Comparative Test of Cronbach's Alpha Reliability Coefficient, Kr-20, Kr-21, And Split-Half Method

Hari Sugiharto Setyaedhi

A R T I C L E   I N F O

Article history:
Received September 15, 2023
Accepted January 20, 2024
Available online February 25, 2024

Keywords:
Empirical Quality, Final Semester Exam, Graphic Media Development.

A B S T R A C T

Final semester exams have an essential role in measuring student learning outcomes. Therefore, the tests must be of good quality. The problem is that lecturers must use reliable measuring tools to measure learning outcomes. This research aims to compare various reliability coefficients to determine which method has the highest reliability. The data collection methods in this research are observation, interviews, and tests based on student responses to the final semester exams. The data analysis technique uses the reliability of the Alpha Cronbach formula, Kuder Richardson 20, 21, and the split-half method consisting of Spearman-Brown, Flanagan, Rulon, and Hoyt. This type of research uses quantitative description. Comparative research designs are ex post facto. The research subjects were 90 undergraduate students. The research results provide several conclusions: first, the odd-even Rulon method has the highest reliability coefficient, namely 0.86; second, all reliable methods have a coefficient above 0.7 as the minimum requirement for acceptance; and third, all reliability methods have relatively the same reliability coefficient, namely around 0.8. Hence, all reliable methods are suitable for measuring tools in the final semester exam. This research implies that lecturers are advised to use the odd-even Rulon reliability coefficient when measuring learning outcomes.

1. INTRODUCTION

The success of the learning process can be measured through student learning outcomes. Student learning outcomes come from the assessment process carried out by the course lecturer (Indriasari et al., 2021). At the end of every semester, universities conduct end-of-semester assessments, or what are usually called final semester examinations (UAS) or final tests (Subarkah et al., 2020; Suradi, 2022). The aim of implementing the UAS is to test student learning outcomes based on what they have learned during one semester (Kadariah, 2018; Purba, 2021; Supriyadi, 2017; Suradi, 2022). Before doing the UAS, the lecturer makes a test instrument. Test instruments are used to measure learning outcomes (Magdalena, Syarijah, et al., 2021). Learning outcome tests are used as an assessment tool to make appropriate decisions about whether the student has achieved the learning objectives (Tri Jampi Setiyorini, Zyah Rochmad Jaelani, 2022; Ulfah et al., 2020). This can be understood because tests are a tool to measure whether educational goals are achieved (Hikmah & Muslimah, 2021). Tests can also determine the success of a learning program (Apsari & Acep Haryudin, 2017). Tests also function as a tool to inform students
about their mastery of the learning material (Wenno et al., 2021). Tests are a planned effort carried out by lecturers to show learning results to students (Kurniawati, 2019). Test instruments as measuring tools must meet the requirements, namely, they must be valid and reliable (Arfin, 2017; Çakir, 2022; Lia et al., 2020). What is worth paying attention to in a test instrument is that it meets the demands of validity and reliability, namely the accuracy of the measurement results and the consistency of the measurement results (Kriegstein et al., 2022; Nengsi & Efrina, 2019). An instrument (a measuring instrument) is said to be of good quality if it has been tested for validity and reliability (Dewi & Sudaryanto, 2020; Puspasari & Puspita, 2022). A test instrument is said to be good if it has evidence of validity and reliability and is suitable for use (Afiatunnisa et al., 2022; Bashooir & Supahar, 2018; Budiantoro & Kurniawan, 2021).

A reliable test is one that will produce relatively the same data even though it is used by different people, times, and places with the same research objectives or research subjects (Lestari & Yudhanegara, 2017). Reliability is the accuracy or stability of the assessment results (Lia et al., 2020; Surucu & Maslakci, 2020). The reliability of a test aims to measure the consistency, precision, repeatability, trustworthiness, etc. of the test (Ajayi, 2013; Nath, 2013; Sarwiningsih, 2017). Reliability is needed to determine whether the measuring instrument, in the form of a test, is reliable or consistent over time (Afah, 2021; Bahri, 2019; Erfan et al., 2020; Sanaky, 2021; Syahfitri et al., 2018). It is hoped that test instruments that have high reliability can serve as a guide or reference for measuring student abilities (Indriasari et al., 2021; Sarwiningsih, 2017). Test instruments as quality measuring tools can produce reliable and trustworthy information (Danni & Tauratiya, 2020). A quality instrument will produce accurate information so that measurement errors can be minimized (Faridah, 2021; Retnawati & Hadi, 2014). Quality instruments have good validity and reliability. This explanation proves that validity and reliability are absolute requirements that must be possessed by quality instruments, including sets of questions made by educators or lecturers (Danni & Tauratiya, 2020; Faridah, 2021). The reliability of assessments after teaching interventions is important to be carried out continuously (Lockhart, 2015). Reliability in quantitative research is very important (Campbell et al., 2013). Generally, before and after data collection, researchers need to consider the reliability of the data. Therefore, various methods are used to increase the validity and reliability of the data (Zohrabi, 2013). There are various types of reliability measurement methods, such as Cronbach’s alpha, KR-20, and KR-21, and split-half methods such as the Spearman-Brown formula, Flanagan formula, Rulon formula, and Hoyt formula (Ajayi, 2013; Ekolu & Quainoo, 2019; Gunartha, 2022; Puspasari & Puspita, 2022).

Test reliability can be influenced by several factors, such as the characteristics of test takers, test conditions, variations in test administration, as well as errors and differences in scoring, the length of the test, the homogeneity of student abilities, and the level of difficulty of test items (Busnawir dalam (Putri & Nahadi, 2019). The research results state that test reliability is closely related to how the test is presented, the mood of the test taker, the test taker’s attitude when facing the test, motivation, the condition of the test room, and so on (Putri & Nahadi, 2019). The procedures that must be carried out in developing a reliable test instrument are: 1) compiling test specifications; 2) writing test questions; 3) analyzing question items qualitatively; 4) conducting test trials; 5) analyzing question items quantitatively; 6) revising the test; 7) compiling the test; 8) implementing the test; and 9) interpreting the test results. (Ndjung & Jediu, 2020). The problem is, lecturers in graphic media development courses actually already know about the use of reliability as a prerequisite for test equipment to become a good measuring instrument, but they don't do it. Instructors must also understand various reliability coefficient formulas, such as Cronbach's alpha, 2), KR-20, 3), KR-21, 4). Spearman Brown, odd-even (5). Spearman Brown, beginning-end, 6), Flanagan odd-even, 7), Flanagan beginning-end (9), Rulon, odd-even, and 10). Rulon beginning-end. Which of the different types of reliability has the highest reliability coefficient? Because the reliability of an instrument is the consistency of an instrument in producing almost the same (relatively similar) data even though it is used by different people, times, and places with the same research objectives or research subjects (Lestari & Yudhanegara, 2017).

Gap analysis occurs between expectations and reality. It is hoped that all lecturers who teach graphic media courses can find out the high test reliability coefficient to determine student learning outcomes and improve the quality of test equipment. The reality in the field shows that: 1) most lecturers do not understand how to determine test reliability; 2) lecturers do not make maximum use of test reliability coefficients; and 3) lecturers do not know the various reliability formulas developed by experts. Ernawati's research shows that the testing tools used so far have been tested qualitatively, but quantitative testing tools have never been implemented (Erawati, 2018). The urgency of this research is that lecturers need to calculate and understand how to determine the reliability coefficient of a test instrument, because accuracy in determining reliability is very necessary to determine the reliability of a test, which will then be used as a measuring tool to evaluate the reliability of the test instrument. This is one of the expected improvements in the quality of education because reliable test instruments have a big
influence on analyzing student abilities as reference material for improving the quality of learning (Sarwinihs, 2017). Testing device reliability is considered very important because it is used as a measuring tool to obtain data and information related to the problem being studied (Ayu & Rosli, 2020). It is important for educators and lecturers to know about reliability testing because it determines how reliable a measurement is because of its reliability (Yusup, 2018). Reliability is very important in measurement, especially to obtain consistent measurement results (Khumaedi, 2012). Some relevant previous research related to the reliability of a test instrument is: Arslan's research shows that a reliable TPACK (Technological Pedagogical Content Knowledge) instrument has an impact on the quality of language questions (Arslan, 2020). Research on the effectiveness of using Padlet in distance learning. As a result, the study recommended the use of padlet (Al Momani & Musa, 2022). Research shows that the professionalism mini-evaluation exercise (P-MEX) has adequate validity and reliability in assessing citizens' professionalism and has a positive impact on education (Taşı et al., 2023). Research on critical thinking skills tests for high school students "produced reliable results in measuring the critical thinking skills of high school students (Orhan & Çeviker AY, 2022). Research aimed at self-assessment among Iranian EFL university students shows that while self-assessment is highly reliable (Manzari, 2023). Research on the Simulation Acceptance Scale (SAS) that was developed can be used because it is reliable (Sezer et al., 2020). Research to provide evidence of the validity and reliability of student assessments of teaching quality using assessments from TIMSS 2019 (Senden et al., 2023).

Research that develops valid and reliable games, good for educators through game-based assessment in physical education (Wilkie et al., 2023). Research aimed at evaluating POMA (Tinetti Performance Oriented Mobility Assessment) stated that POMA has satisfactory reliability and validity among Chinese elderly (Yang et al., 2023). The Canadian Agility Movement Skills Assessment (CAMS A) has instrument reliability and validity so it can be used to assess motor competence in Spanish children (Menescardi et al., 2022). Research that aims to develop a valid and reliable scale to measure students' expectations from the discipline program implemented in their school (Özcan et al., 2020). Research on the Analysis of the National Middle School English Final Examination in View of Validity and Reliability" calculates the reliability results of the Middle School English National Examination using the Kuder-Richardson Formula (KR-20), showing a reliability coefficient value of 0.89 (Sugianto, 2016). The research entitled "Validity and reliability of educational family life measurement instruments” produced a Cronbach’s alpha value of 0.950, indicating that the structure of the child measuring instrument (CMI) and the structural dimensions of the mother's measuring instrument MMI are very reliable. The novelty of this research is that previous research did not widely compare reliability coefficients such as: Cronbach's alpha, 2). KR-20, 3). KR-21, 4). Spearman Brown, odd-even (5). Spearman Brown, beginning-end, 6). Flanagan odd-even, 7). Flanagan beginning-end (9). Rulon, odd-even, and 10). Rulon beginning-end to identify the highest reliability coefficient so that it is appropriate to use as a measuring tool for graphic media UAS. Based on the explanation of the problems above, this research was carried out with the aims of 1) describing several methods for testing the reliability of coefficients, 2) describing several methods for determining reliability coefficients, 3) determining reliability coefficient tests, and 4) determining the reliability coefficients of tests from UAS for media courses graphic. The reliability coefficient compared is 1). Cronbach’s alpha, 2). KR-20, 3). KR-21, 4). Spearman Brown, odd-even 5). Spearman Brown, beginning-end, 6). Flanagan odd-even, 7). Flanagan beginning-end (9). Rulon, odd-even, and 10). Rulon beginning-end.

2. METHODS

This type of research is quantitative, using a quantitative descriptive approach. This research includes comparative research, which is a type of research that aims to compare two or more groups, variables, or conditions to identify differences or similarities between them. Data collection methods include: a) observation. Data was obtained through direct observation according to reality (Dwijj et al., 2020). Observations were carried out in the Unesa Educational Technology department, with the research object being students who were taking graphic media courses. b) interviews aimed at obtaining accurate information from students and lecturers from various questions (Kboridah et al., 2019); c) tests in the form of final exams that aim to obtain data about student learning outcomes. The subjects of this research were 90 undergraduate students in the Department of Educational Curriculum and Technology, class of 2021–2022. The test instrument used is a multiple-choice objective test with 5 answer choices totaling 24. Data analysis uses various reliability formulas assisted by SPSS version 22 and Microsoft Excel. The instrument used in this research is an instrument in the form of a test with graphic design material. The reliability analysis was carried out based on data obtained from the UAS (Cevik & Senturk, 2019). The graphic media Final Semester Examination test instrument grid is presented in Table 1.
Table 1. Final Semester Examination Grid

<table>
<thead>
<tr>
<th>No</th>
<th>Aspect</th>
<th>Indicator</th>
<th>Item</th>
</tr>
</thead>
</table>
| 1  | Get to know bitmap and vector type graphic design | - Students know the characteristics of bitmap and vector graphic design.  
- Students know various bitmap and vector-type graphic designs.  
- Students know the characteristics of bitmap and vector graphic design. | 1, 2 |
| 2  | Shows menus and icons contained in Corel Draw graphic design software | - Students can name color groups in Corel Draw.  
- Students can mention the standard Corel Draw menu.  
- Students can mention terms in the Corel Draw program. | 8, 9, 10, 11, 12, 13 |
| 3  | Distinguish between bitmap and vector | - Students can differentiate bitmap and vector image results.  
- Students can differentiate between the purposes of using bitmaps and vectors. | 14, 15, 16, 17 |
| 4  | Shows menus and icons contained in Adobe Photoshop graphic design software | - Students can mention the standard Adobe Photoshop menu.  
- Students can mention the facilities in Adobe Photoshop. | 20, 21, 22, 23 |
| 5  | Determines applications for creating vector-based graphics | - Students can differentiate between bitmaps and vectors from the software they are made from. | 24 |

To estimate reliability using the split-half method, namely Spearman-Brown, Flanagan, and Rulon (Retnawati, 2017; Steinke & Kopp, 2020). Determining the reliability of learning outcomes tests in the form of objective tests is carried out by "splitting in two," or the Split-Half Technique, or single test, single trial question items, namely odd-even and early-end numbered questions (Betul & Turan, 2019; Jago, 2019; Ossai & Chiekem, 2022; Pröck et al., 2022). The language for reliability is $r$ (Fu et al., 2022). Calculations using the odd-even technique are grouping all the odd-numbered items into one group and giving the name "odd" group, while all even-numbered items are grouped into one group and naming the "even" group (Haq, 2022). The odd group consists of items numbered 1, 3, 5, 7, and so on, while the even group consists of items numbered 2, 4, 6, 8, and so on (Ajayi, 2013; Van Norman & Parker, 2018). The Spearman Brown reliability test uses the product moment correlation formula and after the correlation is found, the Spearman Brown formula is used (Ekolu & Quainoo, 2019; Nath, 2013; Yusup, 2018). Flanagan's reliability test uses the Split Half technique. Odd – even groups and early – late groups use the same formula as follows (Cho & Chun, 2018). Rulon's reliability test uses the Split Half technique. Odd – even groups and early – late groups use the same formula as follows (Cho & Chun, 2018). Formulas 20 and 21 from Kuder and Richardson are abbreviated as KR-20 and KR-21(Foster, 2021). The Kuder Richardson (KR) reliability test uses a single test technique - single trial. Both KR techniques have specific instrument criteria for using the formula. If the instrument has unequal or heterogeneous levels of difficulty for each question item then use the KR-20 formula to test its reliability (Fraenkel et al., 2013). If the instrument has the same or homogeneous level of difficulty for each question item then use the KR-21 formula to test its reliability (Cho & Chun, 2018; Ekolu & Quainoo, 2019; Foster, 2021; Yusup, 2018). C. Hoyt's formula analyzes scores on learning outcomes test items using variance analysis or ANOVA techniques (Magdalena, Fauziah, et al., 2021; Sudijono, 2006). Cronbach's alpha formula can be used to test essays, questionnaires or questionnaires (Yusup, 2018). Cronbach's alpha shows a tendency to measure consistency rather than reliability (Rosli et al., 2021). Cronbach's Alpha Formula to estimate reliability uses the following formula (Rosli et al., 2021). As for the interpretation of the test reliability coefficient ($r_{11}$) can use reliability categorization showed in Table 2.

Table 2. Reliability Categorization

<table>
<thead>
<tr>
<th>Test Reliability Coefficient</th>
<th>Categorization</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.00 &gt; 0.90</td>
<td>Very high reliability</td>
</tr>
<tr>
<td>0.75 - 0.90</td>
<td>Good reliability</td>
</tr>
<tr>
<td>0.50 - 0.75</td>
<td>Medium reliability</td>
</tr>
<tr>
<td>-1.00 &lt; 0.50</td>
<td>Poor reliability</td>
</tr>
</tbody>
</table>

Source: (Chen et al., 2021; Grgic et al., 2021; Ursavaş & Bayrak, 2021).
In general, the reliability coefficient of a test is described numerically in the range between $-1.00 \leq \rho \leq +1.00$. If this value is close to zero, then the item does not differentiate sufficiently between the upper and lower groups (Cakir, 2022; Retnawati, 2017). The test reliability coefficient can be said to be good if it has a coefficient above 0.70 (Alavi & Ghaemi, 2013; Alfiatunnisa et al., 2022; Isa & Azid, 2022; Tobón & Luna, 2021). The higher the coefficient, the smaller the level of measurement error, whereas the lower the coefficient, the greater the measurement error (Retnawati, 2017). The explanation above can give lecturers an idea of the need to compare the accuracy of estimates between reliability formulas so that lecturers can determine the appropriate reliability formula to use for measuring a test. The lecturer will choose a high reliability coefficient for the test meter.

3. RESULT AND DISCUSSION

Results

The results of the split-half calculation using the Spearman Brown odd-even formula produce a reliability coefficient of $r_{11} = 0.84$ (very high category). The results of calculations using the Spearman Brown beginning-end formula produce the following reliability coefficient: $r_{11} = 0.82$ (very high category). The results of calculations using the odd-even Flanagan formula produce a reliability coefficient of $r_{11} = 0.83$ (very high category). The results of calculations using the beginning-end Flanagan formula produce a reliability coefficient of $r_{11} = 0.8$ (very high category). The results of calculations using the odd-even Rulon formula produce a reliability coefficient of $r_{11} = 0.86$ (very high category). The results of calculations using the beginning-end Rulon formula produce a reliability coefficient of $r_{11} = 0.8$ (very high category). The results of calculations using the KR-20 formula produce the following reliability coefficient: $r_{11} = 0.82$ (very high category). The results of calculations using the KR-21 formula produce a reliability coefficient of $r_{11} = 0.8$ (very high category). The results of calculations using the Hoyt formula produce a reliability coefficient of $r_{11} = 0.81$ (very high category). The results of calculations using the Cronbach’s alpha formula produce a reliability coefficient of $r_{11} = 0.82$ (very high category). To make it easier for researchers and readers to analyze comparisons of reliability coefficients, researchers need to create a recapitulation table. The recapitulation is presented in Table 3.

Table 3. Reliability Coefficient Recapitulation

<table>
<thead>
<tr>
<th>Reliability Type</th>
<th>Results</th>
<th>Note</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spearman Brown</td>
<td>Odd – Even</td>
<td>$r_{11} = 0.84 &gt; 0.7$</td>
</tr>
<tr>
<td></td>
<td>Beginning - End</td>
<td>$r_{11} = 0.82 &gt; 0.7$</td>
</tr>
<tr>
<td>Flanagan</td>
<td>Odd – Even</td>
<td>$r_{11} = 0.83 &gt; 0.7$</td>
</tr>
<tr>
<td></td>
<td>Beginning - End</td>
<td>$r_{11} = 0.8 &gt; 0.7$</td>
</tr>
<tr>
<td>Rulon</td>
<td>Odd – Even</td>
<td>$r_{11} = 0.86 &gt; 0.7$</td>
</tr>
<tr>
<td></td>
<td>Beginning - End</td>
<td>$r_{11} = 0.8 &gt; 0.7$</td>
</tr>
<tr>
<td>Kuder Richardson</td>
<td>20</td>
<td>$r_{11} = 0.82 &gt; 0.7$</td>
</tr>
<tr>
<td>(KD)</td>
<td>21</td>
<td>$r_{11} = 0.8 &gt; 0.7$</td>
</tr>
<tr>
<td>Hoyt</td>
<td></td>
<td>$r_{11} = 0.81 &gt; 0.7$</td>
</tr>
<tr>
<td>Alpha Cronbach</td>
<td></td>
<td>$r_{11} = 0.82 &gt; 0.7$</td>
</tr>
</tbody>
</table>

In Table 3, the reliability coefficients use different reliability formulas. The results of all differences in the calculation results between the reliability coefficient formulas used. Research analysis shows that; a) the reliability coefficient in classical test theory has relatively almost the same estimation accuracy, namely 0.8. b) The reliability coefficient has an estimated average of around 0.8; c) All reliability coefficients are above 0.7, which means they are in the very high category; d) All reliability coefficients are suitable for use as a measuring tool for graphic media UAS; e) Reliability coefficient The odd-odd rule has the highest value, namely 0.86. The conclusion is that all reliability coefficients can be used as a tool to test the graphic design UAS test instrument. However, the highest reliability coefficient is obtained by using the odd-even Rulon formula to estimate the reliability coefficient.

Discussion

Reliability is an important psychometric characteristic for all measurements (Steinke & Kopp, 2020). Reliability with the split-half method is the division of question items with odd-even numbers or first and last-numbered questions (Steinke & Kopp, 2020). The comparison test results for all average reliability coefficients are around 0.8 or above 0.7. Research states that the range of reliability coefficients

Hari Sugiharto Setyaedhi / Comparative Test of Cronbach’s Alpha Reliability Coefficient, Kr-20, Kr-21, And Split-Half Method
for a test numerically ranges between $-1.00 \leq \rho \leq +1.00$ (Çakir, 2022; Retnawati, 2017). Correlation coefficient values $< 0.50$ (poor reliability), values between $0.50$ and $0.75$ (medium reliability), values between $0.75$ and $0.90$ (good reliability), and correlation coefficient values $> 0.90$ (very good reliability) (Chen et al., 2021; Grgec et al., 2021; Ursavaş & Bayrak, 2021). Very good ($> 0.90$), good ($0.75 – 0.90$) and poor ($< 0.75$) (Moya-ramon et al., 2022). Other research states that the correlation coefficient is $> 0.60$, sufficient reliability, $0.61 – 0.80$ (good reliability), and $> 0.81$ (very good reliability) (Menescardi et al., 2022). Other research states that the correlation coefficient value is between $0 – 0.39$ (poor reliability), $0.4 - 0.59$ (medium reliability), $0.6 - 0.74$ (good reliability), and $0.75 - 1.0$ (very good reliability) (Keogh et al., 2019; Lee et al., 2021; Matheson, 2019). Other research states that the correlation coefficient value is between $0.00 – 0.25$ (poor reliability), $0.26 – 0.50$, moderate reliability ($0.51 – 0.75$), and good reliability ($0.76 – 1.00$) (Paraskevopoulos et al., 2023). Cronbach’s Alpha is used to measure reliability capabilities (Ford et al., 2015). Cronbach’s alpha ($\alpha$) assesses the internal consistency and reliability of the test instrument (Wafudu et al., 2022). Cronbach’s alpha is recommended to have very good reliability if it ranges between $0.7$ and $0.8$ (Cheung et al., 2023; Derakhshan et al., 2023; Filho et al., 2021; Kulan et al., 2022). Other research says Cronbach’s Alpha is accepted at a value of $0.7$ or $0.6$” (Taber, 2018). Other research says the Alpha value is $> 0.60$ for each variable (Ngaliman et al., 2019). Cronbach’s alpha $0.5 < \alpha < 0.8$ acceptable reliability (Ekolu & Quainoo, 2019). Thus, the results of the recap of the reliability coefficient calculations can be concluded that all reliability calculations are $> 0.70$, which means the reliability coefficient is in the very high category.

Theoretically, measurement results can be trusted if in several measurements on the same group of subjects relatively similar results are obtained, as long as the subjects being measured do not experience changes (Lastari & Yudhanegara, 2017; Syamsuddin, 2017). Stable construction over time when inspected (Krieglstein et al., 2022). Test reliability is closely related to validity because a valid measuring instrument can be guaranteed to be reliable, but a reliable measuring instrument is not necessarily valid (Sugiyono, 2013). Arikunto stated that it is important for a test to have validity and reliability requirements, (Arikunto, 2018). In that test it may be reliable, but it is not valid. On the other hand, a valid test is definitely reliable (Suci Mitra & Helendra, 2022). The quality of the questions used to measure test takers’ abilities needs to be considered, including whether the questions are valid and reliable. Apart from that, a question is said to be good if it is not too easy or too difficult. The questions must be able to differentiate between smart and not-smart test takers, and the effectiveness of the distractors must function well (Friatma & Anhar, 2019).

The results of the comparison test of the reliability coefficient of UAS scores for graphic media courses using various reliability test methods, as shown in Table 2, apparently produced a reliability coefficient with an average of above $0.80$. This is in accordance with several recent previous studies regarding reliability comparisons. Sarwiningsih’s research concluded that various types of reliability coefficients have almost the same accuracy (Sarwiningsih, 2017). Research conducted by Widhiarso and Mardapi (2010) concluded that the split half, KR-20, KR-21, and Cronbach’s alpha tests obtained relatively similar results (Widhiarso & Mardapi, 2010). Studies on the Turkish adaptation, validity, and reliability of the Early Childhood Teacher Academic Competency Assessment Scale produced Spearman Brown reliability coefficients and the Split Half Test. Reliability produced figures that were not much different (Betul & Turan, 2019). The advantage of this research is that the test instrument developed for the final semester exam in graphic media courses has very high reliability. This is proven by the average reliability coefficient, which is above $0.7$. This research aims to compare various reliability coefficients so that it is known which method has the highest reliability. The limitation of this research is that the researcher only tested the test instrument in terms of its reliability, even though a good test instrument must meet the validity and reliability test requirements. The solution to overcome the limitations of this research is that for future research, researchers will use validity and reliability tests with the same samples and subjects.

4. CONCLUSION

Based on the results of calculating the average reliability coefficient, the results show that all reliability formulas show coefficients above the predetermined average, meaning that all reliability coefficients are included in the very high category; thus, all types of reliability meet the requirements and are suitable to be used as measuring instruments. However, if you use reliability with the odd-even Rulon formula, the results are slightly higher compared to other reliabilities, so it would be better to calculate the UAS test for graphic media development courses using the odd-even Rulon formula.
5. REFERENCES


