



Validity and Reliability Contents of Independence Assessment Instruments of Basic Beauty Students for Class X SMK

Dwi Angraini[✉], Muhammad Khumaedi, Trisnani Widowati

Universitas Negeri Semarang, Indonesia

Info article

Published 31 March 2020

Keywords:
Independence
Instrument,
Reliability, Validity

Abstract

This research was motivated by the absence of an instrument for assessing the independence of students in basic beauty subjects that were tested for validity and reliability. The purpose of this study was to determine the validity and reliability of the content of the students' independence assessment instruments in basic beauty subjects at SMK that had previously been compiled. The instruments developed were observation sheets for teachers and self-assessments for students. Analysis of the content validity of the instrument used the Aiken V coefficient and the reliability of the instrument content was determined by the ICC. In this validity and reliability test, it involves 3 expert judgments to assess the instrument. The three expert judgments are development and research experts, beauticians, and beauty practitioners. Based on the results of this research, it shows that the observation sheet and self-assessment are proven valid with the Aiken V coefficient of 0.69 where the score is > 0.3 and reliable with the Alpha Reliability coefficient > 0.874 and $ICC = 0.698$, because $r_{xx} > 0.6$ thus the character assessment instrument for vocational school students can be said to be reliable. Based on the results of this research, the instrument for assessing the independence of students in basic beauty subjects at SMK in skin and hair expertise competency is suitable for use. The benefit of this research is that it can be used as a standard guide in measuring the content of students' independence instruments in learning basic beauty in SMK.

[✉]Correspondence address :
Kelud Utara 3 kampus Pascasarjana UNNES Semarang, Indonesia
E-mail : dwia182@gmail.com

INTRODUCTION

Education in Indonesia tends to be more oriented towards hard skill-based education which is developing on the acquisition of test scores or exams (Aqib and Sujak 2012: 6). The 2013 curriculum emphasizes the competence of attitudes and behaviors in the first order and then the realm of skills and knowledge (Supriati 2014: 23). The formation of national character can be done by implementing character values in every subject taught, agrees with this. honesty and responsibility. With the value of discipline, honesty and teacher responsibility can instill character values through learning.

The values of character education must be instilled in the internalization of learning and real practice in the daily life of students in schools and communities. In this case, besides having religious and nationalist character values, students are required to have good independence, especially vocational students. The character of independence can be used as provisions for SMK students for PKL (Field Work Practices) and in accordance with the mission of the Vocational School, which is to be able to create independent, responsible, polite, hardworking, integrated workforce both as workers (employees) and having their own business.

Based on interviews with Mrs. Kiki Mujiyati (SMK Alhikmah) and Mrs. Devita (SMK Taman Siswa), these teachers have not used valid student character assessment instruments, especially the character of independence when learning basic beauty. The teacher has difficulty making a special character assessment instrument to assess student independence. Because in general, teachers only assess holistically what it is in the class with affective assessment sheets without a standard and clear assessment rubric. This implies the need to develop instruments to measure attitudes and behaviors to assess the implementation of character education in basic beauty learning in Vocational High Schools (SMK). Paltiman Lomban Gaol (2017:3), the independence of students at school is very low, especially when the process of character assessment instruments is not

available that is learning and is used by teacher. The results of the instrument help to greatly assess the character, especially the character of student independence.

Masrukan (2014: 45) defines assessment as a systematic procedure to collect information besides being able to make conclusions about the characteristics of people or objects. The essence of the assessment is to interpret or interpret the measurement data. In interpreting the measurement results data can be compared with the type of benchmark (standard). Regulation of the Minister of National Education number 20 of 2007 concerning national assessment standards provides a benchmark for assessing learning in schools. Based on the policy of assessment regulations, the school and the teachers have the authority to make and develop learning assessments. Rusilowati, Ani (2014: 67) assessment is an assessment that is carried out by observing students doing something. Teachers must have instruments along with their assessment guidelines in order to make more objective assessments.

The instrument used in the assessment has several requirements to be fit for use, namely valid and reliable (Mardapi, 2012: 37). The appropriateness of the assessment instrument must also look at the validity and reliability of the contents of the instrument to be used. Content validity is the extent to which instrument elements are relevant and represent a targeted construct for a specific purpose (Haynes, et al. 1995). While the reliability test is a test to see to what extent a questionnaire can reliably measure a test (Susanto Y, Alfian R, Rahim Z & Karani. 2018: 45). Therefore, it is necessary to test the validity and reliability of the contents of the student's independence assessment instrument. Based on the research of Dewi Zuliani, Totok Sumaryanto and Saiful Ridlo (2017:48), there are 2 instruments developed, namely the observation and self-assessment sheets used during field trials. The benefit of this research is that it can be used as a standard guide in measuring the content of students' independence instruments in learning basic beauty in class X SMK.

METHODS

This research is part of a research on the development of an independent assessment instrument for vocational high school students in basic beauty subjects. There are 8 steps development procedure is 1) compiling instruments specifications (instrument rubric, instrument scale, assessment criteria), 2) analyzing instruments, 3) experiment 1, 4) analyzing instruments, 5) improving instrument, 6) experiment 2, 7) interpreting results of the instrument, 8) final instruments. Instruments that have met the validity and reliability standards can be used for the measurement stage. Before the instrument is tested in the field, the instrument must be tested first to expert judgment to determine the validity and consistency of the contents of the instrument to be used. According to Sugiyono (2015: 33), a valid instrument must have internal and external validity, meaning that internal validity must consist of content and construct validity. The validity of the contents of the instrument is obtained by providing a questionnaire to experts, namely development and research experts, beauticians, and beauty practitioners. Analysis of the content validity of the instrument used the Aiken V coefficient and the reliability of the instrument content was determined by the ICC (Intraclass Correlation Coefficient).

Aiken formulates Aiken's formula to calculate the content-validity coefficient that is based on judgment results of expert panels with n persons to a certain system. It is based on how far the items represent the measured construct (Hendryadi, H., 2017). Aiken proposed the concept of content validity in more detail. This detail can be seen from the validity standard which is influenced by the number of raters and the rating scale used (Aiken, 1985: 133). In proving the validity of content (Azwar, 2012: 113), researchers can determine the number of rating categories they want. The number of rating categories affects the content validity standard set by Aiken. The smallest number of rating categories formulated by Aiken is 2 and the highest is 7 (Aiken, 1985: 134). This study uses 5

rating categories and 3 ratings. The scoring score for each aspect is 5 for the highest score and 1 for the lowest score. The content validity in this study was assessed by 3 Expert Judgments who were analyzed using Aiken's V formula (Azwar, 2014: 134). Aiken's V formula is :

$$v = \frac{\sum s}{n(c - 1)}$$

Ket:

V = Rater's Fit Index

s = Average score - the lowest score in the category

c = Number of Categories

n = Number of Raters

The results of the scoring of the validation sheet are said to be valid in terms of content when checking or consulting the V index obtained by looking at the specified criteria must be > 0.3 (Azwar 2014: 134). If V index > 0.3 then it is declared valid.

The reliability of the content in this instrument uses the ICC approach to calculate the level of agreement between the three (3) Expert Judgment, before using the ICC approach, the reliability value is estimated by the Alpha coefficient. The coefficient value must be $r_{xx} > 0.6$ so that the Intraclass Correlation Coefficient (ICC) analysis can be continued. According to Suharsimi (2008: 75) the assessment instrument is said to be reliable if $r_{xx} > 0.6$. The reliability of the content in this instrument uses the ICC approach to calculate the level of agreement between the three (3) Expert Judgment, before using the ICC approach, the reliability value is estimated with the Alpha coefficient. The coefficient value must be $r_{xx} > 0.6$ so that the Intraclass Correlation Coefficient (ICC) analysis can be continued. According to Suharsimi (2008: 75) the assessment instrument is said to be reliable if $r_{xx} > 0.6$.

RESULTS AND DISCUSSION

The validity of the contents of the instrument is obtained by providing a questionnaire to experts, namely development and research experts, beauticians, and beauty

practitioners. The analysis of the content validity of the instrument used the Aiken coefficient V, where the V index must be > 0.3. The following is table 1 of the results of the content validity test using the Aiken'V formula.

Table 1. Content Validity Results Using the Aiken'V Formula.

Item number	V indeks	Conclusion
		≥ 0.3
1	0.82	Valid
2	0.82	Valid
3	0.55	Valid
4	0.55	Valid
5	0.55	Valid
6	0.73	Valid
7	0.82	Valid
8	0.82	Valid
9	0.73	Valid
10	0.64	Valid
11	0.82	Valid
12	0.55	Valid
13	0.55	Valid
Total	0.69	Valid

(Source: Researcher Data, 2020)

In Table 1, the results of Aiken'V's analysis result in all items declared valid. Because V index $\geq r$ is critical 0.3. This is in accordance with the criteria stated by Azwar (2014: 34) that if the validity coefficient ≥ 0.3 means the item is valid.

Khumaedi (2012: 26) reliability is a coefficient that shows the extent to which an instrument or measuring device can be trusted, meaning that if an instrument is used repeatedly to measure the same thing, the results are relatively stable or consistent. Empirically the level of reliability is indicated by a number called the reliability coefficient. The magnitude of the reliability coefficient ranges from 0-1, where the higher the reliability coefficient means the more consistent the measurement results are, however

empirically the reliability coefficient that reaches 1 is rarely found.

Besides validity, reliability needs to be analyzed carefully. Nugroho, B. S., Djuniadi, D., & Rusilowati, A. (2016) explain that after finding out the validity results based on the experts' judgment, the next step is to calculate the instrument reliability through consistency agreement among the raters. It is due to reliability could show how far the measuring results of the instrument could be trusted (Munadi, S., 2010). Intraclass Correlation Coefficient is used to assess the reliability between two or more observers, as well as test-retest reliability. In essence, ICC is the ratio between the variance between groups and the total variance (Bisma Murti, 2011: 13). The Intraclass Correlation Coefficient shows a comparison between the variation caused by the measured attribute and the variation in the overall measurement. After knowing the results of the content validation from the experts, the next step is the level of agreement (reliability) of the instrument between the 3 expert validators using ICC (Intraclass Correlation Coefficient) using SPSS 16 to estimate the reliability between raters by showing the comparison between variations caused by the attributes measured by the overall measurement variation. The results of the content reliability calculation using SPSS 16 can be seen in table 2 as follows.

Table 2. Expert Reliability Test Results.

Reliability Statistics	
Cronbach's Alpha	N of Items
.874	3

(Source: Researcher Data, 2020)

Based on the results of table 2 of the expert reliability calculations using SPSS 16. Before using the ICC approach, the reliability value was estimated using the Alpha coefficient. According to Khumaedi (2012: 13), the reliability coefficient > 0.5 is sufficient to be accepted as good reliability. Based on Cronbach's Alpha analysis, it shows that the result of agreement between raters, the coefficient of $r_{xx} = 0.874$, means that

the coefficient of rxx is > 0.7 so that it can be continued with the ICC (Intraclass Correlation Coefficient) analysis with the help of SPSS 16.

The guidelines state that, when the reliability coefficient is below 0,4 the level of clinical significance is poor. When it is between 0,4 – 0,59, the level clinical significance is fair. When it is between 0,6 – 0,74, the level of clinical significance is good. When it is between 0,75 – 1,00, the level of clinical significance is excellent (Cicchetti, Domenic V. 1994:286). ICC's (Intraclass Coefficient Correlation) Formula is:

$$ICC = \frac{\sigma_t^2}{\sigma_t^2 + \sigma_{re}^2}$$

Keterangan :

σ_t^2 = the variance of the assessed subjects

σ_{re}^2 = Variance random error.

Following are the results of the ICC test for agreement between raters in the table 3.

Table 3. Intraclass Correlation Coefficient of Agreement between Raters.
Intraclass Correlation Coefficient

	Intraclass Correlation ^a
Single Measures	.698 ^b
Average Measures	.874 ^c

(Source: Researcher Data, 2020)

Suharsimi (2008: 75), the assessment instrument is said to be reliable if $r_{xx} > 0.6$. Based on the results of the ICC output above, seen from Single Measures = 0.698, thus the instrument for assessing the independence of vocational students can be said to be reliable and ready to be used for field testing. Meurut chiche

Based on the results of the validity and reliability tests in tables 1,2 and 3, it is inseparable from the input and suggestions of the 3 experts (expert judgment) to improve the quality of the instruments to be tested. Input and suggestions from experts can be seen in table 4.

Table 4. Expert Judgment Feedback and Suggestions on the Instrument.

No	Expert Judgment	Feedback/Suggestions
1	Expert 1	Some items don't match the observed indicators.
2	Expert 2	Complete self-assessment questionnaire, adjust it to the assessment indicator aspect.
3	Expert 3	All observation indicators are used as statement items.

(Source: Researcher Data, 2020)

The instrument for assessing the independence of vocational high school students for skin and hair expertise competency was declared valid and reliable based on the Aiken V and ICC tests. This is in line with the research of Mause Agreviana, Ani Rusilowati, Bambang Sugeng Suryatna (2018: 145), that content validity using Aiken V and content reliability using ICC resulted in quite high output validity and instrument reliability.

CONCLUSION

The results of the validity of the assessment of the 3 experts with 13 question items to the expert judgment based on the Aiken V formula with a mean score of 0.69. This shows that the instrument items developed are valid.

Assessment reliability test based on 3 experts based on the ICC (Interclass Correlation Coefficient) test with a score of 0.698. Thus the independence assessment instrument of vocational school students in basic beauty subjects has a high enough agreement between rater.

Based on the results of the validity and reliability test, the content of the assessment instruments obtained from the 3 expert judgments was feasible to be used to assess the independence of students in basic beauty subjects for skin and hair skills competencies.

ACKNOWLEDGEMENT

Researchers would like to thank those who have helped during the research process and guided this research, including: 1) expert validator in development and research, (2) expert validator in beauty education, (3) expert practitioner of beauty education. And to the article reviewer who has provided a lot of input and suggestions in this article. Also don't forget to basic beauty teacher at Smk Al Hikmah Jepara, Smk Taman Siswa Kudus, and Smk N 3 Pati who helped in the data collecting process.

REFERENCE

- Aiken, L. R. (1985). Three Coefficients for Analyzing the Reliability and Validity of Ratings. *Educational and Psychological Measurement*, 5(1) hal:131-142.
- Agreviana, Mause., Ani Rusilowati., & Bambang Sugeng S. 2018. Pengembangan Instrumen Penilaian Kinerja Pembelajaran Merias Wajah Sehari-Hari Untuk Jurusan Tata Kecantikan Di Smk. *Journal of Educational Research and Evaluation* 7 (1).
- Aqib, Z dan Sujak. 2012. *Panduan dan Aplikasi Pendidikan Karakter*. Bandung: Yrama Widya.
- Azwar, S. 2015. *Reliabilitas dan Validitas*. Yogyakarta: Pustaka Pelajar.
- _____. 2016. *Penyusunan Skala Psikologi*. Yogyakarta: Pustaka Pelajar.
- Cicchetti, Domenic V. 1994. Guidelines, Criteria, and Rules of Thumb for Evaluating Normed and Standardized Assessment Instruments in Psychology. *Psychological Assessment*. Vol 6 (4). Hal: 284-290.
- Haynes, S. N., Richard, D. C., & Kubany, E. S. (1995). Content Validity in Psychological Assessment: A Functional Approach to Concepts and Methods. *Psychological Assessment*. 7, hal: 238 -247.
- Hendryadi, H. (2017). Validitas isi: tahap awal pengembangan kuesioner. *Jurnal Riset Manajemen dan Bisnis (JRMB) Fakultas Ekonomi UNIAT*. 2(2). Hal:169-178.
- Khumaedi, M. (2012). Reliabilitas Instrumen Penelitian Pendidikan. *Jurnal Pendidikan Teknik Mesin*, 12(1) hal:29.
- Mardapi, D. (2012). *Pengukuran, penilaian dan evaluasi pendidikan*. Yogyakarta: Nuha Medika.
- Masrukan. 2014. *Asesmen Otentik Pembelajaran Matematika, Mencakup Asesmen Afektif dan Karakter*. Semarang: FMIPA Unnes.
- Munadi, S. (2010, June). Analisis Kualitas Soal Untuk Penilaian Aspek Afektif. *Makalah*. Disampaikan pada acara Workshop Penyusunan Instrumen Evaluasi Afeektif mata kuliah Pengembangan Kepribadian.
- Murti, Bisma. 2011. Validitas dan Reliabilitas Pengukuran. *Matrikulasi Program Studi Doktorat*. Fakultas Kedokteran, UNS.
- Nugroho, B. S., Djuniadi, D., & Rusilowati, A. (2016). Pengembangan Penilaian Kinerja Menggambar Teknik Potongan di SMK pada Kurikulum 2013. *Journal of Research and Educational Research Evaluation*. 5(1). Hal:01-07.
- Rusilowati Ani. 2014. *Pengembangan Instrumen Penilaian*. Semarang: University Press.
- Sugiyono, M. (2015). *Penelitian & pengembangan (Research and Development/R&D)*. Bandung: Penerbit Alfabeta.
- Suharsimi, A. 2007. *Dasar-dasar Evaluasi Pendidikan*. Jakarta: Bumi Aksara.
- Sunyoto. 2013. Bentuk Karakter Melalui Pelajaran Matematika. *Makalah*. Seminar Nasional Pendidikan MTK, "Bentuk Karakter Melalui Pelajaran Matematika" 04 Juni 2013.
- Supriati, Y. 2014. "Instrumen Penilaian Otentik". *Makalah*. Seminar Nasional Evaluasi Pendidikan "Implementasi Standar Penilaian dalam Pelaksanaan Kurikulum 2013". Sabtu 8 Maret 2014 di UNJ.
- Susanto Y, Alfian R, Rahim Z & Karani. (2018). Uji Validitas Dan Reliabilitas Kuesioner Eq-5d Bahasa Indonesia Untuk Mengukur Kualitas Hidup Pasien

- Hemodialisa Gagal Ginjal Kronik. *Jurnal ilmiah MANUTUNG*. 4(1).
- Undang-Undang Sistem Pendidikan Nasional No. 20 Tahun 2003 Pasal 58 Ayat 1.
- Zuliani, Dewi, Totok Sumaryanto dan Saiful Ridlo. 2017. Pengembangan Instrumen Penilaian Karakter pada Siswa Kelas IV Sekolah Dasar. *Journal of Educational Research and Evaluation* 6 (1). Hal: 46-54.