JERE 7 (1) (2018) 44 - 51

Journal of Educational Research and Evaluation

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

Development of Performance Assessment Inspection Visual of Acetic Acid (Iva)

Tri Hariyati✉, Sri Ratna Rahayu, M. Khumaedi

Universitas Negeri Semarang

Artikel Info

Article History:
Accepted 17 February 2018
Approved 10 July 2018
Published 10 March 2018

Keywords:
Pemeriksaan IV, pengembangan, penilaian, unjuk kerja

Abstract

Instruments that have valid and reliable scoring guidelines will provide learning evaluation results that match the achievements of social science lesson competencies. This study aims to determine the validity and reliability of the standard essay test instruments to measure the higher order thingking skills of social science of the 8th grade junior high schools developed. This study uses research and development techniques to produce standard essay test instruments using 30 samples on a small-scale test and 105 samples on a large-scale test taken at random. The study used nine steps of instrument development that is 1) compile test specification; 2) write test questions; 3) reviewing test questions; 4) test the test; 5) analyze the items; 6) improve the test; 7) assemble the test; 8) carry out tests; 9) interpret the test results. Data collection techniques used interview techniques, document studies, and tests. The results showed that the instrument has Aiken's V> 0.3, the mean value has a value of 0.8. Estimation of instrument reliability at small scale test is 0.788, large scale test is 0.819 and final test for 12 items is 0.918. The results showed that standard essay test instruments to measure higher order thingking of social skills of junior high school in 8th grade proved to be valid and reliable.

✉ Alamat korespondensi:
Kampus Pascasarjana Jl. Kelud Utara III, Semarang 50237
E-mail: trihariyati94@gmail.com

P-ISSN 2252-6544
E-ISSN 2502-356X

INTRODUCTION

Based on Permenkes No. 29 years 2017 to protect communities from the threat of breast cancer and cervical cancer it is recommended that any women or couples of male sufferers of IMS/ISR are encouraged to conduct an examination of the IVA or papsmear as well as Clinical Examinations in ecstasy (SADANIS). One of the competencies of midwives in the prevention of cervical cancer is an examination of IVA, so that students of midwifery practice in before entering a field already equipped and provided with the evaluation concerning the practice in examination of the IVA. Midwifery students ever provided with sheet Checklist to ease in doing practice examinations IVA. Students pitch in to the place before practice, trained skills in laboratory campus in advance because students may not take care of patients/clients directly before they are declared of the Learning Lab.

Assessment is a process which is very important to learning. The assessment instrument that meets the standards, will measure precisely the end result of a process of learning, so that students can study results will be detected properly and can be used as material for the next learning program evaluation (Juniarta, 20 16, p. 1). This is in line with the Sukestiyarno, and Budi Wachyudi Waluya (2014, p. 217) learning outcomes Assessment or assessment of performance are expected to assess the competence of learners thoroughly. The competence of graduates that includes aspects of affective, cognitive and psychomotor must be developed through the process of learning or assessment. Performance appraisal as an assessment that was done by observing the activities of the students in doing something, so it is considered more authentic performance assessment (Rusilowati, 2013, p. 18). According to Azwar (2011, p. 2) States a good instrument is an instrument that is able to provide reliable information. Measuring instrument or instrument has certain criteria so that the instrument can be called as good or worthy. These criteria are valid, reliability, standards, economical and practical.

Issues related to assessment, Suyanto in Mulyani (2012:265) said that until the current education in all levels in the implementation of the assessment is still more concerned with aspects of cognitive, affective aspects while still would be abandoned. To address the problems related to the assessment of such as have been expressed above, it is important to develop a comprehensive assessment instrument that measures the realm of affective, psychomotor and cognitive during and after the learning process .

According to Mardapi and Andayani (2012, p. m.-1) assessment is the process of forming judgments about the quality and the level of achievement or performance of students. Assessment is no longer merely a separate test and is after the culmination of learning, but is seen as part and parcel of the integration of the learning process.

In realizing the student that has a good basic competence, then needed a performance appraisal have high levels of validity and reliability. If the instrument is used do not have the aspect of performance is judged complete and obvious or does not have the validity of an instrument, as a result of assessment is possible can be subjective, that is, depend on the perception of the lecturer, whether the student's practice is in compliance or not. Even assessments are not implemented in stages in accordance with the steps that must be performed to assess the competence of students ' performance. The reality of the field, assessment instruments used for this is not yet done test validity and reliability in order to find out whether the instrument already meets the standards are good or not.

The purpose of this research is to get an assessment on the performance of students in the examination of midwifery IVA a valid and reliability. So that the development of these instruments is expected to conduct a

comprehensive assessment and can improve the quality of learning and can be used as a reference to making the assessment of student performance in midwifery practices contained in the institution.

METHODS

This study uses to design research development expressed by development model Djemari Mardapi there are ten steps to follow in developing an instrument. (Mardapi, 2016, p. 132). Development procedure in this study refers to 9 out of 10 step development Mardapi year 2016 are modified into 3 main stages, as for the stage 3 preliminary Stage, namely the stage of development, and evaluation Stages.

The source of the data onto this study wes obtained from initial observations, studies and data validation results document experts as well as test result data in the field. The subject of this research is a student of the Academy of Midwifery of the University of Muhammadiyah Semarang and Kudus and Midwifery Academies STIKES Widya Husada Semarang that add up to 110 people. Sample selection technique in this study is either using the technique of total sampling.

Data collection techniques used with interviews, the study documentation. An examination of the instruments by expert or validity of the contents were analyzed using formula V, Aiken's empirical validity of dsn using exploratory factor analysis, while the reliability of the instrument based on the agreement on the ahi with analysis using Two Way Anava and analyzed again by using the formula for Hoyt. reliability of the instrument based on the results of the field is analyzed using Cronbach Alpha with the help of the software SPSS version 16.

RESULTS AND DISCUSSION

Preliminary Studies

Based on the preliminary results of a study which has been carried out using the method of observation, the researcher interviews, documents and studies at the Academy of Midwifery University of Muhammadiyah Semarang and Kudus and Midwifery Academies STIKES Widya Husada Semarang, third The midwifery academies don't have raw judgment that is untested in validity and reliability to assess student performance in the examination of the existing instruments, IVA does not yet have a clear assessment rubrics in the instrument so it is not yet the perception of equality between assessors in assessing inspection IVA. Examination performance assessment IVA developed shaped sheets of observations with the amount of grain as much as 21 grain have three alternatives score i.e. (3) good, (2) and (1) is not good.

The Validity And Reliability Of Contents Based On The Assessment Of Experts

Based on the results of the validation of the contents of the four experts, obtainable score assessment instruments and then analyzed using a formula for Aiken's v. If the validity coefficient $\geq$ 0.30 means items can be said to be valid (Azwar, 2014:34). The results of the validity of the content based of expert assessments are presented in table 1.

Table 1. The Results Of The Validation Of Performance Assessment Of Aiken

No Grain	Expert 1	Expert 2	Expert 3	Expert 4	∑ S	V Index	Description
1	4	3	4	4	15	0,916667	Valid
2	4	3	4	4	15	0,916667	Valid
3	4	4	4	4	16	1	Valid
4	4	4	4	4	16	1	Valid
5	4	3	4	3	14	0,833333	Valid
6	4	2	4	4	14	0,833333	Valid
7	4	3	4	4	15	0,916667	Valid
8	3	3	3	3	12	0,666667	Valid
9	3	3	3	3	12	0,666667	Valid
10	4	3	4	4	15	0,916667	Valid
11	4	3	4	4	15	0,916667	Valid
12	3	3	3	3	12	0,666667	Valid
13	3	3	3	3	12	0,666667	Valid
14	4	4	4	4	16	1	Valid
15	3	3	3	3	12	0,666667	Valid
16	4	3	4	3	14	0,833333	Valid
17	3	3	3	3	12	0,666667	Valid
18	3	3	3	3	12	0,666667	Valid
19	3	4	4	4	15	0,916667	Valid
20	4	4	4	4	16	1	Valid
21	4	4	4	4	16	1	Valid

Validity of the test results by using the formula for Aiken's V indicates that an overall performance assessment of grain 21 rated has a coefficient of Aiken (> 0.30) means assessment of performance examination of the validity of the content have IVA either.

After performing the validation of content of experts, further research results to calculate the level of agreement with the four experts by using reliability test for consistency between appraisers that use the test analysis on the difference between Anava (two two-factor way anava) and further demonstrated through the analysis of a Formula to calculate the coefficients with Hoyt reliability. Two Way Anova analysis results are presented in table 2.

Table 2. The Results Of Calculation Of Reliability Two Way Anova

Tests of Between-Subjects Effects

Dependent Variable:Skor penilai

Source	Type III Sum of Squares	Df	Mean Square
Appraisers	2.381	3	.794
Grain	13.452	20	.673
Appraisers * Grain	7.119	60	.119

Based on the results of a test of reliability testing conducted through two way anova using SPSS 16.0, pointed out that the magnitude of variance between ratter that is expressed by the Mean Square Marker i.e. S2r = 0.119 while the Error variance is expressed by the Mean Square Appraisers * Grains namely S2s = 0.673. The results are calculated using the formula yields the value koefesien Hoyt reliability of 0.82 means inters ratter rate agreed against the suitability of this content and also indicates that the awarding of the score given by each ratter is consistent. In line with research done Astuti, day Wibawanto, Khumaedi (2015, p. 114) results of reliability according to experts point out that for the assessment instrument, jobsheet 22 massage and movement assessment instrument performance indicates that no There are significant differences between the five experts in assessing the content of the assessment instrument performance facial skin care practices.

The Feasibility Of An Instrument Based On The Assessment Of Experts

Testing the validity of content is done by experts not just assess the validity of the content, but in this case the experts are also asked to assess the appropriateness of the assessment of performance examination IVA already developed. The results of the assessment provided by the experts on the feasibility assessment of performance against the inspection percentage amount obtained IVA totals of 94%, so that the assessment instrument performance examination test IVA is said to be very viable and can be used Although with a little revision. Feasibility tests results of table 3 below

Table 3. Interpretation of the Eligibility Criteria

the presentation interval	Description
$0 \geq X \leq 25$	Very not eligible
$25 > X \leq 50$	Not eligible
$50 > X \leq 75$	Eligible
$75 > X \leq 100$	Very decent

This is in line with research done Nugroho, Djuniadi & Rusilowati (2016, p. 4) based on the results of the analysis of the validity of the content and reliability of the expert's agreement shows the nineteenth overall performance assessment instruments this grain already have a value above criteria can be said to have content validity assessment instruments are adequate and have a high level of benefiting from expert and can be used to assess the performance of students in practice to draw the pieces. The Validity And Reliability Of Assessment Invalid Constructs Performance Examination Of IVA Being Developed Based On The Results Of The Test In The Field.

The next stage tested the first instrument in the field, analyzed using factor analysis to look at the correlation value Exploratori every grain, from the results of the analysis showed KMO > 0.5 i.e. of 0.628, data are there already meets the criteria to be analysed further. The results of the analysis of the test of the field the first 4 items found indicate a value that does not meet 0.5 < criteria and cannot be analyzed further. Items have a value of which 0.5 2 grain < value 0.418, 9 grain value 0.459, 12 grains with a value of 0.358 and 13 rounds with a value of 0.487.

Reliability analysis tests results first instrument shows the value of the coefficient Alpha of 0628 0.70 then arguably < not reliability, so do the revision or repair then analyzed back on test on the course II. After a test instrument in the field first and do the repairs the instrument the next step i.e. do a test instrument in the field both on the overall sample as many as 110 midwifery students. The results of the analysis of the test of the feasibility of the instrument on the Test in field II presented in table 4.

Table 4. Test the feasibility of an instrument on the Test field II

Kaiser-Meyer-Olkin Measure of Sampling Adequacy	.904
Bartlett's Test of Sphericity	sig. .000

Based on the results of the analysis of the test instrument in the field of both retrieved value of KMO 0.904 0.000 significance with which means it already meets the criteria of ≥ 0.5, then can deduce the overall 21 grains developed can be analyzed more.

Next to be seen anti-Image analysis results Matrics to see correlations of grain on table Anti Image Matrics, especially at the bottom of the Anti Image Correlation number of digits forming the diagonal MSA marked with "a" (the direction of the diagonal from top left to bottom right), not found indicators of correlation under 0.5, so the results of the analysis of these instruments is already qualified to do an analysis of the factors. The following is a table of Anti

Image Correlation that can be seen in table 5. under this

Table 5. The Results Of Anti Image Correlation

No	Grain	*Anti Image Correlation*
1	Welcome patients by greeting	0.777
2	Introduce yourself politely and friendly	0.779
3	Ask the patient's bio	0.974
4	Ask a history of obstetric history previous screening tests to Asks	0.936
5	Ask the perceived symptoms	0.949
6	Ask the main complaint	0.962
7	Describe the screening tests IVA	0.954
8	Explain the procedure in the action to be performed	0.951
9	Do Informed consent (consent of patient)	0.965
10	Respond to the patient's reactions to a good facial expressions	0.981
11	Respond to the patient's reactions to the patient eye gazing	0.825
12	Set up tools	0.898
13	Prepare materials	0.899
14	Setting up a room that will be used	0.875
15	Menyiapakan patients before examination	0.894
16	Officers prepare	0.882
17	Action inspection of IVA	0.898
18	Do post inspection action IVA	0.613
19	Take action systematically	0.600
20	Maintain the privacy of the patient	0.577
21	Welcomes patients by greeting	0.568

Based on table 5. on columns anti-viral image looks that the value of the correlation between grain gains to value correlation > 0.5 therefore analyses the factors can be passed on inviting all the grains.

Based on the results of the analysis of the factor analysis value by looking at exploratory Total Variance Explained suggests there are 11 items put into the analysis factors summarized into five groups of factors by looking at the values of the characteristics The Eigenvalues >1. The results of Total Variance Explained are depicted in Figure 1. Scree Plot.

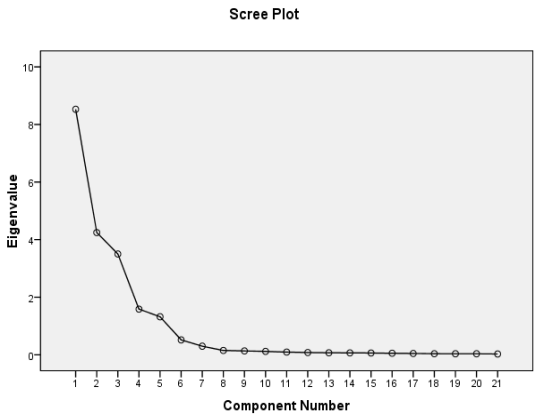

Figure 1. Scree Plot

Based on the Scree Plot above it appears that there are 5 points is above the value 1 and the other points were below a value of 1. This illustrates that there are 5 components that have value of eigenvalue 1. Next do the determination of each grain is going fit into any factor of five factors. Grouping of grains and magnitude of loading factor of a factor viewed from the value of the factor value is loadingnya >0.3.

The results of the grouping of the 21 grains of 5 factors are formed as well as the naming of each of the factors. Components that formed on the factor 1 was named attitude and behavior, grain instrument consisting of grains of 1, 2, 11 and 12. Components that formed on the 2nd factor was named communication instrument consisting of grains of grains, 3, 4, 5, 6, 7, 8, 9 and 10. Components that formed on 3 factors named preparation, grains of instruments consisting of grains of 13, 14, 15, 16 and 17. Components that formed on the factor 4 was given the name of the implementation, the grain

instrument consisting of grains of 18 and 19. Components that formed on the factor 5 was named engineering, grain instrument consisting of grains of 20 and 21.

Reliability analysis based on test instrument on the field both through Alpha Cronbach coefficient shows the value of 0878 > 0.7, so it can be said that the examination performance assessment consistent IVA used in doing the assessment. If the value of Cronbach's and 0.60 Alpha is less than 1, then the instrument has a high correlation or reliability, while if the value of Cronbach's Alpha below 0.50 to down, then the instrument is correlated to low or no reliability (Basuki and Hariya nto, 2014, p. 105). This is in line with the expressed Dragon that reliability coefficient of 0.50 up adequate to accept as a good reliability (khumaedi,2012, p. 13)

In line with research done Suratno, Muhyadi, & Mardapi (2016, p. 112) that EFA analysis (Exploratory Factor Analysis) is to test whether an invalid constructs can be explained by the indicators-charge indicators will. If the indicator variable or invalid construct can form, then the charge is indicated by a high factor (> 0.3) means that the measurement is in compliance data, and the expected value of the KMO is greater than 0.5. As for checking the reliability of an instrument is to look at the value of the coefficient Alpha of Cronbach's Alpha formula, i.e. at least 0.7 as the lowest limit.

Practicality Of Performance Assessment Inspection IVA Developed

Further test the practicality in order to find out how big a level of practicality of instruments already developed. Based on the results of the now charging a percentage amount obtained the practicality of a total of 94%, so that the assessment instrument performance examination test IVA is said to be very practical and usable. In line with the research conducted, Preserved, and Gen. Guidance (2016, p. 17) the practicability of test results do indicate the instrument has a percentage of 88%, the test results showed that practicality made by 15 groups of students have the level of practicality with very high category. Following is the results of the test of practicality in Table 6 below

Table 6. The Interpretation Of The Criteria Of Practicality

the presentation interval	Description
$0 \geq X \leq 25$	Not very practical
$25 > X \leq 50$	Not practical
$50 > X \leq 75$	Practically the very practical
$75 > X \leq 100$	Not very practical

This is in line with research done Setiawan (2017, p. 8) Test is performed with the practicality of two techniques, namely the trial field and the assessment of teachers. The value obtained is quite high, i.e. 95%. The value of high practicality suggests that viable instrument used to record the skills competency because the instruments are easy to use, easy to understand, easy to be processed, and the results reported to the result. Authentic assessment instruments developed stated very practical and feasible in terms of users, namely the teacher to use record of competence skills of grade IV elementary school.

CONCLUSION

Based on the results of research and discussion has been done, it can be concluded that the examination of performance appraisal Form IVA developed i.e. observation sheets shaped by the number of grains as much as 21, declared invalid based on review experts on the retrieved value > 0.3 for each grain and get a total percentage of the overall feasibility of the retrieved value of 94%. The test results by using the analysis of the factor obtained by the value of KMO exploratory 0,904, with the results of this analysis are formed into 5 factor in performance assessment inspection of IVA koefesien with Alpha Cronbach's > value gained 0.70, thus has obtained an assessment of performance examination of the IVA vaid, reliability, reliable and practical.

REFERENCES

Ashari, L, Hasan, Lestari W, Hidayah T. (2016). Instrumen Penilaian Unjuk Kerja Siswa SMP Kelas VIII Dengan Model *Peer Assessment* Berbasis Android Pada Pembelajaran Penjasorkes Dalam Permainan Bola Voli. Journal of Educational Research and Evaluation, *5*(1)

Astuti, Widya Puji, Wibawanto, H, & Khumaedi, M. 2015. Pengembangan Instrumen Penilaian Unjuk Kerja Praktik Perawatan Kulit Wajah Berbasis Kompetensi Di Universitas Negeri Semarang. Journal of Educational Research and Evaluation, 4(2)

Azwar, S. (2011). Reliabilitas dan Validitas.Yogyakarta: Pustaka Pelajar.

Basuki, Ismet, & Hariyanto. (2014). Asesemen pembelajaran. Bandung: PT. Remaja Rosdakarya.

Juniarta, Aswin Try, M. . W. (2016). Pengembangan Instrumen Penilaian Pengetahuan Mata Pelajaran Pendidikan Jasmani Olahraga Dan Kesehatan (Pjok) Kelas Xi Semester Gasal, 1659–1664.

Khumaedi, M. (2012). Reliabilitas Instrumen Penelitian Pendidikan. Jurnal Pendidikan Teknik Mesin Vol. 12.

Mardapi, D. & sri andayani. (2012). Sri Andayani-makalah MCDM semnas 2 Juni Sri Andayani PMAT UNY. *Prosiding Seminar Nasional Penelitian, Pendidikan Dan Penerapan MIPA, Fakultas MIPA, Universitas Negeri Yogyakarta,*.

Mardapi, D. (2016). *Pengukuran Penilaian & Evaluasi Pendidikan*. Yogyakarta: Parama Publshing.

Mulyani, E. (2012).Pengembangan Model Penilaian Komprehensif Berbasis Proyek Pendidikan Kewirausahaan Terintegrasi di SMK. Jurnal Penelitian dan Evaluasi UNY.Tahun 16 No. 2. Hal 252 - 278.

Nugroho, bagus setyo., djuniadi., & A, Rusilowati. (2016). Pengembangan Penilaian Kinerja Menggambar Teknik Potongan Di Smk Pada Kurikulum 2013. Journal of Educational Research and Evaluation, *5*(1), 1–7.

Permenkes NO 29 tahun 2017. (2017). Perubahan Atas Peraturan Menteri Kesehatan Nomor 34 Tahun 2015 Tentang Penanggulangan Kanker Payudara Dan Kanker Leher Rahim, 1–40.

Rusilowati, A. (2013). Pengembangan Instrumen Non-Tes. *Makalah. Seminar Nasional Evaluasi Pendidikan Di Universitas Negeri Semarang*.

Setiawan, H. (2017). Pengembangan instrumen asesmen autentik kompetensi pada ranah keterampilan untuk pembelajaran tematik di sekolah dasar. *Jurnal Pendidikan*, *2*(7), 874–882.

Suratno, A., Keahlian Teknik Kendaraan Ringan, K., Menengah Kejuruan Negeri, S., Barat, C., & Teuku Umar No, J. (2016). Pengembangan Instrumen Penilaian Kompetensi Praktikum Engine Siswa Smk Program Keahlian Teknik Otomotif Developing Assessment Instruments in Competence Practice Engine Student in Smk Automotive Engineering Program. *Agus Suratno VANOS Journal Of Mechanical Engineering Education*, *11*(1), 2528–2700.

Wachyudi, ibnu., Sukestiyarno, YL & Budi Waluya, St. (2014). Pengembangan Instrumen Penilaian Unjuk Kerja Pada Pembelajaran Dengan Model Problem Solving Berbasis Tik. Seminar Nasional Evaluasi Pendidikan Tahun 2014. ISBN 978-602-14215-5-0