JRERE 13 (1) 2024 102-111

Journal of Research and
Educational Research Evaluation

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

Validity of the Ethics Assessment Instrument for Senior High School Students

Ulin Nuha Sun'an[1]✉, Wiwi Isnaeni[2], Muhammad Khumaedi[3]

[1,2,3]Educational Research and Evaluation, Pascasarjana, Universitas Negeri Semarang

Article Info

History Articles
Received:
17 December 2023
Accepted:
19 January2024
Published:
30 March 2024

Keywords:
Content, item, construct, Validity.

Abstract

The basic principle of research instruments is to measure what will be measured. The goal is to ensure that the instrument reflects the theoretical concept. Proof of validity in this study includes instrument content validity, item validity, construct validity. The research subjects were 7 expert judgment and 386 students. Proof of content validity using the Aiken V formula obtained a coefficient index V ranging from 0.810 to 0.905 so that it was concluded to be valid. Proof of item validity using the pearson sorrelation product moment formula 20 instrument items obtained pearson correlation indices between 0.227 to 0.606 so it was concluded valid. Proof of construct validity using EFA that 20 instrument items obtained a KMO and Bartlett's Test of Sphericity index of 0.811 with a significance of 0.000; anti-image correlation index ranged from 0.519 to 0.910; Eigenvalues index between 1.115 to 4.902 with a total of 56.210% variation explained; index loading value between 0.507 to 0.956 so that it was concluded valid. The researcher's suggestion is that establishing a psychological and communicative relationship with prospective research subjects must be done first.

✉Correspondence Address :
Educational Research and Evaluation, Pascasarjana, Universitas Negeri Semarang, Indonesia 50229
E-mail : joe_paol@students.unnes.ac.id

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

INTRODUCTION

In conducting research, the validity of measuring instruments is very important (Reidenbach & Robin, 1988, p. 321, 1990, p. 874). The success of the research is determined by the validity of the measuring instruments used. The validity of measuring instruments is not only an attribute of a series of research procedures, but rather diagnoses the accuracy of measuring instruments, identifies aspects of the measuring object, determines the causal relationship between latent properties of the measuring object. Therefore, the validity of measuring instruments can affect the accuracy of measurement results. Because, the basic principle of instrument development is that it must measure what will be measured. The instrument must reflect or be as accurate as possible with the various real-world problems to be measured (Beck, 2020, p. 2). Validity ensures that instruments reflect theoretical concepts. In instrument development, there are three types of validity that must be proven by the developer of the instrument created, namely content validity, criterion validity, construct validity. Content validity is defined as the extent to which the instrument is able to capture the phenomenon being measured. Content validity provides evidence of instrument construction. Construct validity is the extent to which the instrument measures the targeted construct or the extent of variance in the measures obtained from the construct targeted by the instrument. Criterion validity relates to the relationship of the instrument to specified criteria. Criterion validity consists of predictive and concurrent validity (Rathnasabapathy & Subramani, 2022, p. 79; Shrotryia & Dhanda, 2019, pp. 2–3).

Content validity is also defined as the extent to which the elements of the instrument represent the constructed target of the measurement. The elements of the instrument refer to all aspects in the measurement process and can affect the data to be obtained (Yusoff, 2019, p. 49). In order to obtain instrument elements that can measure the aspects to be measured, the developer must validate the instrument to expert judgment, assess the dimensions and subdimensions of the construct to be measured, collect feedback, analyze the results of the assessment provided by the expert judgment. From the analysis results, a stable instrument will be obtained (Schmitz & Storey, 2020, pp. 805–806; Shrotryia & Dhanda, 2019, p. 3). One way to obtain content validity is to look at the items that make up the instrument. If the items appear to measure what is to be measured then content validity will be met. Thus, content validity ensures the measurement includes a set of aspects that can represent the concept (Sudaryono et al., 2019, p. 3).

The purpose of content validity is to minimize potential errors associated with operationalizing the instrument at an early stage and to increase the probability of obtaining construct validity in favor of a later stage (Shrotryia & Dhanda, 2019, p. 3). Content validity thus provides evidence of the extent to which the elements of the instrument are relevant and at the same time representative of the construct targeted for assessment purposes. At least four components are important in content validity, namely domain definition, domain representation, domain relevance, appropriateness of instrument development procedures. Domain definition indicates that the concept or phenomenon being measured has been operationally defined. This element has a central role in evaluating content validity because it provides information about the theory-based and operational conceptual definitions of the measured constructs. Domain representation refers to the extent to which the instrument is adequate in representing and measuring the domain of the construct being measured. Domain relevance relates to the degree to which each element of the instrument is relevant and the domain being measured. The appropriateness of the instrument development procedure concerns the process used in developing the instrument

and ensuring the representativeness and relevance of the elements (Almanasreh et al., 2019, p. 2).

The purpose of this research is to develop an ethics assessment instrument for high school students. Ethics was developed into 5 indicators, namely: acting according to religious norms, acting according to norms of decency, acting according to norms of decency, acting according to legal norms, acting according to school regulations. Each indicator was developed into two favorable statement items and two unfavorable statement items. The answer choice model uses a frequency scale with five levels of scoring (H. Retnawati, 2015, p. 158). For favorable statement items "always" score 5, "often" score 4, "sometimes" score 3, "ever" score 2, "never" score 1. For unfavorable statement items "always" score 1, "often" score 2, "sometimes" score 3, "ever" score 4, "never" score 5. The instrument model uses self-assessment.

METHODS

This research used research and development method with modified Thiagarajan's 4D model (Sivasailam et al., 1974, pp. 6–9). The research subjects involved 7 expert judgment and 386 high school students. expert judgment to assess the instrument related to the operational definition of ethics, the suitability of the operational definition with the indicators and statement items developed, the suitability of the items with the indicators, the suitability of the items with the ethical aspects measured, the ability of the statement items in describing student ethics. The results of students' responses to the instrument are used to prove the total item validity and construct validity of the instrument developed.

The results of expert judgment will be analyzed using Aiken V formula:

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

lo = Low validity rating

c = High validity rating number

r = Rating given by an validator

n = Number of experts

s = r – lo

(Aiken, 1985, p. 133; Azwar, 2022, p. 113)

The coefficient V according to Aiken (1980, p. 956) that both in large and small samples ranges from 0 to 1. According to Aiken (1985, p. 134) in the validity coefficient V table, if the expert is 7 people and uses a scale of 4, then the coefficient V value is 0.76.

Student response data will be analyzed for total aitem validity using the Pearson correlation product moment formula:

$$r_{iX} = \frac{[\sum iX - \frac{(\sum i)(\sum X)}{n}]}{\sqrt{[\sum i^2 - (\sum i)^2/n]\,[\sum X^2 - (\sum X)^2/n]}}$$

i = Item score

X = Test score

n = Number of subjects

(Azwar, 2022, p. 154)

The value of the pearson correlation analysis results obtained will be confirmed with the r table value based on the number of subjects with a significance of <0.05.

Data from student responses will also be analyzed for construct validity to determine the constructs formed compared to the criteria presented in table 1.

Table 1. Construct Validity Criteria

Kriteria	Value
KMO and Bartlett's Test	> 0.5
Bartlett's Test of Sphericity	< 0.05
Anti image correlation	> 0.5
Measure of sampling adequacy	> 0.5
Initial Eugenvalues Comulative	> 1
Loading value	> 0.4

Source: (Azwar, 2022, p. 125; Maryani et al., 2021, p. 117; R. Retnawati, 2020, pp. 178–179; Retnowati, 2019, p. 82)

RESULTS AND DISCUSSION

Content validity

Expert judgment consists of 7 people. 2 expert judgment experts in philosophy, 2

expert judgment experts in the field of measurement, 3 expert judgment experts in the field of ethics. Ethics was developed into 5 indicators, including: acting according to religious norms, acting according to norms of decency, acting according to norms of decency, acting according to legal norms, acting according to school regulations. Each indicator is represented by two favorable statement items and two unfavorable statement items. The values given by 7 expert judgments on the instruments developed were analyzed using the Aiken V formula with the criteria for the coefficient value V> 0.76 and are presented in table 2.

Table 2. Aiken V Analysis Results

Number	7 Expert Judgement		
	∑S	n(c-1)	V
01	19	21	0.905
02	18	21	0.857
03	19	21	0.905
04	19	21	0.905
05	19	21	0.905
06	19	21	0.905
07	19	21	0.905
08	20	21	0.952
09	17	21	0.810
10	19	21	0.905

Based on table 2. The results of the Aiken V analysis show that each question item gets a coefficient value V> 0.76 with details of Item_01 coefficient value V 0.905; Item_02 coefficient value V 0.857; Item_03 coefficient value V 0.905; Item_04 coefficient value V 0.905; Item_05 coefficient value V 0.905; Item_06 coefficient value V 0.905; Item_07 coefficient value V 0.905; Item_08 coefficient value V 0.952; Item_09 coefficient value V 0.810; Item_10 coefficient value V 0.905. Thus it can be stated that each item is Valid.

All items can be declared content valid. These results are the same as those obtained by Arrosyid et al. (2022, p. 151) in developing an instrument to assess the Pancasila profile, but the V coefficient index uses the criteria > 0.3. Sahrul et al. (2022, p. 86) also used the V inex criteria > 0.3. Wakano et al. (2022, p. 96) also used the V index criteria > 0.3 even though the expert judgment amounted to 4 people and all items were declared valid. Adam et al. (2023, p. 140) also used the criterion index V> 0.3. The criterion index V> 0.3 is also used by Putria et al. (2023, p. 97), Octafia et al. (2021, p. 49), Sholikhin et al. (2020, p. 42) in the development of their instruments with the addition of "greater than equal to" with the aim that if the coefficient V value obtains a value of 0.3, it is considered valid. All items assessed by expert judgment and the results of proving content validity are declared valid also obtained in the research of Emerentiana et al. (2022, p. 119). All items declared valid were also obtained in the research of Muniroh et al. (2022, p. 18). Riza et al. (2022, p. 4) also obtained content validity on all items of the instrument they developed. In contrast to the results of the content validity analysis obtained by Pramesty et al. (2022, p. 74) in their instrument development research, namely there are three items that are declared invalid because the coefficient value V < 0.88 based on the Aiken V table.

Item Validity Analysis

Data on student responses to statement items were analyzed using pearson correlation product moment to prove whether each item was valid or invalid. The results of item analysis using the pearson correlation product moment formula are presented in table 3. The results of the pearson correlation analysis presented in table 3 will be compared with the r table both in terms of the pearson correlation value and its significance value. Based on r table, if the number of respondents is 386 students, the pearson correlation value is > 0.113 and the significance value (2-tailed) < 0.05 (Arnold et al., 1990, p. 105; Hidayanti & Mandalika, 2023, p. 7; Retnowati, 2019, p. 79).

Table 3. Pearson correlation product momen analysis results

Correlations											
		TOTAL			TOTAL			TOTAL			TOTAL
B01	PC	.486**	B06	PC	.538**	B11	PC	.462**	B16	PC	.575**
	Sig.	0.000		Sig.	0.000		Sig.	0.000		Sig.	0.000
	N	386		N	386		N	386		N	386
B02	PC	.518**	B07	PC	.235**	B12	PC	.438**	B17	PC	.606**
	Sig.	0.000		Sig.	0.000		Sig.	0.000		Sig.	0.000
	N	386		N	386		N	386		N	386
B03	PC	.227**	B08	PC	.334**	B13	PC	.305**	B18	PC	.447**
	Sig.	0.000		Sig.	0.000		Sig.	0.000		Sig.	0.000
	N	386		N	386		N	386		N	386
B04	PC	.245**	B09	PC	.567**	B14	PC	.320**	B19	PC	.497**
	Sig.	0.000		Sig.	0.000		Sig.	0.000		Sig.	0.000
	N	386		N	386		N	386		N	386
B05	PC	.541**	B10	PC	.586**	B15	PC	.572**	B20	PC	.587**
	Sig.	0.000		Sig.	0.000		Sig.	0.000		Sig.	0.000
	N	386		N	386		N	386		N	386

Based on table 3. The results of the Pearson correlation product moment analysis show that all statement items, totaling 20 items, can be declared valid. Each item gets a significance value of 0.000 which means <0.05 and each item gets a value> 0.113. The pearson correlation coefficient value of each instrument item is B01 = 0.486 B02 = 0.518 B03 = 0.227 B04 = 0.245 B05 = 0.541 B06 = 0.538 B07 = 0.235 B08 = 0.334 B09 = 0.567 B10 = 0.586 B11 = 0.462 B12 = 0.438 B13 = 0.305 B14 = 0.320 B15 = 0.572 B16 = 0.575 B17 = 0.606 B18 = 0.447 B19 = 0.497 B20 = 0.587.

Ruchliyadi et al. (2022, p. 457) in the research and development of all items also get a pearson correlation coefficient value> 0.138 with a significance of 0.000. Proof of the validity of the instrument items carried out by Alfian & Putra (2017, p. 181) got a pearson correlation coefficient value> 0.396 because the number of respondents was 25 with a significance of 0.000. Proof of the validity of the questionnaire was also carried out by Nuryani (2020, p. 40) and obtained a pearson correlation coefficient value> 0.333 because the number of respondents was 30 with a significance of 0.000. R. Retnawati (2020, p. 178) in research and instrument development analyzes using pearson correlation using the criteria r count> 0.3 then the item is declared valid. 50 Items tested the result is 45 valid items.

The results of the total item correlation analysis by getting a value> 0.113 with a significance <0.05 can also mean that each item has a strong relationship with other items. Hidayanti & Mandalika (2023, p. 9) in their research also said that if the significance value is <0.05 and the correlation coefficient value> r table, it can be interpreted as having a strong and positive correlation. However, Obilor & Amadi (2018, p. 19) say that the product moment correlation coefficient when applied to a population is a measure of the degree, namely the direction and magnitude of the relationship between two variables. Although an important measurement tool in statistics, the product moment correlation coefficient does not measure the significance or otherwise of a linear relationship between variables.

Construct validity

Proof of construct validity of 20 instrument items will be analyzed using

exploratory factor analysis. The results of exploratory factor analysis will be confirmed by table 3, namely the construct validity criteria containing the value of Kaiser-Meyer-Olkin Measure of sampling adequacy> 0.5 Bartlett's Test of Sphericity <0.05 MSA> 0.5 Anti Image Correlation> 0.5 Measure of Sampling Adequacy> 0.5 Initial Eugenvalues Comulatives> 1 Loading Value 0.4. The results of exploratory factor analysis to be seen are Kaiser-Meyer-Olkin Measure of Sampling Adequacy presented in table 4.

Table 4. KMO and Bartlett's Test

Kaiser-Meyer-Olkin Measure of Sampling Adequacy.		0.811
Bartlett's Test of Sphericity	Approx. Chi-Square	2309.893
	Df	190
	Sig.	0.000

Based on table 4. Kaiser-Meyer-Olkin Measure of sampling adequacy has met the requirements> 0.5 because the value obtained is 0.811 and the value of Bartlett's Test of Sphericity <0.05 while the value obtained is 0.000. Then the next analysis can be done, namely the Measures of Sampling Adequacy and Anti-Image Correlation analysis. Measures of Sampling Adequacy value> 0.5 and Anti-Image Correlation value> 0.5. The analysis results are presented in table 5.

Table 5. Anti-image Correlation

Anti-image Matrices Anti-image Correlation							
B01	.519[a]	B06	.890[a]	B11	.855[a]	B16	.897[a]
B02	.534[a]	B07	.615[a]	B12	.794[a]	B17	.903[a]
B03	.731[a]	B08	.647[a]	B13	.777[a]	B18	.824[a]
B04	.696[a]	B09	.899[a]	B14	.692[a]	B19	.871[a]
B05	.887[a]	B10	.910[a]	B15	.892[a]	B20	.899[a]

Based on the results presented in table 5. anti-image correlation analysis, all items obtained a value> 0.5. The Anti-image Correlation value obtained for each instrument item is B01 = 0.519 B02 = 0.534 B03 = 0.731 B04 = 0.696 B05 = 0.887 B06 = 0.890 B07 = 0.615 B08 = 0.647 B09 = 0.899 B10 = 0.910 B11 = 0.855 B12 = 0.794 B13 = 0.777 B14 = 0.692 B15 = 0.892 B16 = 0.897 B17 = 0.903 B18 = 0.824 B19 = 0.871 B20 = 0.899.

Furthermore, it will be analyzed to find out the Comulative Initial Eugenvalues with a value > 1. The analysis results obtained are presented in table 11.

Table 6. Total Variance Explained

Com ponent	Initial Eigenvalues		
	Total	% of Variance	Cumulative %
1	4.902	24.511	24.511
2	2.048	10.241	34.751
3	1.876	9.380	44.131
4	1.301	6.505	50.636
5	1.115	5.574	56.210

Based on the analysis results in table 6. Total Variance Explained, it is known that there are 5 components formed and have initial Eugenvalues> 1, namely component 1 has an Eigenvalues value of 4.902 and is able to explain 24.511% of the variation, component 2 has an Eigenvalues value of 2.048 and is able to explain 10, 2416% variation, component 3 has Eigenvalues 1.876 and is able to explain 9.380% variation, component 4 has Eigenvalues 1.301 and is able to explain 6.505% variation, component 5 has Eigenvalues 1.115 and is able to explain 5.574% variation. The five factors formed are able to explain 56.210% of the variation.

The next analysis is to know the factor loding value with a value> 0.4 clustered in 1 factor, then the analyzed variables will become 1 new factor. The analysis results are presented in table 7.

Table 7. Rotated Component Matrix

Component									
Butir	1	Butir	2	Butir	3	Butir	4	Butir	5
B15	0.747	B12	0.697	B01	0.946	B08	0.811	B03	0.756
B05	0.720	B14	0.628	B02	0.941	B07	0.729	B04	0.740
B06	0.718	B13	0.620			B18	0.507		
B20	0.712	B11	0.596						
B09	0.707								
B16	0.665								
B17	0.665								
B10	0.664								

Based on the analysis results presented in table 12. Rotated Component Matrix, it is known that 8 instrument items, namely B15 with a loading value of 0.747 B05 with a loading value of 0.720 B06 with a loading value of 0.718 B20 with a loading value of 0.712 B09 with a loading value of 0.707 B16 with a loading value of 0.665 B17 with a loading value of 0.665 B10 with a loading value of 0.664, are grouped in the component 1 column and become a new factor. At the beginning of the instrument preparation, items B05 and B06 were in indicator 2, namely acting in accordance with the norms of decency, B09 and B10 were in indicator 3, namely acting according to the norms of decency, B15 and B16 were in indicator 4, namely acting according to legal norms, B17 and B20 were in indicator 5, namely acting according to school regulations. The results of the rotated component matrix analysis grouped these items into one and became a new factor.

In the component 2 column, there are 4 instrument items that group into new factors, namely B12 with a loading value of 0.697, B14 with a loading value of 0.628, B13 with a loading value of 0.620, B11 with a loading value of 0.596. Items B11 and B12 at the beginning of the instrument preparation were in indicator 3, namely acting according to the norms of decency, items B13 and B14 were in indicator 4, namely acting according to legal norms, after factor analysis the items became a new factor group.

In the component 3 column, there are 2 instrument items that group into new factors, namely B01 with a loading value of 0.946 and B02 with a loading value of 0.941. The two items at the beginning of the instrument preparation were in indicator 1, namely acting according to religious norms, after the second factor analysis was carried out, they became a new factor. In the component 4 column, there are 3 instrument items that group into new factors, namely B08 with a loading value of 0.811, B07 with a loading value of 0.729, B18 with a loading value of 0.507. Items B07 and B08 at the beginning of the preparation of the instrument were in indicator 2, namely acting

according to the norms of politeness, item B18 was included in indicator 5, namely acting according to school regulations, after factor analysis these items grouped into one new factor.

In the component 5 column, there are 2 instrument items that cluster into new factors, namely B03 with a loading value of 0.756 and B04 with a loading value of 0.740. At the beginning of the instrument preparation, these two items were in indicator 1, namely acting according to religious norms, after factor analysis the two items became one new factor component. Based on the results of proving the construct validity, it can be stated that the 20 items of the ethics assessment instrument for high school students are valid.

CONCLUSION

Research and development of ethical assessment instruments for high school students based on proving content validity using the Aiken V formula obtained the V coefficient index ranging from 0.810 to 0.905, it can be concluded that it is valid. Based on proving the validity of the items using the pearson sorrelation product moment formula that the 20 items of the instrument obtained the pearson correlation index between 0.227 to 0.606, it can be concluded that it is valid. Based on proving construct validity that 20 instrument items obtained a KMO and Bartlett's Test of Sphericity index of 0.811 with a significance of 0.000; anti-image correlation index ranged from 0.519 to 0.910; Eigenvalues index between 1.115 to 4.902 with a total of 56.210% variation explained; index loading value between 0.507 to 0.956 so it can be concluded valid.

REFERENCES

Adam, M. I., Anirusilowati, A., & Awalya, A. (2023). *Development of An Evaluation Instrument Model for The Counseling Guidance Service Program for Student Discipline*. *12*(2).

Aiken, L. R. (1980). Content validity and reliability of single items or questionnaires. *Educational and Psychological Measurement*, *40*(4), 955–959.

Aiken, L. R. (1985). Three coefficients for analyzing the reliability and validity of ratings, educational and psychological measurument. *Educational and Psychological Measurement*, *45*(1), 131–142.

Alfian, R., & Putra, P. M. A. (2017). Jurnal Ilmiah Ibnu Sina, 2 (2), 176-183 Riza Alfian. *Jurnal Ilmiah Ibnu Sina*, *2*(September), 176–183.

Almanasreh, E., Moles, R., & Chen, T. F. (2019). Evaluation of methods used for estimating content validity. *Research in Social and Administrative Pharmacy*, *15*(2), 1–8.

Arnold, L. E., Votolato, N. A., Kleykamp, D., Baker, G. B., & Bornstein, R. A. (1990). Does hair zinc predict amphetamine improvement of add/hyperactivity? *International Journal of Neuroscience*, *50*(1–2), 103–107.

Arrosyid, M. S., Fakhruddin, F., & Elmubarok, Z. (2022). Developing an attitude assessments instrument and Pancasila student profiles in islamic religious education subject. *Journal of Research and Educational Research Evaluation*, *11*(2).

Azwar, S. (2022). *Reliabilitas dan Validitas* (IV, Cetaka). Pustaka Pelajar.

Beck, K. (2020). Ensuring content validity of psychological and educational tests – The role of experts. *Frontline Learning Research*, *8*(6), 1–37.

Emerentiana, P., Lestari, W., Kurniawati, Y., & Pranoto, S. (2022). Journal of Educational Research and Evaluation Development of Character Assessment Instruments for Elementary School Students in Sikka Regency in Mathematics Subjects Materials for Flat Shapes Based on Android Applications. *Jere*, *11*(2), 2022–2114.

Hidayanti, A. A., & Mandalika, E. N. D. (2023). Analisis Korelasi Pearson Biaya Produksi Terhadap Luas Lahan Petani Garam di Kecamatan Bolo Kabupaten Bima. *Jurnal Inovasi Pendidikan Dan Sains*, *4*(1), 5–10.

Maryani, M., Pardimin, P., & Setiawan, A. (2021). *Pengembangan instrumen penilaian sikap kejujuran siswa sekolah dasar Kecamatan Wadaslintang Development of an instrument of honest attitude assessment of elementary school students in the Wadaslintang District*. *9*(2), 107–120.

Muniroh, N., Rusilowati, A., & Isnaeni, W. (2022). Instrument Development of Science Literacy Assessment with Socio-Sciences Contains in Natural Science Learning for Elementary School. *Journal of Educational Research and Evaluation*, *11*(1), 15–22.

Nuryani, N. (2020). Validitas dan Reliabilitas Kuesioner Pengetahuan, Sikap dan Perilaku Gizi Seimbang Pada Remaja. *Ghidza: Jurnal Gizi Dan Kesehatan*, *3*(2), 37–46.

Obilor, E. I., & Amadi, E. C. (2018). Test for Significance of Pearson's Correlation Coefficient ( ). *International Journal of Innovative Mathematics, Statistics & Energy Policies*, *6*(1), 11–23. www.seahipaj.org

Octafia, D., Supriyadi, S., & Sulhadi, S. (2021). Validity and Reliability Content of Physics Problem Solving Test Instrument Based on Local Wisdom. *Journal of Educational Research and Evaluation*, *9*(1), 46–51.

Pramesty, N., Rusilowati, A., & Khumaedi, M. (2022). Validity and Reliability of Development of Self-confidence Assessment Instruments for Students on Chemistry Subject. *Journal of Research and Educational Research Evaluation*, *11*(1), 70–80.

Putria, A. B., Hartono, R., & Khumaedi, M. (2023). *Validity and Reliability Analysis of Reading Comprehension Skill Test Using Islamic English Texts*. *12*(2).

Rathnasabapathy, M., & Subramani, D. (2022). Content Validity of a Scale Measuring Psychosocial Stress Factors among Infertile Women in Treatment. *International Journal of Infertility and Fetal Medicine*, *13*(2), 76–79.

Reidenbach, R. E., & Robin, D. P. (1988). Some initial steps toward improving the measurement of ethical evaluations of marketing activities. *Journal of Business Ethics: Celebrating the First Thirty Years of Publication*, *7*, 871–879.

Reidenbach, R. E., & Robin, D. P. (1990). Toward the development of a multidimensional scale for improving evaluations of Business Ethics. *Journal of Business Ethics*, *9*(8), 639–653.

Retnawati, H. (2015). Perbandingan Akurasi Penggunaan Skala Likert dan Pilihan Ganda untuk Mengukur Self-Regulated Learning. *Jurnal Kependidikan: Penelitian Inovasi Pembelajaran*, *45*(2), 156–167.

Retnawati, R. (2020). Pengembangan instrumen nilai kejujuran siswa pada pembelajaran ekonomi. *Wiyata Dharma: Jurnal Penelitian Dan Evaluasi Pendidikan*, *8*(2), 173–180.

Retnowati, A. (2019). Pengembangan instrumen penilaian sikap tanggung jawab siswa SMP. *Wiyata Dharma: Jurnal Penelitian Dan Evaluasi Pendidikan*, *7*(1), 76–84.

Riza, M., Kartono, K., & Susilaningsih, E. (2022). Validity and Reliability of Chemical Test Instruments for Acid and Base Solutions Oriented Generic Skills Science. *Journal of Educational Research and Evaluation*, *11*(1), 1–8.

Ruchliyadi, D. A., Adawiah, R., & Aida, N. (2022). Development of Self-Confidence Assessment Instruments for 7th State Junior High School Students in Banjarmasin. *JED (Jurnal Etika Demokrasi)*, *7*(3), 453–460.

Sahrul, Khumaedi, M., & Masrukan. (2022). Development of instruments to measure self-confidence and creative thinking in mathematics learning for vocational high school students. *Journal of Research and Educational Research Evaluation*, *11*(1), 81–92.

Schmitz, K., & Storey, V. C. (2020). Empirical test guidelines for content validity: Wash, rinse, and repeat until clean. *Communications of the Association for Information Systems*, *47*(1), 787–850.

Sholikhin, M. N., Supriyadi, S., & Soegiyanto, S. (2020). Instrument Development Chapter of Learning Game Football Kick The Ball for Elementary School Grade Five. *Journal of Educational Research and Evaluation*, *9*(1), 40–45.

Shrotryia, V. K., & Dhanda, U. (2019). Content Validity of Assessment Instrument for Employee Engagement. *SAGE Open*, *9*(1), 1–7.

Sivasailam, T., Semmel, D. S., & Semmel, M. I. (1974). Instructional development for training teachers of exceptional children: A sourcebook. In *ERIC*. Indiana University Bloomington.

Sudaryono, Rahardja, U., Aini, Q., Isma Graha, Y., & Lutfiani, N. (2019). Validity of Test Instruments. *Journal of Physics: Conference Series*, *1364*(1).

Wakano, R., Isnaeni, W., & Ahmadi, F. (2022). Developing Instruments Assessment of Students' Critical Thinking and Communication Skills in Biology Learning Using Hybrid Learning Models in 3T

Areas. *Journal of Educational Research and Evaluation*, *11*(1), 93–102.

Yusoff, M. S. B. (2019). ABC of Content Validation and Content Validity Index Calculation. *Education in Medicine Journal*, *11*(2), 49–54.