



Universitas Muhammadiyah Malang, East Java, Indonesia

Izdihar : Journal of Arabic Language Teaching, Linguistics, and Literature

p-ISSN: 2622-738X, e-ISSN: 2622-7371 // Vol.6 No.3 Desember 2023, pp. 279-292



<https://doi.org/10.22219/iz.v6i3.27441>



<http://ejournal.umm.ac.id/index.php/izdihar/index>



izdihar.jurnalpba@umm.ac.id

Impact of Portfolio Authentic Assessment on Students' Learning Outcomes of Nahwu at Modern Islamic Boarding School

Ihwan Mahmudi^{a,1}, Aisyah Imani^{a,2}, Chevira Yasin Indriyanti^{a,3}, Rohani Pooteh^{b,4}

^aUniversitas Darussalam Gontor, Indonesia

^bPrince of Songkla University, Thailand

¹ihwanm@unida.gontor.ac.id, ²aisyahimani19@gmail.com,

³cheviraputriindriyanti26@student.pba.unida.gontor.ac.id, ⁴rohani.p@psu.ac.th

ARTICLE INFO

Article History:

Received:26/06/2023

Revised:05/12/2023

Accepted: 28/12/2023

Published:31/12/2023

*Corresponding

Author:

Name: Ihwan

Mahmudi

Email:

ihwanm@unida.gontor.ac.id

ABSTRACT

The study aimed to investigate the potential beneficial impact of portfolio assessment on the learning outcomes of Nahwu lessons of fourth-grade students at the Modern Islamic Boarding School of Al-Iklash Female in Kuningan, West Java. The study used an experimental methodology and a quantitative research approach. The research sample included all fourth graders. The research was carried out during the odd semester of academic year 2022. A quasi-experimental design used to evaluate the effectiveness of portfolio assessment, with the experimental class implementing portfolio assessment while the control class continued to use the standard evaluation method. The learning outcomes of the students in the control class produced an average score of 71.67. Comparing the experimental class' score, which used portfolio assessment, it become clear that the students' comprehension of the Nahwu material had significantly improved. These results showed that portfolio evaluations have a chance to improve fourth-grade students' learning outcomes in Nahwu.

Copyright © 2023, Mahmudi et al

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



Keyword

Nahwu Learning, Nahwu Science, Portfolio Authentic Assessment

مستخلص البحث

مع التركيز بشكل خاص على الصف الرابع تربية المعلمات الإسلامية في المدرسة الداخلية الإسلامية الحديثة في الإخلاص للإناث في كونيبنجان ، جاوة الغربية ، تهدف الدراسة إلى التحقيق في الأثر المفيد المحتمل لتقييم المحفظة على نتائج التعلم لدروس النحو. طلاب الصف الرابع. النتيجة أن 4 تربية المعلمات الإسلامية لها زخم لهذا التحقيق. تم إجراء المجموعة المستهدفة للدراسة ، والتي شملت جميع طلاب الصف الرابع ، خلال الفصل الدراسي الفردي من العام الدراسي 2022. تم تضمين جميع الطلاب في الفصل 4 تربية المعلمات الإسلامية في عينة البحث بفضل طريقة أخذ العينات المشبعة التي تم استخدامها. تم استخدام تصميم شبه تجريبي لتقييم فاعلية تقييم المحفظة ، مع تنفيذ تقييم المحفظة من الفئة التجريبية بينما استمرت فئة الضبط في استخدام طريقة التقييم القياسية. أنتجت نتائج التعلم للطلاب في فئة الضبط متوسط درجة 71.67. عند مقارنة هذه النتيجة مع الدرجة التجريبية ، التي استخدمت تقييم المحفظة ، يتضح أن فهم التلاميذ للمادة الناهوية قد تحسن بشكل ملحوظ. تُظهر هذه النتائج أن تقييمات الحافظة لديها فرصة لتحسين نتائج التعلم لطلاب الصف الرابع في النحو.

تقييم أصيلة؛ مَلَفّ؛ نتائج التعلم النحو

كلمات أساسية

Please cite this article as Mahmudi, I., Imani, A., Indriyanti, C.Y., Pooteh R. (2023). Impact of Portfolio Authentic Assessment on Students' Learning Outcomes of Nahwu at Modern Islamic Boarding School.

Izdihar : Journal of Arabic Language Teaching, Linguistics, and Literature, 6(3), 279-292. DOI:

<https://doi.org/10.22219/iz.v6i3.27441>

INTRODUCTION

Planning, implementation, and assessment are the three essential phases of learning activities (Nurtanto et al., 2021). The assessment step is essential since it assesses how well learning is going and helps to improve the creation of education programs. As a result, these three procedures are crucial for attaining the best learning results. The 2013 curriculum's implementation includes a number of enhancements, particularly in the area of assessment. The curriculum has a strong emphasis on the three categories of learning outcomes identified by Bloom's taxonomy: cognitive (knowledge), emotional (attitude), and psychomotor (skills) (Aheisibwe et al., 2021). The necessity of educators to evaluate students' learning outcomes to track progress and continuously improve them is highlighted by this alignment with Law No. 20 of 2003 Article 58, paragraph 1 (Shaturaev, 2021). As a result, teachers must carefully create tests that take into account students' abilities because the learning process and its result are intertwined. Learning outcomes also typically tend to be more ideal when the learning process is optimized (Mahmudi, 2018).

Short term memory skills are often the product of education that only emphasizes understanding content (Mahmudi & Nadhifah, 2020). This method, therefore, makes it more difficult for students to use what they learn in the scenarios they will face in the real world. As a result, thorough and realistic evaluation is crucial, and teachers should take here factor into account. Realistic assessment was intended to correct traditional evaluations that purposefully ignore the work of productive learners. According to (Hajaroh & Adawiyah, 2018), evaluation entails a full examination of students emotive, cognitive, and psychomotor domains, spanning input, process, and output stages. The aim of learning Arabic for students is to become proficient users of the language (Brosh, 2019). Language includes the sounds that people use to express their meaning (Siregar, 2021), so to achieve this goal, it is very important to use realistic assessments in Arabic language training (Sari & Mahmudi, 2019). Realistic evaluations ensure that students can use their Arabic knowledge in real-world situations.

There are many difficulties in teaching Arabic, which can be divided into linguistic problems (such as language elements and linguistic abilities) and non-linguistic difficulties (Sa'adah, 2019). Non-linguistic difficulties include an overemphasis on evaluation criteria within the cognitive domain, a focus on theoretical elements and application of materials rather than language skills, and instructional materials that do not correspond with real-life students' needs (Albantani, 2018). The science of Nahwu is one of the most important topics in the study of Arabic (Muhammad et al., 2020). In evaluating students' speech and writing, Ahmad Abdel-Qoder raises worries about the students deficiency at all

levels of *Nahwu* science proficiency is to aid students in avoiding linguistic mistakes (Sa'adah, 2019).

The aforementioned problem could be exacerbated by a lack of varied assessments, such portfolio evaluation. The assessment procedures mostly center on cognitive test that are administered after learning and rely on pre-existing inquiries from the textbook (He Xinyu et al., 2020). This constrained strategy might affect how well students learn. Additional variables that may affect the outcomes include attitudes, study habits, learning setting, motivation, interests, and abilities, as well as and perhaps most significantly the teacher's capacity to assess student learning (Mahmudi & Shofro, 2019). The Modern Islamic Boarding School's fourth-grade student learning about *Nahwu*, and this research found that teachers frequently use routine evaluation techniques that are useful and affordable. This technique primarily used written exams as both formative and summative assessments without any alterations or variations (Ismail et al., 2022).

This research intended to experiment with authentic assessment portfolios in the *Nahwu* class of fourth-grade at the Modern Islamic Boarding School of *Al-Ikhlash* Female in order to address the problem of ineffective assessment techniques. According to (Sari & Mahmudi, 2019), authentic assessment methods that help students demonstrate their abilities by finishing tasks, solving problems, or showcasing their knowledge in made-up real-world situations. By implementing authentic assessment portfolios, it is hoped to give students a more thorough and useful evaluation that is in line with real-life circumstances and improves their learning results.

An effective and efficient alternative learning approach has been established as the portfolio authentic assessment (Safriyani, 2012). This concept seeks to involve students in all elements of learning-cognitive, emotional, and psychomotor in order to encourage independent thought, active involvement, and creativity. Additionally, the portfolio model assessment makes sure to provide real information about students' language abilities (Cabrera-Solano, 2020; Sulistyono et al., 2020), enabling comprehensive study of their language growth. Suneetha adds that for educational reforms to be most effective, the curriculum must be improved and real evaluation procedures must be included (Nguyen & Phan, 2020). This suggested that students diverse learning goals ought to be reflective of real live situations (Nurgiyantoro, 2011).

Several studies have concluded that evaluating the practice of Arabic language lessons in schools still experiences difficulties in the practice assessment process (Arifianto et al., 2022; Wanti et al., 2022; Yarun & Markus, 2023), especially with regard to the assessment tools used, because in practical tests teachers have to assess students' performance as a whole so they need specific

test tools that function to avoid subjectivity in assessment (Syafi'i et al., 2023). This is where Authentic Portfolio Assessment (PPA) comes in, offering an innovative approach that focuses on real-world tasks, reflection, and collaboration (Kaukab et al., 2021). Other research also shows that PPA can enhance *Nahwu* learning by encouraging deeper understanding, active engagement, and critical thinking skills (Yarun & Markus, 2023). Despite challenges such as teacher training and limited resources, the future looks bright with continued research and support. PPA has the potential to revolutionize *Nahwu* learning in Islamic boarding schools, empowering students not only to master Arabic, but also to appreciate its beauty and relevance in their life. Therefore, this study aims to investigate the potential beneficial impact of Authentic Portfolio Assessment (PPA) on the learning outcomes of *Nahwu* lessons, especially at the modern Islamic Boarding School.

METHOD

The study was carried out at the Modern Islamic Boarding School of *Al-Ikhlash* Female in Ciomas, Kuningan, West Java. The study used an experimental methodology and a quantitative research approach to examine the association between variables (Mohajan, 2020). The use of portfolio authentic assessment served as the study's independent variable (x), and the students learning outcomes following the evaluation process utilizing the portfolio authentic assessment served as the study's dependent variable (y) (Monib et al., 2020). The study's goal was to ascertain how using portfolio authentic assessment affected students learning outcomes.

A quasi-experimental design was used by this research as a modified form of an actual experimental design (Révész, 2019). The ability of this design to control outside variables that might affect how the experiment is carried out is constrained despite the presence of a control group. The difficulties in finding a control group for the study were a major factor in the choice to utilize a quasi-experimental design (Sugiyono, 2018). The Non-equivalent Control Group Design was used as the research methodology in this study. With the exception that participants are not randomly assigned to the experimental and control groups, it is comparable to a pre-test and post-test design (Sugiyono, 2018). Students from the Modern Islamic Boarding School of *Al-Ikhlash* Female for the academic year 2022 participated in this study as fourth-grade students. Totaled 30 students which divided into 4th B and 4th C classes. All fourth-grade students were chosen by researchers as the relatively small sample size for the study. These students were split into two groups: 15 were placed in the experimental class, which used

the authentic assessment portfolio, and 15 were placed in the control group, which did not.

RESULTS & DISCUSSION

For the purpose of the describing the information gleaned from the experimental investigation, the researchers used descriptive analysis. Presenting different statistics, such as minimum, maximum, mean (average), and standard deviation, is part of descriptive analysis. The researchers used the statistical tool SPSS 26 to carry out this analysis. SPSS 26 enables statistical calculations and data processing, making it easier to provide descriptive statistics to sum up and comprehend the experimental data.

Table 1. Descriptive Test Result

	N	Range	Min.	Max	Sum	Mean	Std. Deviation	Variance
						Statistic	Std.Eror	
PreTest Control	15	30	50	80	1005	67.00	2.430	9.411
PostTest Control	15	70	30	100	1075	71.67	6.204	24.029
PreTest Experiment	15	75	15	90	930	62.00	6.168	23.890
PostTest Experiment	15	50	50	100	1350	90.00	3.684	14.268
Valid N (listwise)	15							

Based on the descriptive analysis table with SPSS 26 in the previous paragraph, that N is the number of samples or research population. Additionally, it has been noted that there are 15 students in the experimental class, 15 students in the control class, and a proper sample size of 15 students. The typical result in the pre-test for the control class was (range) is 30, and the lowest result (minimum) is 50, and the highest (maximum) is 80, and the mode value (mean) is 67.00, and for the mean standard deviation (std. Deviation) worth 9,411, and the total value (sum) 1005. Where for the controller class in the post-test, the average result (range) is 70, and the lowest (minimum) result is 30, and the

highest (maximum) is 100, and the mode value (mean) is 71.67, and for mean standard deviation (std. Deviation) is worth 24.049, and the sum of the values (sum) is 1075.

Then, the average result (range) for the experimental class in the pre-test is 75, and the lowest (minimum) result is 15, and the highest (maximum) is 90, and the mode value (mean) is 62.00, and for the mean standard deviation (std. Deviation) is worth 23,890, and the sum of the values (sum) is 930. However, in the post-test, the average score for the experimental class was (range) is 50, and the lowest (minimum) is 50, and the highest (maximum) is 100, and the mode value (mean) is 90.00, and for the mean standard deviation (std. Deviation) is 14,268, and the total value (sum) is 1350.

This research constructed the following histogram based on the arrangement of the descriptive test table:

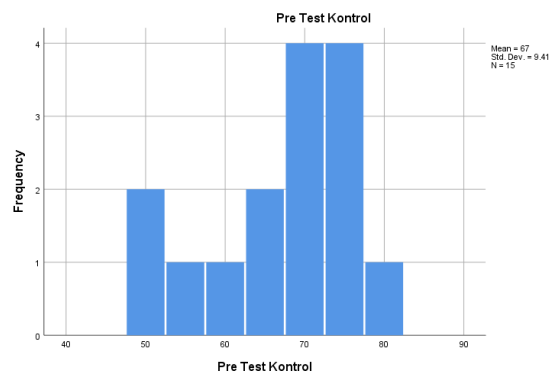


Figure 1. Histogram of Pre-Test Control Class

The [Figure 1](#) illustrates the consistent mastery of *Nahwu* subject by the students in the control class, as indicated by their pre-test scores. Following the description of the descriptive test conducted for the post-test of the control class:

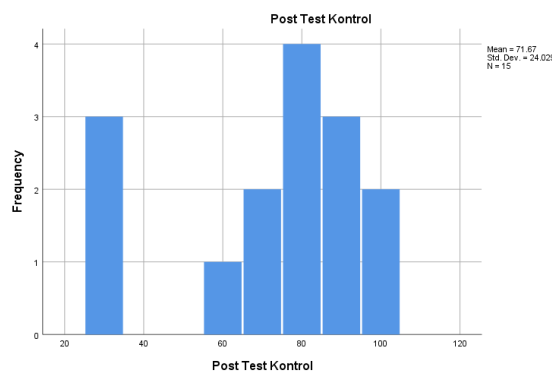


Figure 2. Histogram of Post-Test Control Class

The [Figure 2](#) demonstrates the ongoing correlation between the post-test results of the control class students and their *Nahwu* subject learning outcomes. Following the description of the descriptive test results for the control class post-

test, the researchers present the descriptive test for the experimental class pre-test as follows:

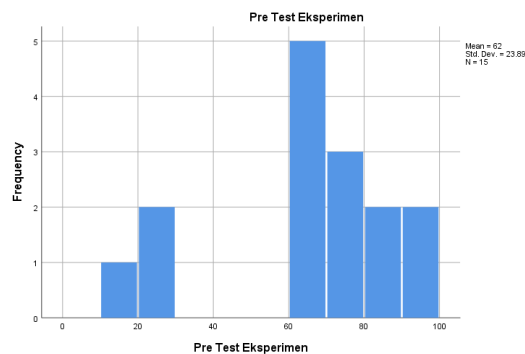


Figure 3. Histogram of Pre-Test Experiment Class

The Figure 3 shows that the experimental class students' learning outcomes for the Nahwu subject were consistent with the pre-test findings. The researchers then went on to discuss the post-test descriptive results for the experimental class in the following manner:

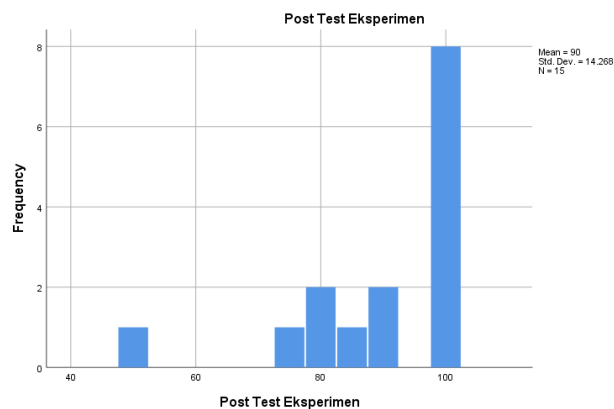


Figure 4. Histogram of Post-Test Experiment Class

The Figure 4 shows that the experimental class students' learning outcomes for the Nahwu subject were consistent with the post-test findings.

Test Requirements Analysis

Normality Test

The results of a normality test are used to assess if a distribution is or is not normally distributed. The Kolmogorov-Smirnov normality test was used in this test using SPSS 26. If a value is significant (sig > 0.05), it is considered to be normally distributed. If a value is significant (sig 0.05), it is considered to be non-normally distributed.

Table 2. Normality Test Result of Pre-Test

	Class	Kolmogrov-Smirnov ^a			Sapiro-Wilk		
		Statistic	df	Sig.	Statistic	df	Sig.
Learning Outcomes	Control	.225	15	.040	.895	15	.080
	Experiment	.267	15	.005	.852	15	.018

The [table 2](#) shows that the significance value (sig) for the experimental class in the pre-test was $0.005 < 0.05$, indicating that the data were not normally distributed, compared to $0.040 < 0.05$ for the control group in the pre-test. So, the non-parametric Mann-Whitney test was employed by the researchers.

Table 3. Normality Test Result of Post-Test

	Class	Kolmogrov-Smirnov ^a			Sapiro-Wilk		
		Statistic	df	Sig.	Statistic	df	Sig.
Learning Outcomes	Control	.236	15	.020	.846	15	.015
	Experiment	.292	15	.001	.746	15	.001

According to the [table 3](#), the significant value (sig) for the experimental class group on the post test is $0.001 < 0.05$, indicating that the data are abnormally distributed, compared to $0.025 < 0.05$ for the controller class group. Then, the researchers apply the non-parametric Mann-Whitney test.

Homogeneity Test

The purpose of using the homogeneity test in this experimental study is to determine that the researchers set a collection of control and experimental class data is homogeneous or heterogeneous, (Usmadi, 2020, p. 51).

Table 4. Homogeneity Test Result for Pre-Test

		Levene Statistic	df1	df2	Sig.
Nahwu learning outcomes	Based on Mean	5.158	1	28	.031
	Based on Median	3.840	1	28	.060
	Based on Mean and with adjusted df	3.840	1	18.244	.065
	Based on trimmed mean	4.589	1	28	.041

According to the output [table 4](#) in the preceding sentence, the results of the significance value (sig) based on mean $0.031 < 0.05$ suggest that the *Nahwu* learning outcomes of the control class group and experimental class in the pre-test is heterogeneous.

Table 5. Homogeneity Test Result for Post-Test

		Levene Statistic	df1	df2	Sig.
<i>Nahwu</i> learning outcomes	Based on Mean	3.360	1	28	.077
	Based on Median	1.347	1	28	.256
	Based on Mean and with adjusted df	1.347	1	26.304	.256
	Based om trimmed mean	3.195	1	28	.085

According to the output [table 5](#) in the preceding sentence, the significance value (sig) of a *Nahwu* learning outcomes for the control class group and experimental class in the post test appears to be homogeneous since the findings of the significance value (sig) based on mean $0.077 > 0.05$.

Hypothesis Test

To evaluate the data, the researchers chose to use the Mann-Whitney test, anon-parametric statistical technique. The normality test's conclusion that the data were not normally distributed led to this a significant difference between two independent samples. It is used in this study to evaluate the achievement gap between the experimental class and the control class in a number of subjects. The test is appropriate when the data does not conform to the normal distribution assumptions. To investigate the potential impact of portfolio authentic assessment on the learning outcomes of *Nahwu* for fourth-grade TMI students Modern Islamic Boarding School of *Al-Ikhlash* Female, the researchers aim to measure and address the research hypothesis using the following hypothesis test: (a) H_a : There was difference between the experimental class value and the control class value. (b) H_o : There was no difference between the experimental class value and the control class value.

To conclude it, it is necessary to pay attention to the formula as follows (1) If the significance value (Ashmp. Sig. (2-tailed)) < 0.05 , then H_o is rejected and H_a is accepted. (2) If the significance value (Ashmp. Sig. (2-tailed)) > 0.05 , then H_o is accepted and H_a is rejected

Table 6. Mann Whitney Test Result for Pre-Test

Test Statistics ^a	
Hasil Belajar Nahwu	
Mann-Whitney U	105.500
Wilcoxon W	225.500
Z	-.293
Asymp.Sig. (2-tailed)	.769
Exact Sig. [2*(1-tailed Sig.)]	.775 ^b
a. Grouping Variable: class	
b. Not Corrected for ties.	

Based on the [table 6](#) provide in the previous sentence, the non-parametric statistical analysis using the Mann-Whitney test for the pre-test indicates that the significance value (p-value) is 0.769, which is greater than the significance level of 0.05. Therefore, based on the acceptance of the null hypothesis (H_0) and rejection of the alternative hypothesis (H_a), it can be concluded that there is no significant difference in the learning outcomes between the control class and the experimental class.

Table 7. Mann Whitney Test Result for Pre-Test

Test Statistics	
<i>Nahwu</i> learning outcomes	
Mann-Whitney U	53.500
Wilcoxon W	173.500
Z	-2.508
Asymp.Sig. (2-tailed)	.012
Exact Sig. [2*(1-tailed Sig.)]	.013 ^b
a. Grouping Variable: Kelas	
b. Not Corrected for ties.	

According to the [table 7](#) in the preceding sentence, non-parametric statistical findings using the Mann Whitney test for post-tests demonstrate that the significant value (Ashmp. Sig. (2-tailed)) of $0.012 < 0.05$ was obtained. There is a difference between the learning outcomes of the control class using the traditional evaluation and the experimental class using the authentic assessment of the portfolio, according to the hypothesis that H_0 is rejected and H_a is accepted. The implementation of the portfolio authentic assessment had an impact on the 4th grade *Nahwu* science curriculum at the Modern Islamic Boarding School Al-*Ikhlah* Female.

Because the learning outcomes of the control class were 71.67 using conventional assessments, which were smaller than the learning outcomes of the experimental class, it is known from the results of non-parametric statistics with the Mann Whitney test that the set results for assessing *Nahwu* material using authentic portfolio assessment are $0.012 < 0.005$. Therefore, the hypothesis H_0 is rejected and H_a is accepted.

This research showed that PPA can significantly influence students' learning outcomes, this is in accordance with research conducted on undergraduate students at Universidad Tecnica Particular de Loja, in southern Ecuador (Cabrera-Solano, 2020). Students who use PPA demonstrate better conceptual understanding, applying *Nahwu's* rules to solve complex linguistic problems. They also demonstrated increased engagement and participation, driven by the relevance and practicality of the PPA task. Additionally, PPA encourages critical thinking, as students analyze and reflect on their work, gaining a deeper understanding of the nuances and applications of *Nahwu* (Yarun & Markus, 2023). Moreover, PPA improves metacognitive skills, allowing students to assess their own learning and become self-directed and self-directed learners. Portfolio Authentic Assessment (PPA) has emerged as a promising approach for assessing *Nahwu* learning in Islamic boarding schools (Kaukab et al., 2021). In contrast to traditional methods that rely on memorization and standardized tests, PPA focuses on real-world assignments, reflection, and collaboration, encouraging deeper understanding, active engagement, and critical thinking skills among students.

Although PPA implementation faces challenges, such as teacher training, curriculum adjustments, and resource limitations, the potential benefits are clear. Future research should focus on longitudinal studies, development of effective PPA models, and exploration of technology integration to improve portfolio creation, feedback, and collaboration.

CONCLUSIONS

It may be inferred from the Mann-Whitney test results, which demonstrated a significance value (sig) of 0.012, that there is a significant difference between the experimental class and the control group's post-test *Nahwu* science scores. There is enough evidence to demonstrate that the implementation of portfolio authentic assessment has an impact on raising students' knowledge of the goal of *Nahwu* science because the significant value (0.012) is smaller than the significance threshold (0.05).

Accordingly, it can be concluded from these results that using the portfolio authentic assessment approach has a beneficial impact on student's knowledge and understanding of the material linked to *Nahwu* science. This implies that using authentic assessment, such portfolio assessment, might help student acquire the material more effectively and achieve better learning outcomes in *Nahwu* science.

ACKNOWLEDGMENT

Thank you to all parties involved in this research. Especially the school that we used as a place for this research.

BIBLIOGRAPHY

- Aheisibwe, I., Kobusigye, L., & Tayebwa, J. (2021). Bridging Education Gap in Higher Institutions of Learning Using Bloom's Taxonomy of Educational Objectives. *African Educational Research Journal*, 9(1), 69–74. <https://doi.org/10.30918/AERJ.91.20.213>
- Albantani, A. M. (2018). Metode Pembelajaran Sastra Arab. *Alfaz (Arabic Literatures for Academic Zealots)*, 6(01), 17. <https://doi.org/10.32678/alfaz.Vol6.Iss01.711>
- Arifianto, M. L., Ahsanuddin, M., Ainin, M., Irhamni, I., Fitria, N., & Rahmah, L. A. (2022). Pelatihan Pengembangan Tes Interaktif Bahasa Arab Untuk Alumni Jurusan Sastra Arab Fakultas Sastra Universitas Negeri Malang. *ABDIMAS UNWAHAS*, 7(1), Article 1. <https://doi.org/10.31942/abd.v7i1.6571>
- Brosh, H. Y. (2019). Arabic language-learning strategy preferences among undergraduate students. *Studies in Second Language Learning and Teaching*, 19(2), 351–377. <https://doi.org/10.14746/ssl.2019.9.2.5>
- Cabrera-Solano, P. (2020). The Use of Digital Portfolios to Enhance English as a Foreign Language Speaking Skills in Higher Education. *International Journal of Emerging Technologies in Learning (IJET)*, 15(24), 159–175. <https://doi.org/10.3991/ijet.v15i24.15103>

- Hajaroh, S., & Adawiyah, R. (2018). *Kesulitan Guru Dalam Mengimplementasikan Penilaian Autentik*.
- He Xinyu, X., Deng, Y., Yu, S., & Wang, H. (2020). The Influence of Context on the Large-Scale Assessment of High School Students' Epistemic Cognition of Scientific Argumentation. *Science & Education*, 29(1), 7–41. <https://doi.org/10.1007/s11191-019-00088-4>
- Ismail, S. M., Rahul, D. R., Patra, I., & Rezvani, E. (2022). Formative vs. Summative assessment: Impacts on academic motivation, attitude toward learning, test anxiety, and self-regulation skill. *Language Testing in Asia*, 12(1), 40. <https://doi.org/10.1186/s40468-022-00191-4>
- Kaukab, M. E., Salamah, & Rahman, R. A. (2021). Penilaian Autentik Dalam Pembelajaran Bahasa Arab. *NIVEDANA: Jurnal Komunikasi Dan Bahasa*, 2(1), Article 1. <https://doi.org/10.53565/nivedana.v2i1.282>
- Mahmudi, I. (2018). Rubrik Analitik Penilaian Hasil Belajar Praktik Pendidikan Agama Islam. *Fikrah: Journal of Islamic Education*, 2(2). <https://doi.org/10.32507/fikrah.v2i2.406>
- Mahmudi, I., & Nadhifah, L. (2020). Peningkatan Hasil Belajar Fiqih Dengan Strategi Pembelajaran Example Non Example Siswi Kelas 1 Pondok Modern Darussalam Gontor Putri Kampus 1. *Al-Tadzkiyyah: Jurnal Pendidikan Islam*, 11(1), Article 1. <https://doi.org/10.24042/atjpi.v11i1.5962>
- Mahmudi, I., & Shofro, M. L. (2019). Peningkatan Hasil Belajar Tarikh Islam Melalui Strategi Pembelajaran Everyone is a Teacher Here pada Siswi Kelas 2 KMI Pondok Modern Darussalam Gontor Putri. *Jurnal Tatsqif*, 17(2), Article 2. <https://doi.org/10.20414/jtq.v17i2.1962>
- Mohajan, H. K. (2020). Quantitative Research: A Successful Investigation in Natural and Social Sciences. *Journal of Economic Development, Environment and People*, 9(4), 50–79. <https://doi.org/10.26458/jedep.v9i4.679>
- Monib, W. K., Karimi, A. Q., & Nijat, N. (2020). Effects of Alternative Assessment in EFL Classroom: A Systematic Review. *American International Journal of Education and Linguistics Research*, 3(2), Article 2. <https://doi.org/10.46545/aijelr.v3i2.195>
- Muhammad, A., Idhan, & Punawan, A. S. Bin. (2020). Nahwu, Origin and Its Urgences in Arabic Learning. In *Online Submission* (Vol. 24, Issue 6, pp. 6894–6900). <https://eric.ed.gov/?Id=ED608786>
- Nguyen, T. T. K., & Phan, H. M. (2020). Authentic Assessment: A Real Life Approach To Writing Skill Development. *International Journal of Applied Research in Social Sciences*, 2(1), Article 1. <https://doi.org/10.51594/ijarss.v2i1.97>
- Nurgiyantoro, B. (2011). Model Penilaian Otentik Dalam Pembelajaran Bahasa. *LITERA*, 10(2), Article 2. <https://doi.org/10.21831/ltr.v10i2.1157>
- Nurtanto, M., Kholifah, N., Masek, A., Sudira, P., & Samsudin, A. (2021). Crucial Problems in Arranged the Lesson Plan of Vocational Teacher. *International Journal of Evaluation and Research in Education*, 10(1), 345–354.

- Révész, J. R., Andrea. (2019). Experimental and quasi-experimental designs. In *The Routledge Handbook of Research Methods in Applied Linguistics*. Routledge. <http://doi.org/10.11591/ijere.v10i1.20604>
- Sa'adah, N. (2019). Problematika Pembelajaran Nahwu bagi Tingkat Pemula Menggunakan Arab Pegon. *Lisanan Arabiya: Jurnal Pendidikan Bahasa Arab*, 3(1), 15–32. <https://doi.org/10.32699/liar.v3i01.995>
- Safriyani, R. (2012). Portofolio as an authentic assessment in teaching English for young learners. *MADRASATUNA (Jurnal Pengembangan Madrasah Ibtidaiyah)*, 4(1), Article 1. <http://madrasatuna.uinsby.ac.id/index.php/madrasatuna/article/view/8>
- Sari, E. D. K., & Mahmudi, I. (2019). The Effect Of Authentic Assessment And Teaching Variation Toward Fiqih Learning Outcomes (An Experiment at Ponpes An Nida Al islamybekasi). *At-Ta'dib*, 14(2), Article 2. <https://doi.org/10.21111/at-tadib.v14i2.3256>
- Shaturaev, J. (2021). Financing and Management of Islamic (Madrasah) Education in Indonesia. *Zeszyty Naukowe Politechniki Częstochowskiej. Zarządzanie*, 42, 57–65. <https://doi.org/10.17512/znpcz.2021.2.05>
- Siregar, I. (2021). Epistemological Challenges Against Sociolinguistics. *International Journal of Linguistics Studies*, 1(2), Article 2. <https://doi.org/10.32996/ijls.2021.1.2.6>
- Sugiyono, S. (2018). *Metode Penelitian Pendidikan, Pendekatan Kuantitatif, Kualitatif, dan R&D*. Alfabeta.
- Sulistyo, T., Eltris, K. P. N., Mafulah, S., Budianto, S., Saiful, & Heriyawati, D. F. (2020). *Portfolio Assessment: Learning Outcomes and Students' Attitudes*. <http://repository.unisma.ac.id/handle/123456789/2349>
- Syafi'i, A. S., Akmal, A., & Bulan, S. (2023). Pemahaman dan Praktik Guru dalam Melaksanakan Penilaian Mata Pelajaran Pendidikan Agama Islam di mts As'adiyah Uloe. *Wahana Karya Ilmiah Pendidikan*, 3(01), Article 01. <https://doi.org/10.35706/wkip.v7i01.8923>
- Usmadi, U. (2020). Pengujian Persyaratan Analisis (Uji Homogenitas Dan Uji Normalitas). *Inovasi Pendidikan*, 3(1). <https://doi.org/10.31869/ip.v7i1.2281>
- Wanti, A. I., a'yuni, Q., & Chamidah, D. (2022). Model Dan Praktik: Asesmen Formatif NonPaper-Based Dalam Pembelajaran Bahasa Arab. *Arabi: Journal of Arabic Studies*, 3(1), Article 1. <https://doi.org/10.24865/ajas.v7i1.419>
- Yarun, A., & Markus, M. (2023). Higher Order Thinking Skill (HOTS) Students in Indonesia and Saudi Arabia (Comparative Study of Assessment Application in Schools). *Journal of Contemporary Islamic Education*, 3(2), Article 2. <https://doi.org/10.25217/jcie.v3i2.3787>