AN NABIGHOH
P-ISSN: 1907-1183 E-ISSN: 2581-2815
Vol. 27, No. 1 (2025): 201-222
http://e-journal.metrouniv.ac.id/an-nabighoh
DOI: https://doi.org/10.32332/an-nabighoh.v27i1.201-222

ITEM ANALYSIS OF INSYA' ASSESSMENT BASED ON BLOOM'S TAXONOMY AT ISLAMIC SENIOR HIGH SCHOOL DARUSSALAM RAJAPOLAH TASIKMALAYA

Cecep Sobar Rochmat[1]*, Fitri Masturoh[2], Fairuz Subakir[3], Zhulfa Halimatus Sadiah[4]
[1,2,3] Universitas Darussalam Gontor, Indonesia
[4] Fatih Sultan Mehmet Vakıf Üniversitesi, Turkey

Article Info

Article History:

Received: March 2, 2025
Revised: April 24, 2025
Accepted: June 23, 2025
Published: June 30, 2025

***Corresponding Author:**

Name:
Cecep Sobar Rochmat
Email:
cecep.rochmat@unida.gontor.ac.id

Abstract

Despite the widespread use of Insya' assessments in educational institutions, comprehensive evaluations of item quality covering validity, reliability, discrimination, difficulty, and cognitive levels based on Revised Bloom's Taxonomy remain limited. Previous studies rarely integrate statistical analysis with cognitive classification, particularly in high school Arabic education. This study addresses that gap by analyzing ten Insya' assessment items from Grade 12 students at Islamic Senior High School Darussalam Tasikmalaya. Data consisted of exam questions, student answers, and scores, supported by literature. Collection methods involved interviews, observation, and documentation, with analysis conducted using SPSS version 30. The results showed 80% of items were valid, with a reliability coefficient of 0.743. Discrimination analysis rated 30% very good, 40% good, 10% acceptable, and 20% poor. Difficulty analysis revealed that 10% of the items were difficult, 70% were moderate, and 20% were easy. Cognitive levels indicated 80% assessed Understanding (C2), 10% Applying (C3), and 10% Creating (C6), with only 10% categorized as higher-order thinking skills (HOTS). These findings highlight the need for improved item design that promotes higher-order thinking and aligns more effectively with cognitive learning goals.

Keywords: Item Analysis; Insya' Assessment; Bloom's Taxonomy; Arabic Lesson; HOTS.

مستخلص البحث

يهدف هذا البحث إلى تحليل جودة أسئلة امتحان لمادة الإنشاء، مؤسّسًا على معايير الصدق، والثبات، والتمييز، ومستوى الصعوبة، والمستويات المعرفية وفقًا لتصنيف بلوم المعدل. وقد ركّز البحث على عشر فقرات من أسئلة الإنشاء لطلاب الصف الثاني عشر في المدرسة الثانوية الإسلامية دار السلام راجاپولاه تاسيكمالايا. تتكوّن البيانات الأساسية من أسئلة الامتحان، وأوراق إجابات الطلاب، ونتائجهم، في حين تتمثّل البيانات الثانوية في الكتب والمراجع العلمية ذات الصلة. جُمعت البيانات باستخدام المقابلات، والملاحظة، والوثائق، ثم حُلّلت باستخدام برنامج SPSS (الإصدار ٣٠). أظهرت النتائج أن ٨٠٪ من الفقرات صالحة، بمعامل ثبات قدره ٠،٧٤٣. وأشارت نتائج التمييز إلى أن ٣٠٪ ممتازة، و ٤٠٪ جيدة، و ١٠٪ مقبولة، و ٢٠٪ ضعيفة. أما من حيث مستوى الصعوبة، فكانت ١٠٪ صعبة، و ٧٠٪ متوسطة، و ٢٠٪ سهلة. ومن حيث المستويات المعرفية، فإن ٨٠٪ من الفقرات تقيس الفهم (ج٢)، و١٠٪ التطبيق (ج٣)، و ١٠٪ الإبداع (ج٦)، مع ١٠٪ فقط تندرج ضمن مهارات التفكير العليا.

كلمات أساسية: تحليل الأسئلة؛ تقييم الإنشاء؛ تصنيف بلوم؛ اللغة العربية؛ مهارات التفكير العليا.

Please cite this article as Cecep Sobar Rochmat, Fitri Masturoh, Fairuz Subakir, Zhulfa Halimatus Sadiah "Item Analysis of Insya' Assessment Based on Bloom's Taxonomy at Islamic Senior High School Darussalam Rajapolah Tasikmalaya." *An Nabighoh* 27, No. 1 (2025): 201-222. DOI: https://doi.org/10.32332/an-nabighoh.v27i1.201-222

Introduction

Languag e is defined as an arbitrary system of sound symbols used to express ideas and feelings.[1] It enables humans to communicate.[2] Among the many languages spoken worldwide, Arabic stands out as one of the most influential.[3] It has maintained a deep and inseparable connection with Islamic teachings. As Islam originated in Arabia, its teachings are best understood within the context of Arab culture and language. Since the time of Prophet Muhammad (SAW), the spread of Islam beyond the Arabian Peninsula has continued to carry Arabic as a key part of its scholarly tradition and the transmission of Islamic values, even across diverse linguistic regions.[4] In conclusion, Arabic's significance extend beyond religion, impacting literature, science, and global communication.

In Arabic language education, Insya' is a crucial component that develops students' ability to compose coherent and creative texts.[5] Insya serves as a platform to develop students' writing skills and enhance their creative thinking abilities.[6] While also fostering critical thinking—one of the essential 21st-century competencies, in which the 4Cs (critical thinking, creativity, collaboration and communication) are essential in this era.[7] This emphasis is reflected in the Islamic and Arabic Language Education Curriculum under KMA No. 183 of 2019, which mandates the implementation of higher-order thinking skills (HOTS)-based assessments.[8] Therefore, Insya' assessment should measure not only linguistic knowledge but also critical thinking and problem-solving abilities.

---

[1] Abdallah Ali Ali Althory, "Linguistic Features Acquired by a Language of A Speaking Society," *Journal of Arabic Language Sciences and Literature* 1, no. 4 (September 30, 2022): 62–77, https://doi.org/10.26389/AJSRP.R150522.

[2] Cecep Sobar Rochmat Cecep et al., "The Analysis of Experiential Learning Method of Dale's Cone Experience Model in Improving the Effectiveness of Arabic Language Learning," *Izdihar : Journal of Arabic Language Teaching, Linguistics, and Literature* 7, no. 1 (April 30, 2024), https://doi.org/10.22219/jiz.v7i1.30823.

[3] Muassomah, "From Global Language Use to Local Meanings: Arabic to Indonesian Absorption," *IAS Journal of Localities* 1, no. 1 (May 24, 2023): 16–29, https://doi.org/10.62033/iasjol.v1i1.10.

[4] Suryadinata et al., "The Challenge of Mastering Arabic Language and Its Relation to the Ability to Read Kitab Kuning among Students in Tapal Kuda Islamic Boarding School," *RETORIKA: Jurnal Ilmu Bahasa* 10, no. 3 (2024), https://ejournal.warmadewa.ac.id/index.php/jret/article/view/10488.

[5] Noza Aflisia, Kasmantoni, and Yunika Febri Yanti, "Writing Errors: A Study of Students Linguistic Errors in Insya Muwajjah Learning," *An Nabighoh* 26, no. 1 (2024): 79–96, https://doi.org/10.32332/annabighoh.v26i1.79-96.

[6] Nidya Ayu, Suharno, and Kartika Chrysti Suryandari, "Exploration: Creative Thinking Skills in Writing Essays Media-Based Image Series," *International Journal of Elementary Education* 7, no. 1 (February 16, 2023): 1–7, https://doi.org/10.23887/ijee.v7i1.54095.

[7] Cecep Sobar Rochmat et al., "The Quality of Education from Islamic Perspective Analysis of The Merdeka Belajar Curriculum in Facing The Society 5.0 Era," *Jurnal Tarbiyatuna* 14, no. 1 (June 30, 2023): 75–93, https://doi.org/10.31603/tarbiyatuna.v14i1.8633.

[8] Susanto Susanto et al., "Improving Students' Creative Thinking In Learning Arabic Through HOTS Based Project Based Learning Model," *An Nabighoh* 24, no. 1 (June 28, 2022): 1, https://doi.org/10.32332/an-nabighoh.v24i1.3924.

This highlights the importance of writing skills as a core component in language learning, especially through Insya'. Writing, defined as the ability to express ideas through structured compositions, is equally important as other language skills.[9] It plays a vital role in conveying structured and meaningful information. Therefore, writers must express ideas accurately in both linguistic structure and clarity. In fact, writing is often viewed as the most complex of the four language skills. It demands the integration of linguistic competence, logical reasoning, and structured expression. These are challenges commonly faced in Arabic language learning, particularly by students unfamiliar with Arabic script.[10]

Well-designed assessments are essential for evaluating whether students have learned. Traditional assessment practices, such as multiple-choice questions, essays, and short answer questions, have been widely used to infer student knowledge and learning.[11] Given its complexity and importance, the assessment of Insya' must be carried out with careful and accurate methods. To effectively assess Insya', a comprehensive evaluation approach is required, wich includes examining the quality of the assessment items. One of the most effective methods for evaluating item quality is item analysis.[12] Item analysis is an integral part of test development and is a necessary for the construction of a test.[13] According to Mahmudi et al. Item analysis involves a series of activities that determine the quality of exam question items, providing essential information for enchancing question quality.[14] However, item analysis is often overlooked, resulting in inaccurate assessments and ineffective measurement of students' true abilities.

As essential aspect of item analysis is aligning assessment item with cognitive levels in Bloom Taxonomy. Bloom's taxonomy is a widely used hierarchy that aligns both curriculum and assessment goals as it describes learning objectives in

[9] Aflisia, Kasmantoni, and Yanti, "Writing Errors."

[10] Zulaeha Zulaeha and Musdalifah Musdalifah, "Strengthening Students' Arabic Writing Skills Through Insya' Muwajjah in Higher Education," *AL-ISHLAH: Jurnal Pendidikan* 14, no. 1 (February 1, 2022): 591–602, https://doi.org/10.35445/alishlah.v14i1.1103.

[11] Sumayyia Marar et al., "Development and Validation of an Instrument to Assess the Knowledge and Perceptions of Predatory Journals," *Heliyon* 9, no. 11 (November 2023): e22270, https://doi.org/10.1016/j.heliyon.2023.e22270.

[12] Ameema Mahroof and Muhammad Saeed, "Evaluation of Question Papers by Board of Intermediate and Secondary Education Using Item Analysis and Blooms Taxonomy," *Bulletin of Education and Research* 43, no. 3 (2021), https://pu.edu.pk/images/journal/ier/PDF-FILES/5_43_3_21.pdf.

[13] Hanwook [Henry] Yoo and Ronald K. Hambleton, "Digital Module 08: Foundations of Operational Item Analysis Https://Ncme.Elevate.Commpartners.Com," *Educational Measurement: Issues and Practice* 38, no. 3 (September 2019): 116–17, https://doi.org/10.1111/emip.12289; Charles I. Mosier, "A Note on Item Analysis and the Criterion of Internal Consistency," *Psychometrika* 1, no. 4 (December 1936): 275–82, https://doi.org/10.1007/BF02287879.

[14] Ihwan Mahmudi et al., "Item Analysis Of Arabic Language Examination," *Ijaz Arabi Journal of Arabic Learning* 6, no. 3 (October 22, 2023), https://doi.org/10.18860/ijazarabi.v6i3.19821.

Please cite this article as Cecep Sobar Rochmat, Fitri Masturoh, Fairuz Subakir, Zhulfa Halimatus Sadiah "Item Analysis of Insya' Assessment Based on Bloom's Taxonomy at Islamic Senior High School Darussalam Rajapolah Tasikmalaya." *An Nabighoh* 27, No. 1 (2025): 201-222. DOI: https://doi.org/10.32332/an-nabighoh.v27i1.201-222

terms of explicit and implicit cognitive skills and abilities.[15] The taxonomy categorizes cognitive processes into six levels: remembering (C1), understanding (C2), applying (C3), analyzing (C4), evaluating (C5), and creating (C6).[16] These levels are further categorized into two groups: Lower Order Thinking Skills (LOTS) and Higher Order Thinking Skills (HOTS).[17]

Previous studies have emphasized the importance of item analysis in educational evaluation. For example, Kumar et al. conducted an analysis of multiple-choice questions to measure validity, reliability, difficulty level, and discrimination index. Their findings improved the quality of the question bank by identifying items that needed revision or elimination.[18] Similarly, Karim et al. analyzed the discrimination power and difficulty level of teacher-made English test items and found that many items needed improvement to enhance assessment effectiveness.[19] Furthermore, Kumar, Singh, and Dhankhar applied data mining to classify exam questions based on Bloom's Taxonomy. Their model successfully categorized cognitive levels, supporting the development of more balanced and effective assessments.[20]

However, these studies primarily focused on multiple-choice items and subjects such as English, physics, and Arabic grammar. Few have addressed Insya' assessments in Arabic writing, particularly with regard to item validity, reliability, difficulty, discrimination power, and alignment with Bloom's cognitive levels. This indicates a gap in the literature and shows that more targeted research is needed in this area. Therefore, this study aims to address this gap by analyzing the quality of Insya' assessment items, thereby contributing to more effective Arabic writing assessments.

---

[15] Syaamantak Das, Shyamal Kumar Das Mandal, and Anupam Basu, "Classification of Action Verbs of Bloom's Taxonomy Cognitive Domain: An Empirical Study," *Journal of Education* 202, no. 4 (October 2022): 554–66, https://doi.org/10.1177/00220574211002199.

[16] Cecep Sobar Rochmat, Syifa Rizki Sholihah, and Shofia Niswah Qonita, "Forming Critical Character With Higher Order Thinking Skill (HOTS) Based Learning Assessment in Islamic Religious Education Subjects," *Educan : Jurnal Pendidikan Islam* 6, no. 2 (August 2, 2022): 236, https://doi.org/10.21111/educan.v6i2.8131.

[17] Ihwan Mahmudi, Fitri Masturoh, and Wiwiek Dwi Febrianti, "Higher Order Thinking Skills (HOTS) Based Assessment: A Proposed Model for Arabic Learning," *Shibghoh: Prosiding Ilmu Kependidikan UNIDA Gontor* 1 (2023), https://ejournal.unida.gontor.ac.id/index.php/shibghoh/article/view/10195.

[18] Dharmendra Kumar et al., "Item Analysis of Multiple Choice Questions: A Quality Assurance Test for an Assessment Tool," *Medical Journal Armed Forces India* 77 (February 2021): S85–89, https://doi.org/10.1016/j.mjafi.2020.11.007.

[19] Sayit Abdul Karim, Suryo Sudiro, and Syarifah Sakinah, "Utilizing Test Items Analysis to Examine the Level of Difficulty and Discriminating Power in a Teacher-Made Test," *EduLite: Journal of English Education, Literature and Culture* 6, no. 2 (2021): 256, https://doi.org/10.30659/e.6.2.256-269.

[20] Amit Kumar, Dinesh Singh, and M. S. Dhankhar, "Analysis of Bloom Taxonomy-Based Examination Data Using Data Mining," *International Journal of Intelligent Systems and Applications in Engineering* 12, no. 4s (2024), https://www.ijisae.org/index.php/IJISAE/article/view/3860.

This study contributes to the field of Arabic language education by analyzing the quality of Insya' assessment items, particularly in measuring higher-order thinking skills (HOTS) and aligning assessments with Bloom's Taxonomy. The findings are expected to support educators and curriculum developers in improving the design and implementation of Insya' assessments that enhances students' critical and creative thinking skills in Arabic writing.

Based on the above background, this study explicitly seeks to answer the following research questions: How valid are the Insya' assessment items? How reliable are the Insya' assessment items? What is the discrimination power of the Insya' assessment items? What is the difficulty level of the Insya' assessment items? And finally, what cognitive levels, according to Bloom's Taxonomy, do the Insya' assessment items measure?

Method

This research used a quantitative descriptive approach. It was conducted at Islamic Senior High School of Darussalam Rajapolah Tasikmalaya, involving all 68 Grade 12 students through saturated sampling. This school was selected because it is one of the leading Islamic senior high schools in the region, accredited with an A (excellent) rating, and has a strong focus on Arabic language education. One of the school's core missions is to enhance students' Arabic language proficiency, including strengthening Insya' (Arabic composition) skills as part of the core curriculum. Additionally, the school regularly conducts Insya' assessments, making it a relevant site for research on item analysis in Arabic writing assessments. Data collection involved documentation of exam questions, answer sheets, and corresponding scores. Interviews with Arabic language teachers were conducted to gain insights into question design and assessment practices. Observation was also carried out to gather information about the Insya' assessment in exams at Islamic Senior High School of Darussalam Rajapolah Tasikmalaya.

Data analysis was performed using SPSS Statistics version 30 with quantitative descriptive analysis. Item validity was measured using Pearson's Product Moment, and reliability was assessed using Cronbach's Alpha. Discrimination power distinguish between high- and low-performing students, and difficulty level was categorized as easy, moderate, or difficult. Cognitive levels were analyzed using Revised Bloom's Taxonomy, focusing on remembering (C1), understanding (C2), applying (C3), analyzing (C4), evaluating (C5), and creating (C6). This method effectively identifies cognitive levels in educational assessments.[21]

[21] Nira Erdiana and Sulastri Panjaitan, "How Is HOTS Integrated into the Indonesian High School English Textbook?," *Studies in English Language and Education* 10, no. 1 (January 31, 2023): 60–77, https://doi.org/10.24815/siele.v10i1.26052.

Result and Discussion

Test instruments are essential for evaluating students' abilities. Their quality depends on several key attributes: validity, reliability, practicality, discrimination, and objectivity.[22] To ensure the effectiveness of these instruments, item analysis is a critical step that provides detailed insights into the quality of each test item.[23] According to Najouan Hrich, et al. item analysis as an statistical technique used to evaluate the quality of test items in educational assessment.[24] It is especially useful for improving items to be reused in future test.[25] This process evaluates each item's validity, reliability, discrimination power, and difficulty level to ensure the overall quality of the test.[26] Additionally, item analysis provides empirical data that helps teacher develop a useful item bank for practical use. Therefore, item analysis involves thoroughly examining test questions along with related components, such as student responses, to improve question quality and help to select the best items by excluding the poor test items.

Item Analysis of Validity, Reliability, Discrimination Power, & Difficulty Level

Validity

Validity is a fundamental consideration in educational measurement. It refers to the extent to which a test accurately measures what it is intended to measure. A valid test ensures that the information used to make decisions about learners is accurate, relevant, and representative.[27] In this context, the purpose of validity is to confirm that the assessment truly reflects the construct it claims to evaluate. Therefore, a test is considered valid when its items effectively measure the intended competencies or learning objectives.[28]

In addition to statistical validity, content validity plays an important role, especially in educational assessments, by involving expert judgment to determine

[22] Maulia Yasminah Zakkiyah et al., "Assessment Design and Analysis of Arabic Reading Skills Instructional Materials," *IJIE International Journal of Islamic Education* 3, no. 1 (June 28, 2024): 31–46, https://doi.org/10.35719/ijie.v3i1.2000.

[23] Memy Wardani Elthia and Rosidah Alawiyah, "An Analysis of the Quality Test Items Used in the Final Semester Assessment at MTs Nurul Yaqin Gili Air," *Jurnal Test 2* 6, no. 1 (November 18, 2023), https://jonedu.org/index.php/joe/article/view/4572.

[24] Najoua Hrich, "Artificial Intelligence Item Analysis Tool for Educational Assessment: Case of Large-Scale Competitive Exams," *International Journal of Information and Education Technology* 14, no. 6 (2024): 822–27, https://doi.org/10.18178/ijiet.2024.14.6.2107.

[25] Z. A. Ashraf and Jaseem K, "Classical and Modern Methods in Item Analysis of Test Tools," *International Journal of Research and Review* 7, no. 5 (2020), https://www.ijrrjournal.com/IJRR_Vol.7_Issue.5_May2020/Abstract_IJRR0058.html.

[26] Nor Hafizi Mohd Khalid et al., "Psychometric Properties of Teacher Classroom Assessment Literacy Instrument Using Rasch Model Analysis," *International Journal of Evaluation and Research in Education (IJERE)* 12, no. 2 (June 1, 2023): 638, https://doi.org/10.11591/ijere.v12i2.23607.

[27] Marar et al., "Development and Validation of an Instrument to Assess the Knowledge and Perceptions of Predatory Journals."

[28] Agung Setiabudi, Mulyadi Mulyadi, and Hilda Puspita, "An Analysis of Validity and Reliability of a Teacher-Made Test," *Journal of English Education and Teaching* 3, no. 4 (December 6, 2019): 522–32, https://doi.org/10.33369/jeet.3.4.522-532.

whether each item is essential and relevant. Lawshe suggested that an item has strong content validity if more than half of the expert panel agree on its essentiality. The higher the number of experts who agree on an item's essentiality, the greater its content validity. This perspective is supported by Bashooir and Supahar, who emphasized the importance of expert involvement from various disciplines, such as linguists, educational technologists, and practitioners, in validating the relevance of test items.[29]

To calculate validity, the Product Moment correlation formula is commonly used, as follows:

$$r_{xy} = \frac{n\sum xy - (\sum x)(\Sigma y)}{\sqrt{[n\sum x^2 - (\sum x)^2][n\sum y^2 - (\sum y)^2]}}$$

The correlation coefficient (r) measures the strength and direction of the relationship between two variables. In the context of item analysis, it is used to determine the validity of test items by examining the correlation between individual item scores and total test scores. The value of the correlation coefficient ranges from -1.00 to +1.00. The interpretation is as follow.[30]

Table 1. The Interpretation of Validity

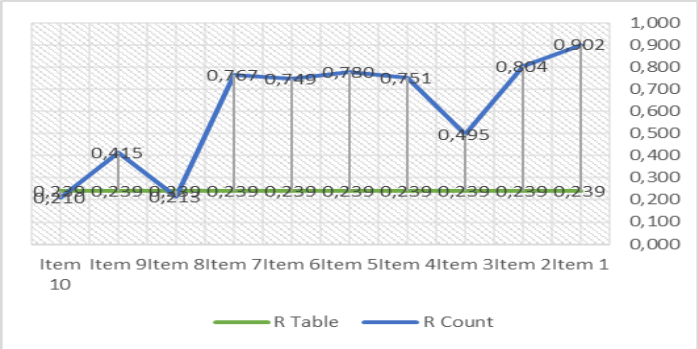
Interval Coefficient	Interpretation
0,800 – 1,00	Very High
0,600 – 0,800	High
0,400 – 0,600	Moderate
0,200 – 0,400	Low
0,00 – 0,200	Very Low

There are two main methods for interpreting the correlation coefficient: Interval Interpretation: Directly interpret the value of r using criteria above, and using r-Table (Product Moment Table): Compare to the correlation to the critical value in the r-table to determine the validity of each test item: If r count > r table, then the item is valid, if r count < r table, then the item is invalid.

The following are the result of correlation coefficients obtained from the Insya' exam questions using SPSS (Version 30). The test consisted of 10 essay questions and was administered to 68 students:

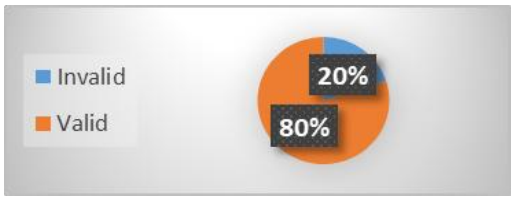
[29] Nur Rahmi Akbarini and Arun Anggrawal, "Validity and Reliability of Test Questions to Measure the Information Literacy Skills of Prospective Teacher Students," *Jurnal Eduscience* 11, no. 1 (May 3, 2024): 163–77, https://doi.org/10.36987/jes.v11i1.5640.

[30] Suharsimi Arikunto, *Dasar-Dasar Evaluasi Pendidikan* (Jakarta: PT Bumi Aksara, 2015).



Graphic 1. Validity Analysis of Insya' Exam Question

Additionally, the percentage distribution of the validity result is presented is presented in the graph below:



Graphic 2. Presentage of Validity Analysis of Insya' Exam Question

Base on the previous graphics, the validity analysis of the Insya' exam questions shows that eight (80%) are valid, specifically items number 1, 2, 3, 4, 5, 6, 7, and 9. Meanwhile, two questions (20%), namely item number 8 and 10, were invalid.

Reliability

Reliability refers to the consistency of the way used to provide data. An item is considered reliable if it consistently yields the same results when administered at different times under similar conditions.[31] The purpose of reliability is to ensure the consistency and stability of the measurement results. The reliability coefficient can be calculated using Cronbach's Alpha formula. Reliability is also defined as the extent to which test scores are free from measurement error. It reflects the degree of stability or internal consistency of an instrument in assessing certain concepts, such as through test-retest, alternate forms, and internal consistency like Cronbach's alpha.[32] The reability coefficient can be calculated using Cronbach's Alpha formula as follows:

[31] Muhammad Faris Imaduddin, Hikmah Maulani, and Ijlal Haziel Taufik, "Test the Validity and Reliability of Arabic Learning Questions," *Arabi : Journal of Arabic Studies* 7, no. 2 (December 21, 2022): 198–207, https://doi.org/10.24865/ajas.v7i2.523.

[32] Nor Hasnida Md Ghazali, "A Reliability and Validity of an Instrument to Evaluate the School-Based Assessment System: A Pilot Study," *International Journal of Evaluation and Research in Education (IJERE)* 5, no. 2 (June 1, 2016): 148, https://doi.org/10.11591/ijere.v5i2.4533.

$$\rho_a = \frac{N}{N-1}\frac{\acute{o}_A^2 - \sum \sigma_i^2}{\sigma_A^2}$$

Table 2. The Interpretation of Cronbach's Alpha Coefficient

Reliability Coefficient	Interpretation
≥ 0.90	Excellent
0.80 – 0.90	Very Good
0.70 – 0.80	Good
0.50 – 0.70	Acceptable
≤ 0.50	Poor

Table 3. Reliability Analysis of Insya Question Exam using SPSS (Version 30)

Reliability Statistics	
Cronbach's Alpha	N of Items
0,743	10

As shown in the table 3, the Cronbach's Alpha coefficient is 0.743, which which exceeds the acceptable threshold of 0.600. Therefore, it can be concluded that the Insya' exam questions demonstrate high reliability. This suggests that the test consistently measures students' writing abilities and produces stable and dependable results across different administrations.

Discrimination Power

The discrimination power refers to the ability of a test item to distinguish between students who have mastered the material and those who have not. In psychometric analysis, the discrimination index (DI) quantifies this capacity by comparing the performance of top- and bottom-scoring groups.[33] High DI values (typically ≥ 0.30) is considered effective for assessment purposes and contributes positively to overall test reliability.[34] Research in medical education shows that items with functioning distractors and moderate difficulty tend to exhibit higher discrimination value. Item writing flaws such as poorly structured stems or implausible distractors are consistently associated with lower discrimination indices.[35] Additionally, newer methods like the Generalized Discrimination Index (GDI) enable accurate measurement of item discrimination in non-binary (Likert-

[33] Karim, Sudiro, and Sakinah, "Utilizing Test Items Analysis to Examine the Level of Difficulty and Discriminating Power in a Teacher-Made Test."

[34] David Hope et al., "Using Differential Item Functioning to Evaluate Potential Bias in a High Stakes Postgraduate Knowledge Based Assessment," *BMC Medical Education* 18, no. 1 (December 2018): 64, https://doi.org/10.1186/s12909-018-1143-0.

[35] Sally A Santen et al., "'Sorry, It's My First Time!' Will Patients Consent to Medical Students Learning Procedures?," *Medical Education* 39, no. 4 (April 2005): 365–69, https://doi.org/10.1111/j.1365-2929.2005.02113.x.

type) scales, broadening application to diverse educational settings.[36] The following formula is used to calculate the discrimination power:

$$t \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\left(\frac{\sum X_1^2 + \sum X_2^2}{n\,(n-1)}\right)}}$$

Once the difference between the mean scores of the top and bottom groups is calculated using the formula above, the resulting value (r-count) is compared against the critical value (r-table). The interpretation is as follows:

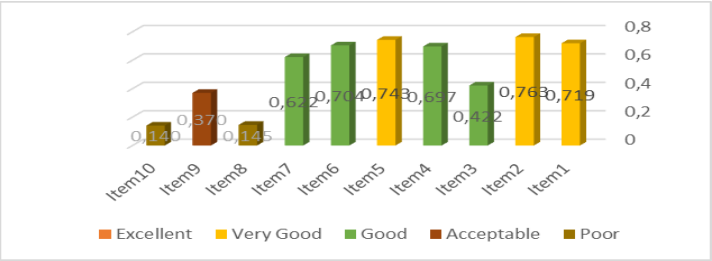
- If r-count > r-table, the item has high discrimination power, effectively distinguishing between high- and low-performing students.
- If r-count < r-table, the item has low or negative discrimination power, indicating it fails to differentiate student abilities effectively.

The higher the discrimination power, the better the item functions within the test. According to Suharsimi Arikunto, good discrimination power ranges from 0.40 to 0.70. The discrimination power can be classified as follows:

Table 4. The Interpretation of Discrimination Power

Discrimination power	Interpretation
0.71 – 1.00	Very Good
0.41 – 0.70	Good
0.21 – 0.40	Acceptable
0.00 – 0.20	Poor

The following are the result of Discrimination of Insya' exam questions. The test consisted of 10 essay questions and was administered to 68 students:



Graphic 3. Discrimination Power Scores of items in Insya' Exam Question

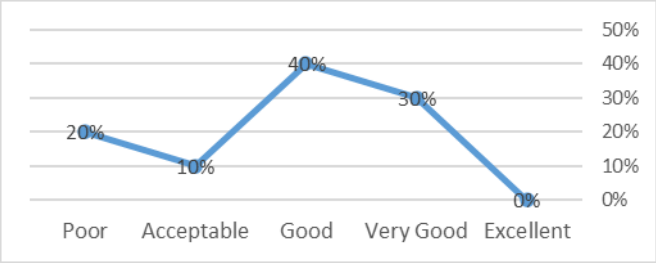
The interpretation of the discrimination power is summarized in Table 5 as follows:

[36] Jari Metsämuuronen, "Generalized Discrimination Index," *International Journal of Educational Methodology* 6, no. 2 (May 15, 2020): 237–58, https://doi.org/10.12973/ijem.6.2.237.

Table 5. The Interpretation of Discriminating Power

Item Quality	Item Number	Frequency	Recommendation
Excellent	-	0	Retein
Very Good	1, 2, 5,	3	Possible for improvement
Good	3, 4, 6, 7	4	Need to check/review
Acceptable	9	1	Discard or review in-deoth
Poor	8, 10	2	Definitely discard

The percentage distribution of discrimination power across these categories is presented in the following graph:



Graphics 4. Presentage of Discrimmination

Based on the table and graphics above, the discrimination power analysis for the Insya' exam questions shows varied results. Questions 1, 2, and 5 have very good discrimination (30%), effectively distinguishing high and low performers, and are recommended to be retained. Questions 3, 4, 6, and 7 are categorized as good (40%) but may benefit from minor revisions. Question 9 is acceptable (10%) and could be improved for better performance. However, Questions 8 and 10 show poor (20%) and are recommended for deletion. These results highlight which items are effective and which require improvement.

Difficulty level

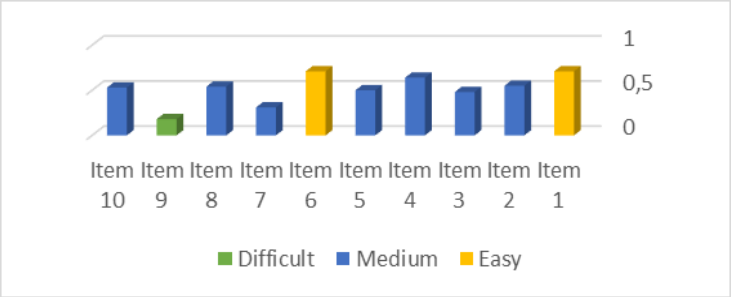
Difficulty level analysis for essay questions requires the students' scores and the maximum score for each question to calculate the difficulty index. This index indicates how easy or difficult an item is: a higher value suggests an easier item, while a lower value indicates a more difficult one. In language assessment contexts, understanding item difficulty is crucial because it affects learners' motivation, fairness, and reliability of the results. Language educators must ensure that questions reflect a balanced distribution of difficulty to differentiate proficiency levels effectively. According to empirical studies, tests with predominantly very easy or very difficult items tend to show reduced discrimination power and weak correlation with actual language ability. Hence, accurate difficulty analysis helps refine assessments for both summative and formative purposes, especially in second language writing and grammar evaluations. The difficulty level can be calculated using the following formula:

$$P = \frac{\Sigma X}{N \times X_{Max}}$$

Table 6. Mean and maximum scores for each item

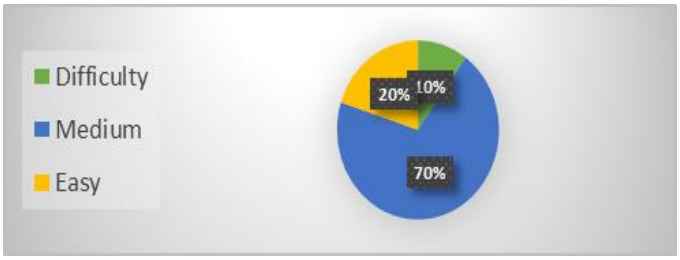
Item	Mean	Maximum	Difficulty Level	Interpretation
1	10.62	15	0.708	Easy
2	1.66	3	0.553	Medium
3	1.43	3	0.477	Medium
4	1.93	3	0.643	Medium
5	1.51	3	0.503	Medium
6	1.41	2	0.705	Easy
7	1.57	5	0.314	Medium
8	0.54	1	0.540	Medium
9	0.18	1	0.180	Difficult
10	0.53	1	0.530	Medium

Based on the data, out of 10 questions, one item (9) is categorized as Difficult, seven items (2, 3, 4, 5, 7, 8, and 10) are Medium, and two items (1 and 6) are Easy. This distribution suggests that the exam includes a balanced range of difficulty levels, which contributes to its overall validity and reliability. The distribution of difficulty levels is illustrated in the following graph:



Graphic 5. Difficulty Level Scores of items in Insya' Exam Question

The percentage distribution of difficulty levels across these categories is presented in the following graph:



Graphic 6. Presentage of Difficulty Level Analysis of Insya' Exam Question

Cognitive Level Analysis Using Revised Bloom's Taxonomy

To categorize the cognitive levels of the Insya' exam questions, the Revised Bloom's Taxonomy operational verbs were utilized, as outlined in the following table:

Table 7. Operational Verb Revised Bloom's Taxonomy (Level Cognitive)

Levels Cognitive	**Operational Verbs** (الأفعال السلوكية)	
Remembering (C1)	Find, cite, locate, recall, highlight, retrieve, search, define, describe, label, list, match, name, reproduce, state	اعثر على، اذكر شواهد (اشتشهد)، حدّد (اشتشهد)، حدّد موقع/مواقع أو مكان/أماكن، استذكر (المعلومات، المفاهيم، الأفكار، النظريات)، سلط الضوء على (النقط المهمة)، استرد (البيانات/المعلومات من ذاكرتك)، فتش، عرّف (بمعنى قم بتعريف)، أوص (اذكر أوصاف/توصيفات)، علّم (بعلامة فارقة)، عدّد (قم بتعداد، أدرج بجدول، جدول)، قابل بين، سمِّ/اذكر (أسماء، مؤشرات)، استرجع (المعلومات)
Understanding (C2)	Annotate, outline, compare, discuss, convert, explain, extend, generalize, exemplify, paraphrase, predict, summarize, translate, research, review, restate	قدّم شروحات على، اكتب الخطوط العريضة، قارن، ناقش، حوّل، اشرح، توسّع في، قم بتعميم، أعطِ مثالاً، عبّر بطريقة أخرى، استشرف، أوجز (اذكر بإيجاز)، انقل (الأفكار)/**ترجم**، اسقصِ علمياً، يراجع، أعد القول (في صيغة جديدة)
Applying (C3)	Apply, articulate, calculate, choose, complete, execute, dramatize, practice, share, change, illustrate, operate, teach, examine, classify, compute, demonstrate, discover, manipulate, prepare, produce, relate, show, solve, use, organize	طبّق، افصِح (في الكلام/النطق)، احسب، انتقِ، أكمل، نفّذ، ضع في قالب تمثيلي، تدرّب، شارك (في)، غيّر (قم بتغيير)، وضّح (بالمثال)، شغّل (قم بتشغيل)، درّس/علّم، تفحّص، صنّف، حَوسب، قدّم بالدليل، اكتشف، عالج، أعِد/قم بإعداد، انتج، اربط، اظهر، حلّ/قدّم حلاً، استخدم/استعمل، **رتّب**
Analyzing (C4)	Analyze, categorize, deduce, edit, investigate, reverse, select, separate, engineer, examine, establish, break down, conclude, diagram, deconstruct, differentiate, discriminate, distinguish, correlate, contrast	حلّل، افرز، استنبط، حرّر، تحقق، اعكس، اختر، افصل، قم بهندسة، تفحّص، أسّس، فصّل، استنتج، ارسم بياناً، فكك البنية، فرّق، ميّز، اظهر التمايز، اربط العلاقة المتبادلة بين، ناظر
Evaluating (C5)	Argue, assess, collaborate, critique, debate, evaluate, hypothesize, judge, moderate, recommend, reflect, test, verify, prioritize, rate, inspect, decide, measure, appraise, conclude, criticize, defend, discriminate, justify, support	قدّم حجاً، ثمّن، بيّن محاسن ومساوئ/انتقد بموضوعية، حاور، قيّم، اقترض، احكم، توسّط، زكِّ/قدّم تزكية، امتحن، رتّب حسب الأولوية، اعطِ معدل، فتّش عن، قرّر، قِس/قم بقياس، احكم على (الموقف)، استنتج من التقييم، إنتقد (بشأن المحاسن أو المساوئ)، دافع (بناء على التقييم)، اعرض بالتباين، برّر، ادعم

Levels Cognitive	Operational Verbs (الأفعال السلوكية)	
Creating (C6)	Integrate, intervene, model, negotiate, plan, progress, rearrange, formulate, construct, reinforce, revise, structure, substitute, validate, assemble, develop, draft, invent, produce, propose, publish, repurpose, upload, write, synthesize, categorize, combine, compile, compose, create, devise, design, generate, organize, reconstruct, reorganize, rewrite, tell, identify	ادمج، تدخّل، صمم نموذجاً، فاوض، خطّط، تقدمّ، أعد ترتيب، قم بتشكيل، قم ببناء، عزّز، نقّح/أعد النظر في، قم بهيكلة، استبدل، تأكد من صحة، قم بتجميع، طوّر، أكتب مسوّدة، اخترع، انتج، اقترح، انشر، أعد صياغة الغرض، حمّل/قم بتحميل، دوّن، قم بتركيب (الأفكار، النظريات)، افرز بطريقة تقييمة، اجمع بأسلوب تقييمي، قم بتجميع تراكمي، **ألّف**، اكتب بأسلوب إبداعي، ابتكر، صمّم، ولّد (أفكاراً)، نظّم، أعِد بناء، أعد تنظيم، قم بإعادة الكتابة، أخبر /قم بموافاة، عين باسلوب مبتكر

The operational verbs used in this analysis are derived from Bloom's Revised Taxonomy. These verbs are commonly applied in leading universities around the world, such as Harvard University and Stanford University. The list has been translated into Arabic and subsequently validated by Arabic language translation experts and educational specialists, particularly those with expertise in Bloom's Taxonomy.[37]

The following are the results of the Insya' exam analysis based on cognitive levels in Bloom's Revised Taxonomy, using the operational verbs listed in the previous table:

Table 8. Insya' Exam Analysis Based on Cognitive Levels in Bloom's Revised Taxonomy

No	Item Question	Verb	Levels Cognitive	HOTS	LOTS
1	ألف/ي مقالة إنشائية بإحدى الموضوعات الآتية بحيث أن لاتقل من ١٥ شطرا (الأدب فوق العلم/سيرة الحياة/ويل لطلاب الدنيا)	ألف	(C6) Creating	✓	
2	ترجم/ي هذه المقالة إلى اللغة العربية Ummahatul Mu'minin Mosque completed in 2009	ترجم	Understanding (C2)		✓
3	ترجم/ي هذه المقالة إلى اللغة العربية The mosque consists of two floors	ترجم	Understanding (C2)		✓
4	ترجم/ي هذه المقالة إلى اللغة العربية ! The mosque is now insufficient for the congregational prayers of all students.	ترجم	Understanding (C2)		✓
5	ترجم/ي هذه المقالة إلى اللغة العربية ! We need an expansion of the mosque	ترجم	Understanding (C2)		✓
6	ترجم/ي هذه المقالة إلى اللغة العربية ! Such is our hope, may Allah grant it	ترجم	Understanding (C2)		✓

[37] Ziad ElJishi et al., "Translating Bloom's Taxonomy Action Verb List into Arabic for Teacher Preparation Programs: Challenges/Problems and Solutions," *International Journal of Education and Literacy Studies* 12, no. 1 (January 27, 2024): 295–303, https://doi.org/10.7575/aiac.ijels.v.12n.1p.295.

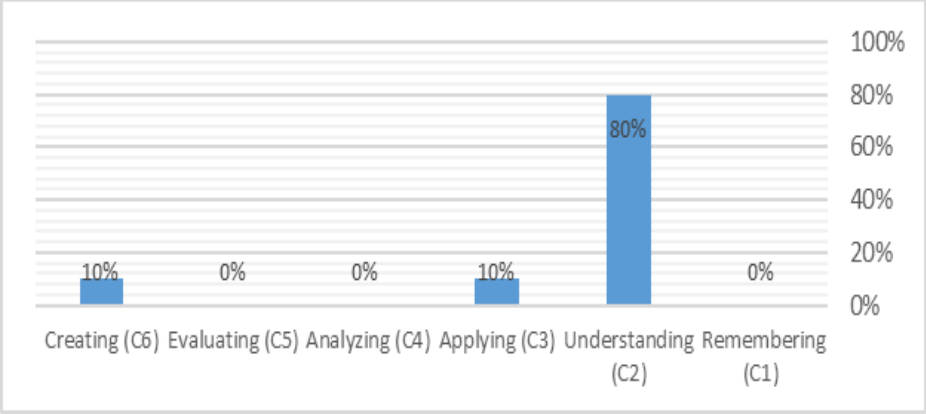
No	Item Question	Verb	Levels Cognitive	HOTS	LOTS
7	رتّب/ي العبارات الآتية حتى تكون مقالة مفهومة . فَتَبِيْضُ كُلَّ يَوْمٍ بَيْضَةً مِنَ الذَّهَبِ . فَفَرِحَ الرَّجُلُ بِهَا . كَانَ لِرَجُلٍ وَزَّةٍ جَمِيْلَةٍ . وَهِيَ تَبِيْضُ بَيْضَةً مِنَ الذَّهَبِ . فَأَ طْعَمُهَا طَعَامًا كَافِيًا لِتَبِيْضَ كُلَّ يَوْمٍ	رتّب	Applying (C3)		✓
8	ترجم/ي هذه المفردات إلى اللغة الاندونيسيا (مِنْظَرٌ)	ترجم	Understanding (C2)		✓
9	ترجم/ي هذه المفردات إلى اللغة الاندونيسيا (مَنْظَرٌ)	ترجم	Understanding (C2)		✓
10	ترجم/ي هذه المفردات إلى اللغة الاندونيسيا (زُرْمَةٌ)	ترجم	Understanding (C2)		✓

Based on the previous data, regarding the percentage distribution the question based on cognitive levels is as follows:

Table 9. Presentage of Cognitive Levels in The Insya' Exam Question

Level	Verbs	Question Numbers	Frequency	Presentage
Remembering	-	-	-	0%
Understanding	ترجم	2, 3, 4, 5, 6, 8, 9, 10	8	80%
Applying	رتب	7	1	10%
Analyzing	–	-	-	0%
Evaluating	–	-	-	0%
Creating	ألف	1	1	10%

These findings are further illustrated in the following graphs:



Graphic 7. Presentage of Cognitive Levels in The Insya' Exam Question

Conclusion

This study analyzed the item quality of Insya' assessment questions based on validity, reliability, discrimination index, difficulty level, and Revised Bloom's Taxonomy. The findings revealed that 80% of the questions were valid while 20% were invalid, with a reliability coefficient of 0.743, indicating high reliability. Additionally, the discrimination analysis showed 30% of the questions had very good discrimination, 40% good, 10% acceptable, and 20% poor. In terms of difficulty, 10% were difficult, 70% moderate, and 20% easy. Regarding cognitive level classification, 80% of the items targeted Understanding (C2), 10% Applying (C3), and 10% Creating (C6). This means that only 10% of the items measured Higher-Order Thinking Skills (HOTS), while the remaining 90% focused on Lower-Order Thinking Skills (LOTS).

These findings contributes to the field of educational assessment by highlighting the need for better alignment between test items and cognitive skill levels, particularly in Insya' assessments. The study demonstrates the importance of incorporating HOTS to enhance critical and creative thinking in students. Additionally, the research emphasizes the need for improving discrimination power to create more effective and fair assessments. This study contributes to modern educational practices by providing empirical evidence on the current state of item quality in Insya' assessments, emphasizing the importance of validity and reliability in evaluating cognitive skills. Educational practitioners and test developers can apply these findings to enhance the quality of Arabic language assessments.

Future research should explore innovative question designs to better measure HOTS and investigate their impact on student performance. Additionally, further studies could expand the sample size and include different educational contexts to generalize the findings. Ongoing research is also recommended to monitor improvements and refine assessment tools to ensure accurate evaluation of students' cognitive abilities.

Acknowledgment

The author would like to express sincere gratitude to all participants, who have contributed their time, knowledge, and support in the research process. A special thanks is extended to the Islamic Senior High School of Darussalam Rajapolah Tasikmalaya for their invaluable cooperation during data collection. We also appreciate the constructive feedback and suggestions from academic mentors and peer reviewers, which have significantly improved the quality of this study. Finally, we hope and pray that this research will benefit the academic community and contribute positively to Arabic language education. We wish success for future researchers, with hopes that their work will surpass what has been achieved in this study, leading to greater innovations and insights.

Author Contribution Statement

CSR conceptualized the study and supervised the research process. FM conducted the literature review and data analysis. FS contributed to the methodology and interpretation of results. ZHS assisted with data collection and manuscript editing. All authors contributed to the writing and approved the final version of the manuscript.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

Aflisia, Noza, Kasmantoni, and Yunika Febri Yanti. "Writing Errors: A Study of Students Linguistic Errors in Insya Muwajjah Learning." *An Nabighoh* 26, no. 1 (2024): 79–96. https://doi.org/10.32332/annabighoh.v26i1.79-96.

Akbarini, Nur Rahmi, and Arun Anggrawal. "Validity and Reliability of Test Questions to Measure the Information Literacy Skills of Prospective Teacher Students." *Jurnal Eduscience* 11, no. 1 (May 3, 2024): 163–77. https://doi.org/10.36987/jes.v11i1.5640.

Althory, Abdallah Ali Ali. "Linguistic Features Acquired by a Language of A Speaking Society." *Journal of Arabic Language Sciences and Literature* 1, no. 4 (September 30, 2022): 62–77. https://doi.org/10.26389/AJSRP.R150522.

Arikunto, Suharsimi. *Dasar-Dasar Evaluasi Pendidikan*. Jakarta: PT Bumi Aksara, 2015.

Ashraf, Z. A., and Jaseem K. "Classical and Modern Methods in Item Analysis of Test Tools." *International Journal of Research and Review* 7, no. 5 (2020). https://www.ijrrjournal.com/IJRR_Vol.7_Issue.5_May2020/Abstract_IJRR0058.html.

Ayu, Nidya, Suharno, and Kartika Chrysti Suryandari. "Exploration: Creative Thinking Skills in Writing Essays Media-Based Image Series." *International Journal of Elementary Education* 7, no. 1 (February 16, 2023): 1–7. https://doi.org/10.23887/ijee.v7i1.54095.

Cecep, Cecep Sobar Rochmat, Amalia Nur Afasa Amalia, Abdul Hafidz Bin Zaid Abdul Hafidz, Moh Alwi Yusron Alwi, and Sania Karimah Liputo Sania. "The Analysis of Experiential Learning Method of Dale's Cone Experience Model in Improving the Effectiveness of Arabic Language Learning." *Izdihar : Journal of Arabic Language Teaching, Linguistics, and Literature* 7, no. 1 (April 30, 2024). https://doi.org/10.22219/jiz.v7i1.30823.

Das, Syaamantak, Shyamal Kumar Das Mandal, and Anupam Basu. "Classification of Action Verbs of Bloom's Taxonomy Cognitive Domain: An Empirical Study."

Journal of Education 202, no. 4 (October 2022): 554–66. https://doi.org/10.1177/00220574211002199.

ElJishi, Ziad, Faten S. M. Abdel-Hameed, Ahmad Khuddro, and Salman Yusuf Zayed. "Translating Bloom's Taxonomy Action Verb List into Arabic for Teacher Preparation Programs: Challenges/Problems and Solutions." *International Journal of Education and Literacy Studies* 12, no. 1 (January 27, 2024): 295–303. https://doi.org/10.7575/aiac.ijels.v.12n.1p.295.

Elthia, Memy Wardani, and Rosidah Alawiyah. "An Analysis of the Quality Test Items Used in the Final Semester Assessment at MTs Nurul Yaqin Gili Air." *Jurnal Test 2* 6, no. 1 (November 18, 2023). https://jonedu.org/index.php/joe/article/view/4572.

Erdiana, Nira, and Sulastri Panjaitan. "How Is HOTS Integrated into the Indonesian High School English Textbook?" *Studies in English Language and Education* 10, no. 1 (January 31, 2023): 60–77. https://doi.org/10.24815/siele.v10i1.26052.

Hope, David, Karen Adamson, I. C. McManus, Liliana Chis, and Andrew Elder. "Using Differential Item Functioning to Evaluate Potential Bias in a High Stakes Postgraduate Knowledge Based Assessment." *BMC Medical Education* 18, no. 1 (December 2018): 64. https://doi.org/10.1186/s12909-018-1143-0.

Hrich, Najoua. "Artificial Intelligence Item Analysis Tool for Educational Assessment: Case of Large-Scale Competitive Exams." *International Journal of Information and Education Technology* 14, no. 6 (2024): 822–27. https://doi.org/10.18178/ijiet.2024.14.6.2107.

Imaduddin, Muhammad Faris, Hikmah Maulani, and Ijlal Haziel Taufik. "Test the Validity and Reliability of Arabic Learning Questions." *Arabi : Journal of Arabic Studies* 7, no. 2 (December 21, 2022): 198–207. https://doi.org/10.24865/ajas.v7i2.523.

Karim, Sayit Abdul, Suryo Sudiro, and Syarifah Sakinah. "Utilizing Test Items Analysis to Examine the Level of Difficulty and Discriminating Power in a Teacher-Made Test." *EduLite: Journal of English Education, Literature and Culture* 6, no. 2 (2021): 256. https://doi.org/10.30659/e.6.2.256-269.

Khalid, Nor Hafizi Mohd, Ibnatul Jalilah Yusof, Adibah Abdul Latif, and Md Daud Md Jani. "Psychometric Properties of Teacher Classroom Assessment Literacy Instrument Using Rasch Model Analysis." *International Journal of Evaluation and Research in Education (IJERE)* 12, no. 2 (June 1, 2023): 638. https://doi.org/10.11591/ijere.v12i2.23607.

Kumar, Amit, Dinesh Singh, and M. S. Dhankhar. "Analysis of Bloom Taxonomy-Based Examination Data Using Data Mining." *International Journal of Intelligent Systems and Applications in Engineering* 12, no. 4s (2024). https://www.ijisae.org/index.php/IJISAE/article/view/3860.

Kumar, Dharmendra, Raksha Jaipurkar, Atul Shekhar, Gaurav Sikri, and V. Srinivas. "Item Analysis of Multiple Choice Questions: A Quality Assurance Test for an Assessment Tool." *Medical Journal Armed Forces India* 77 (February 2021): S85–89. https://doi.org/10.1016/j.mjafi.2020.11.007.

Mahmudi, Ihwan, Fitri Masturoh, and Wiwiek Dwi Febrianti. "Higher Order Thinking Skills (HOTS) Based Assessment: A Proposed Model for Arabic Learning." *Shibghoh: Prosiding Ilmu Kependidikan UNIDA Gontor* 1 (2023). https://ejournal.unida.gontor.ac.id/index.php/shibghoh/article/view/10195 .

Mahmudi, Ihwan, Afni Nurwardah, Siti Nikmatul Rochma, and Agung Nurcholis. "Item Analysis Of Arabic Language Examination." *Ijaz Arabi Journal of Arabic Learning* 6, no. 3 (October 22, 2023). https://doi.org/10.18860/ijazarabi.v6i3.19821.

Mahroof, Ameema, and Muhammad Saeed. "Evaluation of Question Papers by Board of Intermediate and Secondary Education Using Item Analysis and Blooms Taxonomy." *Bulletin of Education and Research* 43, no. 3 (2021). https://pu.edu.pk/images/journal/ier/PDF-FILES/5_43_3_21.pdf.

Marar, Sumayyia, Muaawia A. Hamza, Mohsen Ayyash, and Amani Abu-Shaheen. "Development and Validation of an Instrument to Assess the Knowledge and Perceptions of Predatory Journals." *Heliyon* 9, no. 11 (November 2023): e22270. https://doi.org/10.1016/j.heliyon.2023.e22270.

Md Ghazali, Nor Hasnida. "A Reliability and Validity of an Instrument to Evaluate the School-Based Assessment System: A Pilot Study." *International Journal of Evaluation and Research in Education (IJERE)* 5, no. 2 (June 1, 2016): 148. https://doi.org/10.11591/ijere.v5i2.4533.

Metsämuuronen, Jari. "Generalized Discrimination Index." *International Journal of Educational Methodology* 6, no. 2 (May 15, 2020): 237–58. https://doi.org/10.12973/ijem.6.2.237.

Mosier, Charles I. "A Note on Item Analysis and the Criterion of Internal Consistency." *Psychometrika* 1, no. 4 (December 1936): 275–82. https://doi.org/10.1007/BF02287879.

Muassomah. "From Global Language Use to Local Meanings: Arabic to Indonesian Absorption." *IAS Journal of Localities* 1, no. 1 (May 24, 2023): 16–29. https://doi.org/10.62033/iasjol.v1i1.10.

Rochmat, Cecep Sobar, Syifa Rizki Sholihah, and Shofia Niswah Qonita. "Forming Critical Character With Higher Order Thinking Skill (HOTS) Based Learning Assessment in Islamic Religious Education Subjects." *Educan : Jurnal Pendidikan Islam* 6, no. 2 (August 2, 2022): 236. https://doi.org/10.21111/educan.v6i2.8131.

Rochmat, Cecep Sobar, Angelica Silfana Prisca Yoranita Yoranita, Mispullah Prihatini, and Bagus Asri Wibawa. "The Quality of Education from Islamic

Perspective Analysis of The Merdeka Belajar Curriculum in Facing The Society 5.0 Era." *Jurnal Tarbiyatuna* 14, no. 1 (June 30, 2023): 75–93. https://doi.org/10.31603/tarbiyatuna.v14i1.8633.

Santen, Sally A, Robin R Hemphill, Cindy M Spanier, and Nicholas D Fletcher. "'Sorry, It's My First Time!' Will Patients Consent to Medical Students Learning Procedures?" *Medical Education* 39, no. 4 (April 2005): 365–69. https://doi.org/10.1111/j.1365-2929.2005.02113.x.

Setiabudi, Agung, Mulyadi Mulyadi, and Hilda Puspita. "An Analysis of Validity and Reliability of a Teacher-Made Test." *Journal of English Education and Teaching* 3, no. 4 (December 6, 2019): 522–32. https://doi.org/10.33369/jeet.3.4.522-532.

Suryadinata, Achmad Naufal Irsyadi, Zarkasi, Moh Hasyim Rosyidi, and Yuana Dwi Agustina. "The Challenge of Mastering Arabic Language and Its Relation to the Ability to Read Kitab Kuning among Students in Tapal Kuda Islamic Boarding School." *RETORIKA: Jurnal Ilmu Bahasa* 10, no. 3 (2024). https://ejournal.warmadewa.ac.id/index.php/jret/article/view/10488.

Susanto, Susanto, Ayu Desrani, Apri Wardana Ritonga, and Yohan Rubiyantoro. "Improving Students' Creative Thinking In Learning Arabic Through HOTS Based Project Based Learning Model." *An Nabighoh* 24, no. 1 (June 28, 2022): 1. https://doi.org/10.32332/an-nabighoh.v24i1.3924.

Yoo, Hanwook [Henry], and Ronald K. Hambleton. "Digital Module 08: Foundations of Operational Item Analysis Https://Ncme.Elevate.Commpartners.Com." *Educational Measurement: Issues and Practice* 38, no. 3 (September 2019): 116–17. https://doi.org/10.1111/emip.12289.

Zakkiyah, Maulia Yasminah, Nurriya Maghfirah Fidyahwati, Ahmad Tarajjil Ma'suq, and Novita Anggraini. "Assessment Design and Analysis of Arabic Reading Skills Instructional Materials." *IJIE International Journal of Islamic Education* 3, no. 1 (June 28, 2024): 31–46. https://doi.org/10.35719/ijie.v3i1.2000.

Zulaeha, Zulaeha, and Musdalifah Musdalifah. "Strengthening Students' Arabic Writing Skills Through Insya' Muwajjah in Higher Education." *AL-ISHLAH: Jurnal Pendidikan* 14, no. 1 (February 1, 2022): 591–602. https://doi.org/10.35445/alishlah.v14i1.1103.

Biography of Authors

Cecep Sobar Rochmat is a lecturer at the University of Darussalam Gontor, Indonesia. His academic interests include Islamic Religious Education, Philosophy of Education, Holistic Education, and the Pesantren-based University model. He earned his Bachelor's degree from the Institute for Islamic Studies Darussalam (ISID) Gontor in 2006, his Master's degree from STAIN Datokarama Palu in 2013, and his Doctorate from the State Institute for Islamic Studies (IAIN) Palu in 2020. He is actively engaged in teaching and research, focusing on curriculum development in Islamic education and the integration of pesantren (Islamic boarding school) values into higher education. In addition to research, he mentors students and frequently speaks at academic forums related to Islamic education. He can be contacted via email at: cecep.rochmat@unida.gontor.ac.id.

Fitri Masturoh is a postgraduate student in the Arabic Language Education program at the University of Darussalam Gontor, Indonesia. She holds a Bachelor's degree in Islamic Religious Education froms IAILM Suryalaya Tasikmalaya and is currently pursuing her Master's degree in Arabic Language Education at UNIDA Gontor. Her academic interests include Arabic language teaching and Islamic education. She actively participates in academic conferences and is involved in research focusing on language pedagogy and curriculum development in Islamic educational contexts. she can be contacted via email at: fitrimasturoh@unida.gontor.ac.id.

Fairuz Subakir is a senior lecturer in the Arabic Language Education postgraduate program at the University of Darussalam Gontor, Indonesia. He earned his Bachelor's degree in Da'wah and Islamic Studies from the Islamic University of Madinah, Saudi Arabia, and continued his postgraduate studies in Arabic Language at Punjab University, Pakistan, where he obtained both his Master's and Doctoral degrees. His expertise lies in Arabic language education and Islamic communication. With extensive experience in teaching and research, he plays an active role in academic development within the field of Arabic language studies. He can be contacted via email at: fairuz.subakir@unida.gontor.ac.id.

Zhulfa Halimatus Sadiah is an undergraduate student at Fatih Sultan Mehmet Vakıf University, Turkey, majoring in Islamic Studies. Her academic interests lie in the fields of Islamic education history and Arabic language. She has participated in several academic conferences and research forums focusing on Arabic language learning and the intellectual heritage of Islamic education. She aspires to contribute to the advancement of Arabic pedagogy by integrating historical insights with contemporary educational needs. He can be contacted via email at: zhulfa.sadiah@stu.fsm.edu.tr.

Please cite this article as Cecep Sobar Rochmat, Fitri Masturoh, Fairuz Subakir, Zhulfa Halimatus Sadiah "Item Analysis of Insya' Assessment Based on Bloom's Taxonomy at Islamic Senior High School Darussalam Rajapolah Tasikmalaya." *An Nabighoh* 27, No. 1 (2025): 201-222. DOI: https://doi.org/10.32332/an-nabighoh.v27i1.201-222