

## Validity and Reliability of Instruments: Students' Perception of Google Classroom and Interest in Learning Islamic Religion

### ABSTRACT

The validity and reliability of measurement instruments are important aspects of educational evaluation to ensure accurate and reliable results. Without instruments that meet these standards, evaluation results risk being irrelevant to learning objectives. This study aims to analyze the external validity and reliability of questionnaire instruments for students' perceptions of Google Classroom and interest in religious studies at SMP Islam Ngoro. A quantitative approach was used in this study, with Pearson Product Moment analysis and Cronbach Alpha as statistical methods. The results showed that on variable X, of the 20 items analyzed, there were 11 valid items and 9 invalid items, with a reliability of 0.891. Meanwhile, on variable Y, there were 9 valid items out of 20 items, with a reliability of 0.787 other items were deleted. The instruments on both variables were declared reliable because they met the reliability criteria (value >0.6). This study contributes to ensuring the quality of accurate evaluation instruments, so that they can support the implementation of technology-based learning such as Google Classroom.

**Keywords:** Validity; Reliability; Student Perception; Google Classroom; Learning Interest

**Nur Alfiyah<sup>1</sup>,**  
UIN Maulana Malik Ibrahim Malang

**Naura Diny Chalisah<sup>2</sup>,**  
UIN Maulana Malik Ibrahim Malang

**Indah Aminatus Zuhriyah<sup>3</sup>,**  
UIN Maulana Malik Ibrahim Malang

### Address for Correspondence:

<sup>1</sup> [nuralfiyah17@gmail.com](mailto:nuralfiyah17@gmail.com)

<sup>2</sup> [nauradinikholisoh@gmail.com](mailto:nauradinikholisoh@gmail.com)

<sup>3</sup> [zuhriyah@pgmi.uin-malang.ac.id](mailto:zuhriyah@pgmi.uin-malang.ac.id)

## INTRODUCTION

In today's digital era, technology has changed many aspects of life, including education (Febrianti et al., 2023). One innovation that has been widely implemented is the use of Learning Management System (LMS), such as Google Classroom, to support the teaching and learning process. Recent data shows that digital media has been proven to increase student enthusiasm and the teaching technique skills of Islamic Religious Education teachers, because the integration of digital media, especially Google Classroom, allows for effective and flexible delivery of materials (Nurqozin & Putra, 2023). However, the implementation of this technology requires a thorough evaluation to ensure its effectiveness in improving the quality of learning (Sholihin et al., 2024). Recent research recommends that future researchers explore the interactions between other variables that may influence the effectiveness of online learning, as well as conduct longitudinal analyses to assess the long-term impact of digital platform use (Fauzan, 2025). In the context of Islamic Religious Education (PAI), such evaluation becomes even more important considering that this subject aims to shape students' understanding, attitudes, and religious behavior holistically. Therefore, the existence of accurate and consistent measurement instruments is necessary so that the evaluation process can accurately describe students' perceptions of Google Classroom as well as their interest in learning religion. Thus, the need for valid and reliable evaluation instruments becomes a major issue in the effort to strengthen the role of technology in Islamic religious education in schools.

This study aims to analyze the external validity and reliability of the questionnaire instruments of students' perceptions of Google Classroom and interest in religious studies at Ngoro Islamic Junior High School (SMP Islam Ngoro). Through a quantitative approach, this

study used Pearson Product Moment and Cronbach Alpha statistical analysis to test the instruments. Pearson Product Moment correlation was used to test the validity of the instrument by measuring the linear relationship between two variables (Salsabila et al., 2023). In the validation process, Pearson Product Moment analysis will be used to test the linear relationship between the instrument items and the main variables (Utama, 2024), thus ensuring that each question in the questionnaire is truly relevant and capable of measuring the expected concept. Meanwhile Cronbach Alpha is applied to measure reliability, which describes the internal consistency of the instrument in measuring the same variable (Anggraini et al., 2022). The reliability of the instrument will be assessed through the calculation of the Cronbach Alpha coefficient, which aims to assess the internal consistency of each item and ensure the stability of the measurement results in various application conditions. This is because the function of the reliability test is to find out how much trust the research instrument has as a data collection tool (Rosidin et al., 2024). By evaluating these two variables, the research is expected to provide an accurate and reliable picture of how students perceive faith-based learning technologies. The results of this analysis are also expected to contribute to the development of better evaluation instruments to support technology-based Islamic religious learning in schools.

Based on the researcher's analysis of several relevant scientific literatures, the tendency of research on evaluation instruments is more directed at the substance of science and technology subjects, and the types of evaluation tools used; so that not many have focused on studying how research instruments on student perceptions and their learning interests in religious learning. In fact, in the development of measurement instruments, validity and reliability are two fundamental concepts for all subjects (Subhaktiyasa, 2024). Validity refers to the extent to which an instrument is able to measure what it is supposed to measure, while reliability relates to the consistency of measurement results (Sürücü & Maslakci, 2020). Some previous studies emphasize the importance of validity and reliability in technology-based learning evaluation. According to Suharsis Arikunto, a valid instrument will produce relevant data and support appropriate decision-making, while a reliable instrument ensures the stability of measurement results in various situations. Recent research by Nwake (2023) shows that validity is not only related to the accuracy of measurement but also to the consequential impact of evaluation result-based decisions (Nkwake, 2023). In the context of PAI, the validity and reliability of the instrument becomes more complex because it must cover the cognitive, affective, and psychomotor dimensions of students related to the teachings of Islam. Thus, this research will be an alternative study related to how the validity and reliability tests of research instruments on student perceptions and their learning interests are applied in the context of religious learning in schools.

Validity and reliability are the main foundations that ensure the credibility of an evaluation instrument (Ahmed & Ishtiaq, 2021). Without validity, the measurement results cannot be considered relevant to the objectives to be achieved (Mulyana et al., 2024); and without reliability, the measurement results cannot be relied on for decision making (Defrijon, 2025). Therefore, this study emphasizes the importance of systematic validity and reliability testing in order to produce instruments that are able to represent the actual conditions of student understanding. The findings of this study contribute to the development of more accurate and standardized evaluation instruments to measure the effectiveness of technology-based learning in Islamic religious education. In addition, the results of this study provide insight for schools and educators in choosing and designing digital learning methods that are more appropriate to students' needs, so that they can increase their involvement and understanding of religious material in the digital era. Thus, the analysis of the questionnaire instrument in this study is expected to not only support a more credible learning evaluation, but also provide practical benefits in designing instruments that are appropriate to the context

of Islamic religious education and contribute to the policies of educational institutions that integrate technology in teaching.

## METHOD

This research uses a quantitative approach, where data in the form of numbers are then described (Sugiyono, 2017). As part of the instrument testing, an external test was conducted at Ngoro Islamic Junior High School (SMP Islam Ngoro), with respondents consisting of 20 seventh grade students. The instrument was designed to measure students' perceptions of Google Classroom and their interest in learning Religious Studies. Although the instrument was tested at Ngoro Islamic Junior High School, the main research will be conducted at SMPN 1 Ngoro. This external test is intended to ensure that the instruments to be used are relevant and reliable.

Validity and reliability are very important in the development of Islamic religious education tests (Alfiyah et al., 2025). In this study, the external validity of the instrument was tested using Pearson Product Moment correlation analysis to evaluate the relationship of each questionnaire item with the total score, to ensure that each item can represent the construct being measured. The reliability was tested using the Cronbach Alpha coefficient to determine the consistency of the measurement results (Forester et al., 2024). Data were analyzed using SPSS statistical software. Tests were conducted with a 5% significance level to provide accurate and relevant results (Duwi Priyatno, 2010).

Respondents used in the external test of this instrument were selected by purposive sampling. Purposive sampling is a sampling method where researchers choose/determine specific identities in respondents with certain considerations so that they are in accordance with research objectives and are expected to be able to respond to research cases better (Lenaini, 2021). Although the number of respondents was limited to 20 students, this instrument test is expected to provide a valid picture of the quality of the instrument before it is applied to a larger population. The data collected was analyzed descriptively to describe students' perceptions of the use of Google Classroom and their interest in learning religion in general.

## RESULTS AND DISCUSSION

A learning outcome test is a measuring tool used to make measurements to obtain learning outcome data. Because it is a measuring tool, it must meet the requirements. Based on the analysis of related references, a good measuring instrument is one that meets 2 criteria, namely valid and reliable. Therefore, before use, an instrument needs to be tested for validity and reliability (Purwanto, 2021).

According to Cronbach, test validation is a process by which test developers gather evidence to support the kinds of inferences that can be appropriately drawn from test scores (Osterlind, 2002). In this definition, more emphasis is placed on the interpretation of the scores generated by a test. Meanwhile, according to AERA/APA/NCME standards, validity is defined as the most important consideration in test development; which refers to the appropriateness, meaningfulness, and usefulness of specific inferences made from test scores. Validity is a unitary concept that refers to the extent to which evidence supports the inferences made by the scores. Validity indicates the extent to which specific items on a tool accurately assess the concept being measured, to ensure that the questions asked allow valid inferences to be made (Masuwai et al., 2024). Validity testing involves the calculation of correlation statistics, and the concept of correlation is related to the relationship between variables X (independent) and Y (dependent). The independent variable is the variable that causes the dependent variable to occur, while the dependent variable is the one caused by the independent variable. The level of correlation is denoted by the coefficient  $r$  (*relation*), so

that the relationship between X and Y becomes  $r_{xy}$ . The interpretation and significance level of the relationship between X and Y can be compared with  $r_{table}$ .  
 The product moment correlation formula:

$$r_{xy} = \frac{N \sum xy - (\sum x) (\sum y)}{\sqrt{\{N \sum x^2 - (\sum x)^2\} \{N \sum y^2 - (\sum y)^2\}}}$$

Where:

- $r$  = Correlation index number, "r" product moment
- $N$  = Number of cases
- $\sum xy$  = The sum of the product of x and y scores
- $\sum x$  = Total number of scores x
- $\sum y$  = Total number of y scores

Reliability is a concept that refers to the reliability and consistency of a measuring instrument or measurement. In this context, reliability refers to the extent to which measurement results can be trusted and provide the same results when carried out repeatedly under the same conditions (Mabruri Pudyas Salim, n.d.) . Anderson argues that validity is more important than reliability in the context of measurement and evaluation (Arikunto, 2013) . Validity refers to the extent to which a test or instrument actually measures what it is intended to measure, while reliability is concerned with the consistency of measurement results. According to Anderson, although reliability is necessary to support validity, a test can be considered reliable but not valid. Conversely, a valid test will usually also be reliable (Tentama & ., 2018) .

### I. Validity Test Result Data

After going through the expert judgment stage to conduct internal validation of the instrument, the researcher conducted an external validity test. The external test was applied to 20 students of Ngoro Islamic Junior High School. The data obtained was then analyzed using the Pearson Product Moment correlation formula. According to Sugiyono, the criteria so that an item is said to be valid if  $r_{count} > r_{table}$  (Sugiyono, 2017). Because the number of students sampled is 20, the  $r_{table}$  is 0.444. The following is a table of instrument validity test results:

**Table I.** Validity Test Results of Questionnaire Instrument Variable X  
 (Student Perceptions of Google Classroom)

No. Item	$r_{count}$ (Correlations)	$r_{table}$ ( $\alpha = 5\%$ )	Description
1.	0,658	0,444	Valid
2.	0,635	0,444	Valid
3.	0,489	0,444	Valid
4.	0,467	0,444	Valid
5.	0,109	0,444	Invalid
6.	0,118	0,444	Invalid
7.	0,542	0,444	Valid
8.	0,372	0,444	Invalid
9.	0,338	0,444	Invalid
10.	0,122	0,444	Invalid
11.	0,069	0,444	Invalid
12.	0,140	0,444	Invalid
13.	0,293	0,444	Invalid
14.	0,455	0,444	Valid
15.	0,329	0,444	Invalid
16.	0,589	0,444	Valid
17.	0,716	0,444	Valid
18.	0,719	0,444	Valid
19.	0,700	0,444	Valid

No. Item	$r_{count}$ (Correlations)	$r_{table}$ ( $\alpha = 5\%$ )	Description
20.	0,515	0,444	Valid

Based on the data from the calculation, it is known that there are 11 valid items and 9 invalid items.

**Table 2** Validity Test Results of Questionnaire Instrument Variable Y (Islamic Religious Learning Interest)

No. Item	$r_{count}$ (Correlations)	$r_{table}$ ( $\alpha = 5\%$ )	Description
1.	0,558	0,444	Valid
2.	0,535	0,444	Valid
3.	0,598	0,444	Valid
4.	0,241	0,444	Invalid
5.	0,083	0,444	Invalid
6.	0,241	0,444	Invalid
7.	0,064	0,444	Invalid
8.	0,392	0,444	Invalid
9.	0,444	0,444	Invalid
10.	0,521	0,444	Valid
11.	0,101	0,444	Invalid
12.	0,528	0,444	Valid
13.	0,222	0,444	Invalid
14.	0,439	0,444	Invalid
15.	0,494	0,444	Valid
16.	0,619	0,444	Valid
17.	0,313	0,444	Invalid
18.	0,454	0,444	Valid
19.	0,498	0,444	Valid
20.	0,358	0,444	Invalid

Based on the data from the calculation, it is known that there are 9 valid items and 11 invalid items.

## 2. Reliability Test Result Data

In general, a valid instrument will be reliable. To measure the consistency of the instrument, a reliability test was carried out using the Cronbach Alpha method. The criteria for an instrument is said to be reliable if the reliability coefficient value is  $> 0.6$ . The results of the instrument reliability test can be seen in the following table:

**Table 3** Reliability Test Results of Questionnaire Instrument Variable X (Student Perceptions of Google Classroom)

Reliability Statistics		
Cronbach's Alpha	Cronbach's Alpha Based on Standardized Items	N of Items
.891	.893	11

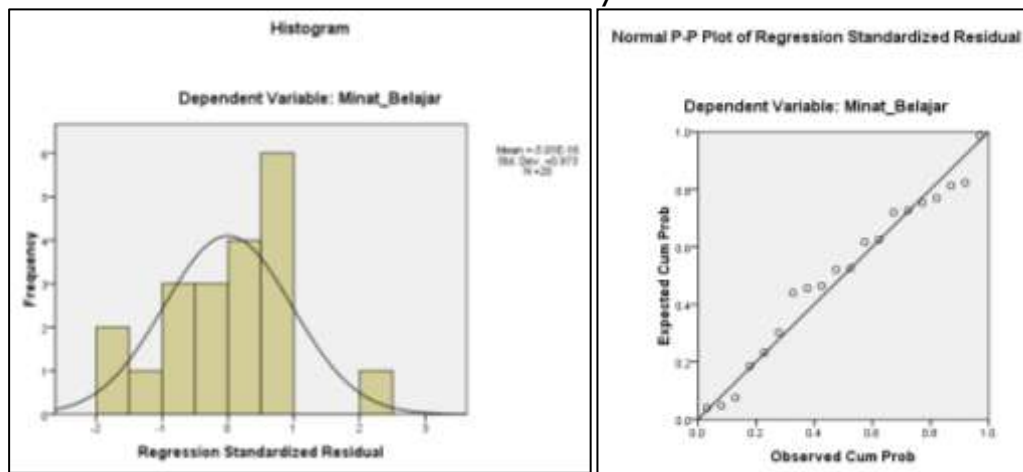
Based on the calculated data, it is known that the Cronbach Alpha reliability coefficient of the student perception instrument about Google classroom is 0.891.

**Table 4** Reliability Test Results of Questionnaire Instrument Variable Y (Religious Learning Interest)

Reliability Statistics		
Cronbach's Alpha	Cronbach's Alpha Based on Standardized Items	N of Items
.787	.821	9

Based on the calculation data, it is known that the Cronbach Alpha reliability coefficient of the Religious learning interest instrument is 0.787.

**Table 5** Normality Test



**Tests of Normality**

	Kolmogorov-Smirnov <sup>a</sup>			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
Google_Classroom	.102	20	.200*	.980	20	.929
Minat_Belajar	.164	20	.163	.961	20	.571

a. Lilliefors Significance Correction

\*. This is a lower bound of the true significance.

### 3. Instrument Validity Discussion

Based on the validity test results shown in Table 1 for the instrument of student perceptions of the use of Google Classroom (variable X), it is known that of the 20 items tested, 11 items were declared valid, while 9 other items were invalid. Invalid items will be discarded (dropped) to ensure the instrument used only includes items that can measure exactly the intended construct. This decrease in the number of invalid items indicates that the instrument used is increasingly relevant and focused in measuring student perceptions.

In Table 2, for the instrument measuring interest in religious learning (variable Y), of the 20 items tested, 9 items proved to be valid, while 11 other items were invalid. Of the 11 invalid items, 9 items will be discarded and 2 items, namely item numbers 9 and 14, will be revised because the calculated r value is close to the r table. This action is taken to ensure that the instrument can measure religious learning interest more precisely, especially on items that show a tendency of marginal validity.

### 4. Instrument Reliability Discussion

Reliability analysis was conducted using Cronbach's Alpha coefficient to evaluate the internal consistency of both instruments. Based on the results presented in Table 3, the student perception instrument towards the use of Google Classroom (variable X) obtained a Cronbach's Alpha value of 0.891. This value is higher than the recommended minimum limit of 0.6, which indicates that this instrument has an excellent level of internal consistency and can be relied upon to measure student perceptions.

Similarly, in Table 4, the instrument to measure religious learning interest (variable Y) obtained a Cronbach's Alpha value of 0.787, which is also greater than 0.6, indicating that this instrument has adequate reliability and can be used to measure students' religious learning interest consistently.

### Discussion

The results of the validity and reliability analyses show that the instruments used in this study have met the validity and reliability criteria for both variables measured. The high level of validity, with some items removed or revised, ensures that the instrument is able to accurately measure the intended constructs, namely students' perceptions of Google Classroom and their interest in learning religion. In addition, high reliability scores reflect good consistency of measurement, so the instrument can be trusted in assessing the factors that are the focus of the study.

Instruments that have gone through validity and reliability tests are then tested for normality, so that the type of analysis that can be used can then be determined. Based on the histogram and table of normality test results, it is known that the data is normally distributed. This is indicated by the shape of the curve which resembles a bell. Thus, in further research, researchers can use parametric tests to analyse the data.

Therefore, this tested instrument is not only relevant to be used in similar research in the future, but also has the potential to become a reference in the development of more comprehensive measurements, especially related to students' perceptions and their interest in learning religion in the context of Islamic religious education.

## CONCLUSION

Based on the analysis of the questionnaire instruments of students' perceptions of Google Classroom and interest in learning religion, it is concluded that both instruments have met the criteria of adequate validity and reliability. The student perception instrument consists of 11 valid items with a high level of reliability, indicated by the Cronbach Alpha coefficient of 0.891. Meanwhile, the religious learning interest instrument has 9 valid items with a reliability of 0.787. For invalid items, the steps taken were to delete the item and revise some items that were close to the validity threshold. Therefore, these two instruments are considered valid and reliable to be used in further research, especially in measuring students' perceptions and interest in learning religion. This finding confirms the importance of validity and reliability tests in the process of developing educational measurement instruments, to ensure accurate and consistent evaluation results.

Based on the results of this study, the researcher recommends that the development of educational evaluation instruments (especially in technology-based learning such as Google Classroom) can continue to be carried out with a more comprehensive approach. Further researchers can also consider external validation with a wider and more diverse sample to increase the generalizability of the findings. In addition, it is necessary to explore other variables that are thought to contribute to the effectiveness of technology-based learning in Islamic religious education, such as teacher involvement, learning environment factors, and students' intrinsic motivation. By continuing to refine measurement instruments and making adjustments to technological developments and student needs, learning evaluation will be more accurate and meaningful in supporting the improvement of the quality of religious education in the digital era.

## REFERENCE

- Ahmed, I., & Ishtiaq, S. (2021). Reliability and Validity: Importance in medical research. *Methods*, 12(1), 2401–2406.
- Alfiyah, N., Zuhriyah, I. A., & Chalishah, N. D. (2025). Validity and Reliability Analysis: Measuring the Accuracy and Consistency of Instruments for the Evaluation of Islamic Religious Education Tests. *JIP-Jurnal Ilmiah Ilmu Pendidikan*, 8(3), 2672–2679.
- Anggraini, F. D. P., Aprianti, A., Setyawati, V. A. V., & Hartanto, A. A. (2022). Pembelajaran statistika menggunakan software SPSS untuk uji validitas dan reliabilitas. *Jurnal Basicedu*, 6(4), 6491–6504.

- Arikunto, P. D. S. (2013). *Dasar-Dasar Evaluasi Pendidikan* (R. Damayanti (ed.)).
- Defrijon. (2025). Reliabilitas alat ukur, jenis-jenis dan faktor- faktor yang mempengaruhinya. *Jurnal Studi Multidisipliner*, 9(1), 77–86.
- Duwi Priyatno. (2010). *Teknik Mudah Dan Cepat Melakukan Analisis Data Penelitian*. Gava Media.
- Fauzan, A. (2025). Evaluasi Efektivitas Google Classroom Sebagai Learning Management System dalam Meningkatkan Pemahaman Mahasiswa di Pendidikan Jarak Jauh. ... : *Jurnal Masalah*, 6(1), 50–67. <http://jurnal.iailm.ac.id/index.php/thoriqotuna/article/view/690%0Ahttp://jurnal.iailm.ac.id/index.php/thoriqotuna/article/download/690/370>
- Febrianti, I., Tuffahati, J., Rifai, A., Affandi, R. H., Pradita, S., Akmalia, R., & Siahaan, A. (2023). Pengaruh Penggunaan Teknologi Informasi Dalam Manajemen Perencanaan Pendidikan Untuk Meningkatkan Efisiensi Pendidikan. *Academy of Education Journal*, 14(2), 506–522.
- Forester, B. J., Idris, A., Khater, A., Afgani, M. W., Isnaini, M., Islam, U., Raden, N., & Palembang, F. (2024). Penelitian Kuantitatif : Uji Reliabilitas Quantitative Research : Data Reliability Test. *Edu Society: Jurnal Pendidikan, Ilmu Sosial, Dan Pengabdian Kepada Masyarakat*, 4(3), 1812–1820.
- Lenani, I. (2021). Teknik Pengambilan Sampel Purposive Dan Snowball Sampling. *HISTORIS: Jurnal Kajian, Penelitian & Pengembangan Pendidikan Sejarah*, 6(1), 33–39. <http://journal.ummat.ac.id/index.php/historis>
- Mabruri Pudyas Salim. (n.d.). Reliabilitas Adalah Keandalan, Simak Ciri-Ciri, Jenis, dan Cara Menguji yang Tepat. *Liputan 6*, <https://www.liputan6.com/hot/read/5366998/reliabil>.
- Masuwai, A., Zulkifli, H., & Hamzah, M. I. (2024). Evaluation of content validity and face validity of secondary school Islamic education teacher self-assessment instrument. *Cogent Education*, 11(1). <https://doi.org/10.1080/2331186X.2024.2308410>
- Mulyana, A., Vidiati, C., Danarahmanto, P. A., Agussalim, A., Apriani, W., Fiansi, F., Fitra, F., Aryawati, N. P. A., Ridha, N. A. N., & Milasari, L. A. (2024). *Metode penelitian kualitatif*. Penerbit Widina.
- Nkwake, A. M. (2023). Credibility, validity, and assumptions in program evaluation methodology. In *Credibility, Validity, and Assumptions in Program Evaluation Methodology*. <https://doi.org/10.1007/978-3-031-45614-5>
- Nurqozin, M., & Putra, D. (2023). Pembelajaran Berbasis Media Digital dalam Meningkatkan Kualitas Pembelajaran Pendidikan Agama Islam pada Sekolah Menengah Kejuruan Islam Terpadu Tebuireng III Indragiri Hilir Riau. *Didaktika: Jurnal Kependidikan*, 12(4), 637–646. <https://jurnaldidaktika.org>
- Osterlind, S. J. (2002). *Constructing Test Items: Multiple-Choice, Constructed-Response, Performance, and Other Formats* (Second). Kluwer Academic Publishers.
- Purwanto. (2021). *Evaluasi Hasil Belajar* (VIII). Pustaka Pelajar.
- Rosidin, U., Widyastuti, W., Kadaritna, N., & Handoko. (2024). The Implementation of Dynamic Society Oriented 6c Skills Assessment Instruments in Case Method and Team Based Project Learning. *JISAE: Journal of Indonesian Student Assessment and Evaluation*, 10(1), 22–31. <https://doi.org/10.21009/jisae.v10i1.43879>
- Salsabila, Y., Fatah, A., & Jaenudin, J. (2023). Hubungan antara literasi numerasi terhadap kemampuan berpikir kritis dan kreatif siswa SMP di kecamatan Curug. *EQUALS: Jurnal Ilmiah Pendidikan Matematika*, 6(1), 42–54.
- Sholihin, S., Supardi, S., & Lubna, L. (2024). Analisis Kurikulum K13 Untuk Mengetahui Evaluasi Hasil Akhir Belajar Siswa di MTs NW Selaparang Lombok Barat. *Jurnal Ilmiah Profesi Pendidikan*, 9(3), 1704–1709.
- Subhaktiyasa, P. G. (2024). Evaluasi Validitas dan Reliabilitas Instrumen Penelitian Kuantitatif : Sebuah Studi Pustaka. *Journal of Education Research*, 5(4), 5599–5609.
- Sugiyono. (2017). *Metode Penelitian Pendidikan : Pendekatan Kuantitatif, Kualitatif, dan R&D*. Alfabeta.
- Sürücü, L., & Maslakci, A. (2020). Validity and reliability in quantitative research. *Business &*

- Management Studies: An International Journal*, 8(3), 2694–2726.
- Tentama, F., & . S. (2018). Pengujian Validitas dan Reliabilitas Konstruk pada Organizational Citizenship Behavior. *Humanitas*, 15(1), 62.  
<https://doi.org/10.26555/humanitas.v15i1.5282>
- Utama, R. W. (2024). Hubungan antara Asesmen Kompetensi Minimum dan Kemampuan Berpikir Tingkat Tinggi pada Pembelajaran Matematika Siswa Kelas V di Sekolah Dasar. *Review Pendidikan Dasar: Jurnal Kajian Pendidikan Dan Hasil Penelitian*, 10(03), 187–196.