

## Comprehensive Assessment: Balancing Test and Non-Test Methods in Islamic Education Evaluation

**Asep Abdurahman**

SD Plus Intan Al-Sali, Bandung, Jawa Barat, 40614

Email: [abdurahmanasep16@gmail.com](mailto:abdurahmanasep16@gmail.com)

**Mohamad Erihadiana**

Sunan Gunung Djati State Islamic University Bandung

Email: [erihadiana@uinsgd.ac.id](mailto:erihadiana@uinsgd.ac.id)

**Abstract:** Test and non-test assessment forms are crucial topics in education, particularly regarding the fair and accurate measurement of students' abilities. This study aims to examine various forms of test and non-test assessments in Islamic religious education evaluation. The method used is a literature review, with data sourced from relevant books and scientific articles. Test assessments include questions or tasks answered by students, divided into objective and subjective (essay) tests. Conversely, non-test techniques are designed to evaluate students' personalities holistically, covering attitudes, behaviors, social traits, speech, and life history. Non-test forms include observation, interviews, attitude scales, questionnaires, and personal reports. This study emphasizes the importance of combining both methods to achieve a comprehensive evaluation, providing deeper insights into students' academic abilities and character development. This combination supports educators in making better decisions to foster students' holistic development.

**Keywords:** Assessment; Islamic Religious Education; Non-Tests, Tests

**Abstrak:** Bentuk penilaian tes dan non-tes menjadi topik penting dalam dunia pendidikan, khususnya terkait pengukuran kemampuan siswa secara adil dan akurat. Penelitian ini bertujuan mengkaji berbagai bentuk penilaian tes dan non-tes dalam evaluasi pendidikan agama Islam. Metode yang digunakan adalah studi kepustakaan dengan sumber data dari buku dan artikel ilmiah. Penilaian tes mencakup pertanyaan atau tugas yang dijawab siswa, terbagi menjadi tes objektif dan subjektif (essay). Sebaliknya, teknik non-tes bertujuan menilai kepribadian siswa secara menyeluruh, meliputi sikap, tingkah laku, sifat sosial, ucapan, dan riwayat hidup. Bentuk non-tes meliputi observasi, wawancara, skala sikap, angket, dan laporan pribadi. Penelitian ini menegaskan pentingnya menggabungkan kedua metode penilaian untuk menghasilkan evaluasi yang komprehensif, memberikan pemahaman lebih dalam tentang kemampuan akademik dan perkembangan karakter siswa. Kombinasi ini membantu pendidik membuat keputusan yang lebih baik dalam mendukung perkembangan siswa secara holistik.

**Kata Kunci:** Non-Tes; Pendidikan Agama Islam; Penilaian; Tes

**DOI:** <http://dx.doi.org/10.15575/jipai.v4i2.40752>

Received: 10, 2024; Accepted: 11, 2024; Published: 12, 2024

**Copyright:** © 2024. The Authors.

**Licensee:** JIPAI: Jurnal Inovasi Pendidikan Agama Islam is licensed under the Creative Commons Attribution License.

## INTRODUCTION

In the realm of education, it is widely recognized that assessment is an integral part of any educational program, regardless of its type or format, conducted at specific intervals during the educational journey (Abdillah et al., 2023; Nurjanah, 2017). Assessment and evaluation in education play a crucial role in measuring students' understanding and abilities (Akram et al., 2021).

Evaluation can be defined as the process of assessing learning outcomes or measuring students' comprehension of material concepts studied during the teaching and learning process (Hamilton et al., 2021). Evaluation must be conducted systematically and continuously to accurately reflect the abilities of the students being assessed. Moreover, evaluating is an obligation for teachers to determine the extent of competencies mastered by students after the learning process. Additionally, assessment activities can serve as benchmarks for teachers' success in conducting the educational process (Qomariyah et al., 2023).

Broadly speaking, evaluation methods in education can be categorized into two forms: test and non-test assessments (Febyronita & Giyanto, 2016). Currently, the forms of test and non-test assessments have sparked significant debate in the education sector. Issues surrounding assessment methods, particularly test and non-test formats, include the complexity of measuring and evaluating various aspects of student abilities fairly and accurately.

A test-dominated assessment system often prioritizes measurable academic knowledge but frequently fails to capture students' overall abilities in real-life contexts. On the other hand, non-test assessments tend to emphasize the application of knowledge and skills in real situations but may face challenges regarding consistency, objectivity, and clear evaluation standards. These differences in assessment approaches create a dilemma regarding effectiveness, validity, and fairness in comprehensively evaluating students' abilities and supporting their holistic development.

Therefore, to achieve effective evaluation goals, it is essential to master various methods for conducting relevant assessments. The results of evaluations can assist educators in making accurate judgments regarding subsequent actions. Moreover, evaluation outcomes can serve as motivation for students to improve their performance in future endeavors (Ariefky & Inayati, 2023).

In the realm of education, assessment is an integral part of every educational program and is conducted to ensure the success of the learning process (Nurjanah, 2017). Assessment not only measures students' understanding of the material but also serves as a tool to evaluate the

effectiveness of teachers' instruction. These assessment methods are divided into two main categories: test-based and non-test-based assessments (Febyronita & Giyanto, 2016).

Previous research has highlighted the importance of test-based assessments in measuring students' cognitive aspects that can be evaluated through standardized methods, such as the findings of (Irawati et al., 2018), which emphasize the advantages of tests in providing consistent and comparable data. However, this approach often overlooks students' affective and psychomotor aspects, which are more challenging to measure. On the other hand, non-test assessments focus more on the real-world application of students' abilities, as described by (Septikasari et al., 2023). However, this research also acknowledges challenges in maintaining consistency and objectivity in non-test assessments.

Despite extensive research, a gap remains in integrating these two approaches to create a comprehensive, fair, and relevant assessment system that meets students' needs. The current educational landscape demonstrates that the dominance of one assessment type can impact the validity of evaluation results overall, from both the teacher's and students' perspectives.

This study contributes novel insights by deeply exploring the integration of test-based and non-test-based assessments in the context of Islamic Religious Education evaluation, an approach rarely addressed in previous research. This method emphasizes the pursuit of holistic evaluation by combining the unique strengths of both methods: the objectivity and standards of test-based assessments, along with the comprehensive evaluation of affective and contextual aspects provided by non-test-based assessments. The research also offers practical insights for educators, particularly in designing evaluation methods that not only measure cognitive competencies but also encompass character and religious values, thereby enhancing the overall effectiveness of learning evaluations.

This research aims to delve deeper into the differences and advantages of test-based and non-test-based assessments and explore how these approaches can be integrated to achieve more holistic evaluation goals. By understanding the unique characteristics of each assessment type, this study is expected to provide practical insights for educators to improve the effectiveness of Islamic Religious Education evaluation.

## **RESEARCH METHODS**

This paper discusses the forms of test-based and non-test-based assessments in the evaluation of Islamic Religious Education. The purpose of this paper is to identify various forms of test-based and non-test-based assessments

used in evaluating Islamic Religious Education. It employs a literature study method by gathering data from books and scholarly articles relevant to the topic of discussion. The data is then processed through descriptive analysis, which is systematically and objectively presented by the author (Zed, 2008). The data was obtained through documentation by extracting information from various journals, books, and other related resources. Subsequently, the author examined several journals, articles, papers, books, and other sources relevant to this study. This paper employs a literature-based approach, with literature serving as the theoretical foundation to discover data related to the essence of Islamic educational institutions. The collected data is systematically analyzed and objectively concluded to provide insights into the discussed topic.

## **RESEARCH RESULTS AND DISCUSSION**

### **Research Results**

#### **Definition of Test and Non-Test Assessment Forms**

Literally, the word "test" comes from Old French: *\*testum\**, meaning "a plate for separating precious metals" (meaning that by using this plate-like tool, one could obtain high-value types of precious metals). In English, it is written as "test," which in Indonesian is translated as "tes," "ujian," or "percobaan." In Arabic, it is written as امتحان (Rizqiyah, 2018).

There are several terms that require explanation in relation to the above description, namely: *\*test\** is a tool or procedure used for measurement and assessment; *\*testing\** refers to the process or event of conducting the measurement and assessment; *\*tester\** refers to the person who administers the test, the test maker, or the experimenter, i.e., the person conducting the experiment; while *\*testee\** (singular) and *\*testees\** (plural) refer to the person being tested (test participants, exam takers), or the person who is undergoing the task (the subject of the experiment).

Furthermore, it can be explained that a *\*test technique\** is a measurement method used to assess someone's ability, knowledge, or skills by providing certain questions or tasks. The purpose of the test technique is to evaluate someone's understanding and performance in a specific field or subject.

In addition, a test technique can also be understood as a method that involves administering a test in the form of questions that need to be answered, questions that need to be responded to, or tasks that need to be carried out by the person being tested. A test is a systematic and objective tool or procedure to obtain the desired data or information about an individual, in a way that can be considered accurate and efficient. In the context of learning outcomes, the test measures the learner's ability to master the material, covering aspects of knowledge and skills (Syahrudin, 2021).

Next is the form of non-test assessment. Non-test evaluation techniques refer to conducting assessments without using tests. This evaluation technique is generally used to assess a child's personality as a whole, including attitude, behavior, characteristics, social behavior, speech, life history, and others. It relates to learning activities in education, both individually and in groups. Non-test evaluation instruments are assessments or evaluations of student learning outcomes conducted using systematic observations, interviews, questionnaires, documentary analysis, and others (Syahrudin, 2021).

According to (Sudijono, 2009) "non-test techniques generally play an important role in evaluating students' learning outcomes in terms of the affective domain and psychomotor domain." Meanwhile, according to (Uno, 2013) non-test instruments are generally used in several assessment techniques, including: (a) performance assessment, (b) product assessment, (c) project assessment, (d) portfolio, and (e) attitude scale. In agreement with Widiyoko in Maulia (2013), non-test evaluation techniques are usually used to measure learning outcomes related to soft skills, especially those related to what students can create or perform (Magdalena et al., 2021).

### **Various Forms of Test Assessment**

**The form of the test can be broadly divided into two types of tests, namely objective and subjective tests (essay).**

#### **1. Objective Tests**

An objective test is a type of learning test consisting of questions that can be answered by students by selecting one of several possible answers provided or by writing their answer by choosing specific codes representing the available alternatives. The answers to objective tests are "certain," meaning there is only one correct answer. If a student does not answer "correctly," the response is considered incorrect. Since the answer is definite, the student's correct response to a particular question will be marked as correct by the evaluator. Because the results of the student's work, when examined by anyone, will produce the same score, this is called an objective test (Ghufron & Sutarna, n.d.). Objective tests can be classified into:

##### **a. Objective test of the form of true-false test**

A true-false type test is a test whose question items consist of statements accompanied by alternative answers, namely true and false answers and questions. The test taker marked each answer, choosing "B" if the answer or statement was considered correct in his opinion and selecting "S" if the answer or statement was considered false in his opinion (Sanusi & Aziez, 2021).

##### **b. Objective test of matchmaking form (*Matching Test*)**

Matching, an objective test in the form of matching or often known as a matchmaking test is arranged in two groups or lists of which each contains words, terms or sentences placed next to each other (Magdalena et al., 2021). In the form of a matching test, students are required to match, match, adjust, or connect between the two statements provided.

c. Objective test of complementary form (*Completion Test*)

Test is an objective test commonly known as a complementary or perfecting test. Item completion consists of a statement or sentence with one or several words at the points provided. Item completion can also be in the form of a direct question. This test is in the form of an incomplete statement, where students are asked to complete the statement with a sentence or number (Purwanto, 2000).

d. Objective test of the form of short fill (*Fill-In Test*)

Fill-in (filling in the dots in a blank sentence), the objective test of the fill-in form is very similar to the objective test of the completion form. The difference is that, in the objective test, the fill-in form of the material tested is a single story, while in the objective test, the completion form does not have to be so. In other words, in this objective test form of completion, the test question items can be made different from one another (Magdalena et al., 2021).

e. Multiple-choice form objective test (*Multiple Choice Test*)

Multiple choice is an objective test consisting of questions or statements that have not been completed, and to complete them, one (or more) must be selected from several possible answers that have been provided in each of the questions concerned. Or in other words, a multiple-choice test is a form of objective test consisting of a question or statement (stem) and followed by a number of alternative answers (options), the testee's task is to choose the most appropriate alternative answer. The possible answer can be a word, phrase, place name, character name, symbol or sentence that is certain (Rosyidi, 2020).

## 2. Tes Subjektif

A subjective test is a type of test in which, in addition to being influenced by the answers/responses of the test takers, the scoring is also influenced by the subjectivity of the examiner. A test with the same questions and answers, when evaluated by different examiners, will produce different scores. Subjective tests are generally in the form of essays (Widoyoko, 2014).

A subjective test (essay test) is a form of question that requires students to answer in the form of an essay using their own language. In an essay test, students are expected to think critically and use their knowledge to address

the question at hand. This type of test allows students the freedom to structure and express their own answers, enabling them to demonstrate their ability to apply knowledge to analyze, connect, and evaluate the issues presented (Ghufron & Sutama, 2011).

There are two types of essay tests: 1) limited or structured essay tests, and 2) free-response essay tests. Limited or structured essay tests, also called objective essay tests, have answers that are constrained (already directed) both in terms of content and responses. Scoring for limited essay tests tends to be more consistent and objective (Rosyidi, 2020).

a. Limited/Structured Essay Test

To answer questions in the form of a limited essay, students must address specific aspects as the boundaries for their responses. Although the wording of the students' answers may vary, there must still be essential points within the structure of their answers that align with the boundaries outlined and required in the question.

b. Free Response Essay

A free response essay is a form of essay test that requires a detailed (long) answer. This type of essay allows the test taker to express their answer in writing or an essay form. The correctness of the test taker's answer can only be assessed by an experienced teacher. In this form, students are free to answer the question using their own methods and structure. Students can express their opinions according to their abilities. Therefore, each student may have different ways and structures for their responses. However, the teacher must still have a reference or guideline for evaluating the students' answers later (Putri et al., 2022).

### **Various Forms of Non-Test Assessment**

There are several kinds of tools used to evaluate learning with non-test techniques. However, here the author only explains in general several types of non-test assessments in the learning process.

1. Observation

Observation is a method or way of analyzing and systematically recording behavior by seeing or observing individuals or groups directly (Purwanto, 2019). This is in accordance with the statement of Drs. Zaenal Arifin, M.Pd in his book entitled *Learning Evaluation*, that: "Observation is a process of observing and recording systematically, logically, objectively and rationally about various phenomena, both in real situations and in artificial situations to achieve certain goals"(Arifin, 2009).

Observation is a process of collecting data by directly observing and systematically recording the observed subject. According to Sutrisno Hadi

(1981), as cited by Zaenal Arifin in his book *Evaluasi Pembelajaran*, there are three types of observation: participant observation, systematic observation, and experimental observation.

Participant observation occurs when the observer becomes involved in the situation being observed (Nurkencana, 1986). Systematic observation is an observation conducted by the Islamic Religious Education teacher after thorough planning. The main features of systematic observation include a framework that categorizes factors beforehand and the specific characteristics of each factor within those categories. Experimental observation refers to a situation where the Islamic Religious Education teacher is not involved in the students' activities, but the students' behavior or actions are controlled by the teacher to meet the desired objectives.

The instruments used by the Islamic Religious Education teacher for observation can include checklists, anecdotal records, or rating scales.

a. Check List

A check list is an observation guide containing a list of all aspects to be observed, so the Islamic Religious Education teacher only needs to mark a check (√) next to the aspects being observed, i.e., the students. There are two types of checklists: individual and group. An individual checklist is used to record the presence or absence of an evaluated aspect in a single person. In contrast, a group checklist is used to record the activities of individuals within a group.

b. *Anekdotol Record*

An anecdotal record is an observation tool used to document extraordinary events that are considered important. It is recommended to take the anecdotal record during the event as it occurs. This record can be used when a student engages in either positive or negative behaviors. The process of taking an anecdotal record includes: preparing the format for the anecdotal record, determining the position of observation, observing, and recording the behavior of individual students.

c. Rating Scale

A rating scale is an observation method that contains a list of all observed aspects, which are then broken down into a scale. The rating scale consists of two parts: (a) statements about the quality of something's existence, and (b) guidelines for assessing these statements. There are three forms of rating scales: categorical, numerical, and graphic.

## 2. Interview

Interview is a method used to get answers from respondents by way of one-sided questions and answers (Arikunto, 2021). According to Johnson and Johnson (2001) as quoted by Sukardi in his book *Evaluation of Principles*



and Operational Education, it states that: "An interview is a personal interaction between the interviewer (teacher) and the interviewee (the student) where verbal questions are asked to them". There are three steps that must be considered in conducting an interview, namely:

a. Preparing the Interview Situation

This stage aims to create a comfortable atmosphere for the interview, fostering a sense of familiarity so that students do not feel intimidated and are encouraged to express their opinions freely and honestly.

b. Questioning

Once the situation is conducive, questions should be asked systematically, following the interview objectives based on the interview guide.

c. Recording the Interview Results

It is advisable to take notes during the interview as it is happening, so the results can be recorded promptly and accurately.

### 3. Attitude Scale

According to Muhajir, attitude is an affectionate tendency to like or dislike a social object. In measuring students' attitudes, especially in the learning process of Islamic Religious Education, Islamic Religious Education teachers should pay attention to three components of attitudes, namely:

a. Cognition, which refers to the students' knowledge about the object.

b. Affect, which refers to the students' feelings towards the object.

c. Conation, which refers to the students' tendency to behave towards the object.

The Islamic Religious Education teacher can use a model that is often used to measure students' attitudes in the learning process, namely the model developed by Likert, known as the Likert Scale.

### 4. Questionnaire

A questionnaire is a list of questions that must be filled out by the person being measured (the respondent). From the perspective of how to answer, questionnaires are divided into two types: closed questionnaires and open questionnaires. A closed questionnaire is one that provides answer choices, allowing the respondent (student) to simply mark the selected answer. On the other hand, an open questionnaire is one that does not provide answer choices, giving the respondent the freedom to provide their own answers. Open questionnaires usually ask for the respondent's opinions (Arikunto, 1999).

### 5. Personal Report

This technique is used when teachers want to assess students' opinions or comments on problems, circumstances or things that are the object of attitudes,

such as about "clashes between schools". From the reviews given by students, we can see the tendency of attitudes possessed by students.

## **Discussion**

Assessment is a systematic process to measure students' learning achievements across various aspects, including cognitive, affective, and psychomotor domains. In the context of Islamic Religious Education (PAI), test assessments are commonly used to evaluate cognitive aspects, such as understanding the concepts of faith, worship, and morality. Tests in the form of multiple-choice questions, essays, or oral exams often serve as primary tools to assess how well students comprehend the subject matter conceptually. On the other hand, non-test assessments, such as observations, portfolios, and reflective journals, provide opportunities to evaluate the affective and psychomotor aspects of students, such as the internalization of religious values and the implementation of worship in daily life (Diniati, 2024).

Tests have advantages in terms of objectivity, efficiency, and ease of data analysis. Test formats, particularly written tests, allow teachers to assess students' learning achievements in a short period and generate quantitative data that can be compared in a standardized manner. However, this approach is often limited to measuring cognitive abilities, and thus may not reflect students' understanding holistically. In the context of Islamic Religious Education (PAI), for instance, written tests cannot directly measure how students apply religious values in their daily lives, which can create a gap between test results and students' actual behavior (Arikunto, 2021).

Non-test assessments, such as observations and interviews, allow teachers to evaluate the affective and psychomotor dimensions more deeply. In PAI, these methods are crucial for observing student behaviors, such as honesty, discipline, and social piety, which cannot be measured with standard tests. Although they have these advantages, non-test assessments often face challenges in maintaining consistency and objectivity. The subjectivity of the teacher in assigning grades can affect the validity of the evaluation results. Therefore, clear guidelines and standardized assessment instruments are needed to reduce bias in non-test assessments (Nurdiana, 2024).

To achieve a holistic evaluation, it is important for educators to integrate both test and non-test assessments. In PAI, this integration can be achieved by combining written tests to evaluate students' understanding of religious concepts with non-test methods such as portfolios or practical worship projects. For example, after teaching the procedures for prayer, the teacher may give a written test to measure students' theoretical understanding and complement it with observation of prayer practice to assess psychomotor skills. With this approach,

the assessment covers not only academic achievement but also the internalization and application of religious values in students' lives (Nurjanah, 2017).

The implementation of this combined assessment system provides an opportunity for teachers to gain a more comprehensive picture of students' development. However, the main challenge lies in the teachers' readiness to design and implement relevant assessments and their ability to manage time and resources. Furthermore, the use of technology in education can be a solution to improve efficiency, such as digital assessment applications that can automatically record and analyze data. With adequate training, PAI teachers can utilize various assessment methods to support more effective and inclusive learning (Wusthoa & Fadilah, 2024).

## **CONCLUSION AND IMPLICATION**

### **Conclusion**

The test technique can be defined as a method used by administering tests in the form of questions that must be answered, questions that need to be responded to, or tasks that must be completed by the person being tested. A test is a systematic and objective tool or procedure used to obtain the desired data or information about someone, in a manner that is both accurate and efficient. In the case of learning outcome tests, the focus is on measuring the students' ability to master the lessons taught, which includes aspects of knowledge and skills. Broadly speaking, test assessments can be divided into two types: objective tests and subjective tests (essays).

On the other hand, non-test evaluation techniques refer to conducting assessments without using tests. This assessment method is generally used to evaluate the child's personality comprehensively, including aspects such as attitude, behavior, traits, social attitudes, speech, life history, and others. These evaluations are related to learning activities in education, whether on an individual or group basis. Forms of non-test assessments include observation, interviews, attitude scales, questionnaires, and personal reports.

### **Implication**

The results of this study have significant implications both theoretically and practically. Theoretically, this research reinforces the understanding of the diversity of test and non-test assessment techniques in educational evaluation, particularly in Islamic education. The findings can serve as a reference for the development of a more comprehensive evaluation concept that considers the cognitive, affective, and psychomotor aspects of students. This study also contributes to the academic literature on the need for a combination of test and non-test methods to ensure validity and fairness in evaluations.

Practically, the implications of this research are relevant for educators, especially Islamic education teachers, in designing a more comprehensive evaluation system. Teachers can use test methods to measure students' cognitive understanding and non-test techniques to assess their personality, attitudes, and social skills. The combination of these two methods allows teachers to provide assessments that are not only accurate but also support the holistic development of students.

## REFERENCE

- Abdillah, F., Azmi, K., Hafizah, C. V., Anisha, D., Bintang, N. D., & Mulyani, S. (2023). Strategi Pelaksanaan Evaluasi Program Pendidikan Terhadap Kualitas Belajar Siswa di Sekolah. *Jurnal Bintang Pendidikan Indonesia*, 1(2), 13–23. <https://doi.org/https://doi.org/10.55606/jubpi.v1i2.1190>
- Akram, H., Yingxiu, Y., Al-Adwan, A. S., & Alkhalifah, A. (2021). Technology Integration in Higher Education During COVID-19: An Assessment of Online Teaching Competencies Through Technological Pedagogical Content Knowledge Model. *Frontiers in Psychology*, 12, 736522. <https://doi.org/https://doi.org/10.3389/fpsyg.2021.736522>
- Ariefky, M. M., & Inayati, N. L. (2023). Peran Guru PAI dalam Pelaksanaan Evaluasi Pembelajaran Sumatif Siswa di SMK Negeri 6 Sukoharjo. *EDUKASIA: Jurnal Pendidikan Dan Pembelajaran*, 4(2), 2343–2350. <https://doi.org/https://doi.org/10.62775/edukasia.v4i2.586>
- Arifin, Z. (2009). *Evaluasi Pembelajaran*. Dirjen Pendidikan Islam, Depag.
- Arikunto, S. (1999). *Dasar-dasar Evaluasi Pendidikan (Edisi Revisi)/Suharsmi Arikunto*.
- Arikunto, S. (2021). *Dasar-dasar evaluasi pendidikan edisi 3*. Bumi aksara.
- Diniati, S. W. (2024). Challenges of Using Information Technology in the Needs Analysis Design Process for Islamic Religious Education Learning. *Jurnal Inovasi Pendidikan Agama Islam (JIPAI)*, 3(2), 128–139. <https://doi.org/10.15575/jipai.v3i2.31096>
- Febyronita, D., & Giyanto. (2016). Survei Tingkat Kemampuan Siswa dalam Mengerjakan Tes Berbentuk Jawaban Singkat (Short Answer Test) Pada Mata Pelajaran IPS Terpadu (Geografi) Kelas VII di SMP Negeri 1 Mesuji Tahun Pelajaran 2015/2016. *Jurnal Swarnabhumi*, 1(1), 17–21.
- Ghufron, A., & Utama. (n.d.). *Tes, Pengukuran, Asesmen, dan Evaluasi, Peran dan Fungsinya dalam Pembelajaran*.
- Ghufron, A., & Utama. (2011). Tes, Pengukuran, Asesmen, dan Evaluasi, Peran dan Fungsinya dalam Pembelajaran. *Evaluasi Pembelajaran Matematika*, 1–27.
- Hamilton, D., McKechnie, J., Edgerton, E., & Wilson, C. (2021). Immersive virtual reality as a pedagogical tool in education: a systematic literature review of quantitative learning outcomes and experimental design. *Journal of*

- Computers in Education*, 8(1), 1–32.  
<https://doi.org/https://doi.org/10.1007/s40692-020-00169-2>
- Irawati, H., Saifuddin, M. F., & Ma'rifah, D. R. (2018). Pengembangan Instrumen Tes Dan Non Tes Dalam Rangka Menyiapkan Penilaian Autentik Pada Kurikulum 2013 Di Smp/Mts Muhammadiyah Se-Kabupaten Bantul. *Jurnal Pemberdayaan: Publikasi Hasil Pengabdian Kepada Masyarakat*, 1(2), 503.  
<https://doi.org/10.12928/jp.v1i2.362>
- Magdalena, I., Ismawati, A., & Amelia, S. A. (2021). Penggunaan Evaluasi Non-Tes Dan Kesulitannya Di Sdn Gempol Sari. *PENSA: Jurnal Pendidikan Dan Ilmu Sosial*, 3(2), 187–199.
- Magdalena, I., Syariah, E. N., Mahromiyati, M., & Nurkamilah, S. (2021). Analisis Instrumen Tes sebagai Alat Evaluasi pada Mata Pelajaran SBDP Siswa Kelas II SDN Duri Kosambi 06 Pagi. *Jurnal Pendidikan Dan Ilmu Sosial*, 3(2), 276–287.
- Nurdiana, A. (2024). Implementation of Religious Culture and Its Implications for Students' Religious Development in Junior High School 10 Bandung. *Jurnal Inovasi Pendidikan Agama Islam (JIPAI)*, 3(2), 79–94.  
<https://doi.org/10.15575/jipai.v3i2.31117>
- Nurjanah, E. (2017). Teknik dan Instrumen Asesmen PAUD Ranah Afektif: Teknik Non Tes. *At-Ta'lim: Jurnal Pendidikan*, 3(1), 25–38.
- Nurkencana, W. (1986). Evaluasi Pendidikan. *Usaha Nasional*, 47.
- Purwanto, M. N. (2000). *Prinsip-prinsip dan Teknik Evaluasi Pengajaran*. PT. Remaja Rosdakarya.
- Purwanto, M. N. (2019). *Prinsip-prinsip dan teknik evaluasi pengajaran*.
- Putri, H., Susiani, D., Wandani, N. S., & Putri, F. A. (2022). Instrumen Penilaian Hasil Pembelajaran Kognitif pada Tes Uraian dan Tes Objektif. *Jurnal Papeda: Jurnal Publikasi Pendidikan Dasar*, 4(2), 139–148.  
<https://doi.org/10.36232/jurnalpendidikandasar.v4i2.2649>
- Qomariyah, S., Darmayanti, R., Rosyidah, U., & Ayuwanti, I. (2023). Indicators and Essay Problem Grids on Three-Dimensional Material: Development of Instruments for Measuring High School Students' Mathematical Problem-Solving Ability. *JEMS: Jurnal Edukasi Matematika Dan Sains*, 11(1), 261–274.  
<https://doi.org/https://doi.org/10.25273/jems.v11i1.14708>
- Rizqiyah, L. (2018). Evaluasi Pembelajaran PAI FAI VI D | 1. *Evaluasi Pembelajaran PAI*, 1530304667, 1–14.
- Rosyidi, D. (2020). Teknik dan instrumen asesmen ranah kognitif. *Tasyri': Jurnal Tarbiyah-Syari'ah Islamiyah*, 27(1), 1–13.
- Sanusi, R. N. A., & Aziez, F. (2021). Analisis Butir Soal Tes Objektif dan Subjektif untuk Keterampilan Membaca Pemahaman pada Kelas VII SMP N 3 Kalibagor. *Metafora: Jurnal Pembelajaran Bahasa Dan Sastra*, 8(1), 99.  
<https://doi.org/10.30595/mtf.v8i1.8501>

- Septikasari, R., Inayah, F., Husniyah, N. A., & Rini, R. M. (2023). Teknik Penilaian Tes dan Non Tes. *Jurnal Ilmiah Multidisiplin*, 1(11), 761–764.
- Sudijono, A. (2009). *Pengantar Evaluasi*. PT. Raja Grafindo Persada.
- Syahrudin, H. (2021). *Buku Ajar Evaluasi Pembelajaran*. EUREKA MEDIA AKSARA.
- Uno, H. B. (2013). *Assessment Pembelajaran*. Bumi Aksara.
- Widoyoko, E. P. (2014). Penilaian hasil pembelajaran di sekolah. *Yogyakarta: Pustaka Pelajar*, 1(2), 8.
- Wusthoa, J., & Fadilah, F. (2024). Tantangan Guru Pendidikan Agama Islam dalam Kurikulum Merdeka Belajar di Kelas X SMAN 20 Gowa. *Cendikia Pendidikan*, 6(3), 21–30.  
<https://doi.org/https://doi.org/10.9644/sindoro.v6i3.5403>
- Zed, M. (2008). *Metode Penelitian Kepustakaan*. Yayasan Obor Indonesia.