



## Analysis of the Effectiveness of Instruments and Implementation of Summative Evaluation in Tahfizh Education: Case Study of MTs Al-Fatah Boyolali

Efri Joni<sup>1\*</sup>, Ari Anshori<sup>2</sup>

<sup>1\*</sup>Universitas Muhammadiyah Surakarta, Kota Surakarta, Indonesia

\*Co-email: [o100230064@student.ums.ac.id](mailto:o100230064@student.ums.ac.id) [aa112@ums.ac.id](mailto:aa112@ums.ac.id)

### ARTICLE INFO

#### Article History:

Received : 17-Jul-2025

Revised : 05-Aug-2025

Accepted : 17-Sep-2025

Available online : 30-Sep-2025

#### Keyword:

summative evaluation;  
tahfizh education;  
assessment instruments.

### ABSTRACT

This research analyzes the quality of instruments and the effectiveness of summative evaluation implementation in tahfizh education at MTs Al-Fatah Boyolali. This research uses a qualitative research method with a descriptive approach. In collecting data, researchers used observation, interview, and documentation methods. The research results showed that the oral assessment rubric, which includes memorization fluency (40%), tajweed and makhraj (40%), and time accuracy (10%), is valid for measuring memorization, but does not include understanding the meaning of the verse due to the design of the tahfizh program which focuses on memorization. Time limitations, especially in the randomized test (mubādarah) which covers 70% of the evaluation, affect the accuracy of the assessment. The implementation of summative evaluation is effective in measuring students discipline, supported by graduation motivation, with an average score of 75 (odd semester) and 80 (even semester) for 75 students. The distribution of even semester grades showed 20 students scored A (91-100), 30 students scored B (81-90), 20 students scored C (71-80), and 5 students scored D (<70), with 27 passing and 3 not passing out of 30 students. Implementation constraints include student anxiety, stress-induced thrush, and teacher or student absenteeism. This research recommends the addition of meaning comprehension assessment, better time management, and psychological support to improve the effectiveness of tahfizh evaluation.

This is an open access article under the [CC-BY-SA](https://creativecommons.org/licenses/by-sa/4.0/) license



## INTRODUCTION

Introduction is a part of a scientific article that brings readers or other people to understand the problems to be discussed in a scientific article in a clear, detailed, and orderly manner. In the introduction, the author or researcher can include sufficient citations.

The things that should be included in the introduction of an article are as follows: 1) research context and exposure to the forefront of scientific developments related to the topic under study from the results of reviewing previous research findings published in the latest published journals of international repute; 2) theoretical basis; 3) the results of previous studies which show gaps; 4) insight into problem-solving plans and / or scientific contributions that are "promised"; 5) Novelty of the research and 6) the formulation of research objectives.

Qur'anic memorization education is one of the important pillars in Islamic education, aiming to form a generation that is not only able to memorize the Qur'an but also understand and practice its values (S. Jamil et al., 2024). Summative evaluation in tahfizh education plays a crucial role in measuring students' achievement in memorizing the Qur'an, ensuring the quality of memorization, and supporting holistic educational goals, such as the formation of noble morals and spiritual understanding. However, the success of summative evaluation depends on the quality of the assessment instrument (validity, reliability, relevance) and the effectiveness of its implementation in the field (Earle, 2020).

MTs Al-Fatah Boyolali, as one of the Islamic educational institutions in Central Java, runs a tahfizh program that focuses on memorizing the Qur'an, with summative evaluations conducted twice a year to assess students' abilities. Evaluation instruments in the form of oral assessment rubrics are used to measure fluency, tajweed, and timing accuracy, but there are indications that understanding the meaning of the verse is not assessed, as is common in memorization-oriented tahfizh programs (Mohd Saad et al., 2021; Rahim et al., 2018). In addition, the implementation of summative evaluations faces challenges such as time constraints, student anxiety, and technical constraints such as teacher or student absenteeism, which can affect their effectiveness (Ford & Hewitt, 2020).

This research focuses on analyzing the effectiveness of summative evaluation in tahfizh education by evaluating the quality of the instruments used and their implementation in the field. The promised scholarly contribution is to provide comprehensive empirical insights into the challenges and opportunities of developing summative evaluations that fit the characteristics of Qur'anic memorization and Islamic values. The novelty of this research lies in the use of a qualitative approach to explore the experiences of students and teachers in depth, which is rarely done in previous research that tends to be quantitative or descriptive in nature.

The objectives of this research are: (1) to analyze the quality of summative evaluation instruments in tahfizh education in terms of accuracy, relevance, and suitability to the characteristics of memorizing the Qur'an; and (2) to evaluate the effectiveness of summative evaluation implementation in the field, including

challenges and opportunities faced by students and teachers at MTs Al-Fatah. This research is expected to provide insight into the strengths and weaknesses of summative evaluation, as well as recommendations for improving the quality of tahfizh education.

## LITERATURE REVIEW

### Tahfizh Education and the Memorization Process

Tahfizh education aims to form individuals who are able to memorize the Qur'an accurately and understand the meaning and practice its teachings (Sulastini & Zamili, 2019). The process of memorizing the Qur'an involves strengthening long-term memory (Parveen, 2021) through systematic methods such as talaqqi (hearing and imitating the recitation from the teacher) and muraja'ah (periodic repetition of memorization) (Sabrina et al., 2022; Ulumiyah et al., 2021).

### Summative Evaluation and Instrument Quality

Summative evaluation is an assessment conducted at the end of a learning period to measure student achievement of specific learning objectives (Adom et al., 2020; Kibble, 2017). In tahfizh education, summative evaluation is often through oral exams (S. R. Jamil & Othman, 2023), written exams (Bani et al., 2017), or recitation assessments to measure memorization accuracy and tajweed mastery (Wildan & Prastowo, 2022). However, the quality of summative evaluation instruments is a key factor in determining their effectiveness (Garger et al., 2019). In summative evaluation, evaluation instruments must have validity (measuring what should be measured) and reliability (consistency of measurement results) (Mohajan, 2017; Souza et al., 2017). In the context of tahfizh, a valid instrument should be able to measure not only the number of verses memorized, but also the accuracy of recitation and understanding of meaning.

### Implementation of Evaluation in Islamic Education

The implementation of summative evaluation in Islamic educational institutions, including tahfizh madrasas, faces a number of challenges, such as lack of teacher training in designing evaluation instruments, limited resources, and different perceptions about the purpose of evaluation (Hanapi & Rayhani, 2024). In evaluation, many teachers in madrasahs use conventional evaluation approaches, such as written tests, which are less suitable for the characteristics of tahfizh education that emphasizes oral memorization and mastery of tajweed (Muawanah et al., 2022).

## METHOD

This research used a qualitative approach with a descriptive case study design to analyze the effectiveness of summative evaluation at MTs Al-Fatah Boyolali, Central Java, focusing on the quality of the instrument and its implementation. The research subjects included 5 tahfizh teachers ( $\geq 2$  years of experience) and 10 students (5 grade 8, 5 grade 9), selected through purposive sampling. Data were collected through face-to-face semi-structured interviews with 5 teachers and 10 students, focusing on instrument quality (validity, reliability, relevance) and

implementation challenges, recorded for thematic analysis. Observations were conducted on three tahfizh class sessions during summative exams, observing oral evaluations, teacher-student interactions, and technical constraints.

## FINDINGS

Document analysis shows that the oral assessment rubric at MTs Al-Fatah Boyolali includes three aspects: fluency of memorization (40%), tajweed and makhraj (40%), and time accuracy (10%). Scores are given on a scale of 1-100 and recorded on an official form. This rubric does not include an assessment of understanding the meaning of the verse, as the tahfizh program at this institution is designed to focus only on memorizing the Qur'an. The absence of assessment of understanding the meaning of the verses limits the achievement of holistic tahfizh education goals, such as the formation of spiritual understanding and noble morals. Interviews with students revealed that the time for oral evaluation, especially randomized tests (mubādarah) which account for 70% of the test, is often insufficient, thus affecting the accuracy of the assessment. The tahfizh teacher (WA, Head of Tahfizh Program) stated that the rubric is valid for measuring the quantity and quality of memorization, but the absence of assessing the understanding of verse meaning is a major weakness in the broader context of tahfizh education.

Summative evaluation is held twice a year (December for odd semester, June for even semester) for 2-3 days, depending on the number of memorization targets. The evaluation format consists of a random test (mubādarah) of 70% for students with memorization of 4 juz and above and a full memorization test of 30% for students with memorization below 4 juz, conducted through direct deposit to the examiner. The tahfizh teacher stated that this test format is still very relevant for measuring the students' memorization ability. Observations show that teacher-student interaction during the evaluation supports the learning atmosphere, but students often feel nervous, experience thrush due to stress, feel sad if they have not reached the memorization target, but happy when the target is met and enthusiastic because tahfizh graduation is considered an achievement for parents. The tahfizh teacher (NJ, Memorization Examiner) stated that students with consistent daily deposits tend to perform well, demonstrating the effectiveness of evaluation in measuring discipline. Another teacher (AT, Hafizh Examiner) added that students' anxiety is mainly experienced by those who often do not attend the tahfizh program. Time constraints are the main obstacle, as reported by students and reinforced by the teacher (HF, Tahfizh Examiner), who mentioned that evaluation time is disrupted by the absence of teachers or students due to illness or sudden needs. The head of madrasah (SA) emphasized that tahfizh graduation is "the most beautiful gift that students dream for their parents," demonstrating the motivational value of summative evaluation.

Based on the 2024/2025 semester data, 75 students were tested with an average score of 75 (odd semester) and 80 (even semester). The distribution of even semester grades shows that 20 students received A grades (91-100), 30 students

received B grades (81-90), 20 students received C grades (71-80), and 5 students received D grades (<70). Of the 30 students in the even semester, 27 passed (score  $\geq 75$ ) and 3 did not pass, required to follow remedial guidance for one month before re-evaluation. This remedial program supports grade promotion and graduation.

Summative evaluation instruments at MTs Al-Fatah Boyolali are valid for measuring fluency, tajweed, and discipline, but do not include understanding the meaning of verses, according to the design of the tahfizh program which only focuses on memorization, thus limiting the holistic goals of tahfizh education. The implementation of evaluation is effective in measuring the consistent achievement of students, supported by the motivation for graduation and the relevance of the test format (70% mubādarah, 30% full memorization), but is constrained by limited time, student anxiety and anxiety (especially those who rarely attend), and teacher or student absence due to illness or sudden needs.

## DISCUSSION

The research findings regarding the quality of summative evaluation instruments and the effectiveness of their implementation in tahfizh education at MTs Al-Fatah Boyolali, and compares them with the literature to draw theoretical and practical implications.

The findings show that the oral assessment rubric at MTs Al-Fatah Boyolali, which includes memorization fluency (40%), tajweed and makhraj (40%), and time accuracy (10%), is valid and reliable for measuring the quantity and quality of memorization, as confirmed by the tahfizh teacher (WA, Head of Tahfizh Program). The validity of an instrument depends on its ability to measure the desired construct, in this case the ability to memorize the Qur'an. A well-designed summative evaluation should reflect clear learning objectives, including technical aspects such as memorization and tajweed. However, the absence of assessment of understanding the meaning of verses, which is a design decision of the tahfizh program in this institution, is a weakness that needs to be followed up.

The implementation of summative evaluation at MTs Al-Fatah Boyolali is effective in measuring the discipline and memorization ability of students, especially through a relevant test format (70% mubādarah, 30% full memorization). The tahfizh teacher stated that this test is relevant for evaluating memorization skills, and students who are consistent in daily deposits tend to perform well, as stated by NJ. The relevance of this test format is in line with the learning objectives and can increase student accountability. However, constraints such as time constraints, student anxiety, stress-induced mouth ulcers, and teacher or student absence due to illness or sudden need reduce the effectiveness of implementation. According to AT, anxiety is mainly experienced by students who rarely attend, suggesting that consistency of attendance affects performance. The motivation of tahfizh graduation, which SA referred to as the "most beautiful gift" for students and parents, was a strong driving factor.

However, students reported emotional impacts such as sadness when targets were not met and mouth ulcers due to stress, which is in line with the literature on psychological distress in oral evaluation (Krahwinkel et al., 2004). A one-month remedial program for students who did not pass (score <75) was shown to support grade promotion and graduation, suggesting structured follow-up. However, teacher or student absenteeism, as reported by HF, adds to the technical challenges of implementing evaluation in educational contexts.

Theoretically, this research reinforces the importance of aligning evaluation instruments with educational objectives (Tractenberg, 2021). The focus on memorization without understanding the meaning of the verse indicates the need for a more holistic approach to tahfizh evaluation (Colina & Listiana, 2021; Mohd Saad et al., 2021). Practically, MTs Al-Fatah Boyolali can consider adding a component of understanding the meaning of the verse in the assessment rubric, although maintaining the focus on memorization, extending the evaluation time, especially for the mubādarah test, to reduce time pressure, providing stress management training for students to overcome anxiety and thrush, preparing a backup schedule to anticipate teacher or student absence.

## CONCLUSION

This research analyzes the quality of instruments and the effectiveness of summative evaluation implementation in tahfizh education at MTs Al-Fatah Boyolali. Based on the findings, the summative evaluation instrument, which consists of three assessment aspects (fluency of memorization 40%, tajweed and makhraj 40%, time accuracy 10%), proved to be valid and reliable to measure the quantity and quality of memorization of the Qur'an. However, this rubric does not include an assessment of understanding the meaning of the verse, because the tahfizh program at this institution is designed to focus only on memorization. The absence of the aspect of understanding the meaning is a major weakness, because it limits the achievement of holistic tahfizh education goals, such as the formation of spiritual understanding and noble character. In addition, time constraints, especially on randomized tests (mubādarah) which account for 70% of the evaluation, affect the accuracy of the assessment, as reported by students.

In terms of implementation, summative evaluation, which is carried out twice a year (December and June) with a test format of 70% mubādarah and 30% full memorization, is considered relevant by tahfizh teachers to measure memorization ability. This evaluation is effective in measuring the discipline of students, especially those who are consistent in daily deposits, and is supported by the motivation of tahfizh graduation, which is considered the "most beautiful gift" for students and parents. The 2024/2025 semester data shows 75 students were tested with an average score of 75 (odd semester) and 80 (even semester), with an even semester score distribution: 20 students with grade A (91-100), 30 students with grade B (81-90), 20 students with grade C (71-80), and 5 students with grade D (<70). Of the 30 students in the even semester, 27 passed (score ≥75) and 3 did not pass, with a one-



month remedial program that supports grade promotion and graduation. However, implementation is constrained by time constraints, student anxiety (especially those who rarely attend), stress-induced thrush, and teacher or student absences due to illness or sudden needs.

Based on the findings, several suggestions are proposed to improve the quality and effectiveness of summative evaluation in tahfizh education at MTs Al-Fatah Boyolali:

Instrument refinement, by considering the addition of an assessment component for understanding the meaning of deep verses, extending the duration of the evaluation, especially for the mubādarah test, conducting stress management training or counseling sessions for students, and preparing a backup or replacement schedule for examiners to anticipate the absence of teachers or students due to illness or sudden needs.

Further research can be conducted to develop a tahfizh evaluation model that integrates understanding the meaning of verses without reducing the focus on memorization, as well as exploring strategies to increase student attendance in the tahfizh program.

## REFERENCES

- Adom, D., Mensah, J. A., & Dake, D. A. (2020). Test, measurement, and evaluation: Understanding and use of the concepts in education. *International Journal of Evaluation and Research in Education*, 9(1), 109–119.
- Bani, H., Jaaffar, M. Y., Katan, M., & Mohd Noor, A. H. (2017). An Overview of Governance and Accountability of Tahfiz Institutions in Malaysian: Religious Councils Perspective. *36SHS Web of Conferences*, 00028. <https://doi.org/10.1051/shsconf/20173600028>
- Colina, N., & Listiana, A. (2021). *Al-Quran Based Learning in Early Childhood Education BT - Proceedings of the 5th International Conference on Early Childhood Education (ICECE 2020)*. 19–22. <https://doi.org/10.2991/assehr.k.210322.005>
- Earle, S. (2020). Balancing the demands of validity and reliability in practice: Case study of a changing system of primary science summative assessment. *18London Review of Education*, (2), 221–235. <https://doi.org/10.14324/LRE.18.2.06>
- Ford, T. G., & Hewitt, K. K. (2020). Better Integrating Summative and Formative Goals in the Design of Next Generation Teacher Evaluation Systems. *28Education Policy Analysis Archives*, , 63. <https://api.semanticscholar.org/CorpusID:216394438>
- Garger, J., Jacques, P. H., Gastle, B. W., & Connolly, C. M. (2019). Threats of common method variance in student assessment of instruction instruments. *Higher Education Evaluation and Development*, 13(1), 2–17.
- Hanapi, W., & Rayhani, A. (2024). PENGEMBANGAN ALAT EVALUASI DALAM PENDIDIKAN AGAMA ISLAM: TANTANGAN DAN SOLUSI. *Jurnal Studi Multidisipliner*, 8(6).
- Jamil, S., Makhsin, M., & Mohamad, A. (2024). The Formation Of Huffaz Through The Application Of Sufism Concepts In Tahfiz Education. *6International Journal of*

- Modern Education*, , 1-14. <https://doi.org/10.35631/IJMOE.623001>
- Jamil, S. R., & Othman, M. K. H. (2023). ASSESSMENT AND RECOMMENDATIONS ON THE IMPLEMENTATION OF TAHFIZ INTEGRATED CURRICULUM (KBT). *International Journal of Modern Education*. <https://api.semanticscholar.org/CorpusID:261690832>
- Kibble, J. D. (2017). Best practices in summative assessment. *41Advances in Physiology Education*, (1), 110-119. <https://doi.org/10.1152/advan.00116.2016>
- Krahwinkel, T., Nastali, S., Azrak, B., & Willershausen, B. (2004). The effect of examination stress conditions on the cortisol content of saliva-a study of students from clinical semesters. *European Journal of Medical Research*, 9, 256–260.
- Mohajan, H. K. (2017). Two criteria for good measurements in research: Validity and reliability. *Annals of Spiru Haret University. Economic Series*, 17(4), 59–82.
- Mohd Saad, M. F., Mohamad, S., Ishak, H., Hussin, H., & Zainol, N. Z. N. (2021). Implementation of Tadabbur Element in Quran Memorization Process. *11International Journal of Academic Research in Business and Social Sciences*, (7). <https://doi.org/10.6007/IJARBSS/v11-i7/10381>
- Muawanah, S., Said, A., Furqoni, R., Muzayanah, U., & Mustolehudin, M. (2022). Evaluating Mandatory Tahfız Quran Program Implementation at Madrasah. *Journal of Islamic Education*. <https://api.semanticscholar.org/CorpusID:255730352>
- Parveen, N. (2021). Investigating the Memorization of the Quran Using the Grounded Theory Methodology. *The Qualitative Report*. <https://doi.org/10.46743/2160-3715>.
- Rahim, M. M. A., Borham, A. H., Hashim, A., & Hashim, I. (2018). Ability in Comprehending Meaning of the Quranic Verses among Tahfiz Students. *8International Journal of Academic Research in Business and Social Sciences*, (11). <https://doi.org/10.6007/IJARBSS/v8-i11/5338>
- Sabrina, V., Oktavia, G., Albizar, A., Susanti, H., AR, F. M., & Suryani, Y. (2022). Eight Supporting Factors for Students Success in Quran Memorization. *Khalifa: Journal of Islamic Education*, 6(1), 73.
- Souza, A. C. de, Alexandre, N. M. C., & Guirardello, E. de B. (2017). Psychometric properties in instruments evaluation of reliability and validity. *Epidemiologia e Servicos de Saude*, 26, 649–659.
- Sulastini, F., & Zamili, M. (2019). Efektivitas Program Tahfidzul Qur'an dalam Pengembangan Karakter Qur'ani. *Jurnal Pendidikan Islam Indonesia*. <https://api.semanticscholar.org/CorpusID:214392500>
- Tractenberg, R. E. (2021). The Assessment Evaluation Rubric: Promoting Learning and Learner-Centered Teaching through Assessment in Face-to-Face or Distanced Higher Education. *Education Sciences*. <https://api.semanticscholar.org/CorpusID:238699570>
- Ulumiyah, M., Maarif, M. A., & Zamroni, M. A. (2021). Implementation of the Tallaqi, Tafahhum, Tiktirar and Murajaah (3T+ 1M) Method in the Tahfidz Palace Learning Program. *Jurnal Pendidikan Islam Indonesia*, 6(1), 23–33.
- Wildan, M., & Prastowo, A. (2022). The Technique of Processing the Results of Skills Domain Assessment in the Field of Tahfidz at Islamic Boarding Schools. *Journal Tarbiyatuna*. <https://api.semanticscholar.org/CorpusID:251089428>