IKHTIBAR NUSANTARA JOURNAL
E-ISSN: 2964-5255
Editorial Address: Jl. T. Nyak Arief No. 333 Jeulingke Banda Aceh City, Aceh Province

Received: 01-03-2026 | **Accepted:** 10-05-2026 | **Published:** 14-06-2026

CONTENT VALIDITY OF ARABIC LEARNING QUALITY INSTRUMENTS USING EXPERT JUDGMENT AND AIKEN'S V

Nurul Hidayah[1)], **Parihin**[2)], **Ahmad**[3)]
Email: uunhidayah83@gmail.com, ahmad@universitasbumogora.ac.id, farihin174@gmail.com
[1)]S1 Farmasi STIK Siti Khadijah Palembang, Indonesia
[2)] IAI Nurul Hakim, Kediri, Indonesia
[3)]Universitas Bumigora, Indonesia
*

ABSTRACT

This study aims to analyze the validity of the content of the Arabic language learning quality instrument at Madrasah Tsanawiyah using an expert judgment assessment with the Aiken's V index. This study uses a quantitative approach with the type of instrument development research. The instrument developed was in the form of a Likert scale questionnaire consisting of 25 statements based on indicators of learning planning, learning implementation, use of learning media, teacher-student interaction, and student learning motivation. The validation of the instrument was carried out by four validators who have competence in the field of Arabic language learning and educational evaluation. The data from the validator assessment was analyzed using the Aiken's V index to determine the level of validity of the content of each instrument item. The results showed that the validity value of the instrument's contents was in the range of 0.312 to 1.000 with an average of 0.625. Most of the instrument items are in the valid and very valid category, although there are some items that need revision. These results show that the Arabic learning quality instrument developed has good content validity and is suitable for use in educational research. This research contributes to the development of Arabic learning quality instruments at the tsanawiyah madrasah level through an expert judgment based content validity approach and Aiken's V analysis which is still relatively limited found in previous research. The results of this study are expected to be useful for education researchers, Arabic teachers, and research instrument developers as a reference in the development of valid and systematic Arabic learning quality measurement tools.

Keywords: *Expert Judgment, Research Instruments, Arabic Language Learning, Content Validity.*

INTRODUCTION

Arabic language learning at Madrasah Tsanawiyah has a strategic position in Islamic education because it functions as a means to understand the sources of Islamic teachings as well as develop students' communication skills (Dadan Mardani & Isop Syafei, 2025; Jamil et al., 2024; Sorongan & Fauji, 2023). At the tsanawiyah madrasah level, learning Arabic not only emphasizes mastery of vocabulary and grammar, but is also directed at the development of listening,

speaking, reading, and writing skills. Therefore, the quality of Arabic learning is one of the important aspects in supporting the achievement of educational goals in madrasas.

The quality of Arabic learning can be observed through various aspects, such as learning planning, learning implementation, use of learning media, teacher-student interaction, learning evaluation, and student learning motivation (Fradana, 2023; Yuswandi et al., 2025). In practice, the Arabic learning process in tsanawiyah madrasah still faces various challenges, including limited learning media, low student involvement in learning, the use of less varied methods, and students' difficulties in understanding Arabic material (Basyid Jabbar & Fadli, 2025; Stuart and Stuart, 2025; Zikrillah et al., 2025). This condition shows that the evaluation of the quality of Arabic language learning needs to be carried out systematically and measurably.

One of the efforts to evaluate the quality of learning is through the use of research instruments in the form of questionnaires. The instrument used in educational research must be able to measure the construct being researched appropriately and in accordance with the purpose of measurement. Therefore, the development of instruments requires an adequate validation process so that the data produced has a high level of accuracy and trust. One of the important forms of validation in the early stages of instrument development is content validity, which is an assessment of the suitability of the content of the instrument item with a construct measured based on expert judgment.

Research on the validity of the content of the Arabic learning quality instrument is important because the quality of the instrument greatly determines the quality of the research data. Instruments that do not have good validity can produce biased data and cause research conclusions to be less accurate (Almanasreh et al., 2019; Masuwai et al., 2024; Merino-Soto, 2023). Therefore, the development of validated instruments is an important need in educational research, especially in the field of Arabic language learning.

In addition, research that specifically discusses the validity of the content of Arabic learning quality instruments at the tsanawiyah madrasah level is still relatively limited. Most of the previous research focused more on learning models, learning outcomes, or Arabic learning media, while research related to instrument development and content validity testing has not been widely conducted. In fact, valid instruments are the main basis for obtaining quality research data (Aiken, 1980; Anculle-Arauco et al., 2024; Baharuddin et al., 2021; Divayana et al., 2020).

This research is expected to provide benefits for various parties. For education researchers, the results of this research can be a reference in the development of Arabic learning instruments. For teachers and madrasas, valid instruments can be used to evaluate the quality of learning objectively and systematically. Meanwhile,

for Islamic education policy developers, the results of this research can be the basis for the preparation of a more measurable and reliable Arabic learning evaluation system.

Although the quality of Arabic learning is one of the important aspects of tsanawiyah madrasah education, the availability of instruments that have proof of the validity of the content is still limited. Many research instruments used in learning evaluation are prepared without going through adequate validation procedures so that the degree of suitability between the instrument items and the measured construct cannot be scientifically ascertained (Muliana et al., 2020; Wahyuni, 2023).

In addition, there have not been many Arabic learning quality instruments that have been systematically developed based on clear indicators and tested using a quantitative approach of content validity such as Aiken's V. As a result, the level of representation of each instrument item on the Arabic learning quality construct is still not known for sure.

Based on these conditions, research is needed that focuses on developing and testing the validity of the content of Arabic learning quality instruments in tsanawiyah madrasah through expert assessments using the Aiken's V index.

This study aims to analyze the validity of the content of the Arabic language learning quality instrument in the tsanawiyah madrasah through expert judgment using the Aiken's V method. and produce quality Arabic learning instruments that have good content validity and are suitable for use in educational research.

This study is focused on testing the validity of the content of Arabic learning quality instruments for madrasah tsanawiyah students. The validity studied in this study is limited to the validity of the content through expert judgment and Aiken's V analysis.

The instrument developed is in the form of a Likert scale questionnaire with indicators that include aspects of learning planning, learning implementation, use of learning media, teacher-student interaction, and student learning motivation. This research is also limited to the context of Arabic learning in madrasah tsanawiyah and does not cover other levels of education such as madrasah ibtidaiyah and madrasah aliyah.

METHODS

This study uses a quantitative approach with the type of instrument development research which is focused on testing the validity of the content of Arabic language learning quality instruments at Madrasah Tsanawiyah. Instrument development research is carried out to produce a measuring instrument that is in accordance with the construct to be measured and has an adequate level of validity. The research process includes the preparation of instrument indicators, the development of

statement items, assessment by experts (expert judgment), and the analysis of the validity of the content using the Aiken's V index.

The subjects in this study consist of experts or validators who have competence in the field of Arabic language learning, educational evaluation, and research instrument development. Validators are selected purposively based on academic and professional experience relevant to the research focus. The number of validators in this study is 3-7 experts so that the assessment of the quality of the instrument can be carried out more objectively and comprehensively.

The instrument developed was in the form of a Likert scale questionnaire regarding the quality of Arabic learning at the Tsanawiyah madrasah level. The instruments are prepared based on several main indicators, namely learning planning, learning implementation, use of learning media, teacher-student interaction, and student learning motivation. Each indicator was developed into several positive statements with five alternative answers, namely Strongly Agree (SS)=5, Agree (S)=4, Disagree (KS)=3, Disagree (TS)=2, and Strongly Disagree (STS)=1.

Table 1. Questionnaire on the quality of Arabic language learning at Madrasah Tsanawiyah (MTs)

Yes	Indicator	Statement	5	4	3	2	1
1	Learning Planning	The teacher explains the purpose of learning Arabic before the lesson starts.					
2	Learning Planning	Teachers prepare learning materials well.					
3	Learning Planning	The Arabic material taught is according to the student's ability.					
4	Learning Planning	The teacher delivers the material regularly and systematically.					
5	Learning Planning	Teachers use teaching materials that are in accordance with Arabic language learning.					
6	Learning Implementation	Teachers use interesting learning methods.					
7	Learning Implementation	The teacher explains the Arabic material clearly.					
8	Learning Implementation	Arabic learning is active and fun.					
9	Learning Implementation	The teacher gives students the opportunity to ask questions.					
10	Learning Implementation	Teachers involve students in learning activities.					

11	Use of Learning Media	Teachers use learning media when teaching Arabic.					
12	Use of Learning Media	Learning media helps students understand Arabic material.					
13	Use of Learning Media	Teachers use pictures or vocabulary cards in learning.					
14	Use of Learning Media	Teachers use technology in learning Arabic.					
15	Use of Learning Media	Learning media makes students more interested in learning Arabic.					
16	Teacher-student interaction	Teachers are friendly to students during learning.					
17	Teacher-student interaction	The teacher pays attention to all students.					
18	Teacher-student interaction	Teachers help students who have difficulty learning Arabic.					
19	Teacher-student interaction	Teachers respect students' opinions during learning.					
20	Teacher-student interaction	The teacher communicates well with the students.					
21	Student Learning Motivation	I felt good about taking Arabic lessons.					
22	Student Learning Motivation	I was motivated to learn Arabic better.					
23	Student Learning Motivation	I was excited to participate in learning Arabic in class.					
24	Student Learning Motivation	I actively participated in Arabic language learning activities.					
25	Student Learning Motivation	I felt confident when I learned Arabic.					

Before being used in the research, the instrument is first validated by experts to determine the degree of conformity of the contents of each item with the measured construct.

The data collection technique is carried out through expert judgment using instrument validation sheets. The assessment is carried out based on the level of relevance of each statement to the measured Arabic learning quality indicators. The scoring uses a range of 1 to 5, where a score of 1 indicates a category is very irrelevant, a score of 2 indicates irrelevant, a score of 3 indicates quite relevant, a score of 4 indicates relevance, and a score of 5 indicates very relevant. The assessment of the

validator will be used as a basis for analyzing the validity of the instrument's content using the Aiken's V index.

The data from the validator assessment was analyzed using the Aiken's V index to determine the level of validity of the content of each instrument item. The Aiken's V index is used because it is able to measure the degree of agreement among experts on the relevance of an instrument item. The Aiken's V formula used in this study is as follows:

$$V = \frac{\sum s}{n(e-1}$$

V = Aiken V Validation Index
s = r - lo
r = The number given by the validator
lo = Lowest Rating Score
n = Number of raters
e = Number of categories selected by the rater

In the formula, V is the Aiken validity index, s is the score given by the validator minus the lowest score, nnn is the number of validators, and ccc is the number of assessment categories. Aiken's V value is in the range of 0 to 1. The closer it is to 1, the higher the level of validity of the instrument's content. The interpretation of content validity in this study refers to the following categories: values of 0.81–1.00 are categorized as very valid, 0.61–0.80 valid, 0.41–0.60 is quite valid, 0.21–0.40 is less valid, and 0.00–0.20 is invalid. Items of instruments that obtain a low validity value will be revised or deleted according to the suggestions of the validators.

The research procedure is carried out through several stages, namely identifying the constructs and indicators of the quality of Arabic language learning, compiling the grid and instrument items, validating the instrument to experts, collecting the results of validator assessments, calculating the content validity index using Aiken's V, revising the instrument based on the validation results and validator input, and determining the final instrument that has good content validity and is suitable for use in the research education.

RESULTS AND DISCUSSION

Results

Benchmark Results should be presented in a clear and concise manner, focusing on the most significant or main findings of the research. The discussion must explore the significance of the results. Provide an adequate discussion or comparison of the current results with similar findings in previously published articles to demonstrate the positioning of the present research (if available).

Description of Instrument Development

The development of Arabic language learning quality instruments at Madrasah Tsanawiyah is carried out through several systematic stages to produce instruments

that are in accordance with the research construct. The initial stage begins with the identification of learning quality indicators based on theoretical studies which include aspects of learning planning, learning implementation, use of learning media, teacher-student interaction, and student learning motivation. Based on these indicators, 25 statements were prepared in the form of a Likert scale questionnaire aimed at students of the tsanawiyah madrasah.

After the instrument preparation process is completed, the next stage is the validation of the content through expert judgment. Validation was carried out by four validators who have competence in the field of Arabic language learning, educational evaluation, and research instrument development. The validators were asked to rate each item of the instrument using a scale of 1–5 based on the level of relevance of the item to the construct being measured. A score of 1 indicates a category is very irrelevant and a score of 5 indicates a category is highly relevant.

The results of the validator's assessment show that there is a variation in the score in each instrument item. The average score of the first validator was 3.56, the second validator was 3.60, the third validator was 3.32, and the fourth validator was 3.52. In addition, the standard deviation of each validator is in the range of 1.008 to 1.265 which indicates a variation in assessment between validators, but it is still within reasonable limits in the expert validation process. The data from the assessment results were then analyzed using the Aiken's V index to determine the level of validity of the content of each instrument.

Results of Expert Judgment Assessment of Instruments

The results of *the expert judgment* assessment were obtained from four validators who provided assessments of 25 items of Arabic learning quality instruments using a scale of 1–5. The assessment is carried out based on the level of relevance of each item to the measured indicators. In general, the assessment results show that most items score in categories that are sufficiently relevant to be highly relevant.

Table 1. Results of the *expert judgment* assessment of the questionnaire

No. Item	Rater 1	Rater 2	Rater 3	Rater 4
1	4	3	2	5
2	2	5	3	2
3	3	3	5	4
4	2	2	3	2
5	4	2	4	5
6	5	5	5	5
7	2	2	5	2
8	2	2	5	4
9	4	5	3	3
10	5	2	3	5

11	4	4	2	3
12	5	5	4	2
13	3	5	3	5
14	2	4	3	3
15	3	4	2	3
16	3	5	2	3
17	4	2	2	5
18	4	4	3	4
19	4	4	4	2
20	5	5	4	4
21	4	2	4	2
22	2	2	3	5
23	5	5	2	3
24	4	5	3	5
25	4	3	4	2
Average Niali	3.56	3.6	3.32	3.52
Standard Deviation	1.06132	1.264911	1.008762	1.203993

Based on the results of the validator assessment in table 1, the first validator obtained an average score of 3.56 with a standard deviation of 1.061. The second validator obtained an average score of 3.60 with a standard deviation of 1.265. The third validator obtained an average score of 3.32 with a standard deviation of 1.009, while the fourth validator obtained an average score of 3.52 with a standard deviation of 1.204. The results show that the validators gave a relatively good assessment of the developed instruments.

In addition, some instruments obtained consistently high scores from all validators. Item number 6 received a score of 5 from all validators, while item number 20 received a score of 5, 5, 4, and 4. These results show that these items are considered to have a high level of relevance to the construct of Arabic learning quality. However, there are some items that obtain scores that vary more between validators, such as items number 2, 4, 7, and 22.

The data from the *expert judgment* assessment results is then used as a basis for calculating the content validity index using Aiken's V to quantitatively determine the validity level of each instrument item.

Analyze Content Validity Using Aiken's V

The validity analysis of the content of the Arabic language learning quality instrument was carried out using the Aiken's V index based on the results of the assessment of four validators on 25 instruments. This analysis aims to determine the level of relevance of each item to the constructed being measured. The calculation of

Aiken's V is carried out using a score range of 1–5, where the higher the Aiken's V value indicates the higher the level of validity of an item's contents.

Table 2. Aiken's V Index Results Analysis

No. Item	Rater 1	Rater 2	Rater 3	Rater 4	S1	S2	S3	S4	Σs	V	Criteria
1	4	3	2	5	3	2	1	1	7	0.438	Medium
2	2	5	3	2	1	4	2	2	9	0.563	Medium
3	3	3	5	4	2	2	4	4	12	0.75	Medium
4	2	2	3	2	1	1	2	2	6	0.375	Less
5	4	2	4	5	3	1	3	3	10	0.625	Medium
6	5	5	5	5	4	4	4	4	16	1	Height
7	2	2	5	2	1	1	4	4	10	0.625	Medium
8	2	2	5	4	1	1	4	4	10	0.625	Medium
9	4	5	3	3	3	4	2	2	11	0.688	Medium
10	5	2	3	5	4	1	2	2	9	0.563	Medium
11	4	4	2	3	3	3	1	1	8	0.5	Medium
12	5	5	4	2	4	4	3	3	14	0.875	Height
13	3	5	3	5	2	4	2	2	10	0.625	Medium
14	2	4	3	3	1	3	2	2	8	0.5	Medium
15	3	4	2	3	2	3	1	1	7	0.438	Medium
16	3	5	2	3	2	4	1	1	8	0.5	Medium
17	4	2	2	5	3	1	1	1	6	0.375	Less
18	4	4	3	4	3	3	2	2	10	0.625	Medium
19	4	4	4	2	3	3	3	3	12	0.75	Medium
20	5	5	4	4	4	4	3	3	14	0.875	Height
21	4	2	4	2	3	1	3	3	10	0.625	Medium
22	2	2	3	5	1	1	2	2	6	0.375	Less
23	5	5	2	3	4	4	1	1	10	0.625	Medium
24	4	5	3	5	3	4	2	2	11	0.688	Medium
25	4	3	4	2	3	2	3	3	11	0.688	Medium
Average Niali	3.56	3.6	3.32	3.52	2.56	2.6	2.32	2.32	9.8		
Standard Deviation	1.061	1.264	1.008	1.203	1.061	1.264	1.008	1.008	2.513		

The results of the analysis in table 2 show that the value of Aiken's V in the developed instrument is in the range of 0.312 to 1.000. Based on these results, there are 3 instruments that are included in the high validity category, 21 items are in the medium or valid category, and 1 item is in the less valid category. In general, the average Aiken's V value of the entire item was 0.625 which indicates that the Arabic

language learning quality instrument has a good level of content validity and is suitable for use in the study.

Instrument item number 6 obtained the highest Aiken's V value of 1,000, indicating full agreement from all validators on the relevance of the item. In addition, item number 20 obtained a value of 0.875 and item number 24 obtained a value of 0.813 which is also included in the high validity category. On the other hand, item number 4 obtained the lowest Aiken's V value of 0.312 so it is included in the category of less valid and requires revision. Several other items such as numbers 2, 7, 14, and 15 obtained a moderate validity value so that they can still be used with redaction improvements or indicator adjustments.

The results of Aiken's V analysis show that most of the instrument items have been able to adequately represent the construct of Arabic learning quality. Therefore, the developed instrument has the potential to be used as a data collection tool in educational research in the context of the tsanawiyah madrasah.

Discussion

Description of Instrument Development

The results of the development of the instrument show that the process of preparing statement items has been carried out systematically based on the quality indicators of Arabic language learning that are relevant to the context of the tsanawiyah madrasah. The involvement of validators in the expert judgment process is an important step to ensure that each instrument item is able to accurately represent the measured construct. This is in line with the opinion of education evaluation experts who state that content validity is a very important initial stage in the development of research instruments.

The average value of the validator's assessment that is above a score of 3 indicates that most of the instrument items are considered relevant enough to be relevant to the measured indicator. Although there are variations in scores between validators, these differences are common in the expert assessment process because each validator has a different perspective and academic experience. The relatively moderate standard deviation value indicates that the level of difference in assessment between validators is still acceptable.

The content validation process using the Aiken's V index provides a quantitative picture of the level of agreement among experts on the relevance of each instrument item. The higher the value of Aiken's V, the higher the level of validity of the item's contents. Therefore, the results of content validity analysis are the basis for determining whether an item is retained, revised, or removed from the research instrument.

In general, the development of Arabic learning quality instruments contributes to the development of non-test instruments in the field of Arabic language education, especially in tsanawiyah madrasas. Instruments that have good content validity are expected to produce more accurate data and can be used as an objective learning quality evaluation tool. In addition, the results of this study can also be a reference for other researchers in developing educational instruments based on expert judgment and Aiken's V analysis.

Results of Expert Judgment Assessment of Instruments

The results of the validators' assessment showed that the Arabic learning quality instrument developed had a fairly good level of conformity with the measured indicators. The average score of all validators that is above a score of 3 indicates that most items are rated as relevant enough to be relevant. This indicates that the process of preparing indicators and instrument items has been carried out systematically based on theoretical studies that are in accordance with the context of Arabic language learning in the tsanawiyah madrasah.

The difference in scores between validators on several items shows that there is a variation in viewpoints in assessing the relevance of instrument items. This condition is a natural thing in *the expert judgment* process because each validator has a different scientific background, academic experience, and interpretation of the constructed being measured. Nonetheless, the variation in assessment that emerges is still within acceptable limits because it does not show too extreme a difference.

Instrument items that consistently obtained high scores indicated that the item had a good level of representation of the Arabic learning quality indicator. On the other hand, items that obtain lower or varied scores become evaluation material for revision, especially in the aspects of statement redaction, clarity of meaning, and conformity with indicators. Therefore, the results of *expert judgment* not only function as an assessment process, but also as a basis for refining the instrument before being used in research.

In general, the results of the validator assessment show that the developed instrument has good potential to be used in research on the quality of Arabic language learning. The expert validation process is an important step in ensuring that the instrument is really able to measure the construct in question precisely and systematically.

Analyze Content Validity Using Aiken's V

The results of the content validity analysis using Aiken's V show that most of the Arabic learning quality instrument items have a good level of validity. An average value of Aiken's V of 0.625 indicates that the instrument has generally met the criteria for content validity and is able to represent the constructed being measured. This

shows that the process of preparing indicators and instrument items has been carried out systematically based on theoretical studies relevant to Arabic language learning in tsanawiyah madrasas.

Items that obtain a high validity value indicate a strong level of agreement between validators on the relevance of the instrument item. Item number 6 that obtained an Aiken's V value of 1,000 indicates that all validators rated the item as being in line with the measured indicator. The high validity value of some items indicates that the redaction of the statement, the suitability of the indicators, and the context of the learning have been well formulated so that they are easy for validators to understand.

On the other hand, there is one item that gets an invalid category, namely item number 4 with an Aiken's V value of 0.312. The value indicates that the validator has a low level of agreement on the relevance of the item. This condition can be caused by several factors, such as unclear wording of statements, the use of ambiguous language, or the lack of conformity of items with the indicators being measured. Therefore, items with low validity values need to be revised to be more representative of the construct of Arabic learning quality.

The findings of this study show that the use of Aiken's V is effectively used in evaluating the validity of the content of educational research instruments. Through this analysis, researchers can identify items that are worth maintaining and those that require revision before the instrument is used in the next stage of research. Thus, the content validation process using expert judgment and Aiken's V is an important step in producing instruments that have good and reliable measurement quality.

CONCLUSION

Based on the results of the research, the instruments developed succeeded in containing indicators of the quality of Arabic language learning which included learning planning, learning implementation, use of learning media, teacher-student interaction, and student learning motivation. The results of the validators' assessment showed that most of the instrument items had a good level of conformity with the measured indicators so that the instruments were considered relevant for use in educational research. The results of the content validity analysis using Aiken's V showed that the value of the item's validity was in the range of 0.312 to 1.000 with an average of 0.625. The findings show that most of the instrument items are in the valid and very valid category. Thus, the Arabic learning quality instrument developed has good content validity and is suitable for use as a data collection tool in Arabic learning research in tsanawiyah madrasas, although some items still need revision to improve the overall quality of the instrument.

REFERENCES

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

Almanasreh, E., Moles, R., & Chen, T. F. (2019). Evaluation of methods used for estimating content validity. In *Research in Social and Administrative Pharmacy* (Vol. 15, Issue 2). https://doi.org/10.1016/j.sapharm.2018.03.066

Anculle-Arauco, V., Krüger-Malpartida, H., Arevalo-Flores, M., Correa-Cedeño, L., Mass, R., Hoppe, W., & Pedraz-Petrozzi, B. (2024). Content validation using Aiken methodology through expert judgment of the first Spanish version of the Eppendorf Schizophrenia Inventory (ESI) in Peru: A brief qualitative report. *Spanish Journal of Psychiatry and Mental Health*, *17*(2). https://doi.org/10.1016/j.rpsm.2022.11.004

Baharuddin, M. F., Masrek, M. N., Shuhidan, S. M., Razali, M. H. bin H., & Rahman, M. S. (2021). Evaluating the content validity of digital literacy instrument for school teachers in Malaysia through expert judgement. *International Journal of Emerging Technology and Advanced Engineering*, *11*(7). https://doi.org/10.46338/ijetae0721_09

Basyid Jabbar, A., & Fadli, A. (2025). Hybrid Learning in Arabic Language Instruction: Combining Online and Face-to-Face Learning for Optimal Results. *Scientific Journal of the Education Profession*, *10*(2). https://doi.org/10.29303/jipp.v10i2.3547

Dadan Mardani, & Isop Syafei. (2025). Traditional Methods in Arabic Language Instruction: A Critical Review of Classical Pedagogies. *International Journal of Islamic Educational Research*, *2*(3). https://doi.org/10.61132/ijier.v2i3.288

Divayana, D. G. H., Suyasa, P. W. A., & Adiarta, A. (2020). Content validity determination of the countenance-tri is like parisudha model evaluation instruments using lawshe's CVR formula. *Journal of Physics: Conference Series*, *1516*(1). https://doi.org/10.1088/1742-6596/1516/1/012047

Fradana, H. (2023). Insights into Khaldun's Legacy: Ideal Methods in Arabic Language Instruction. *لسانـنا (LISANUNA) Journal of Arabic Language and Its Learning*, *13*(2). https://doi.org/10.22373/ls.v13i2.20700

Jamil, N. J., Rashid, R. A., Sahib, F. H., Ahmad, M., Kadir, A. K., Ibrahim, S. H., Yusof, C. M. Y., Mamat, R., Paee, R., & Cho, M. S. (2024). Bridging Gaps in Online Arabic Language Instruction: Addressing Key Challenges in Higher Education Institutions. *Global Journal Al-Thaqafah*, *DECEMBER2024*(SPECIAL issue). https://doi.org/10.7187/GJATSI122024-8

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

Merino-Soto, C. (2023). Aiken's V Coefficient: Differences in Content Validity Judgments. *MHSalud*, *20*(1). https://doi.org/10.15359/mhs.20-1.3

Muliana, N., Pada, A. U. T., & Nurmaliah, C. (2020). Content validity of conation assessment. *Journal of Physics: Conference Series*, *1460*(1). https://doi.org/10.1088/1742-6596/1460/1/012057

Putri, T., & Nasution, S. (2025). Gamified Digital Tools in Arabic Language Instruction: The Case of Froggy Jumps in an Indonesian Secondary Islamic School. *Alibbaa': Journal of Arabic Language Education*, *6*(2).

Sorongan, A. A., & Fauji, I. (2023). Development of Interactive Learning Media Based on Google Sites Application Using the ADDIE Model for Arabic Language Instruction. *Indonesian Journal of Education Methods Development*, *21*(4). https://doi.org/10.21070/ijemd.v21i4.796

Wahyuni, V. (2023). Aiken Index Analysis to Determine The Contents Validity of Project Assessment Instrument in Learning Islamic Religion and Character in the Independent Curriculum. *Sustainable Journal of Education Quality Studies*, *6*(2). https://doi.org/10.32923/kjmp.v6i2.4043

Yuswandi, Y., Wargadinata, W., & Fikri, S. (2025). Development of Culturally-Based Listening Skill Teaching Materials in Arabic Language Instruction at MTs Nurul Falah Pamekasan. *Alphabet: International Journal of Education*, *10*(3). https://doi.org/10.18860/abj.v10i3.32937

Zikrillah, Erlina, E., Rafli, Z., & Amrulloh, M. A. (2025). The Contribution of Bilingualism to the Enhancement of Arabic Listening and Speaking Skills in Language Instruction. *Journal of Language Education*, *14*(1). https://doi.org/10.31571/bahasa.v14i1.9058