Vol. 6, No. 3, March 2025
E-ISSN: 2723 - 6692
P-ISSN: 2723 - 6595
http://jiss.publikasiindonesia.id/

Instrument Reliability Analysis

Fendik Saputra
STIBA Ar Raayah Sukabumi Jawa Barat, Indonesia
Email: fendikibnusyakiran95@gmail.com
Correspondence: fendikibnusyakiran95@gmail.com*

KEYWORDS
Instrument Reliability; Quality Measurement; Research Instruments; Validity

ABSTRACT
This study aims to analyze the reliability of research instruments used to measure the quality of education. The research instrument used was a questionnaire consisting of 20 items. Data were collected from 100 randomly selected respondents. Reliability analysis was conducted using the Cronbach's Alpha method. The results showed that the research instrument had high reliability, with a Cronbach's Alpha value of 0.85. This means that the research instruments are reliable for measuring education quality. This study has important implications for educational researchers and practitioners in developing reliable research instruments.

Introduction

Research in education, psychology, and social fields often uses research instruments to collect data. The research instruments must be highly reliable to ensure reliable results. The reliability of research instruments shows how consistent and reliable the instrument is in measuring the variables under study. However, many studies still do not pay attention to the reliability of research instruments, so the research results become inaccurate. Therefore, it is necessary to analyse the reliability of the research instrument to ensure it is reliable.

This study aims to analyse the reliability of research instruments used to measure the quality of education. The research instrument used is a questionnaire consisting of 20 items. The reliability analysis results are expected to provide information about the consistency and reliability of the research instrument.

In this study, the authors will focus on discussing the reliability of assessment instruments to understand better what instrument reliability is and how an assessment tool is said to have good quality. The criteria for an assessment instrument of high quality is that the instrument has been tested for validity, reliability, difficulty level, differentiating power, and distractors (Amalia & Widayati, 2012).

Material and Method

This research uses a quantitative research design with an instrument reliability analysis method. It aims to analyze the reliability of research instruments used to measure the quality of education. The research instrument used was a questionnaire consisting of 20 items. This questionnaire was used to measure the quality of education based on several indicators, such as teaching ability, ability to manage class, and ability to develop curriculum.

The population of this study is teachers who teach in senior high schools in X City. The sample of this study was 100 randomly selected teachers. Data was collected through questionnaires distributed to the respondents. Respondents were asked to complete the questionnaire honestly and based on their experience. The data were analyzed using the instrument reliability analysis method using SPSS software. Instrument reliability analysis was conducted to determine how much consistency and reliability of the research instrument.

The reliability test was carried out using the Cronbach's Alpha method. The Cronbach's Alpha value obtained was used to determine the reliability of the research instrument.

Results and Discussion

1. Definition of Instrument Reliability

Reliability, which comes from the word reliability, means the extent to which the results of a measurement can be trusted. A measurement result can only be trusted if several times, the implementation of measurements on the same group of subjects, relatively the same measurement results are obtained, as long as the aspects measured in the subject have not changed (Mansur, 2019). In line with that, Suharsimi(2013) argues that an instrument that is said to be reliable will produce when used several times to measure the same object.

Reliability is the extent to which a measurement shows consistent results when repeated measurements are made with the same measuring instrument (Sudaryono, 2012). Some of these opinions conclude that an instrument can be said to have a high level of confidence if it can provide fixed results. So, instrument reliability relates to the fixity of the results.

The reliability of a measuring instrument relates to the ability of the measuring instrument to provide consistent and stable results when repeated measurements are made at different times on the same object (Sugiyono, 2019). Suppose the condition of student A is initially lower than student B, then if a re-measurement is held, student A is also lower than B. That is said to be stable, namely the same in the state of students among other groups. That is said to be fixed (steady), namely the same in the state of students among other groups. The amount of constancy shows the high reliability of the instrument (Arkunto, 2006).

2. Types of Instrument Reliability

According to Djaali quoted by Zulkipli, the reliability of the instrument can be divided into two, namely the reliability of the consistency of responses, and the reliability of the consistency of the combined items.

a. The reliability of response consistency is whether the respondent or measuring object responds to the instrument is good or consistent. In this case, if the instrument is used to measure the respondent, then the measurement is made again to the same respondent, then whether the results remain the same or not with the results of the previous measurement. If the results of the second measurement show inconsistency, it is clear that the measurement results do not reflect the true state of the respondent (Janti, 2014).

b. While the reliability of the combined consistency of items is related to the stability between the items of a test. This can be expressed with a statement, whether for the same respondent, one item shows the same measurement results as the other items, then the instrument can be said to be consistent and can be said to be reliable. However, if on the contrary, on the same respondent, one item with another produces contradictions or is inconsistent, the instrument cannot be trusted in the sense that the instrument is not reliable (Janti, 2014).

3. Analysis of the reliability of the description form test

Testing the reliability of the description form test, generally uses the Alfa Crombach formula. Because the scoring model in the description form is not just right or wrong, if it is correct it is given a value of 1 and if it is wrong it is given a value of 0. Assessing the question of the description form cannot be done like that, but is more of a continuum, for example the assessment is 1-5, or 1-10, and so on (Sukiman, 2012). The Alfa Crombach formula is:

$$r_i = \frac{n}{(n-1)}\left\{1 - \frac{\sum s_i^2}{s_t^2}\right\}$$

Description:

n = Number of questions

$\sum s_i^2$ = The square sum of the scores for each item

s_i^2 = Varians total

Formula for total variance:

$$s_i^2 - \frac{\sum s_i^2 \frac{(\sum xt)^2}{n}}{n}$$

Formula for item variance

$$s_i^2 - \frac{\sum s_i^2 \frac{(\sum xt)^2}{n}}{n}$$

For example, the data from the pilot test is as follows:

Table 1. the data from the pilot test

No. Res.	Item Number					x	x2
	1	2	3	4	5		
1	10	4	8	10	10	**42**	**1764**
2	6	2	4	6	5	**23**	**529**
3	8	3	6	7	8	**32**	**1024**

4	7	5	7	6	6	**31**	**961**
5	0	5	3	4	4	**16**	**256**
6	2	4	2	6	8	**22**	**484**
7	4	3	6	6	6	**25**	**625**
8	5	5	5	7	7	**29**	**841**
9	5	5	3	6	5	**24**	**576**
10	3	6	4	4	6	**23**	**529**
Total	50	42	48	62	65	**267**	**7589**
Sum of Squares	328	190	264	410	451	**1643**	

4. Reliability analysis of objective form tests

Determining the reliability of objective form tests can be done through one of three approaches, namely test-retest reliability, equivalent- form reliability and internal consistency.

a. Test-Retest Reliability

The reliability of the instrument tested by test-retest is carried out by trying the instrument several times on the respondent. So in this case the instrument is the same, the respondents are the same, but the time is different. Reliability is measured by the correlation coefficient between the first trial and the next. If the correlation coefficient is positive and significant then the instrument has been declared reliable (Surapranata, 2019). An example of implementation is as follows:

1) The first trial will be conducted in January on grade X students.
2) Some time later after the first test, for example in February, the test will be conducted again on class X students. ∑
3) Then, after that, the first and second acquisition scores will be correlated to determine their reliability. To determine the correlation of the first score with the second by using the product moment correlation formula as follows:

$$r_{\mathrm{xy}} = \frac{N\sum - (\sum X)(\sum \mathrm{Y})}{\sqrt{\{N\sum \mathrm{X}^2 - (\sum \mathrm{X})^2\}\{N\sum \mathrm{X}^2 - (\sum \mathrm{Y})^2\}}}$$

It is best if the first test is not too long and not too close. If it is too long, the test taker's knowledge and experience could change. Moreover, if it is too close, the results may be influenced by the memory of the answers given during the first measurement, which can affect reliability (Surapranata, 2019).

b. Equivalent-Form Reliability

Equivalent instruments are questions that are linguistically different but mean the same thing. For example, how many years of work experience have you had at this institution? This question can be equivalent to the question, what year did you start working at this institution?

Testing the instrument's reliability in this way is done once, but the instrument is two, on the same respondent, at the same time, different instruments. Instrument reliability is calculated by correlating the data of one instrument with the data of the instrument used as an equivalent. If the correlation is positive and significant, the instrument can be declared reliable (Arifin, 2017; Mohajan, 2017).

c. Internal Consistensy

Testing reliability with internal consistency is done by trying the instrument once, and then the data obtained is analysed with certain techniques. The analysis results can be used to predict the instrument's reliability. Instrument reliability testing can be done in the following way:

1) Spearman Brown formula

In this case, respondents are only given one test. The question items will be divided into two comparable parts: odd question items and even question items, beginning and end. Each question item will be corrected for the results, and the scores from the two parts will be correlated to find the correlation coefficient (Matondang, 2009). The Spearman-Brown formula is:

$$r_i = \frac{2r_b}{1 + r_b}$$

Description:

ri : Internal reliability of the entire instrument

rb : product-moment correlation between the first and second hemispheres.[14]

Example of reliability analysis with odd-even halves.

Table 2. reliability analysis with odd-even

No.	Name	Odd Items X	Even Item Y	X^2	Y^2	XY
1	A	5	3	25	9	15
2	B	3	2	9	4	6
3	C	0	4	0	16	0
4	D	3	2	9	4	6
5	E	3	3	9	9	9
6	F	4	0	16	0	0
7	G	4	3	16	9	12
8	H	3	5	9	25	15
Total		$\sum X = 25$	$\sum Y = 22$	$\sum X^2 = 93$	$\sum Y^2 = 76$	$\sum XY = 63$

After calculating with the product moment correlation formula with rough numbers, it is known thatrxy = -0.3786. After finding the correlation, it is then entered into the spearman-brown formula above.

$$r_i = \frac{2\ (-0.3786)}{1+(-0,3786)} = \frac{-0.7572}{1,3786} = -0.5493$$

2) K-R formula. 20

$$r_i = \frac{k}{(k-1)}\{\frac{s_t^2 \Sigma \text{pq}}{s_t^2}\}$$

Description:
k : Number of items in the instrument
p : proportion of the number of subjects who answered on item 1
q : 1-p
s_t^2 : Total variance
Example : an instrument that will be used for assessment, will be tested for reliability. Because the score used in the instrument produces a dichotomous score (1 and 0), the reliability of the instrument will be analyzed by the KR formula.
20. Experiments with 10 respondents yielded the following data:

Table 3. Experiments with 10 respondents yielded

No. Res	**No. Item**										**x**	x^2
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**		
1	1	1	0	0	0	1	0	0	1	0	4	16
2	1	1	1	0	0	1	0	0	0	0	4	16
3	1	0	0	1	0	1	0	0	0	0	3	9
4	1	1	1	0	0	1	0	1	0	0	5	25
5	1	1	1	1	1	1	1	1	1	0	9	81
6	1	1	1	1	1	0	1	1	1	0	8	64
7	1	1	1	1	1	0	1	1	1	0	8	64
8	1	0	1	0	1	0	1	0	0	1	5	25
9	1	1	1	1	0	0	1	0	0	1	6	36
10	0	1	0	1	1	0	1	0	0	1	5	25
Np	**9**	**8**	**7**	**6**	**5**	**5**	**6**	**4**	**4**	**3**	**57**	**361**
P	**0,9**	**0,8**	**0,7**	**0,6**	**0,5**	**0,5**	**0,6**	**0,4**	**0,4**	**0,3**		
Q	**0,1**	**0,2**	**0,3**	**0,4**	**0,5**	**0,5**	**0,4**	**0,6**	**0,6**	**0,7**		
Pq	**0,09**	**0,16**	**0,21**	**0,2**	**0,25**	**0,25**	**0,24**	**0,24**	**0,2**	**0,21**	**2,13**	

Before these prices are entered into the formula, we must calculate the total variance first.

$$s_t^2 = \frac{x}{n}$$

n = number of respondents
$x2$ =$\sum x$ 2-($\sum x$)2 = 361-(57)2 = 36.1
n 10
$s2$ =x 2 = 361.1 = 3.61
t n 10

Then we input the price into formula KR.20:

$$r_i = \frac{10}{(10-1)}\{\frac{36{,}1 - 2{,}13}{36{,}1}\} = 0456$$

K-R formula. 21

$$r_i = \frac{k}{(k-1)}\{1 - \frac{M\ (\text{K–M})}{s_t^2}\}$$

Description:
k: Number of items in the Instrument
M : mean total score
s_t^2: Total variance

Example. If the data in table 1.2. will be tested for reliability with KR. 21, and then we just need to calculate the value of M (mean total score). The formula calculates the price of M: Calculating the value of M (mean total score) in the reliability test with KR-21 is:
The formula for Calculating the M Value
M = (ΣX) / N
Description
- M = mean total score
- ΣX = sum of total scores
- N = number of respondents
However, in the KR-21 reliability test, the M value can also be calculated by formula:
Alternative Formula for Calculating M Values
M = (ΣPi) / k
Description
- M = mean total score
- ΣPi = the sum of the proportion of items answered correctly
- k = number of items
Determine the correct formula based on the data.

After obtaining the price of M with the existing item analysis table, the priceri can be calculated.

$$r_i = \frac{k}{(k-1)}\{1 - \frac{M\ (\text{K–M})}{s_t^2}\}$$

$$r_i = \frac{10}{(10-1)}\{1 - \frac{5,7\ (10-5,7)}{10\ (3,6)}\}$$
$$= 0,357$$

It turns out thatr_i from the calculation with the KR formula. 21 is lower than the pricer_i with the KR 20 formula. Indeed, there is a reliability calculation using the KR 20 formula that tends to give a higher price. However, the disadvantage is that the researcher must find the value of $\sum pq$ which is a longer step and needs accuracy.

Hoiyt analysis of variance (hoyt anova)

$$r_i = 1 - \frac{MKe}{MKs}$$

To t:

MKs = mean square between subjects MKe = mean square error

Example: if the data is to be tested for reliability using the Hoyt Anova formula, the following calculation steps are required:

a) Calculating the total sum of squares

$$JK_{tot} = \sum x_{tot} - \frac{(\sum x_{tot^2}}{n}$$

b) Calculating the sum of squares of items

$\Sigma X^2 = \Sigma(Xi^2)$ Or with the formula $\Sigma X^2 = (\Sigma X)^2 / k + \sigma^2$

c) Calculating the subject's sum of squares

$\Sigma X^2 = \Sigma(Xi^2)$ Or with the formula $\Sigma X^2 = (\Sigma X)^2 / N + \sigma^2$

d) Calculating the sum of squares of interactions (item x subject)

$JKint = JKtot - JKite - JKsub$

Then sum of squares The entered into the Variance analysis table as follows:

Table 4. The variance analysis

Source Variance	JK	dk	MK
Between items	JK_i	k-1	MK_i
(I)			
Between Subject (S)	JK_s	n-1	MK_s
Interaction (IxS)	JK_{int}	(k-1)(n-1)	MK_e
Total	JK_{tot}	(n-k-1)	-

Based on the item analysis table in table1.2 these prices can be calculated.

$$JK_{tot} = 57 - \frac{(570)^2}{100} = 24{,}51$$

$$JK_{item} = \frac{9^2 + 8^2 + 7^2 + 6^2 + 5^2 + 5^2 + 6^2 + 4^2 + 4^2 + 3^2}{} \frac{(570)^2}{100} =$$

$$JK_{sub} = \frac{4^2 + 4^2 + 3^2 + 5^2 + 9^2 + 8^2 + 5^2 + +6^2 + 5^2}{} \frac{(57)^2}{100} =$$

$$JK_{int} = \frac{4^2 + 4^2 + 3^2 + 5^2 + 9^2 + 8^2 + 5^2 + +6^2 + 5^2}{} \frac{(57)^2}{100} =$$

$$JK_{int} = 24{,}5 - 3{,}21 - 3{,}61 = 17{,}69$$

Source Variance	JK	dk	MK
Between items (I)	3,21	10-1= 9	$\frac{3{,}21}{9}$ = 0,357
Between Subject (S)	3,61	10-1= 9	$\frac{3.61}{9}$= 0.401
Attractions (IxS)	17,69	(10-1)(10-1)= 81	$\frac{17.69}{81}$= 0.218
Total	JK_{tot}	(n-k-1)	-

Then, entered in the Hoyt Anova formula

$r_i = 1 - \frac{MKe}{MKs} = 1 - \frac{0{,}218}{0{,}401} = 0{,}456$ so the reliability coefficient is **0.456**

5. Non-test reliability analysis (questionnaire)

To determine the level of reliability of the questionnaire instrument given to students, it can be analyzed with internal consistency. One technique is the even-odd two-split technique, which is then analyzed using the product moment correlation formula (Gavrilenko, 2016). The results are then entered into the Spearman-Brown formula. An example of the results of the questionnaire instrument that has been obtained is listed in the table below.

Table 5. The results of the questionnaire instrument

No Res.	Item 1	Item 2	Item 3	Item 4	Item 5	Item 6	Item 7	Item 8
1	4	3	2	4	3	3	4	4
2	3	4	3	3	4	2	3	3
3	4	4	2	4	4	3	4	2
4	1	2	4	1	1	1	2	2
5	4	4	3	3	4	2	4	3

Next, it is analyzed by the predetermined formula, namely the Spearman brown umus. The steps are as follows:

a. First, we group the odd and even item scores.

Table 6. Odd Items

Item 1	Item 3	Item 5	Item 7	Total
4	2	3	4	13
3	3	4	3	13
4	2	4	4	14
1	4	1	2	8
4	3	4	4	15

Table 7. Even Items

Item 2	Item 4	Item 6	Item 8	total
3	4	3	4	14
4	3	2	3	12
4	4	3	2	13
2	1	1	2	6
4	3	2	3	12

b. Then look for the correlation coefficient

Table 8. The correlation coefficient

No Res.	Total Item Score	Odd	Even	X2	Y2	XY
		X	Y			
1	27	13	14	169	196	182
2	25	13	12	169	144	156
3	27	14	13	196	169	182
4	14	8	6	64	36	48
5	27	15	12	225	144	180
Total	120	63	57	823	689	748

$$r_{xy} = \frac{n\sum XY - (\sum X)(\sum Y)}{\sqrt{\{(n\sum X^2 - (\sum X)^2\}\{n\sum Y^2 - (\sum Y)^2)\}}}$$

Unknown:

N= 5	$\sum X^2$= 823
$\sum X$= 63	$\sum Y^2$= 689
$\sum Y$= 57	$\sum XY$= 748

$$r_{xy} = \frac{3740 - 3591}{\sqrt{\{4115 - 3969\}\{3444 - 3249\}}}$$

$$r_{xy} = \frac{149}{\sqrt{\{146\}\{196\}}}$$

$$r_{xy} = \frac{149}{\sqrt{28641}}$$

$r_{xy} = \frac{149}{169{,}16} = 0{,}88$ so the coefficient is 0.88

c. Then enter it into the Spearman Brown formula.
   $r_i = \frac{2r_b}{1+r_b} = \frac{2x0{,}88}{1+0{,}88} = \frac{1{,}76}{1{,}88} = \mathbf{0{,}94}$ so the reliability result of the questionnaire instrument is 0.94 and greater than 0.7. So that the questionnaire instrument above can be said to be reliable.

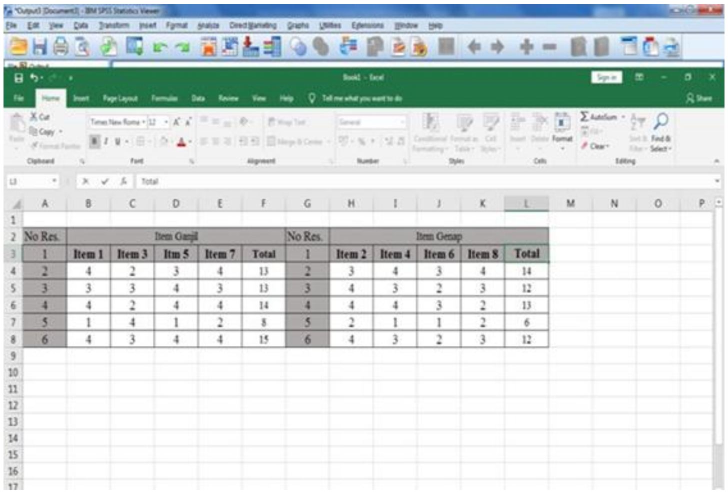
6. Factors affecting the reliability of assessment instruments

a. The atmosphere and conditions during the test. The atmosphere and conditions when the test is taking place will also affect its reliability results, such as a calm atmosphere, noisy conditions, and many disturbances.
b. The size of the sample taken. The more samples taken, the more reliable an instrument will be.
c. Differences in the talents and abilities of the respondents being tested. Tests given to respondents with different talents and abilities will produce different reliabilities.[18]

7. Instrument Reliability Analysis with Excel

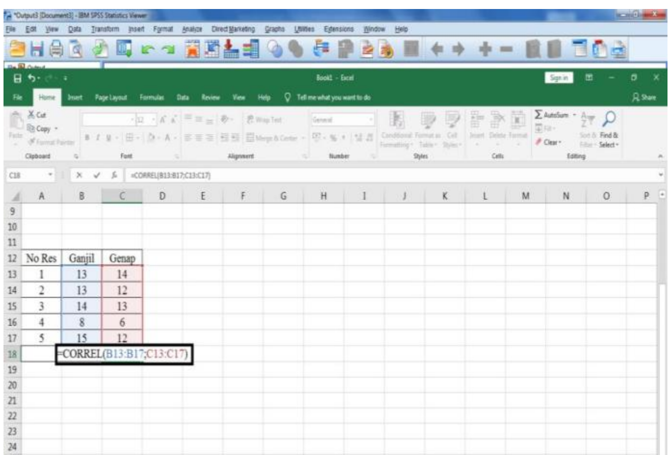
The steps for analyzing instrument reliability using Excel are as follows:

a. Separate respondents' answers with odd numbers and even numbers



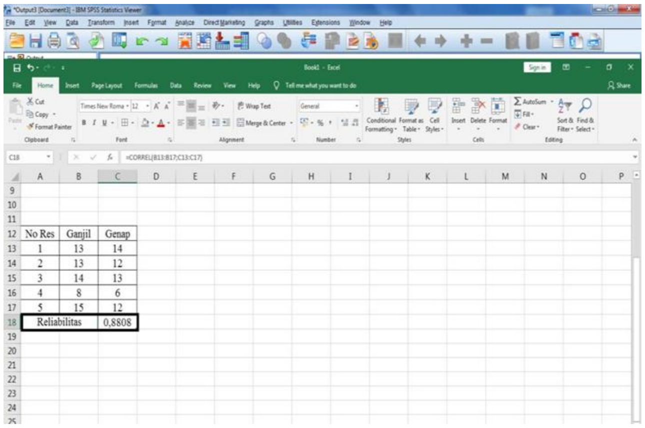
No Res.	Item Ganjil					No Res.	Item Genap				
1	Item 1	Item 3	Itm 5	Item 7	Total	1	Item 2	Item 4	Item 6	Item 8	Total
2	4	2	3	4	13	2	3	4	3	4	14
3	3	3	4	3	13	3	4	3	2	3	12
4	4	2	4	4	14	4	4	4	3	2	13
5	1	4	1	2	8	5	2	1	1	2	6
6	4	3	4	4	15	6	4	3	2	3	12

b. Copy the odd total score and even total score in new cells, as shown below. Then type equal (=), choose correl => array1 for odd items => array2 for even items, and enter.



No Res	Ganjil	Genap
1	13	14
2	13	12
3	14	13
4	8	6
5	15	12
=CORREL(B13:B17;C13:C17)		

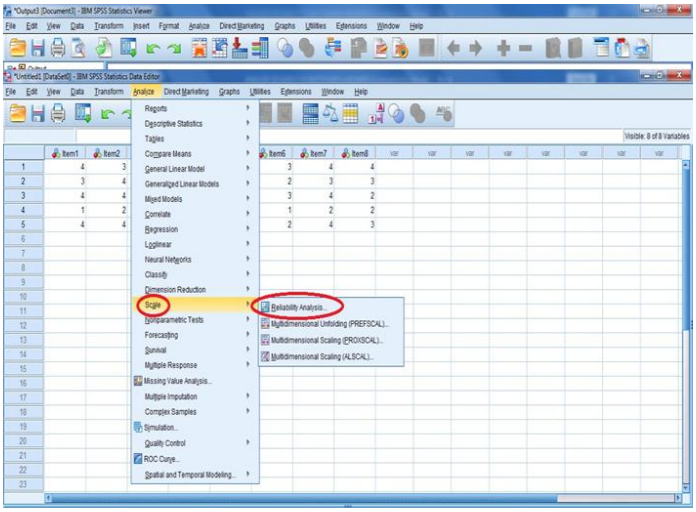
c. So that the results of the instrument reliability test are obtained as follows in the figure



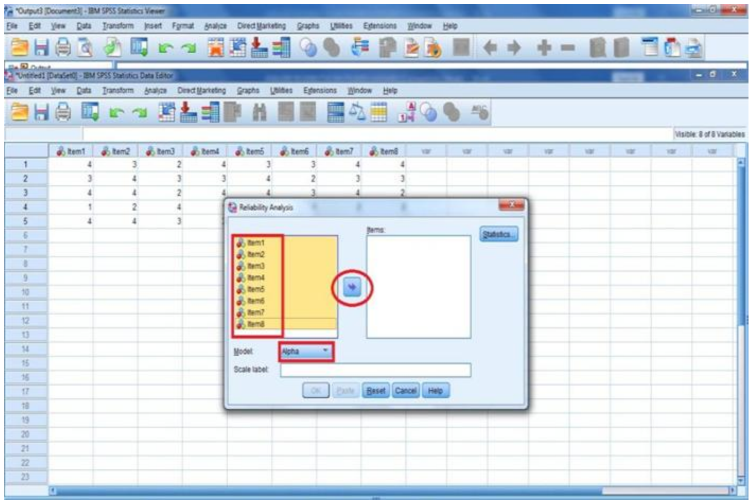
8. Instrument Reliability Analysis with SPSS

The steps in analyzing instrument reliability using SPSS are as follows:

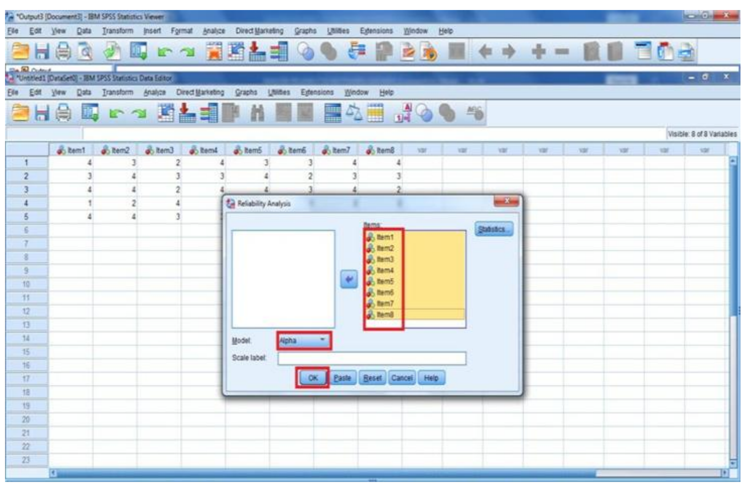
a. Click the analyze => scale => reliability analyze menu



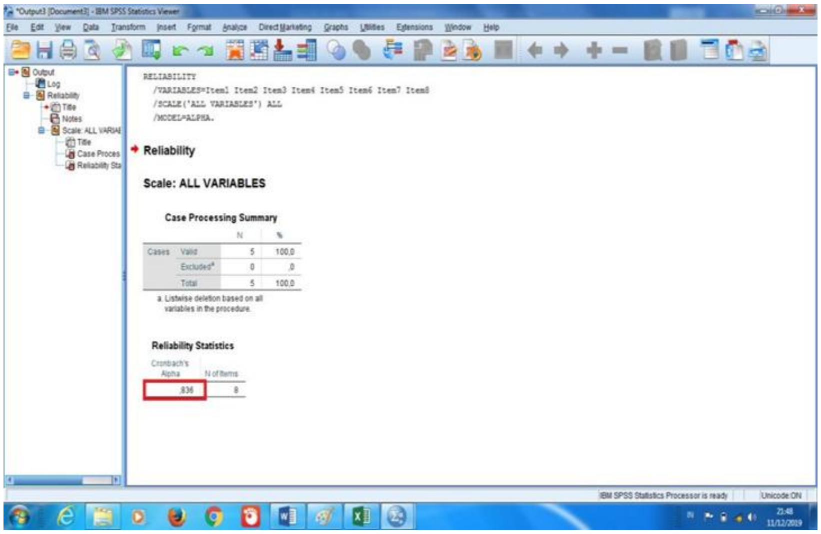
b. So it appears like the following image



c. Move the data from item 1 to item 8, ensure it is in alpha mode, and then click okay.



d. So that the reliability test results will appear as shown below



Conclusion

Instrument reliability refers to the extent to which a measurement produces consistent results when repeated measurements are taken with the same instrument. The types of instrument reliability include response consistency reliability and combined item consistency reliability. Based on the results of the instrument reliability analysis, it can be concluded that the research instrument used in this study has high reliability, with a reliability coefficient of 0.85. The instruments are reliable for measuring the variables under study and can serve as a reference for the development of more effective research instruments in the future. The implications of these findings are significant for improving research instrument reliability, and this study may serve as a reference for other researchers aiming to develop more reliable instruments. The authors suggest the development of better and more reliable instruments, the use of instruments tested for reliability, and further research to test the validity of the instruments. However, the study has limitations, such as a relatively small sample size and the need to improve the research instruments. Testing the instruments' validity and developing more reliable instruments are recommended for future research.

References

Amalia, A. N., & Widayati, A. (2012). Analisis butir soal tes kendali mutu kelas XII SMA mata pelajaran ekonomi akuntansi di kota Yogyakarta tahun 2012. *Jurnal Pendidikan Akuntansi Indonesia*, *10*(1).

Arifin, Z. (2017). Kriteria instrumen dalam suatu penelitian. *Jurnal Theorems (the Original Research of Mathematics)*, *2*(1), 28–36.

Arkunto, S. (2006). *Dasar-Dasar Evaluasi Pendidikan*. Bumi Aksara.

Arkunto, S. (2013). *Prosedur Penelitian Suatu Pendekatan Praktek*. Rineka Cipta.

Gavrilenko, V. I. (2016). *Optics of Nanomaterials* (1st Edition). Jenny Stanford Publishing. https://doi.org/10.1201/9781315272894

Janti, S. (2014). Analisis validitas dan reliabilitas dengan skala likert terhadap pengembangan si/ti dalam penentuan pengambilan keputusan penerapan strategic planning pada industri garmen. *Prosiding Snast*, 155–160.

Mansur, H. R. (2019). *Asesmen Pembelajaran di Sekolah Panduan Bagi Guru dan Calon Guru*. Pustaka Pelajar.

Matondang, Z. (2009). Validitas dan Reliabilitas Suatu Instrumen Penelitian. *Jurnal Tabularasa*, *6*(1), 87–97.

Mohajan, H. K. (2017). Two criteria for good measurements in research: Validity and reliability. *Annals of Spiru Haret University. Economic Series*, *17*(4), 59–82.

Sudaryono, S. (2012). *Dasar-Dasar Evaluasi Pembelajaran*. Graha Ilmu.

Sugiyono. (2019). *Statistik Untuk Penelitian*. CV. Alfabeta.

Sukiman, S. (2012). *Pengembangan Sistem Evaluasi*. Insan Madani.

Surapranata, S. (2019). *Analisis Validitas, Reliabilitas, dan Interpretasi Hasil tes Implementasi Kurikulum 2004*. PT. Remaja Rosdakarya.