *Penilaian Autentik (Penilaian Hasil Belajar Peserta Didik Berdasarkan Kurikulum 2013) Suatu Pendekatan Praktis*. Jakarta: PT. RajaGrafindo Persada.

Majid, Nur K., Tri Joko R., & Supriyadi. 2014. "Journal of Educational Research and Evaluation." *Journal of Educational Research and Evaluation* 3 (2):36–40.

Mardapi, Djemari. 2016. *Pengukuran Penilaian & Evaluasi Pendidikan*. Yogyakarta: Pustaka Pelajar.

Rahmawan, Eris F., Sumaryanto., & Supriyadi. 2016. "Journal of Educational Research and Evaluation." *Journal of Educational Research and Evaluation* 3 (2):36–40.

Rusilowati, Prof. Ani. 2017. *Pengembangan Instrumen Penilaian*. Semarang: UNNES Press.

Sujarwanto, and Ani Rosilawati. 2015. "Pengembangan Instrumen Performance Assessment Pendekatan Scientific Pada Tema Kalor Dan Perpindahan." *Unnes Science Education Journal* 4 (1):1–9.

Suranto, Muhyadi, & Mardapi D. 2014. "Pengembangan Instrumen Evaluasi Uji Kompetensi Keahlian (UKK) Administrasi Perkantoran Di SMK." *Jurnal Penelitian & Evaluasi Pendidikan* 18 (1):98–114.

Wachyudi, Ibnu., & Sukestiyarno., and Budi Waluya. 2015. "Pengembangan Instrumen Penilaian Unjuk Kerja Pada Pembelajaran Dengan Model Problem Solving Berbasis TIK." *Journal of Educational Research and Evaluation* 4 (1):23–25.

Widoyoko, Eko Putro. 2016. *Penilaian Hasil Pembelajaran Di Sekolah*. Yogyakarta: Pustaka Pelajar.

Yudha, Rivo Panji, Masrukan, and Djuniadi. 2014. "Pengembangan Instrumen Asesmen Otentik Unjuk Kerja Materi Bangun Ruang Di Sekolah Dasar." *Journal Of Educational Research and Evaluation* 3 (2):63–67.