Development of Performance Assessment of Learning Instruments Apply Daily Make-Up for Make Up Department at Vocational High School

Mausa Agrevinna✉, Ani Rusilowati, Bambang Sugeng Suryatna

Universitas Negeri Semarang

Article Info	Abstract
	This research has not been effected by the presence of learning performance assessment instruments apply daily makeup standards, haven't tested the validity or reliability. Moreover the assessment instrument Competency Skills Exam of BSNP one indicator is less detailed, less range and scale criteria in section is too long so that the teachers feel difficulty in judging. The purpose of the research is to produce a performance assessment of learning instruments apply daily makeup valid, reliability and practical. This research is the development of research conducted at SMK Perwari Kendal, SMKN 6 Semarang, SMK Taman Siswa Kudus, SMK PGRI 1 Mejobo Kudus. The subject of this research is the 95 respondents. Test the validity of the content using the formula of Aiken's V and test reliability using the ICC. At trial, reliability, analyzed using Cronbach Alpha. Test the validity of invalid constructs using Confirmatory Factor Analysis (CFA) was formed of 3 factors i.e. factor 1 (pra practice), a factor of 2 (practice) and factor 3 (post practice). Test the practicality of the instrument indicates that the performance assessment instruments are very practical for use. The profile shows the performance capability of the students has the capability of excellent performance. Based on the results of the analysis, this research produces a valid performance assessment instruments, reliability and practical. The benefits of this research is a guide that can be used raw in measuring student performance in learning daily make-up.

✉Correspondence Addres:
Kampus Pascasarjana Jl. Kelud Utara III, Semarang 50237
E-mail: mausaagrevinna@gmail.com

INTRODUCTION

Thobroni and Arif (2012:18) defines learning is a change in behavior is relatively fixed and is the result of repeated practice. Apply daily makeup practices is a basic competency in the subjects Basic Beauty skin on Skin Beauty Department. Basic subjects Beauty Skin i.e. the subjects decide to go to skin beauty subjects, so that these very important subjects to master students. Apply daily makeup practices can measure student skills at the same time can improve the performance of students at the time of practice. Apply daily makeup practices have learning objectives, namely: (1) knowledge to understand the basic concepts apply daily makeup. (2) showing the attitude in practice apply daily makeup. (3) showing performance through capabilities. It shows learning apply makeup daily covers the whole realm of learning cognitive, psychomotor or affective. With regard to the third realm, learning to know the results of the study then conducted an assessment of performance.

According to Mardapi (2012:19-20) performance assessment is the process of collecting data to make decisions by way of observation systematically. Then in the process of assessment of practice required reference criteria assessment principle that has an open, objective, integrated, sustainable and fair.

Based on the observation at SMK PGRI Mejobo Kudus, the teacher has not informed the aspects that will be assessed in practice, so the teacher pay less attention to three aspects, namely pre-practice, practice and post-practice, so that the assessment of daily makeup practice is *holistic*.

This is evidenced based on the observation of the researchers during the Field Work Practice, the assessing teacher has not used the instrument sheet with the rubric, so the assessment process is subjective or according to teacher's perception. Similarly, based on the results of observations and interviews with productive teachers in SMK PGRI 1 Mejobo Kudus Beauty Skincare expertise on December 5, 2017, that the instrument of learning practice daily makeup faces used by teachers has not been there. During this time the teacher uses a self-made assessment system, so has not been tested the validity and reliability.

In addition to the assessment of learning, the results of interview researchers with teachers SMK PGRI 1 Mejobo Kudus dated December 5, 2017 that the assessment instrument of BSNP to assess the competency test students SMK, there is one indicator that is not detailed, for example make makeup with proper techniques. The meaning of the word "exact" is not yet detailed. In addition, the assessment instrument of the UKK (Skills Competency Expert) of BSNP uses a less strict span of scale, for example: for the achievement of competent competencies using the range scale of 7 , 0 -7.9, so that the assessor finds difficulty in determining the value. In addition, the criteria in the rubric is too long so that teachers feel difficulty in assessing.

In Law Number 20 Year 2003, Article 58 paragraph (1) explained that the evaluation of the learning process and learning outcomes of students in a systematic and continuous, so it is necessary to prepare a standardized assessment instrument sheet, so it can be stated that the instrument can affect the validity of the assessment technique and can affect the prevalence and in the objectives of the assessment that is informing the learning achievement of the learners. Therefore, the assessment instrument becomes one of the most important part in the process of student performance learning, so that the researcher is interested to develop a valid, reliable and practical performance appraisal instrument to assess the daily makeup learning for the department of Beauty in SMK. Based on the description of the background of the problem, then formulated the following problems: (1) How are the characteristics of the assessment instrument of learning performance make up daily faces developed ?; (2) How does the validity and reliability of the performance appraisal apparatus make up the faces daily ?; (3) how is the practicality of the guidebook of the assessment of learning performance make up daily faces ?; (4) What is the profile of students' performance ability in learning face daily makeup ?.

A good instrument is required of standard instrument development steps, as proposed by Borg & Gall with 9 modified steps, namely: (1) introduction; (2) planning; (3) early product development; (4) initial product trials;

(5) improvement of the initial product; (6) broad field trials; (7) product improvement; (8) product practicality test; (9) the final product.

Collision (2014) defines a ses- sion as a systematic procedure for collecting information but can be used to make the conclusion of the characteristic of a person or an object . The essence of the assessment is to interpret or interpret the measurement data. In interpreting the data the measurement results can be compared with the standard type. Regulation of the Minister of National Education number 20 of 2007 on the national assessment standard provides a benchmark in the assessment of learning in schools. Based on the regulatory assessment policy, the school and its teachers have the authority to create and develop a learning assessment.

Stiggins (1994) defines an assessment performance is a form of test in which the student is asked to perform special activities under the supervision of the examiner (teacher) who will observe his performance and make decisions about the quality of the learning he has shown .

Suharsimi (2006: 26) defines instrument assessment instruments also called evaluation tools is something that can be used to carry out the task or melakasanakan goals more effectively and efficiently. In the activity evaluation tool function also to obtain better results in accordance with the reality being evaluated .

The development of instruments is an activity to develop existing instruments to be better than ever valid and reliable. Development of this instrument becomes important in determining the quality or information of a public service, especially in puskesmas on the quality of services provided. Development of the instrument must be through the stages of good development in order to obtain a quality instrument (Rusilowati, 2013).

According to Kustanti (2008: 452), makeup is the art of decorating the face, covering the existing deficiencies in the face and make the appearance becomes more *fresh* so the purpose of makeup to beautify and enhance the appearance of the face. According to Maya and Esther (2014) , the purpose of daily makeup is to (a) correct the face shape and the near-ideal face (b) the face looks fresh , beautiful and healthy. Facial makeup consists of various types,

including daily makeup, makeup, stage makeup, makeup cikatri, makeup geriatri, makeup, and bridal makeup . However, makeup every day is divided into two: morning makeup and night makeup.

METHODS

This research is a development research that produces the instrument of performance appraisal of daily makeup make up which is manifested in manual form.

The research was conducted at SMK Perwari Kendal, SMKN 6 Semarang, SMK Taman Siswa Kudus and SMK PGRI 1 Mejobo Kudus during April-May 2018. Sampling technique using *random sampling* .

This development model using a modified Borg and Gall consisting of 9 steps are modified, namely: (1) introduction; (2) planning; (3) early product development; (4) initial product trials; (5) improvement of the initial product; (6) broad field trials; (7) product improvement; (8) product practicality test; (9) the final product.

The data source, data collecting techniques used in this study with interview techniques, observation and document studies. Test method of content validity using expert valuation method and analyzed with Aiken's V formula while reliability from expert study using *Intraclass Correlation Coeficient* (ICC) analysis . In small-scale trials, reliability was analyzed using the *Cronbach Alpha* method , while the instruments were tested for readability using a legibility questionnaire. In large-scale trials of construct validity was analyzed using analysis *Confirmatory Factor Analysis* (CFA). Testing instrument reliability using *Alpha Cronbach* method . In addition to testing validity and reliability, developed instruments will also be tested for practicality in their use. In addition, from small-scale pilot analysis and large-scale trials, the students' performance profiles were analyzed.

RESULTS AND DISCUSSION

The results of this study consisted of instrument characteristics, content validity test, construct validity and reliability of the instrument, instrument practicality test and

performance profile of students' performance in daily makeup practice.

Characteristics of Performance Appraisal Instruments

The instrument is designed from the preparation of the instrument grille based on the theory of daily facial makeup, the performance appraisal instrument designed for use by SMK teachers. The grid begins from the syllabus of the competency standard and then downgraded to the basic competencies, so that the indicators appear to be assessed. The instrument of appraisal of learning performance make up daily face in the form of an observation sheet consisting of 22 indicators derived from the stages of pre-practice, practice and post-practice process. The pre-process stage consists of 6 indicators, while the practical process stages consist of 12 indicators, while the post-practice process stages consist of 4 indicators, so the total is 22 indicators / observations. Grading scale used in the assessment tool in the form of sheets of observation is a scale range, where the highest score of 4 and the lowest score is 1.

Expert Review, Small Scale Test, Large Scale Test, Practicality Test and Performance Capability Profile

1. Expert Review

Performance appraisal instruments that have been developed , then reviewed by 3 experts who are competent in the field of research, education evaluation and in the field of beauty. Experts rate every grain of a performance appraisal instrument, with scores ranging from 1-4. Then, the results of the three experts, were analyzed using Aiken's V formula. The results of Aiken V analysis can be seen in Table 1.

Table 1. Results of Content Validity Analysis Aiken V Performance Appraisal Instrument

No Item	The value of Aiken V	Criteria
1	1.00	Valid
2	0.78	Valid
3	1.00	Valid
4	1.00	Valid
5	0.67	Valid
6	0.67	Valid
7	0.67	Valid
8	0.78	Valid
9	0.56	Valid
10	0.67	Valid
11	1.00	Valid
12	0.67	Valid
13	1.00	Valid
14	0.67	Valid
15	0.78	Valid
16	0.89	Valid
17	0.67	Valid
18	1.00	Valid
19	0.33	Valid
20	0.67	Valid
21	1.00	Valid
22	0.89	Valid

In Table 1, the result of the analysis yields all the items (22 items) are valid. Because r ≥ r critical 0.3 . This is in accordance with the criteria proposed by Azwar (2014: 34) that if the validity coefficient ≥ 0.3 means the item is declared valid.

After passing the validity test, then performed the reliability test using ICC approach. Before using the ICC approach, reliability values are estimated by Alpha coefficients. Based on the analysis show the coefficient value $r_{xx} = 0.899$, means the coefficient value $r_{xx} > 0.7$ so that can be continued analysis of *Intraclass Correlation Coeficient (ICC)*. According Suharsimi (2008: 75) assessment instrument is said to be reliable if $r_{xx} > 0.6$. Based on the calculation of ICC values seen from single measures 0.747, thus the instrument of assessment of learning performance daily makeup face is said to be reliable and ready to use. This is in line with Susanti's research (2015) which obtain that the reliability coefficient is in the high category (> 0.6) so that the instrument is declared reliable.

2. Small Scale Trial

Grains were tested on a small scale of 22 grains, with a sample of 40 respondents, consisting of 10 students from SMK Perwari Kendal, 10 students from SMKN 6 Semarang, 10 students from SMK Taman Siswa Kudus and 10 students from SMK PGRI 1 Mejobo Kudus. Test the reliability of the performance appraisal appraisal instrument on the face of everyday life on small-scale trials using *Alpha Cronbach* through SPSS 16.0 program. Estimated instrument reliability in small scale tests with Alpha values $r_{xx} = 0.894 ≥ 0.7$. This means that the instrument has a high degree of consistency. Furthermore the resulting data is said to be

feasible or can be continued to be tested for validity if it meets the requirements of KMO MSA ≥ 0.5. Terms are met can be continued to see the value of correlation between multivariate variables can be seen in the Output Anti Image Correlation in Table 2.

Table 2. Output of Anti Image Correlation

No	Item	Anti Image Correlation
1	Creating jobsheet	0.472
2	Make personal preparations	0.507
3	Preparing work areas, tools	0.368
4	Preparing of materials and cosmetics	0.609
5	Conduct customer/ client preparation	0.460
6	Communicate service services	0.893
7	Diagnose skin type and face correction	0.826
8	Perform facial cleansing and refreshment	0.575
9	Apply moisturizer	0.718
10	Apply foundation	0.956
11	Apply shading	0.761
12	Apply powder	0.570
13	Apply solid powder	0.788
14	Make eyebrows	0.690
15	Apply eye shadow	0.514
16	Apply pencil eye liner	0.672
17	Apply blush on	0.647
18	Apply lipstick and lipgloss	0.643
19	Time allocation	0.336
20	Work result	0.306
21	Work attitude	0.942
22	Packing	0.614

Table 2 indicates that items 1, 3, and 5 have corelation values of anti-image <0.5 are not eligible for further test because the variables are unpredictable and can not be further analyzed or removed from other variables, so the instrument must be revised. The next step is to match the results of the analysis with the test legibility. Test this legibility through a questionnaire filled by teachers daily makeup makeup. 8 teachers who fill the questionnaire of the legibility of the instrument, the analysis of the legibility test can be summarized for improvement, from the pre-practice aspect of indidkator 1 is short, the pre-practice aspects of the indicator 3 are less assertive and the pre-practice aspects of the indicator 3 are less assertive, after practice indicators 1 the criterion is confusing and post-

practice aspects of the 2nd indicator are not considered assertive, so the instrument needs to be improved.

3. Large Scale Trial

Grains were tested on a large scale of 22 grains, and a sample of 95 students, consisting of 12 students from SMK Perwari Kendal, 53 students from SMKN 6 Semarang, 14 students from SMK Taman Siswa Kudus and 16 students from SMK PGRI 1 Mejobo Kudus. Large-scale trials were analyzed using construct validity and *Alpha Cronbach* reliability . Valid data based on confirmatory factor analysis procedure, if data qualify *Keizer-Meyer-Olkin Measure of Sampling Adequacy* (KMO MSA)> 0,5 Based on large scale field trial data obtained KMO value> 0.5 that is 0,847 and *Chi-Square* sig value 0.000, so that from sample sufficiency have been fulfilled, so can be continued further test. Sig. 0.000 ≤ 0.05 can be called any correlation between multivariate variables. The following table of Anti Image Correlation can be seen in Table 3.

Table 3. Results of Anti Image Correlation

No	Item	Anti Image Correlation
1	Creating jobsheet	0.812
2	Make personal preparations	0.784
3	Preparing work areas, tools	0.718
4	Preparing of materials and cosmetics	0.815
5	Conduct customer/ client preparation	0.746
6	Communicate service services	0.820
7	Diagnose skin type and face correction	0.866
8	Perform facial cleansing and refreshment	0.939
9	Apply moisturizer	0.905
10	Apply foundation	0.934
11	Apply shading	0.880
12	Apply powder	0.933
13	Apply solid powder	0.909
14	Make eyebrows	0.867
15	Apply eye shadow	0.922
16	Apply pencil eye liner	0.908
17	Apply blush on	0.970
18	Apply lipstick and lipgloss	0.957
19	Time allocation	0.653
20	Work result	0.901
21	Work attitude	0.742
22	Packing	0.762

Based on table 3 column *anti image correlation* it can be seen that the value of correlation between grains obtained correlation value> 0.5 therefore factor analysis can be continued by including all items.

The next step is to see how many factors may be formed in the factor analysis with the number of samples 95 *through the total variance explained* , after confirmatory factor analysis and 3 components are formed, of 22 items are 6 point 1, 12 point 2 and 4 buti r factor 3. The number of factors can be seen in Figure 1 .
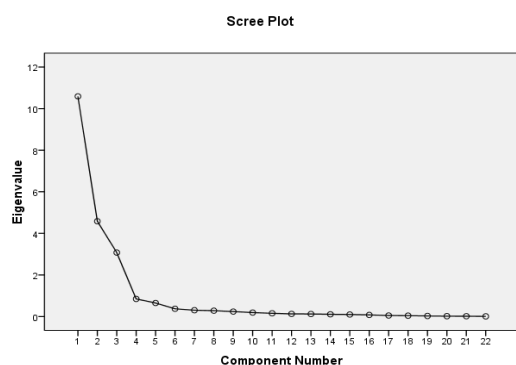
Figure 1. Scree Plot

From Figure 1 the scree plot from point 1 to 3 decreases eigenvalues but is still above 1, whereas point 4 and so on eigenvalues is below 1, so that 3 factors are formed ie factor 1 (pre-practice), factor 2 (practice) and factor 3 (post-practice).

Reliability measurement results obtained value Alpha = 0.906 ≥ 0.7, so it can be concluded that the performance appraisal instrument has a very high level of reliability on large-scale trials .

4.   Test of Practicality of Instruments

Practicality of the instrument shows the extent to which the instrument can be used well and easily by educational practitioners such as the subject of Basic Beauty Skin teachers learning daily makeup faces. The calculation results obtained information that the assessment instrument for all aspects has the lowest practicality score 46 and the highest 49 of 8 teachers. Based on the scoring criteria, the instrument of performance appraisal learning make up the everyday face is very practical to use.

5.   Performance Ability Profile

Based on the result of the analysis of the lowest practice values of the students on a small scale trial that is 70 and the highest value is 98. The value of practice is interpreted into percentages in Table 4.

Table 4. Percentage of Practice Value on Small Scale Trial

No	Student Value Interval	Frequency	Percentage	Categori
1	82-100	34	85	Very good
2	63-81	6	15	Good
3	44-62	0	0	Enough
4	25-43	0	0	Less

Percentage of practice value obtained 34 students have excellent performance ability category with percentage of 85%, while 6 students have good performance ability category with percentage 15%.

Based on the results of the lowest student practice score on large-scale trials is 72 and the highest score is 100. Practice values are interpreted into percentages in Table 5.

Table 5. Percentage of Practice Value on Large Scale Trial

No	Student Value Interval	Frequency	Percentage	Categori
1	82-100	77	81	Very good
2	63-81	18	19	Good
3	44-62	0	0	Enough
4	25-43	0	0	Less

Based on Table 5 the percentage of practice score obtained 77 students have excellent performance ability category with the percentage of 81%, while 18 students have performance performance category either with percentage 19%.

Research conducted by Sarjono (2015) shows that performance appraisal includes the assessment of knowledge, behavior and interaction among learners through direct observation. Performance assessment is very influential on the achievement of competence in learning. In theory, performance assessments need to be done because it can see students'

ability during the learning process, in addition to helping teachers to easily observe and assess students in learning, thus can be obtained strengths and weaknesses of students in practice (Stiggins, 1994). The same thing is shown by Astuti (2015) that the results of the assessment by using the instrument is made to be able to inform comprehensively the performance of students during implementing practical learning.

CONCLUSION

Based on the results and discussion of research that has been described, can dita rik conclusion as follows:

Instruments generated in this study is a performance appraisal instrument to assess the ability of learning performance makeup facial everyday for students of class X SMK beauty majors with 22 items of valid and reliable instrument. The validity of the content of the expert judgments resulted in the whole grains (22 points) declared valid, with the validity of 0.79> 0.3 with the medium criterion and reliability of 0.747> 0.6.

Small-scale trials were tested on 40 respondents from SMK Perwari Kendal, SMKN 6 Semarang, SMK Taman Siswa Kudus and SMK PGRI 1 Mejobo Kudus. Reliability on small-scale trials using the reliability test *Alpha Cronbach* otherwise reliable with $r_{xx} = 0.894 > 0.7$. The next step is the test legibility, from 8 teachers who fill the questionnaire of the legibility of the instrument, it can be concluded the instrument is improved with details of pre-practice aspects of indicator 1 is too short, the pre- practice aspects of indicator 3 have less assertive aspects, the practice of indicator 5 has less assertive aspects, practice indicator 1 has a confusing criteria of assessment and post-practice aspects of indicator 2 have a less assertive aspect.

Large-scale trials were tested on 95 respondents. Test the validity of constructs using CFA (*Confirmatory Factor Analysis*). Items tested on a large scale of 22 grains . Based on large scale field trial data obtained KMO value> 0.5 is 0.847 and *Chi-Square* value of sig 0,000, so that from the adequacy of the sample has been met, so it can be continued further test. Sig. 0.000 ≤ 0.05 can be called any correlation between multivariate variables . S fter a confirmatory factor analysis and obtained 3 components are

formed, of 22 grains is a 6 point factor 1 (pre-practice), 12 grains of factor 2 (practice) and 4 grains factor 3 (after practice). Reliability measurement results obtained value Alpha = 0.906 ≥ 0.7, so it can be concluded that the performance appraisal instrument has a very high level of reliability on large-scale trials.

The results of the calculation of the practicality test obtained information that the assessment instrument for all aspects has the lowest practicality score 46 and the highest 49 of 8 teachers. Based on the scoring criteria, the instrument of performance appraisal learning make up the everyday face is very practical to use.

Profile of students' performance ability, on small-scale trial percentage of practice score obtained 34 students have excellent performance ability category with percentage of 85%, while 6 students have good performance ability category with percentage 15%. In the large-scale trial the percentage of practice score obtained 77 students have excellent performance ability category with the percentage of 81%, while 18 students have a category of performance ability either with percentage 19%.

THANK-YOU NOTE

REFERENCES

Arikunto, Suharsimi. (2006). *Prosedur Penelitian Suatu Pendekatan Praktik*. Jakarta: Rineka Cipta.

Arikunto, Suharsimi. (2008). Penelitian Tindakan Kelas. Jakarta: Bumi Aksara

Astuti, Widya Puji, Hari Wibawanto, M. Khumaedi. (2015). Pengembangan Instrumen Penilaian Unjuk Kerja Praktik Perawatan Kulit Wajah Berbasis Kompetensi di Universitas Negeri Semarang. *Journal of Educational Research and Evaluation*, *4*(2), 109–115.

Azwar, S. (2014). *Penyusunan Skala Psikologi*. Yogyakarta: Pustaka Belajar.

Depdiknas. (2017). Badan Standar Nasional Pendidikan (BSNP), Standar Penilaian PP No. 19/2005, Pasal 63-72, Kepmendiknas No.20/2007.

Kustanti, Herni. (2008). *Tata Kecantikan Kulit* . Jakarta PT. Gramedia Pustaka

Mardapi Djemari, (2012). *Pengukuran Penilaian & Evaluasi Pendidikan.* Yogyakarta : Nuha Medika

Mardapi, D.(2008). *Teknik Penyusunan Instrumen Tes dan Non Tes*. Yogyakarta: Mitra Cendekia Press.

Masrukan. (2014). *Asesmen Otentik Pembelajaran Matematika, Mencakup Asesmen Afektif dan Karakter*. Semarang: FMIPA Unnes.

Maya Dian Sari, Esther Restie Sipahutar. (2014). *Pengembangan Media Pembelajaran Interaktif Berbasis Blog Rias Wajah Sehari-Hari Tata Kecantikan Kulit*. Jurnal Teknologi Informasi & Komunikasi dalam Pendidikan 1(1), 63-74

Peraturan Menteri Pendidikan Nasional Republik Indonesia Nomor 20 Tahun 2007 tentang Standar Penilaian Pendidikan

Rusilowati. (2013). Pengembangan Instrumen Non Tes. *Makalah*. Seminar Nasional Evaluasi Pendidikan I di Universitas Negeri Semarang. Semarang, 2013.

Sarjono. (2015). *Penilaian Unjuk Kerja dalam Praktikum Fisika*. Jurnal Madaniyah ISSN: 2086-3462 VIII, 109-130.

Stiggins, R. J. (1994). *Student-centered classroom assessment.* New York: Macmillan Publishing Company.

Susanti. (2015). "Pengaruh Appointment Registration System terhadap Waktu Tunggu dan Kepuasan Pasien". Global Medical and Health Communication, 3(1): 40-47.

Thobroni, Muhammad dan Arif Mustafa. (2012). *Belajar dan Pembelajaran.* Yogyakarta: Ar-Ruzz Media

Undang-Undang Sistem Pendidikan Nasional No. 20 Tahun 2003 Pasal 58 Ayat 1