SCHOLA
ISSN (Online): 2988-7100
Vol. 1, No. 1, Mei 2023, Hal. 17-22
Available Online at https://journal2.upgris.ac.id/index.php/schola/

Analysis of the Validity and Reliability of Indonesian Vocabulary Mastery Instrument Items Using AIKEN'S Model Calculations

Kartika Rismawati
Master of Basic Education, University of PGRI Semarang

ABSTRAK

Penelitian ini merupakan bagian dari penelitian pengembangan media pembelajaran kartu domino untuk meningkatkan penguasaan kosakata bahasa Indonesia siswa SD. Tujuan dari penelitian ini adalah untuk mengetahui validitas isi dan reabilitas dari butir instrument penguasaan kosakata Bahasa Indonesia yang telah disusun berupa isian. Analisis isi instrument menggunakan Koefesien V oleh Aiken's sedangkan Reliabelitas menggunakan perhitungan Cronbach Alpha. Hasil penelitian ini menunjukkan bahwa: setiap butir instrument penguasaan kosakata Bahasa Indonesia yang telah disusun terbukti valid , tertinggi dengan koefisien V Aiken's 0,94. Namun dalam hal ini intrumen yang disusun masih rendah dalam uji realibilitas dengan koefisien Cronbach's Alpha sebesar 0,309. Sehingga instrument yang telah disusun perlu adanya revisi sebelum digunakan.

Kata kunci : validitas isi, model AIKEN'S, reliabilitas

ABSTRACT

This research is part of the research on the development of domino card learning media to improve elementary school students' mastery of Indonesian vocabulary. The purpose of this study was to determine the content validity and reliability of the items of the Indonesian vocabulary mastery instrument that had been prepared in the form of entries. The analysis of the instrument content uses Coefficient V by Aiken's while Reliability uses the Cronbach Alpha calculation. The results of this study indicate that: each item of the Indonesian vocabulary mastery instrument that has been compiled is proven to be valid, the highest with Aiken's V coefficient of 0.94. However, in this case the instrument compiled is still low in the reliability test with the Cronbach's Alpha coefficient of 0.309. So that the instrument that has been prepared needs to be revised before being used.

Keywords: content validity, AIKEN'S model, reliability

INTRODUCTION

A test can be said to be good as a measuring tool that must meet the main criteria/principles of preparing a test in order to obtain a good and correct test tool, so that the test results describe the actual state of the object being measured. Measurement instruments, for example in the form of tests or questionnaires are proven valid if the expert (expert) believes that the instrument measures the abilities defined in the domain or also the psychological construct being measured. In order to obtain content validity, test items must describe important skills.

In everyday life, the term valid is no longer foreign to our ears, especially those related to the measurement process. For example, suppose if we want to measure body weight, then the proper measuring instrument used to measure it is a scale, not a ruler. In this case, the scale can be called a valid measuring instrument because it can obtain the correct measurement results in units of kilograms, while the ruler is definitely not valid because it cannot show the correct measurement results with the measurement results in centimeters which incidentally is a unit of length measurement. Another example is that when measuring height, a valid tool to use is a tape measure, not a ruler measuring 30 cm, even though the tape measure and ruler both actually produce measurement results in centimeters. The meter is more valid than

the 30 cm ruler, because the height measurement stage only requires one measurement process if using a tape measure, whereas using a 30 cm ruler requires more than one measurement process because the standard height exceeds the 30 cm ruler. the.

In research in the field of education, the term valid cannot be separated from instrument items, because the validity of instrument items is a determinant of success/accuracy in the process of measuring objects studied in the field of education. The instrument items can be said to be good if they are valid. Valid means that the items of the instrument can correctly be used to measure the object to be measured with the measurement results reflecting the characteristics of the object precisely. This statement is also in accordance with several research results conducted by Matondang [1], Sugiharni [2-3], Siswanto [4], Divayana, et al [5], Amala [6], Setyawati, Happy, and Murtianto [7] , Sugiharni and Setiasih [8], Hendryadi [9], Diponegoro [10], which states that basically the measurement result of the measurement process is a quantity that is able to reflect accurately and with certainty about the actual state/facts of what is to be measured through items -instrument items.

The research instrument is a tool used to measure observed natural and social phenomena (Sugiyono, 2009: 147). The definition of an instrument in the realm of evaluation can be stated as a device for measuring children's learning outcomes which includes learning outcomes in the cognitive, effective and psychomotor domains. Instruments that have been made must be analyzed first in order to determine the quality of the instrument. the instrument is said to be qualified if it meets the characteristics of the assessment so that it is said to be valid and reliable.

Gronlund (2009) argues that a good instrument when used as an evaluation tool should meet the demands for validity, meaning that it only measures one aspect, and demands for reliability, namely being reliable in measurement. This reliability includes the accuracy of measurement results and the constancy of measurement results. Guided by this view, when viewed in terms of empirical test validity analysis calculations, as a whole it meets the expected qualifications.

Specifically in evaluative research, a valid instrument is very important as a measuring tool to obtain an overview of the level of effectiveness of the object being evaluated. Based on these needs, it is necessary to carry out an appropriate calculation process in determining the validity of the instrument using several formulas, including: Gregory, Aiken, Point Biserial, and Product Moment Correlation.

Likewise, in this study it is hoped that a valid evaluation instrument will be produced. One of the formulas that has been used is the Gregory formula, but this formula is still not optimal for measuring the validity of each item of the instrument because the Gregory formula only displays the overall content validity of the evaluation instrument. To get around this problem, the process of calculating the validity of the instrument items uses the Aiken formula because this formula can carry out the calculation process to determine the validity of each instrument item with the assessment of more than 2 experts, so that the calculation results will be more accurate.

Several studies are the background of this study

RESEARCH METHODS

This study uses a quantitative approach because the expected results in this study are to show the process of calculating instrument item validation. The subjects involved in conducting the instrument item validation test were 3 validators who were co-authors. The object of this study was an instrument for mastering Indonesian vocabulary for elementary school students in the form of fillings. The location of this research is in

SCHOLA
ISSN (Online): 2988-7100
Vol. 1, No. 1, Mei 2023, Hal. 17-22
Available Online at https://journal2.upgris.ac.id/index.php/schola/

Pekalongan. Furthermore, instruments that have met the validity and reliability standards can be used for the measurement stage.

To manage the existing data, quantitative descriptive analysis is used. The stages of management and analysis are as follows:

1. The level of validity (validity of the questions)

The way to calculate the validity of a test is by using the "aiken's V validity coefficient" with the formula:

$$V = \sum S/[n(c-1)]$$

- lo = angka penilaian validitas terendah
- c = angka penilaian validitas tertinggi
- r = angka yang diberikan penilai
- s = r – lo

To determine the value of validity classification can be determined based on the following criteria:

Tabel 1. Klasifikasi Koefisien Validitas Tes

Kriteria Skala	Klasifikasi
0,00 – 0,20	Sangat rendah
0,21 – 0,40	Rendah
0,41 – 0,60	Cukup
0,61 – 0,80	Tinggi
0,81 – 1,00	Sangat Tinggi

(Sumber: Arikunto, 2009)

2. Determine the value of reliability

To determine the value of reliability used the formula r_{11} as follows :

$$r_{11} = \frac{k}{k-1} \; x \left\{1 - \frac{\sum S_i}{S_t}\right\}$$

Dimana :
r_{11} = Nilai reliabilitas
$\sum S_i$ = Jumlah varians skor tiap-tiap item
S_t = Varians total
k = Jumlah item

To determine the value of reliability classification can be determined based on the following criteria :

Tabel 2. Klasifikasi Reliabilitas Tes

Kriteria Skala	Klasifikasi
0,00 – 0,20	Sangat rendah
0,21 – 0,40	Rendah
0,41 – 0,60	Cukup
0,61 – 0,80	Tinggi
0,81 – 1,00	Sangat Tinggi

(Sumber: Arikunto, 2009)

If the instrument items belong to the very high, high, and sufficient categories, then these items are used. However, if the evaluation instrument items belong to the low and very low categories, then the instrument items are discarded/not used. If some of the items or even all of the items fall into the low category, it is necessary to review and even re-create the instrument items in each aspect with in-depth analysis and validation of the material content.

SCHOLA
ISSN (Online): 2988-7100
Vol. 1, No. 1, Mei 2023, Hal. 17-22
Available Online at https://journal2.upgris.ac.id/index.php/schola/

RESULTS AND DISCUSSION

a. validity

The validity of the instrument here is in terms of the validity of the instrument items. Analyzing the validity of the instrument can be done in a logical way and in an empirical way. The logical way of validating the instrument means that the instrument is analyzed in a rational way, namely by analyzing the suitability of the instrument with the material and demands of the applicable curriculum. Meanwhile, analyzing the instrument in an empirical way means that the instrument is analyzed according to experience.

In this study, the results were obtained in the form of a table of validity coefficients using Aiken's V calculations. Aiken's V validity coefficients were based on the results of an expert panel's assessment of n people on an item regarding the extent to which the item represented a construct. The following is the result of the classification of Aiken's V validity coefficient calculation for each item.

Table 3
Aiken's V statistical validation results

Scale criteria	No. items	Classification
0,81 – 1,00	1, 2, 3, 4, 5, 6 ,7, 8, 9, 10, 11, 12, 14, 15, 16, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35	Very high
0,61 – 0,80	13, 17, 18, 25, 27	High
0,41 – 0,60	-	Enough
0,21 – 0,40	-	Low
0,00 – 0,20	-	Very low

Hasil analisis validasi dengan Aiken's V menunjukkan bahwa ada 30 butir soal dengan very high classification and 5 high classification. It can also be seen that there are no items with sufficient, low and very low criteria. So it can be concluded that all items totaling 35 items are valid.

This can also be interpreted that all item questions can be used and nothing is dropped. This is because from the calculation of the Aiken's V coefficient it meets the specified criteria. Items that are in a very high classification include numbers 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 14, 15, 16, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35. The remaining 5 items, namely numbers 13, 17, 18, 25, 27 are in the high classification.

b. Reliability

the reliability test on this instrument uses the calculation of the Cronbach alpha formula. The results of calculating the coefficients with the Cronbach alpha formula can be seen in the following table.

SCHOLA
ISSN (Online): 2988-7100
Vol. 1, No. 1, Mei 2023, Hal. 17-22
Available Online at https://journal2.upgris.ac.id/index.php/schola/

Table 4
Reliability Results with Cronbach Alpha Statistics

Scale criteria	Item Calculation Results (using the Cronbach Alpha formula)	Classification
0,81 – 1,00		Very high
0,61 – 0,80		High
0,41 – 0,60		Enough
0,21 – 0,40	0,309	Low
0,00 – 0,20		Very low

From the results of these calculations it can be concluded that the coefficient obtained is in the low classification. So that this instrument can be said to be unreliable and needs improvement before being tested.

CLOSING

Based on the results of the analysis of the validity and reliability of the test items on the test instrument for increasing Indonesian vocabulary for elementary students, the conclusions are obtained:

1. The level of validity of the items on the Indonesian vocabulary mastery test has a very high and high validity value, with a total of 30 items with very high scores while 5 items with high scores. This har means that all item items are declared valid.
2. The level of reliability of Indonesian vocabulary mastery test questions has a low reliability value, with a coefficient value of 0.309. This means that this instrument is not reliable and requires improvement.

Suggestions for researchers who wish to conduct similar research are as follows (a) The more experts who study the instrument, the better the quality of the instrument in terms of content validity, (b) it is important to prepare more than one for each indicator in order to avoid underrepresentation of an indicator if certain items fail in the validation process (c) it is important to make good items so that they meet the required validity and reliability so that it is easy to get the expected research data.

THANK-YOU NOTE

Thanks to Dr. Sumarno, M.Pd and Dr. Muhtarom, M.Pd as a Lecturer in the Educational Assessment Course, also to her husband and family for their support and friends - friends of postgraduate students in arms for their support. Don't forget to thank both parents for all their best wishes.

REFERENSI

Aiken, L. R. (1980). Content validity and reliability of single items or questionnaires. *Educational and Psychological Measurement*, 40 (4), 955–959. https://doi.org/10.1177/001316448004000419

Anita, dkk. 2018. *Analisis Butir Soal Fisika Kelas X Sekolah Menengah Atas*. Vol. 16 No. 1 Asrul, dkk. 2014. Evaluasi Pembelajaran. Medan: Perdana Mulya Sarana.

Arikunto, Suharsimi. 2009. *Dasar - Dasar Evaluasi Pendidikan (Edisi Revisi)*. Jakarta: Bumi Aksara.

Azizah, R N. 2017. *Analisis Butir Soal Pilihan Ganda Ujian Akhir Semester Ganjil Tahun 2016/2017 Mata Pelajaran Ekonomi Kelas XII Di SMA Swasta Se-Surakarta*. ISSN : 2503-4855

Jannah, Miftahul. 2019. *Analisis Butir Soal Ujian Akhir Semester (UAS) Mata Pelajaran Fisika Pada Ujian Semester Genap Kelas XI Tahun Ajaran 2017/2018 Di SMAN 16 Banda Aceh*. Skripsi : Universitas UIN Ar-Raniry.