ATLANTIS PRESS

Atlantis Highlights in Engineering (AHE), volume 1

International Conference on Science and Technology (ICST 2018)

The Effect Of Performance-Based Practical Assessment Model Towards Students' Competency Level At Civil Engineering Workshop In Kupang State Polytechnic

Wayan Sri Kristinayanti[1]
Civil Engineering Department
Bali State Polytechnic
Bali, Indonesia
[1]srikristinayanti@pnb.ac.id

I Made Anom Santiana[3]
Civil Engineering Department
Bali State Polytechnic
Bali, Indonesia
[3]madeanomsantiana@pnb.ac.id

Wira Ditta Lokantara[2]
Economics Department
Universitas Gadjah Mada (UGM)
Yogyakarta, Indonesia
[2]wiraditta@gmail.com

N K Suciani[4]
Business Administration Department
Bali State Polytechnic
Bali, Indonesia
[4]suciani@pnb.ac.id

***Abstract* - Assessment in the learning process in formal education in terms of assessing student competence does not only cover cognitive aspects and affective aspects, but also psychomotor aspects of students. The purpose of this study is to improve student competency and performance based on the development of a performance-based assessment rubric in the practicum conducted at Kupang State Polytechnic. This study involved a large group of students in the wooden practicum at the civil engineering workshop in Kupang State Polytechnic, and carried out in several stages, namely needs analysis, material development, instrument development, drafting of models, instrument testing, and calculation of learning effects. Performance-based assessment in this study uses rubric in assessing student competencies in the civil engineering workshop in terms of cognitive aspects, affective aspects and psychomotor aspects, and then tested the validity of the assessment rubric. Based on the validity test, 35 valid statement items from the 35 items in the performance-based assessment rubric were produced, this was indicated by the Cronbach alpha value equal to 0.981. In the expert trial involving 8 experts in providing assessment, the CVI value = 0.80, which means that the assessment instrument can be used in the trial group. After learning the practicum students have a treatment effect, the results show that there are significant differences in student test results for untreated (using old method) with those treated (using a new assessment tool). The score generally increases by 20.4%. It was concluded that assessment methods that were improved using new instruments could improve students' competence during the workshop. Therefore, it is recommended that this new instrument for assessment be used in future workshops.**

Keywords— Practicum assessment; performance-based; within-treatment; students' competency level.

I. INTRODUCTION

Assessment is an integral part of learning. This practice appears in various forms in vocational education, which includes the assessment of practical classes. Significant contributions can be generated in learning by compiling good and correct practicum assessments and motivating lecturers to develop knowledge and asking students to explain and implement what students have learned [1]. The existence of practicum can motivate students to develop knowledge, skills and provide opportunities for students to apply knowledge, test theory and after that can result in changes in student understanding. Students in understanding the practicum have difficulty expressing these things and the possibility of exploring knowledge that occurs outside, also in the assessment requires better ways [1,2].

Most lecturers and educators in general, are not aware and interested in performance-based assessments as they argue that its time and energy consuming while being too expensive, which is discussed a lot in many researches [3]. Preliminary studies were carried out to show that 65% of students and 62% of practicum lecturers dissatisfied and feeling that the assessments were incomplete. The assessments were not able to describe the diverse abilities of students due to inappropriate and invalid methods and assessment instruments that mainly caused by limited time, and lecturers' lack of knowledge. Based on these conditions, it is necessary to assess both the process and learning outcomes. Performance-based assessment is a form of assessment with a new paradigm that more precisely and comprehensively measures students' abilities, compared to assessment types that have dominated the assessment [2,4]. The Education Minister Regulation No. 20 year 2007 states that in providing a good

assessment for learning, it is necessary to establish valid assessment criteria and determine the various skills and knowledge acquired by students [5].

We propose the introduction of this performance-based assessment method with a rubric for civil engineering workshops, and test its validity with two methods: theoretical validity by conducting expert tests and empirical validity by conducting field trials. A rubric is an instrument that contains criteria used to assess a particular job or performance and as a result can explain more than one class or sign [2]. Assessment using rubric is very helpful in subjective assessment to be more objective and more measurable towards the ability of students including the required task elements so as to provide timely feedback to students and improve student skills [2,5,6].

Brookhart stated that there are two general conditions in the rubric, namely the number of criteria and assessments that are analytical and holistic and the type of process that will be assessed as general or specific [7]. We will make an assessment by developing a task rubric with a special and analytical type of process that is consistent with the characteristics of the practicum work in a civil engineering workshop, where students need to carry out special tasks in wood, steel, concrete and scaffold lab work, which requires practical preparation, painting, welding, cutting, assembly, mounting, finishing [8]. The Kupang State Polytechnic to date in the assessment in the workshop still uses an assessment format that refers to the 2007 Curriculum, namely the existence of an assessment of theory, practice and attitude processes, where the assessment contains high subjectivity and unmeasurable assessment [8]. In assessing the construction work in the workshop that has been mentioned previously, requires a valid instrument and can be used to improve student competence in practicum [9].

II. RESEARCH METHODOLOGY

In this study, validity is an important factor in choosing or implementing an instrument, which then can measure an instrument's properties of the construct under study [10]. With regards to the three forms of general validity, which include content, construction, and validity related to criteria [11], content validity received the greatest priority during instrument development. Validity must be shown in each study used by the instruments as it proves the appropriate use of them to the assessment [12]. Internal and external validity will be used to determine the characteristics of the assessment model developed while achieving a theoretically valid (Lawshe's Expert test) and empirically valid results of effect of the application of assessment models on student competencies through a within-treatment test result on six classes at Kupang State Polytechnic where three classes acts as a control (group 1) and three others as the treated group (group 2). Preliminary and post-tests were done to measure the effect of a new assessments' instruments towards students' competency in the workshop.

The tests used in this study were the 35-instruments rubrics test, where the reliability and validity will be revealed in the results tables. Purposive sampling technique was used for both the 1) experts test are included in internal validity test, which involves two vocational education, two language, two timber steel and scaffolding expert, and two construction practitioners, and 2) within-treatment test for students' competency level assessment. The validation of the rubrics model used the Lawshe's Formula to test on the expert judges. The data from students through each point in the rubric's assessment is also tested with Pearson Correlation test, and external validity test is tested using Cronbach's alpha score in all 70 students' samples across all three classes who are treated (group 2). Statistical analysis was carried out descriptively on competency level of the students who had gone through the assessments' explanation (group 2) and the ones who had not or controlled (group 1).

III. RESULTS AND DISCUSSIONS

A. *Instruments Validity Test*

- *Draft Performance Based Practical Assessment Rubric*

The resulting assessment instrument was a performance-based assessment rubric on wood, steel, scaffold at Civil Engineering Workshop in Bali State Polytechnic, which contained 35 statement items, completed with an explanation of the observed value, indicators, description of each item statement and score given to the description.

- *Internal Validity Test (Expert Test)*

To test the validity as a whole can be calculated using the Content Validity Index in each statement item, where there are 35 statement items. The value of the content validity obtained reflects the entire assessment item produced. This study tested the content validity that is looking for the value of CVR (Content Validity Ratio), if the value obtained in each item of CVR statement> 0 (positive) means that all items declared valid by experts. The value generated in each statement item is around 0.75 - 1.00. According to Lawshe, if using 8 experts, the minimum CVR value that must be produced is 0.582. It can be concluded that the CVR value has fulfilled so that it can be continued with the calculation of CVI (Content Validity Index) as follows:

TABLE I. CATEGORY OF CALCULATION RESULT (CVI)

Range	Category
0.00 – 0.33	Statement item that is not appropriate
0.34 – 0.67	Appropriate statement item
0.68 – 1.00	Very appropriate statement item

where, the calculation of would be:

$$CVI = 28 / 35$$
$$= 0.80$$

Based on the above calculation, the CVI value ranges from 0.68 to 1.00 so that it falls into the very appropriate category, it can be concluded that the instrument in the form of this performance-based assessment rubric can be used as a whole.

- *External Validity Test (Empirical)*

The external validation test in this study was a field trial involving a larger group consisting of 210 students (6 classes) who carried out wood, steel and scaffold practicums using 5 assessors (assessors), the assessment results showed the following reliability and validity:

TABLE II. CRONBACH'S RELIABILITY TEST

Reliability Statistics

Cronbach's Alpha	N of Items
0.981	35

The degree of reliability of the test according to Guilford as follows: (≤ 0.20: Very Low; 0.20 ≤ 0.40: Low; 0.40 ≤ 0.60: Moderate; 0.60 ≤ 0.80: Height; 0.80 ≤ 1.00: Very High).

Based on the reliability test, the Cronbach's Alpha value of 0.981 indicates a very high level of reliability.

TABLE III. CORRELATION VALIDITY TEST RESULTS

Correlations

		Total			Total			Total
x1	Pearson Correlation	.890**	x13	Pearson Correlation	.890**	x25	Pearson Correlation	.842**
	Sig. (2-tailed)	.000		Sig. (2-tailed)	.000		Sig. (2-tailed)	.000
	N	70		N	70		N	70
x2	Pearson Correlation	.895**	x14	Pearson Correlation	.921**	x26	Pearson Correlation	.623**
	Sig. (2-tailed)	.000		Sig. (2-tailed)	.000		Sig. (2-tailed)	.000
	N	70		N	70		N	70
x3	Pearson Correlation	.842**	x15	Pearson Correlation	.842**	x27	Pearson Correlation	.492**
	Sig. (2-tailed)	.000		Sig. (2-tailed)	.000		Sig. (2-tailed)	.000
	N	70		N	70		N	70
x4	Pearson Correlation	.842**	x16	Pearson Correlation	.633**	x28	Pearson Correlation	.549**
	Sig. (2-tailed)	.000		Sig. (2-tailed)	.000		Sig. (2-tailed)	.000
	N	70		N	70		N	70
x5	Pearson Correlation	.921**	x17	Pearson Correlation	.921**	x29	Pearson Correlation	.894**
	Sig. (2-tailed)	.000		Sig. (2-tailed)	.000		Sig. (2-tailed)	.000
	N	70		N	70		N	70
x6	Pearson Correlation	.890**	x18	Pearson Correlation	.842**	x30	Pearson Correlation	.895**
	Sig. (2-tailed)	.000		Sig. (2-tailed)	.000		Sig. (2-tailed)	.000
	N	70		N	70		N	70
x7	Pearson Correlation	.874**	x19	Pearson Correlation	.850**	x31	Pearson Correlation	.895**
	Sig. (2-tailed)	.000		Sig. (2-tailed)	.000		Sig. (2-tailed)	.000
	N	70		N	70		N	70
x8	Pearson Correlation	.874**	x20	Pearson Correlation	.503**	x32	Pearson Correlation	.503**
	Sig. (2-tailed)	.000		Sig. (2-tailed)	.000		Sig. (2-tailed)	.000
	N	70		N	70		N	70
x9	Pearson Correlation	.890**	x21	Pearson Correlation	.503**	x33	Pearson Correlation	.890**
	Sig. (2-tailed)	.000		Sig. (2-tailed)	.000		Sig. (2-tailed)	.000
	N	70		N	70		N	70
x10	Pearson Correlation	.921**	x22	Pearson Correlation	.503**	x34	Pearson Correlation	.921**
	Sig. (2-tailed)	.000		Sig. (2-tailed)	.000		Sig. (2-tailed)	.000
	N	70		N	70		N	70
x11	Pearson Correlation	.809**	x23	Pearson Correlation	.850**	x35	Pearson Correlation	.891**
	Sig. (2-tailed)	.000		Sig. (2-tailed)	.000		Sig. (2-tailed)	.000
	N	70		N	70		N	70
x12	Pearson Correlation	.809**	x24	Pearson Correlation	.503**	Total	Pearson Correlation	1
	Sig. (2-tailed)	.000		Sig. (2-tailed)	.000		Sig. (2-tailed)	
	N	70		N	70		N	70

***. Correlation is significant at the 0.01 level (2-tailed).*
**. Correlation is significant at the 0.05 level (2-tailed).*

Overall all 35 instruments are valid through Pearson Correlation test, which are significant at the 0.01 level (α<0.01/1%). The calculation of the validation and reliability test above states that there are 35 items of statements that are declared valid theoretically and empirically, so that the performance-based assessment rubric can be used to assess the level of competence of students in wood, steel and scaffold at civil engineering workshop.

B. *Effect on Overall Competency Level*

After looking at the validity of the test through instruments'-based criteria, we then look at the fundamental purpose of the assessment itself, which is to improve the students' competency level at the workshop through a better understanding of the required skills and examinations.

TABLE IV. PAIRED SAMPLES CORRELATION

		N	Correlation	Sig.
Group 1	After & Before	350	0.411	0.000
Group 2	After & Before	350	0.184	0.001

We used the paired samples correlations to see the relationship between the assessment and the score the students had for the two groups which were the group 1 (who had not gone through the course with a specified assessment explanation period) and group 2 (who had gone through the explanation and performance-based skills period). We had positive correlation for both group which means that all the students will likely to improve their score from the pre-test in the post-test assessment. Group 1 had 0.411 correlation while group 2 had 0.184 which is shown in Table 4.

Moreover we also looked at how much did they improve over the course of workshop in which we had seen the two group treated differently. The two group exactly has the same number of students with also three classes on both sides. From a simple paired samples statistics we could see a similar mean of pre-test score from both group which were 55.17 for the control-group and 55.28 for the treated group as seen in Table 5.

TABLE V. PAIRED SAMPLES STATISTICS

		Mean	N	Std. Deviation	Std. Error Mean
Group 1	Before	55.17	350	4.995	0.267
	After	58.15	350	6.471	0.346
Group 2	Before	55.28	350	4.999	0.267
	After	79.67	350	14.703	0.786

It is important to look at the pre-test score and to see that it has quite a similar mean and standard deviation, we could safely say that they are at the same level even before the

course was given. Therefore we could quite see the difference in these two groups after the treatement on the post-test which then is reflected in Table 6 below.

TABLE VI. PAIRED SAMPLES TEST

		Paired Differences					t	df	Sig. (2-tailed
		Mean	Std. Deviation	Std. Error Mean	95% Confidence of the Difference				
					Lower	Upper			
Grup 1	After-Before	2.980	6.345	0.339	-3.647	-2.313	-8.787	349	0.000
Grup 2	After-Before	24.397	14.633	0.782	-25.935	-22.859	-31.192	349	0.000

From the paired samples test from group 1 and 2 we could see the paired difference from group 1 was 2.980 which was significant at 5% level, and from group 2 was 24.397 (significant at 5% level). This reflect the strong performance of group 2 after treatment was given, which was to give clear understanding of the skills required during the assessment and bringing better results and competency in the end from Table 6.

Also from Fig. 1 we could see the illustration of results from pre-test (time frame 1) and post-test (time frame 2) that the superior score results is seen. The improvement of 20.4% is seen from an overall maximum score of 105 in those 350 students.

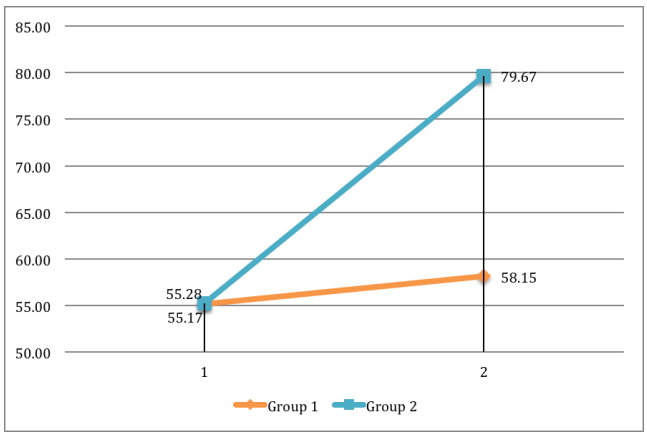

Fig. 1. Paired Difference Test Results In Graphic

IV. CONCLUSION AND SUGGESTIONS

Based on the results, it is concluded that the validity of the instruments used in the assessment are further analyzed through the within effect paired difference treatment analysis in performance-based practicum assessment instruments on timber workshop at Civil Engineering Department in Kupang State Polytechnic. The instrument with the assessment rubric consists of 35 statement items, and after an internal validation test (expert test), the result is that 35 items are declared valid with CVR value> 0, the CVR value ranges between 0.75 - 1.00 and CVI = 0.80. Based on this value it can be concluded that the instrument in the form of an assessment rubric is very appropriate and can be used to conduct an assessment [11]. The results of the external consistency reliability test were analyzed using the Alpha Cronbach formula, obtained by the reliability coefficient value of 0.981 classified as very high reliability. Validity test shows that 35 instruments are valid at a theoretical significance level of 1%. In addition, the overall score increased by 20.4%. It was concluded that an improved assessment method using new instruments had increased students' competencies during the workshop through superior test results. Therefore, it is recommended that this new instrument for assessment be used in future workshops.

ACKNOWLEDMENT

The authors would like to thank all those who supported this research, at the Bali State Polytechnic so that this research can be completed properly and on time, thanks also for the support of LP3M PNB for their guidance and cooperation regarding the development of this paper. We would like to thank the Indonesian Ministry of Research and Technology for the funds that have been given for the completion of this research.

REFERENCES

[1] S. Toohey, G. Ryan, C. Hughes, "*Assessing the practicum,*" Journal Assessment & Evaluation in Higher Education, 1996, 21(3)

[2] D.D. Stevens, A.J. Levi, "Introduction to rubrics: an assessment tool to save grading time, convey effective feedback, and promote student learning," Sterling, VA: Stylus, 2005.

[3] D.L. Gabel, "Handbook of research on science teaching and learning," New York: Maccmillan Company, 1993.

[4] MOE, "Indonesian minister of education regulation number 20 year 2007," Jakarta, 2007.

[5] J. Arter, J. McTighe, "Scoring rubrics in the classroom: Using performance criteria for assessing and improving student performance," Thousand Oaks, CA: Corwin Press, Inc, 2001.

[6] M. Wilson, "Rethinking rubrics in writing assessment," Portsmouth, NH: Heinemann, 2006.

[7] S. Brookhart, A. Nitko, "Assessment and grading in classrooms," Upper Saddle River, NJ: Pearson Education, vol. 1, 2008, pp 201.

[8] KSP, "Civil engineering workshop teaching manual," Politeknik Negeri Kupang, 2013.

[9] D. Rubio, M. Berg-Weger, S. Tebb, E. Lee, S. Rauch, "Objectifying content validity: conducting a content validity study in social work research," Social Work Research, Vol 2, 2003, pp 27:94-104.

[10] H. DeVon, M.E. Block, et al, "A psychometric toolbox for testing validity and reliability," J Nurs Scholarship, vol 2, 2007, pp:155–64.

[11] M.R. Lynn,"Determination and quantification of content validity," Nurs Res, vol 6, 1986, pp:35:382–5.

[12] C.F. Waltz, O. Strickland, E.R. Lenz, "Measurement in nursing and health research," 4th ed, New York: Springer Publishing Company, 2010.