

**BUKTI KORESPONDENSI**  
**ARTIKEL JURNAL SINTA 2**

JUDUL ARTIKEL	Impact of Portfolio Authentic Assessment on the Learning Outcomes of Nahwu at Modern Islamic Boarding School
JURNAL	Izdihar: Journal of Arabic Language Teaching, Linguistics, and Literature . Vol 6. No. 3 (2023)
PENULIS	Ihwan Mahmudi, Aisyah Imani, Chevira Yamin Indriyani, Rohani Poteh

No	Perihal	Tanggal
1	Submit	26 June 2023
2	Pre-assessment	29 Juli
3	Notifikasi Review	8 November
4	Submit Revisi	5 Desember
5	Accepted	22 Desember
6	Publish	4 Januari 2024

## **BUKTI SUBMIT 26 JUNI 2023**

27441 / Mahmudi et al. / Impact of Portfolio Authentic Assessment on Students' Learning Outcomes of Nahwu at Modern Islamic Boa

Libr

Workflow

Publication

Submission

Review

Copyediting

Production

### Submission Files

Q Search

▶	 91295	Impact of Authentic Assessment Portfolio on The Learning Results of Nahwu Science 4th grade TMI in the Modern Islamic Boarding School of Al-Ikhlash Female Kuningan West Java.docx	June 26, 2023	Article Text
---	---	--	---------------	--------------



**Universitas Muhammadiyah Malang, East Java, Indonesia**

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

p-ISSN: [2622-738X](#), e-ISSN: [2622-7371](#) // Vol. x No. y Month 20xy, pp. xx-xx



<https://doi.org/10.22219/iz.Volxx.Nov.iii-iii>



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



[izdihar.jurnalpba@umm.ac.id](mailto:izdihar.jurnalpba@umm.ac.id)

# **Impact of Authentic Assessment Portfolio on The Learning Results of Nahwu Science 4th grade TMI in the Modern Islamic Boarding School of Al-Ikhlash Female Kuningan West Java**

**Ihwan Mahmudi<sup>a,1</sup>, Aisyah Imani<sup>a,2</sup>, Chevira Yasin Indriyanti<sup>a,3</sup>, Rohani Pooteh<sup>b,4</sup>**

<sup>a</sup> **Universitas Darussalam Gontor, Ponorogo**

<sup>b</sup> **Prince of Songkla University, Thailand**

<sup>1</sup>[ihwanm@unida.gontor.ac.id](mailto:ihwanm@unida.gontor.ac.id), <sup>2</sup>[aisyahimani19@gmail.com](mailto:aisyahimani19@gmail.com),

<sup>3</sup>[cheviraputriindriyanti26@student.pba.unida.gontor.ac.id](mailto:cheviraputriindriyanti26@student.pba.unida.gontor.ac.id), <sup>4</sup>[rohani.p@psu.th](mailto:rohani.p@psu.th)

## **ARTICLE INFO**

### **Article History:**

Received:

Revised:

Accepted:

Published:

### **\*Corresponding**

#### **Author:**

Name:

Email:

Phone/WA:

## **ABSTRACT**

Specifically focusing on class 4 TMI (Tarbiyah Muallimat al-Islamiyah) at the Modern Islamic Boarding School of Al-Ikhlash Female in Kuningan, West Java, the study aims to investigate the potential beneficial impact of portfolio assessment on the learning outcomes of Nahwu lessons for fourth-grade students. The finding that 4 TMI has impetus for this investigation. The target population for the study, which included all fourth graders, was carried out during the odd semester of the 2022 academic year. All students in class 4 TMI were included in the research sample thanks to the saturated sampling method that was used. A quasi-experimental design was 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. When comparing this score to that of the experimental class, which used portfolio assessment, it became clear that the pupils' comprehension of the Nahwu material had significantly improved. These results show that portfolio evaluations have a chance to improve fourth-grade students' learning outcomes in Nahwu.

Copyright © 20xy, First Author et al

This is an open access article under the [CC-BY-SA](#) license



## مستخلص البحث

مع التركيز بشكل خاص على الصف الرابع تربية المعلمات الإسلامية في المدرسة الداخلية الإسلامية الحديثة في الإخلاص للإناث في كونينجان ، جاوة الغربية ، تهدف الدراسة إلى التحقيق في الأثر المفيد المحتمل لتقييم المحفظة على نتائج التعلم لدروس النهو. 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 النهو

## INTRODUCTION

Planning, implementation, and assessment are the three essential phases of learning activities are highlighted in the paper (Suwardi, 2008, p. 29). The assessment step is essential since it assesses how well learning is going and helps to improve the creation of education programs (2006, p. 10). 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). The necessity of educators evaluating student learning outcomes to track progress and continuously improve them is highlighted by this alignment with Law No. 20 of 2003 Article 58, paragraph 1 (Majid, 2013, pp. 32-33). 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 (Ihwan Mahmudi, 2018, p. 4).

Short term memory skills are often the product of education that only emphasizes understanding content. 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 are intended to correct traditional evaluations that purposefully ignore the work of productive learners. According to Hajroh and Raudatul Adawiyah (2018), p. 132, specific evaluation entails a full examination of students emotive, cognitive, and psychomotor domains, spanning input, process, and output stages.

The main goal of teaching Arabic is for students to become proficient users of the language. Language, according to Ibn Jinni, includes the sounds people use to express their intents (2001, p. 45), in order to achieve this goal, it is crucial to use realistic assessment in Arabic language training. Realistic evaluations make sure that students can use their knowledge of Arabic 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. 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 student needs (Hidayat, 2012, p. 85). The science of nahwu is one of the most important topics in the study of Arabic. In evaluating student 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, p. 17).

The aforementioned problem could be exacerbated by a lack of varied assessments, such portfolio evaluation. Sukanti (2010, p. 33) claims that the assessment procedures mostly center on cognitive test that are administered after learning and rely on pre-existing inquiries from the textbook. 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 teachers capacity to assess student learning. The Modern Islamic Boarding School's fourth-grade student learning about nahwu, and researchers found that teachers frequently use routine evaluation techniques that are useful and affordable. These technique primarily use written exams as both formative and summative assessments without any alterations or variations (Wijayanti, 2022, p. 108).

Researchers intend to experiment with authentic assessment portfolios in the nahwu class for fourth-graders at the Modern Islamic Boarding School of Al-Ikhlash Female in order to address the problem of ineffective assessment techniques. According to Eva Dwi and Ihwan Mahmudi (2019, p. 6), 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 authentic assessment portfolio (Safriyani, 2012, p. 2). 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 student language abilities, enabling comprehensive study of their language growth. Dr. Suneetha adds that for educational reforms to be most effective, the curriculum must be improved and real evaluation procedures must be included (Kim Nguyen & Minh Phan, n.d., p. 21). this suggests that students diverse learning goals ought to be reflective of real live situations (Nurgiyantoro & Suyanta, 2011, p. 7).

## 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 to variables. The use of authentic assessment portfolio served as the study's independent variable (x), and the students learning outcomes following the evaluation process utilizing the authentic assessment portfolio served as the study's dependent variable (y) (Riadi, 2017, p. 54). The study's goal was to ascertain how using authentic assessment portfolio affected students learning outcomes.

A quasi-experimental design was used by the researcher in this study as a modified form of an actual experimental design. 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, p. 114).

The Nonequivalent 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, p. 116).

Students from the Modern Islamic Boarding School of Al-Ikhlash Female for the academic year 2022 participated in this study as fourth-semester students. 30 students in total, divided into 4<sup>th</sup> B and 4<sup>th</sup> C classes. All 30 of the fourth-grade students were chosen by researcher 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 A.1**  
**Descriptive Test Results**

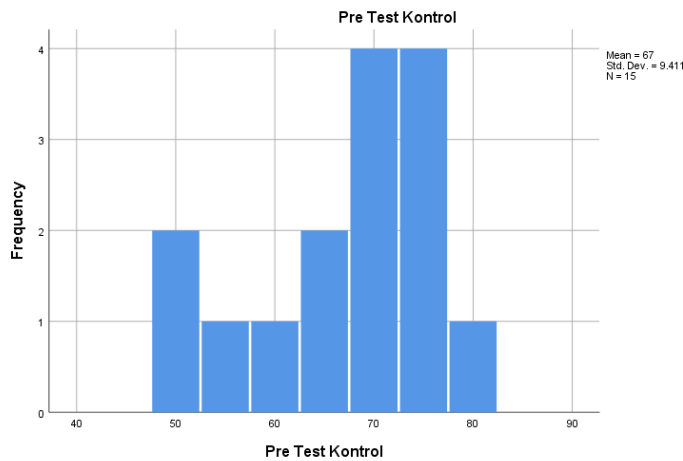
	N	Range	Minimum	Maximum	Sum	Mean		Std. Deviation	Variance
	Statistic	Statistic	Statistic	Statistic	Statistic	Statistic	Std. Error	Statistic	Statistic
Pre Test Kontrol	15	30	50	80	1005	67.00	2.430	9.411	88.571
Post Test Kontrol	15	70	30	100	1075	71.67	6.204	24.029	577.381
Pre Test Eksperimen	15	75	15	90	930	62.00	6.168	23.890	570.714
Post Test Eksperimen	15	50	50	100	1350	90.00	3.684	14.268	203.571
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.

And for the experimental class in the pre-test, the average result (range) 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.

The researcher constructed the following histogram based on the arrangement of the descriptive test table:

### A.2 Histogram of Pre-Test Control Class

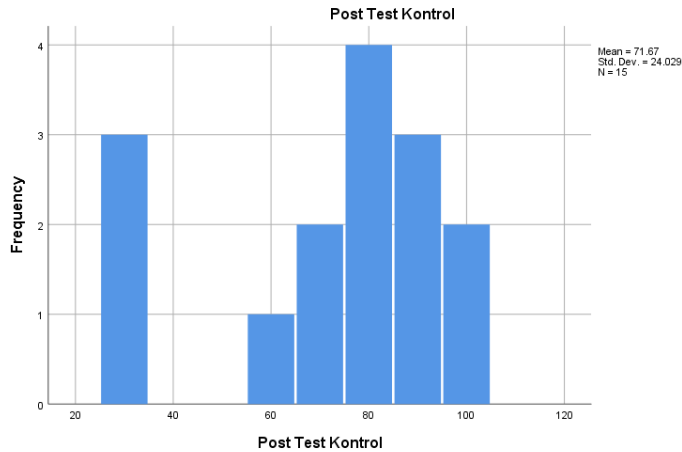


The table illustrates the consistent mastery of nahwu subject by the students in the control class, as indicated by their pre-test scores. Following the description of the descriptive test conducted for the post-test of the control class:

### A.3



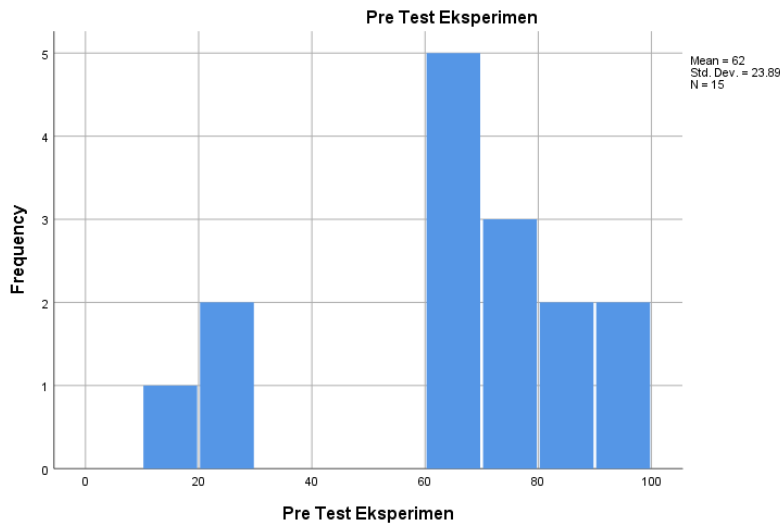
### Histogram of Post-Test Control Class



The table 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 researcher presents the descriptive test for the experimental class pre-test as follows:

#### A.4

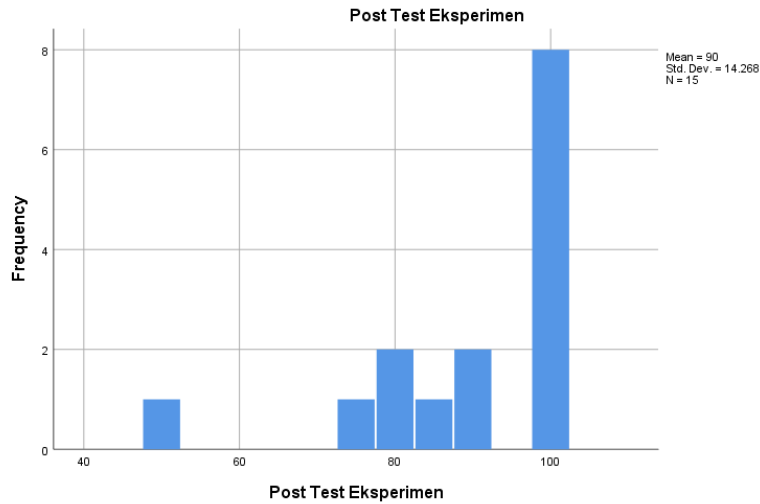
### Histogram of Pre-Test Experiment Class



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

#### A.5

## Histogram of Post-Test Experiment Class



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

### 1. Test Requirements Analysis

#### a) 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 ( $\text{sig} > 0.05$ ), it is considered to be normally distributed. If a value is significant ( $\text{sig} < 0.05$ ), it is considered to be non-normally distributed.

## A.6

### Normality Test Results for Pre-Test

	Kelas	Kolmogorov-Smirnov <sup>a</sup>			Shapiro-Wilk		
		Statistic	df	Sig.	Statistic	df	Sig.
Hasil Belajar Santri	Pre Test Kontrol	.225	15	.040	.895	15	.080
	Post Test Eksperimen	.267	15	.005	.852	15	.018

The table shows that the significance value ( $\text{sig}$ ) 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.

## A.7

### Normality Test Results for Post-Test

	Kelas	Kolmogorov-Smirnov <sup>a</sup>			Shapiro-Wilk		
		Statistic	df	Sig.	Statistic	df	Sig.
Hasil Belajar Santri	PostTest Kontrol	.236	15	.025	.846	15	.015
	PostTest Eksperimen	.292	15	.001	.746	15	.001

According to the table, 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 researcher applies the non-parametric Mann-Whitney test.

#### b). Homogeneity Test

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

- a. A significant value  $< 0.05$  indicates that the population's data has not the same/not homogeneous.
- b. A significant value  $> 0.05$  indicates that the population's data has the same variance/homogeneous.

### A.8

#### Homogeneity Test Results for Pre-Test

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

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

### A.9

### Homogeneity Test Results for Post-Test

		Levene Statistic	df1	df2	Sig.
Hasil Belajar Nahwu	Based on Mean	3.360	1	28	.077
	Based on Median	1.347	1	28	.256
	Based on Median and with adjusted df	1.347	1	26.304	.256
	Based on trimmed mean	3.195	1	28	.085

According to the output table in the preceding sentence, the significance value (sig) of a Nahwu learning outcome 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  $0.077 > 0.05$ .

#### 2. Hypothesis Test

To evaluate the data, the researchers chose to use the Mann-Whitney test, a non-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 (Yanti, 2007, p. 55).

To investigate the potential impact of authentic assessment portfolio on the learning outcomes of nahwu for fourth-grade TMI students Modern Islamic Boarding School of Al-Ikhlah Female, the researchers aim to measure and address the research hypothesis using the following hypothesis test:

$$H_0: \mu_1 = \mu_2$$

$$H_a: \mu_1 \neq \mu_2$$

$H_a$  : There was difference between the experimental class value and the control class value.

$H_0$  : 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  $H_0$  is rejected and  $H_a$  is accepted.
- 2) If the significance value (Ashmp. Sig. (2-tailed))  $> 0.05$ , then  $H_0$  is accepted and  $H_a$  is rejected

**A.10**  
**Mann Whitney Test Results for Pre-Test**

Test Statistics <sup>a</sup>	
	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 <sup>b</sup>
a. Grouping Variable: Kelas	
b. Not corrected for ties.	

Based on the table 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.

**A.11**  
**Mann Whitney Test Results for Post-Test**

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

According to the table 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 authentic assessment portfolio had an impact on the 4<sup>th</sup> grade nahwu science curriculum at the Modern Islamic Boarding School Al-Ikhlash 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.

## 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 authentic assessment portfolio 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 authentic assessment portfolio 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.

## BIBLIOGRAPHY

Basrowi, & Suwandi. (2008). *Memahami Penelitian Kualitatif*. Jakarta, Indonesia: Rineka Cipta.

Hajroh, S., & Raudatul Adawiyah. (2018). Kesulitan Guru dalam Mengimplementasikan Penilaian Autentik. *Jurnal PGMI*, 10.

Hidayat, N. S. (2012). Problematika Pembelajaran Bahasa Arab. *Jurnal Pemikiran Islam*, 37.

Kim Nguyen, T. T., & Minh Phan, H. (n.d.). AUTHENTIC ASSESSMENT: A REAL LIFE APPROACH TO WRITING SKILL DEVELOPMENT. *International Journal of Applied Research in Social Sciences*, 2.

- M. Aini. (2006). *Evaluasi dalam Pembelajaran Bahasa Arab*. Malang: Misykat.
- Mahmudi, I. (2019). RUBRIK ANALITIK PENILAIAN HASIL BELAJAR PRAKTIK PENDIDIKAN AGAMA ISLAM. *Fikrah: Journal of Islamic Education*, 2(2), 136-154.
- Majid, A. (2013). *Penilaian Autentik Proses dan Hasil Belajar*. Bandung: PT. Remaja Rosdakarya.
- Nurdiyantoro, B., & Suyata, P. (2011). Model Otentik Dalam Pembelajaran Bahasa. *Jurnal LITERA*, 10.
- Riadi, A. (2017). KOMPETENSI GURU DALAM PELAKSANAAN EVALUASI PEMBELAJARAN, 15.
- Sa'adah, N. (2019). Problematika pembelajaran nahwu bagi tingkat pemula menggunakan Arab pegon. *Jurnal Pendidikan Bahasa Arab*, 3.
- Safriyani, R. (2012). PORTOFOLIO AS AN AUTHENTIC ASSESSMENT IN TEACHING ENGLISH FOR YOUNG LEARNERS, 04.
- 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 Islamy Bekasi). *At-Ta'dib*, 14(2), 14-25.
- Sugiyono. (2018). *Metode Penelitian Pendidikan, Pendekatan Kuantitatif, Kualitatif, dan R&D* (28th ed.). Bandung: Alfabeta.
- Sukanti, S. (2010). PEMANFAATAN PENILAIAN PORTOFOLIO DALAM MENINGKATKAN HASIL BELAJAR AKUNTANSI. *Jurnal Pendidikan Akuntansi Indonesia*, 8(2). Retrieved 24 January 2023 from <https://doi.org/10.21831/jpai.v8i2.952>
- Usmadi, U. (2020). PENGUJIAN PERSYARATAN ANALISIS (UJI HOMOGENITAS DAN UJI NORMALITAS). *Inovasi Pendidikan*, 7(1). Retrieved 24 January 2023 from <https://doi.org/10.31869/ip.v7i1.2281>
- Wijayanti, M. D. (2022). STRATEGI GURU DALAM MENINGKATKAN SIKAP NASIONALISME MELALUI PEMBELAJARAN PENDIDIKAN PANCASILA DAN KEWARGANEGARAAN DI SMAN 1 NGADIROJO PACITAN, 10.
- Yanti, T. S. (2007). Uji Rank Mann-Whitney Dua Tahap. *STATISTIKA*, 7(1). Retrieved 24 January 2023 from <https://doi.org/10.29313/jstat.v7i1.954>

أحمد, م., & القادر, ع. (1986). طرق تعليم اللغة العربية. القاهرة: مكتبة النهضة المصرية.

داود, م. (2001). العربية وعلم اللغة الحديثة. القاهرة: دار غريب

## **BUKTI PRE-ASSESSMENT 29 JULI 2023**

AAS and CTA Notification of Journal Izdihar External Inbox x



Jurnal PBA UMM <izdihar.jurnalpba@umm.ac.id>

Jul 29, 2023, 9:32 AM



to me, aisyahimani19, chevraputriindriyanti26 ▾

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

Thank you for submitting an article to our Journal with the title

After a pre-assessment by the editorial team, we have decided that your article deserves to be continued in the review process.

Please fill in and send the Author Agreement Statement (AAS) and the Copyright Transfer Agreement (CTA) if you agree to our terms. The review process will be carried out after you send the required documents and the proof of payment of the Article Processing Charges (APC). If your manuscript is rejected after double-blind peer review processes, the APCs will be sent back fully.

please transfer the APC and send the proof of payment to this email. The payment is sent to the following account:

## **BUKTI REVIEW DARI REVIEWER**

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

to Submissions

### Notifications

<a href="#">[Iz] Editor Decision</a>	2023-11-08 09:10 AM
<a href="#">[Iz] Editor Decision</a>	2023-12-22 01:08 PM

### Reviewer's Attachments

[Q Search](#)

96238 21.docx	September 4, 2023
101574 portfolio.pdf	November 8, 2023

### Revisions

[Q Search](#) [Upload File](#)

103827 Inwan Mahmudi_Izdihar_revisi.docx	December 5, 2023	Article Text
--	------------------	--------------

### Review Discussions

[Add discussion](#)

Name	From	Last Reply	Replies	Closed
No Items				



We have reached a decision regarding your submission to Izdihar : Journal of Arabic Language Teaching, Linguistics, and Literature, "Impact of Authentic Assessment Portfolio on The Learning Results of Nahwu Science 4th grade TMI in the Modern Islamic Boarding School of Al-Ikhlah Female Kuningan West Java".

Our decision is: Revisions Required

-----  
Reviewer A:

Recommendation: Revisions Required

-----  
Reviewer B:

The author needs to revise the parts that we have mentioned above. This paper meets the basic rules of scientific publication. However, the writer needs to emphasize a number of things: (1) the novelty of the writing or research; (2) the urgency of research results; (3) benefits that can be given to academics regarding research findings.

Activate Windows  
Go to Settings to activate Windows.

## Review Form Jurnal Izdihar

**No. Article File** : 21

**Title** : **Impact of Authentic Assessment Portfolio on The Learning**

**Results of**

**Nahwu Science 4th grade TMI in the Modern Islamic Boarding  
School of Al-Ikhlash Female Kuningan West Java**

No	Pernyataan	Nilai					Komentar
		Insufficient	Weak	Good	Very Good	Excellent	
1	Is the topic relevant to the journal area of interest? Is it contemporary and interesting for researchers?			x			Researchers need to explain the urgency of research and its relevance in the world of Arabic education
2	<b>Abstract + Keywords</b> Are all required components included in the abstract? Are the keywords appropriately chosen?			x			The author needs to explain the research steps and include significant findings
3	<b>Goal</b> Is the goal explicitly stated in the Introduction? Is its formulation clear and unambiguous?			x			The aim of the research has been well stated. The introduction needs to be complemented by several research gaps that have been carried out previously.
4	<b>Structure</b> Is the paper's structure coherent? Is it in coherence with the goal of the paper?			x			It is in accordance with the purpose of writing
5	<b>Tools and Methods</b> Are methods the author uses adequate and well used?			x			In the research methods section, it is necessary to explain the instruments and data analysis. To assess the quality of the impact of an authentic portfolio assessment, any supporting indicators are needed, including expert validation of pretest and posttest questions

6	<b>Discussion and Conclusion</b> Is it related to the results presented before? Do you consider them as coherent?			x			This discussion is well written. The author needs to complete it with the results of a comparison of the level of portfolio effectiveness with several other assessment models.
7	<b>Literature</b> Does the author utilize relevant literature?		x				The literature is relevant to the topics discussed. It is necessary to add various international journals of the same type as this article and used a reference manager (Zotero/mendelay).
8	<b>Author's knowledge</b> What is the level of author's knowledge? Does he/she utilize all recent contributions relevant to the topic?			x			The author has the right knowledge and capacity to research this topic.
9	<b>Length</b> Is the length of the paper adequate to the significance of the topic? Do you suggest shortening the paper without losing its value?			x			It complies with the provisions of the journal
10	<b>Figures and Tables</b> Does the author use them suitably? Are legend and notations clear?			x			Images need to be enlarged proportionally to be seen clearly
11	<b>Writing style</b> Is it clear and understandable?			x			Language needs to be checked again so as not to cause misinterpretation
12	<b>Further comments on the paper</b> Comments	The author needs to revise the parts that we have mentioned above. This paper meets the basic rules of scientific publication. However, the writer needs to emphasize a number of things: (1) the novelty of the writing or research; (2) the urgency of research results; (3) benefits that can be given to academics regarding research findings.					
13	<b>Recommendation</b>	Revision					

Malang, 11 September 2023  
Reviewer,



Universitas Muhammadiyah Malang, East Java, Indonesia  
**Izdihar : Journal of Arabic Language Teaching,  
 Linguistics, and Literature**  
 p-ISSN: 2622-738X, e-ISSN: 2622-7371 // Vol. x No. y Month 20xy, pp. xx-xx



<https://doi.org/10.22219/jiz.Volxx.Nov.iii-iii>



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



[izdihar.jurnalpba@umm.ac.id](mailto:izdihar.jurnalpba@umm.ac.id)

## Impact of Authentic Assessment Portfolio on The Learning Results of Nahwu Science 4th grade TMI in the Modern Islamic Boarding School of Al-Iklash Female Kuningan West Java

Commented [MFF1]: 1.Judul terlalu panjang. Maksimal 15 kata  
 2>Nama institusi tidak perlu disebut dalam judul

### ARTICLE INFO

#### Article History:

Received:

Revised:

Accepted:

Published:

#### \*Corresponding

Author:

Name:

Email:

Phone/WA:

### ABSTRACT

Specifically focusing on class 4 TMI (Tarbiyah Muallimat al-Islamiyah) at the Modern Islamic Boarding School of Al-Iklash Female in Kuningan, West Java, the study aims to investigate the potential beneficial impact of portfolio assessment on the learning outcomes of Nahwu lessons for fourth-grade students. The finding that 4 TMI has impetus for this investigation. The target population for the study, which included all fourth graders, was carried out during the odd semester of the 2022 academic year. All students in class 4 TMI were included in the research sample thanks to the saturated sampling method that was used. A quasi-experimental design was 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. When comparing this score to that of the experimental class, which used portfolio assessment, it become clear that the pupils' comprehension of the Nahwu material had significantly improved. These results show that portfolio evaluations have a chance to improve fourth-grade students' learning outcomes in Nahwu.

Copyright © 20xy, First Author et al

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



### Keyword

Nahwu Science, Portfolio Authentic Assessment, Nahwu Learning

### مستخلص البحث

مع التركيز بشكل خاص على الصف الرابع تربية المعلمات الإسلامية في المدرسة الداخلية الإسلامية الحديثة في الإخلاص للإناث في كونيغان ، 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

Please cite this article as Author's Name. (Year). Title. *Izdihar : Journal of Arabic Language Teaching, Linguistics, and Literature*, Vol(Issue), page. DOI: <https://doi.org/10.22219/jiz.xxx.xxx>

التلاميذ للمادة الناهوية قد تحسن بشكل ملحوظ. تُظهر هذه النتائج أن تقييمات الحافظة لديها فرصة لتحسين نتائج التعلم لطلاب الصف الرابع في النهو

## INTRODUCTION

Planning, implementation, and assessment are the three essential phases of learning activities are highlighted in the paper (Suwardi, 2008, p. 29). The assessment step is essential since it assesses how well learning is going and helps to improve the creation of education programs (2006, p. 10). As a result, these three procedures are crucial for attaining the best learning results.

Commented [MFF2]: Poor language

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). The necessity of educators evaluating student learning outcomes to track progress and continuously improve them is highlighted by this alignment with Law No. 20 of 2003 Article 58, paragraph 1 (Majid, 2013, pp. 32-33). 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 (Ihwan Mahmudi, 2018, p. 4).

Commented [MFF3]: Perhatikan lagi cara menulis sitasi yang benar

Short term memory skills are often the product of education that only emphasizes understanding content. 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 are intended to correct traditional evaluations that purposefully ignore the work of productive learners. According to Hajroh and Raudatul Adawiyah (2018), p. 132, specific evaluation entails a full examination of students emotive, cognitive, and psychomotor domains, spanning input, process, and output stages.

Commented [MFF4]: Perhatikan lagi cara menulis sitasi yang benar

Commented [MFF5]: Paragraf ini tidak koheren tidak koheren dengan paragraf sebelumnya

The main goal of teaching Arabic is for students to become proficient users of the language. Language, according to Ibn Jinni, includes the sounds people use to express their intents (2001, p. 45), in order to achieve this goal, it is crucial to use realistic assessment in Arabic language training. Realistic evaluations make sure that students can use their knowledge of Arabic in real world situations.

Commented [MFF6]: Perhatikan lagi cara menulis sitasi yang benar

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. 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 student needs (Hidayat, 2012, p. 85). The science of nahwu is one of the most important topics in the study of Arabic. In evaluating student 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, p. 17).

The aforementioned problem could be exacerbated by a lack of varied assessments, such portfolio evaluation. Sukanti (2010, p. 33) claims that the assessment procedures mostly center on cognitive test that are administered after learning and rely on pre-existing inquiries from the textbook. 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 teachers capacity to assess student learning. The Modern Islamic Boarding School's fourth-grade student learning about nahwu, and researchers found that teachers frequently use routine evaluation techniques that are useful and affordable. These technique primarily use written exams as both formative and summative assessments without any alterations or variations (Wijayanti, 2022, p. 108).

Researchers intend to experiment with authentic assessment portfolios in the nahwu class for fourth-graders at the Modern Islamic Boarding School of Al-Ikhlash Female in order to address the problem of ineffective assessment techniques. According to Eva Dwi and Ihwan Mahmudi (2019, p. 6), 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 authentic assessment portfolio (Safriyani, 2012, p. 2). 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 student language abilities, enabling comprehensive study of their language growth. Dr. Suneetha adds that for educational reforms to be most effective, the curriculum must be improved and

real evaluation procedures must be included (Kim Nguyen & Minh Phan, n.d., p. 21). this suggests that students diverse learning goals ought to be reflective of real live situations (Nurgiyantoro & Suyanta, 2011, p. 7).

## METHOD

The study was carried out at the Modern Islamic Boarding School of Al-Ikhlas Female in Ciomas, Kuningan, West Java. The study used an experimental methodology and a quantitative research approach to examine the association between to variables. The use of authentic assessment portfolio served as the study's independent variable (x), and the students learning outcomes following the evaluation process utilizing the authentic assessment portfolio served as the study's dependent variable (y) (Riadi, 2017, p. 54). The study's goal was to ascertain how using authentic assessment portfolio affected students learning outcomes.

A quasi-experimental design was used by the researcher in this study as a modified form of an actual experimental design. 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, p. 114).

The Nonequivalent 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, p. 116).

Students from the Modern Islamic Boarding School of Al-Ikhlas Female for the academic year 2022 participated in this study as fourth-semester students. 30 students in total, divided into 4<sup>th</sup> B and 4<sup>th</sup> C classes. All 30 of the fourth-grade students were chosen by researcher 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

**Commented [MFF7]:** Catatan:

- 1.Gap analisis tidak ada
- 2.Sense of crisis tidak terasa
- 3.Tujuan tidak ada
- 4.Novelty tidak ada
- 5.Narasi yang dibangun kurang mencerminkan urgensi penelitian ini dengan topik



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 is easier to provide descriptive statistics to sum up and comprehend the experimental data.

**Table A.1**  
**Descriptive Test Results**

	N	Range	Minimum	Maximum	Sum	Mean	Std. Deviation	Variance
	Statistic	Statistic	Statistic	Statistic	Statistic	Statistic	Statistic	Statistic
Pre Test Kontrol	15	30	50	80	1005	67.00	2.430	9.411
Post Test Kontrol	15	70	30	100	1075	71.67	6.204	24.029
Pre Test Eksperimen	15	75	15	90	930	62.00	6.168	23.890
Post Test Eksperimen	15	50	50	100	1350	90.00	3.684	14.268
Valid N (listwise)	15							

**Commented [MFF8]:** 1.Tabel tidak sesuai dengan model APA  
2.Font tidak sesuai author guidelines  
3.Judul tabel tidak sesuai format

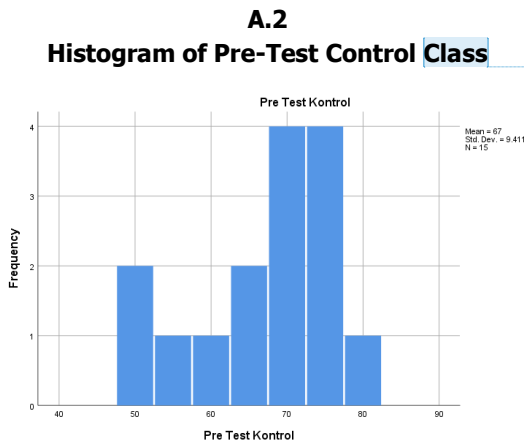
Based on the descriptive analysis table with SPSS 26 in the previous paragraph, that N is the number of samples or research population. Additionally, it has been noted that are 15 students in the experimental class, 15 students in the control class, and a proper sample size of 15 students. The typical 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.

And for the experimental class in the pre-test, the average result (range) 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

**Commented [MFF9]:** Tidak boleh di awal paragraf

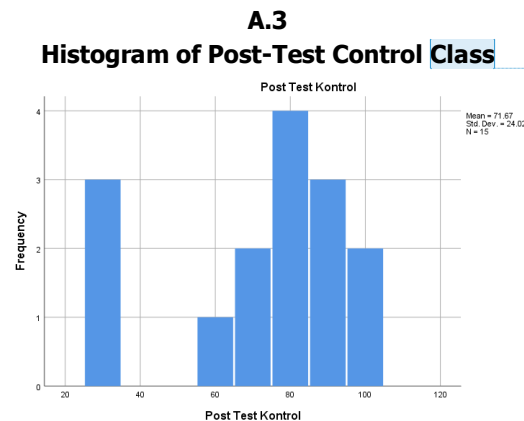
value (mean) is 90.00, and for the mean standard deviation (std. Deviation) is 14,268, and the total value (sum) is 1350.

The researcher constructed the following histogram based on the arrangement of the descriptive test table:



**Commented [MFF10]:** 1.Tabel tidak sesuai dengan model APA  
2.Font tidak sesuai author guidelines  
3.Judul tabel tidak sesuai format

The table 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:

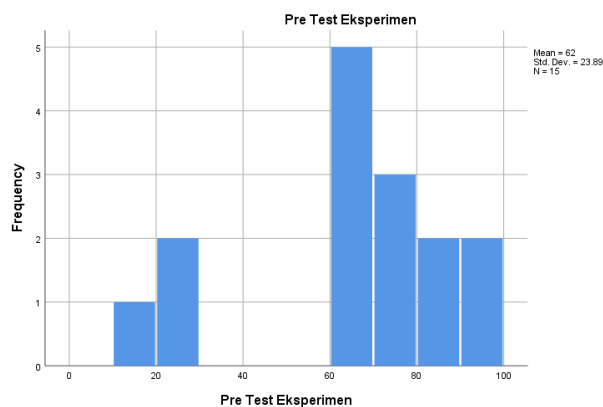


**Commented [MFF11]:** 4.Tabel tidak sesuai dengan model APA  
5.Font tidak sesuai author guidelines  
6.Judul tabel tidak sesuai format

The table 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 researcher presents the descriptive test for the experimental class pre-test as follows:

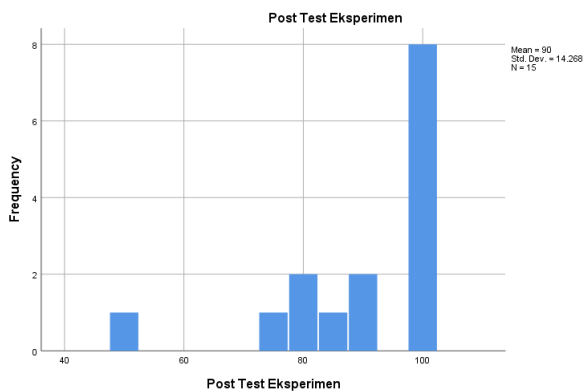
**A.4**  
**Histogram of Pre-Test Experiment Class**



**Commented [MFF12]:** 7. Tabel tidak sesuai dengan model APA  
8. Font tidak sesuai author guidelines  
9. Judul tabel tidak sesuai format

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

**A.5**  
**Histogram of Post-Test Experiment Class**



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

## 1. Test Requirements Analysis

### a) 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 ( $\text{sig} > 0.05$ ), it is considered to be normally distributed. If a value is significant ( $\text{sig} < 0.05$ ), it is considered to be non-normally distributed.

### A.6 Normality Test Results for Pre-Test

	Kelas	Kolmogorov-Smirnov <sup>a</sup>			Shapiro-Wilk		
		Statistic	df	Sig.	Statistic	df	Sig.
Hasil Belajar Santri	Pre Test Kontrol	.225	15	.040	.895	15	.080
	Post Test Eksperimen	.267	15	.005	.852	15	.018

**Commented [MFF13]:** 10.Tabel tidak sesuai dengan model APA  
11.Font tidak sesuai author guidelines  
12.Judul tabel tidak sesuai format

The table shows that the significance value ( $\text{sig}$ ) 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.

### A.7 Normality Test Results for Post-Test

	Kelas	Kolmogorov-Smirnov <sup>a</sup>			Shapiro-Wilk		
		Statistic	df	Sig.	Statistic	df	Sig.
Hasil Belajar Santri	PostTest Kontrol	.236	15	.025	.846	15	.015
	PostTest Eksperimen	.292	15	.001	.746	15	.001

**Commented [MFF14]:** 13.Tabel tidak sesuai dengan model APA  
14.Font tidak sesuai author guidelines  
15.Judul tabel tidak sesuai format

According to the table, the significant value ( $\text{sig}$ ) 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 researcher applies the non-parametric Mann-Whitney test.

### b). Homogeneity Test

The purpose of using the homogeneity test in this experimental study is to determine that the researcher sets a collection of control and experimental

class data is homogeneous or heterogeneous, (Usmadi, 2020, p. 51) with the following criteria:

- a. A significant value  $< 0.05$  indicates that the population's data has not the same/not homogeneous.
- b. A significant value  $> 0.05$  indicates that the population's data has the same variance/homogeneous.

### A.8

#### Homogeneity Test Results for Pre-Test

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

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

### A.9

#### Homogeneity Test Results for Post-Test

		Levene Statistic	df1	df2	Sig.
Hasil Belajar Nahwu	Based on Mean	3.360	1	28	.077
	Based on Median	1.347	1	28	.256
	Based on Median and with adjusted df	1.347	1	26.304	.256
	Based on trimmed mean	3.195	1	28	.085

According to the output table in the preceding sentence, the significance value (sig) of a Nahwu learning outcome 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  $0.077 > 0.05$ .

#### 2. Hypothesis Test

To evaluate the data, the researchers chose to use the Mann-Whitney test, a non-parametric statistical technique. The normality test's conclusion that the data were not normally distributed led to this 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 (Yanti, 2007, p. 55).

To investigate the potential impact of authentic assessment portfolio 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:

$$H_0: \mu_1 = \mu_2$$

$$H_a: \mu_1 \neq \mu_2$$

$H_a$  : There was difference between the experimental class value and the control class value.

$H_0$  : 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 (Asymp. Sig. (2-tailed)) < 0.05, then  $H_0$  is rejected and  $H_a$  is accepted.
- 2) If the significance value (Asymp. Sig. (2-tailed)) > 0.05, then  $H_0$  is accepted and  $H_a$  is rejected

#### A.10

##### Mann Whitney Test Results for Pre-Test

Test Statistics <sup>a</sup>	
	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 <sup>b</sup>
a. Grouping Variable: Kelas	
b. Not corrected for ties.	

Based on the table 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.

**A.11  
Mann Whitney Test Results for Post-Test**

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

According to the table 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 authentic assessment portfolio had an impact on the 4<sup>th</sup> grade nahwu science curriculum at the Modern Islamic Boarding School Al-Ikhlash 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.

**Commented [MFF15]:** 1.Tidak ada pembahasan sama sekali  
2.Tidak menjawab tujuan penelitian

## 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 authentic assessment portfolio 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 authentic assessment portfolio 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.

## BIBLIOGRAPHY

- Basrowi, & Suwandi. (2008). *Memahami Penelitian Kualitatif*. Jakarta, Indonesia: Rineka Cipta.
- Hajroh, S., & Raudatul Adawiyah. (2018). Kesulitan Guru dalam Mengimplementasikan Penilaian Autentik. *Jurnal PGMI*, 10.
- Hidayat, N. S. (2012). Problematika Pembelajaran Bahasa Arab. *Jurnal Pemikiran Islam*, 37.
- Kim Nguyen, T. T., & Minh Phan, H. (n.d.). AUTHENTIC ASSESSMENT: A REAL LIFE APPROACH TO WRITING SKILL DEVELOPMENT. *International Journal of Applied Research in Social Sciences*, 2.
- M. Aini. (2006). *Evaluasi dalam Pembelajaran Bahasa Arab*. Malang: Misykat.
- Mahmudi, I. (2019). RUBRIK ANALITIK PENILAIAN HASIL BELAJAR PRAKTIK PENDIDIKAN AGAMA ISLAM. *Fikrah: Journal of Islamic Education*, 2(2), 136-154.
- Majid, A. (2013). *Penilaian Autentik Proses dan Hasil Belajar*. Bandung: PT. Remaja Rosdakarya.
- Nurgiyantoro, B., & Suyata, P. (2011). Model Otentik Dalam Pembelajaran Bahasa. *Jurnal LITERA*, 10.



- Riadi, A. (2017). KOMPETENSI GURU DALAM PELAKSANAAN EVALUASI PEMBELAJARAN, 15.
- Sa'adah, N. (2019). Problematika pembelajaran nahwu bagi tingkat pemula menggunakan Arab pegon. *Jurnal Pendidikan Bahasa Arab*, 3.
- Safriyani, R. (2012). PORTOFOLIO AS AN AUTHENTIC ASSESSMENT IN TEACHING ENGLISH FOR YOUNG LEARNERS, 04.
- Sari, E. D. K., & Mahmudi, I. (2019). THE EFFECT OF AUTHENTIC ASSESSMENT AND TEACHING VARIATION TOWARD FIQH LEARNING OUTCOMES (An Experiment at Ponpes An Nida Al Islami Bekasi). *At-Ta'dib*, 14(2), 14-25.
- Sugiyono. (2018). *Metode Penelitian Pendidikan, Pendekatan Kuantitatif, Kualitatif, dan R&D* (28th ed.). Bandung: Alfabeta.
- Sukanti, S. (2010). PEMANFAATAN PENILAIAN PORTOFOLIO DALAM MENINGKATKAN HASIL BELAJAR AKUNTANSI. *Jurnal Pendidikan Akuntansi Indonesia*, 8(2). Retrieved 24 January 2023 from <https://doi.org/10.21831/jpai.v8i2.952>
- Usmadi, U. (2020). PENGUJIAN PERSYARATAN ANALISIS (UJI HOMOGENITAS DAN UJI NORMALITAS). *Inovasi Pendidikan*, 7(1). Retrieved 24 January 2023 from <https://doi.org/10.31869/ip.v7i1.2281>
- Wijayanti, M. D. (2022). STRATEGI GURU DALAM MENINGKATKAN SIKAP NASIONALISME MELALUI PEMBELAJARAN PENDIDIKAN PANCASILA DAN KEWARGANEGARAAN DI SMAN 1 NGADIROJO PACITAN, 10.
- Yanti, T. S. (2007). Uji Rank Mann-Whitney Dua Tahap. *STATISTIKA*, 7(1). Retrieved 24 January 2023 from <https://doi.org/10.29313/jstat.v7i1.954>
- أحمد, م., & القادر, ع. (1986). طرق تعليم اللغة العربية. القاهرة: مكتبة النهضة المصرية.  
داود, م. (2001). العربية وعلم اللغة الحديثة. القاهرة: دار غريب

**Commented [MFF16]:** Catatan:

1. Penulisan judul di rujukan dengan model APA 7 ditulis dengan model **Sentence case** bukan **Capitalize Each Word, UPPER CASE** atau **lainnya**. Kapital hanya digunakan untuk partikel tertentu
2. Masih ada penulisan nama author yang tidak konsisten
3. Sudahkan rujukan disitasi semua begitu juga sebaliknya?
4. Rujukan berbahasa Arab tidak ditransliterasi

# BUKTI SUBMIT REVISI

## TANGGAL 5 DESEMBER

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

to Submissions

### Notifications

<a href="#">[Iz] Editor Decision</a>	2023-11-08 09:10 AM
<a href="#">[Iz] Editor Decision</a>	2023-12-22 01:08 PM

### Reviewer's Attachments

[Q Search](#)

96238 21.docx	September 4, 2023
101574 portfolio.pdf	November 8, 2023

### Revisions

[Q Search](#) [Upload File](#)

103827 Inwan Mahmudi_Izdihar_revisi.docx	December 5, 2023	Article Text
--	------------------	--------------

### Review Discussions

[Add discussion](#)

Name	From	Last Reply	Replies	Closed
<i>No Items</i>				



Universitas Muhammadiyah Malang, East Java, Indonesia

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

p-ISSN: 2622-738X, e-ISSN: 2622-7371 // Vol. x No. y Month 20xy, pp. xx-xx



<https://doi.org/10.22219/jiz.Volxx.Nov.iii-iii>



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



[izdihar.jurnalpba@umm.ac.id](mailto:izdihar.jurnalpba@umm.ac.id)

## Impact of Portfolio Authentic Assessment on the Learning Outcomes of Nahwu at Modern Islamic Boarding School

Ihwan Mahmudi<sup>a, 1</sup>, Aisyah Imani<sup>a, 2</sup>, Chevira Yasin Indriyanti<sup>a, 3</sup> Rohani Poteh<sup>b, 4</sup>  
<sup>a</sup>Universitas Darussalam Gontor, Indonesia  
<sup>a</sup>Prince of Songkla University, Thailand

<sup>1</sup>[ihwanm@unida.gontor.ac.id](mailto:ihwanm@unida.gontor.ac.id), <sup>2</sup>[aisyahimani19@gmail.com](mailto:aisyahimani19@gmail.com),  
<sup>3</sup>[chevireputriindriyanti26@student.pba.unida.gontor.ac.id](mailto:chevireputriindriyanti26@student.pba.unida.gontor.ac.id) <sup>4</sup>[rohani.p@psu.ac.th](mailto:rohani.p@psu.ac.th) \*

### ARTICLE INFO

#### Article History:

Received:  
Revised:  
Accepted:  
Published:

#### \*Corresponding

#### Author:

Name: Ihwan Mahmudi

Email:

[ihwanm@unida.gontor.ac.id](mailto:ihwanm@unida.gontor.ac.id)

### ABSTRACT

Specifically focusing on class 4 TMI (Tarbiyah Muallimat al-Islamiyah) at the Modern Islamic Boarding School of Al-Ikhlash Female in Kuningan, West Java. The study aims to investigate the potential beneficial impact of portfolio assessment on the learning outcomes of Nahwu lessons for fourth-grade students. The finding that 4 TMI has impetus for this investigation. The target population for the study, which included all fourth graders, was carried out during the odd semester of the 2022 academic year. All students in class 4 TMI were included in the research sample thanks to the saturated sampling method that was used. A quasi-experimental design was 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. When comparing this score to that of the experimental class, which used portfolio assessment, it become clear that the pupils' comprehension of the Nahwu material had significantly improved. These results show that portfolio evaluations have a chance to improve fourth-grade students' learning outcomes in Nahwu.

Copyright © 20xy, First Author et al  
This is an open access article under the [CC-BY-SA](https://creativecommons.org/licenses/by-sa/4.0/) license



### Keyword

Authentic Assessment; Portfolio; Learning Outcome of Nahwu

### مستخلص البحث

مع التركيز بشكل خاص على الصف الرابع تربية المعلمات الإسلامية في المدرسة الداخلية الإسلامية الحديثة في الإخلاص للإناث في كونيغان، 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؛ مَلَفَ: نتائج التعلم النحو

كلمات أساسية

## 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 evaluating student 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 student 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 student 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 intend to experiment with authentic assessment portfolios in the nahwu class for fourth-graders at the Modern Islamic Boarding School of Al-Ikhlah 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 student language abilities (Cabrera-Solano, 2020; Sulisty 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 suggests that students' diverse learning goals ought to be reflective of real-life 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 student 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 lives. 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-semester students. 30 students in total, divided into 4<sup>th</sup> B and 4<sup>th</sup> C classes. All 30 of the fourth-grade students were chosen by researcher 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 is easier to provide descriptive statistics to sum up and comprehend the experimental data.

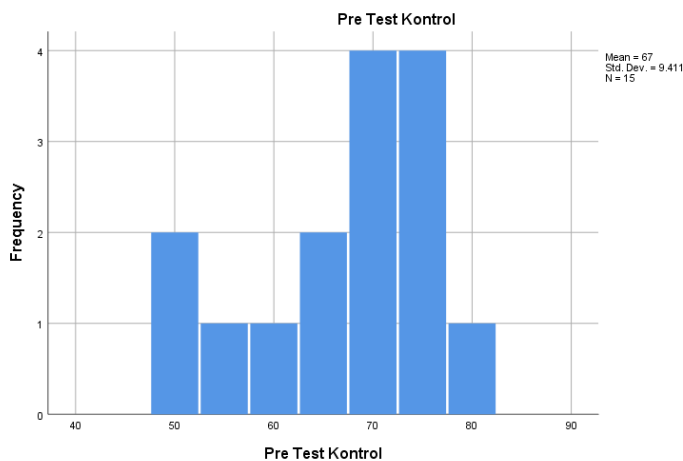
**Table 1.** Descriptive Test Result

	<b>N</b>	<b>Range</b>	<b>Min.</b>	<b>Max</b>	<b>Sum</b>	<b>Mean</b>	<b>Std. Deviation</b>	<b>Variance</b>
						<b>Statistic</b>	<b>Std.Error</b>	
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 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.

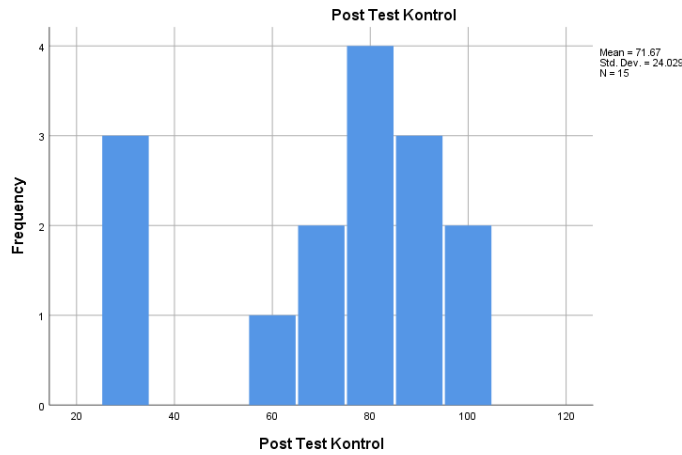
This research constructed the following histogram based on the arrangement of the descriptive test table:



**Figure 1.** Histogram of Pre-Test Control Class

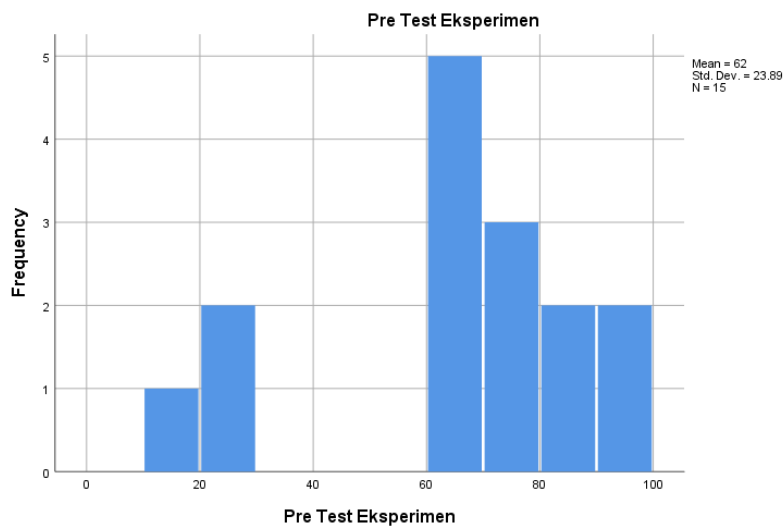


The table illustrates the consistent mastery of nahwu subject by the students in the control class, as indicated by their pre-test scores. Following the description of the descriptive test conducted for the post-test of the control class:



**Figure 2.** Histogram of Post-Test Control Class

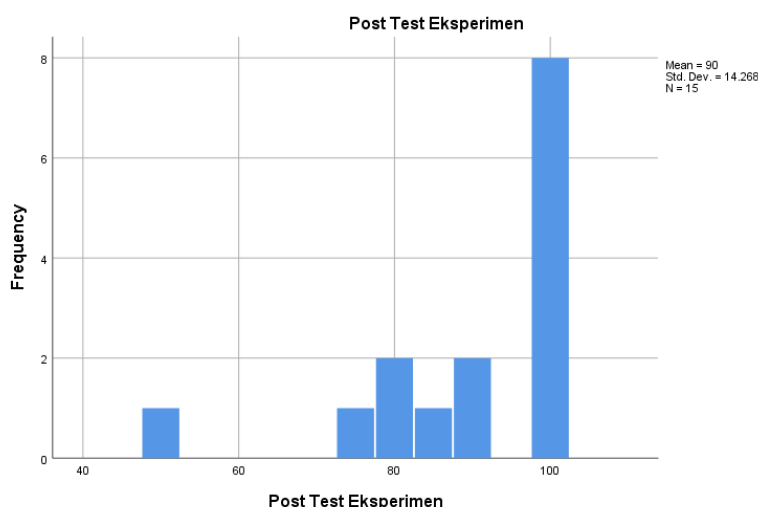
The table demonstrates the ongoing correlation between the post-test results of the control class students and their nahwu subject learning outcomes. Following the description of the descriptive test results for the control class post-test, the researcher presents the descriptive test for the experimental class pre-test as follows:



**Figure 3.** Histogram of Pre-Test Experiment Class

The table shows that the experimental class students' learning outcomes for the nahwu subject were consistent with the pre-test findings. The researcher

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 table shows that the experimental class students' learning outcomes for the nahwu subject were consistent with the post-test findings.

#### 1. Test Requirements Analysis

##### a) 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 ( $\text{sig} > 0.05$ ), it is considered to be normally distributed. If a value is significant ( $\text{sig} 0.05$ ), it is considered to be non-normally distributed.

**Table 2.** Normality Test Results of Pre-Test

	Class	Kolmogorav-Smirnov <sup>a</sup>			Sapiro-Wilk		
		Statistic	Df	Sig.	Statistic	df	Sig.
Learning Outcome	Control	.225	15	.040	.895	15	.080
	Experiment	.267	15	.005	.852	15	.018

The table shows that the significance value ( $\text{sig}$ ) 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 Results of Post-Test

	Class	Kolmogorov-Smirnov <sup>a</sup>			Sapiro-Wilk		
		Statistic	Df	Sig.	Statistic	df	Sig.
Learning Outcome	Control	.236	15	.020	.846	15	.015
	Experiment	.292	15	.001	.746	15	.001

According to the table, 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 researcher applies the non-parametric Mann-Whitney test.

b). Homogeneity Test

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

- a. A significant value  $< 0.05$  indicates that the population's data has not the same/not homogeneous.
- b. A significant value  $> 0.05$  indicates that the population's data has the same variance/homogeneous.

**Table 4.** Homogeneity Test Results for Pre-Test

		Levene Statistic	df1	df2	Sig.
Hasil Belajar Nahwu	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 in the preceding sentence, the results of the significance value (sig) based on mean  $0.031 < 0.05$  suggest that the nahwu learning outcome of the control class group and experimental class in the pre-test is heterogeneous.

**Table 5.** Homogeneity Test Results for Post-Test

		Levene Statistic	df1	df2	Sig.
Hasil Belajar Nahwu	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 in the preceding sentence, the significance value (sig) of a Nahwu learning outcome 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  $0.077 > 0.05$ .

## 2. Hypothesis Test

To evaluate the data, the researchers chose to use the Mann-Whitney test, a non-parametric statistical technique. The normality test's conclusion that the data were not normally distributed led to this 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-Ikhlah Female, the researchers aim to measure and address the research hypothesis using the following hypothesis test:

$$H_0: \mu_1 = \mu_2$$

$$H_a: \mu_1 \neq \mu_2$$

$H_a$  : There was difference between the experimental class value and the control class value.

$H_0$  : 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  $H_0$  is rejected and  $H_a$  is accepted.

- 2) If the significance value (Ashmp. Sig. (2-tailed)) > 0.05, then  $H_0$  is accepted and  $H_a$  is rejected

**Table 6.** Mann Whitney Test Results for Pre-Test

Test Statistics <sup>a</sup>	
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 <sup>b</sup>
a. Grouping Variable: Kelas	
b. Not Corrected for ties.	

Based on the table 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 Results for Pre-Test

Test Statistics <sup>a</sup>	
Hasil Belajar Nahwu	
Mann-Whitney U	53.500
Wilcoxon W	173.500
Z	-2.508
Asymp.Sig. (2-tailed)	.012

Exact Sig. [2*(1-tailed Sig.)]	.013 <sup>b</sup>
--------------------------------	-------------------

a. Grouping Variable: Kelas
--------------------------------

b. Not Corrected for ties.
----------------------------

According to the table 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 4<sup>th</sup> grade nahwu science curriculum at the Modern Islamic Boarding School Al-Ikhlash 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 shows that PPA can significantly influence student 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.

## 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.
- 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*, IX(2), 351–377.

- 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.
- 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).
- 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.
- 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.
- 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.
- 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 Non Paper-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.

## **BUKTI ACCEPTED**

[Jlz] Editor Decision External Inbox x



Moh. Fery Fauzi <noreply-ejournal@umm.ac.id>

Fri, Dec 22, 2023, 1:08 PM



to me, Aisyah, Chevira ▾

Ihwan Mahmudi, Aisyah Imani, Chevira Yasin Indriyanti:

We have reached a decision regarding your submission to Izdihar : Journal of Arabic Language Teaching, Linguistics, and Literature, "Impact of Authentic Assessment Portfolio on The Learning Results of Nahwu Science 4th grade TMI in the Modern Islamic Boarding School of Al-Ikhlash Female Kuningan West Java".

Our decision is to: Accept Submission

## **BUKTI PUBLISH**

Dashboard 0 About ▾ Current Archives Announcements Statistic Lockss Q Search

Home / Archives / Vol. 6 No. 3 (2023): Izdihar: Journal of Arabic Language Teaching, Linguistics, and Literature / Articles

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

**Ihwan Mahmudi**  
Universitas Darussalam Gontor, Indonesia

**Aisyah Imani**  
Universitas Darussalam Gontor, Indonesia

**Chevira Yasin Indriyanti**  
Universitas Darussalam Gontor, Indonesia

**Rohani Pooteh**  
Prince of Songkla University, Thailand

DOI: <https://doi.org/10.22219/jiz.v6i3.27441>

**Keywords:** Nahwu Science, Portfolio Authentic Assessment, Nahwu Learning

READ PDF 161

Published  
2024-01-04

How to Cite  
I. Mahmudi, A. Imani, C. Yasin Indriyanti, and R. Pooteh, "Impact of Portfolio Authentic Assessment on Students' Learning Outcomes of Nahwu at Modern Islamic Boarding School", *Jlz*, vol. 6, no. 3, pp. 279–292, Jan. 2024.

More Citation Formats ▾

- New Submission
- Editorial Board
- Peer Reviewers
- Focus and scope
- Publication Ethics
- Plagiarism Policy
- Open Access Policy
- Copyright & License Notice
- Peer Review Process
- Author Guideline activate

