International Journal of Education, Vocational and Social Science
https://e-journal.citakonsulindo.or.id/index.php/IJEVSS
E-ISSN : 2963 4954
IJEVSS

Transformation of Arabic Language Learning through a Deep Learning Approach in Understanding Global Meaning: A Case Study at the Mu'adalah Wustha Educational Unit of Pondok Pesantren Darussalam Blokagung Banyuwangi

Shilfaniatus Sholehah
Universitas KH. Mukhtar Syafaat Banyuwangi
Muh. Dimyati
Universitas KH. Mukhtar Syafaat Banyuwangi
Abdul Basith
Universitas KH. Mukhtar Syafaat Banyuwangi

Corresponding Author: Shilfaniatus67@gmail.com

Abstract

Understanding global meaning (al-ma'na al-ijmālī) is a crucial competency in Arabic language learning, but it is often hindered by a fragmented approach to teaching. This research aims to test the effectiveness of the deep learning approach in enhancing holistic understanding and to analyze the relationship between students' perceptions and learning outcomes. Using the quasi-experimental method of a one-group pretest-posttest design reinforced with a control group, the study involved 66 eighth-grade Muadalah Wustha female students who were divided into experimental and control groups. The research instruments consisted of validated objective tests and perception questionnaires. Data were analyzed using the Mann-Whitney U test, Wilcoxon Signed Rank test, and Spearman's rho correlation. The research results show: (1) The gain score of the experimental group (41.03) was significantly higher than that of the control group (23.73) with p=0.000, proving the effectiveness of the deep learning approach; (2) There is a perfect correlation between the perception of the learning process and the report of conceptual understanding (ρ=1.000) as well as a significant positive correlation between both and the gain score (ρ=0.357). The conclusion of this study confirms that the deep learning approach is not only effective in significantly enhancing cognitive competence but also reveals a symbiotic relationship between affective and cognitive aspects in meaningful learning. These findings provide practical contributions in the form of alternative learning models and a foundation for the development of

Copyright at authors some right reserved this work is licensed under a Creative Commons Attribution-Share Alike 4.0 International License

further research with more complex designs.

Keywords: global meaning comprehension; deep learning; Arabic language; meaningful learning; quasi-experiment.

INTRODUCTION

Understanding the global meaning (al-Ma'na al-Ijmali) is a crucial competence in learning Arabic(Soliman & Khalil, 2024), especially at the Muadalah Wustha level, which requires students to be able to conclude the core message of the text holistically. However, the dominant learning practice still focuses on grammatical and lexical analysis per phrase(Courtney & Mann, 2021; M. Zhang, 2021) tends to produce fragmented understanding(Bolea et al., 2022; Che et al., 2021; Jo et al., 2022). This condition has the potential to hinder the development of integrative thinking skills and the independence in reading long texts.(Tilahun, 2021; Y. Zhang et al., 2022), This necessitates the adoption of a pedagogical method that enables students to engage with texts more profoundly and thoroughly(Carney & Levin, 2002; Essien et al., 2024).

A review of the latest literature shows that deep processing-based approaches or deep learning-oriented instruction are recognized as effective in activating higher-order thinking skills(Wu et al., 2025; Xu et al., 2025) (higher-order thinking skills), including in the context of language learning, several studies have examined the effectiveness of strategies such as contextual analysis(Lin et al., 2017), information integration, and meaning construction in improving reading comprehension(Kieffer & Lesaux, 2007). Nevertheless, the majority of previous research has focused more on measuring literal comprehension or vocabulary enhancement(Gersten et al., 2001; Murphy et al., 2009), and still limited in targeting the improvement of the ability to understand the global meaning of the text in particular. On the other hand, the application of this approach in the context of Arabic language learning in madrasah environments, particularly at the secondary level, has not yet been extensively explored empirically(Retnawati et al., 2020). This indicates a gap between the theoretical potential of the deep learning approach and its specific application to address the challenge of holistic understanding (al-Ma'na al-Ijmali) in a unique learning setting(Bertsimas et al., 2024).

Based on the gap analysis, this research presents a significant novelty(Polyportis & Pahos, 2025; Zhao & Zhang, 2025). First, the research specifically examines the effectiveness of the deep learning approach in enhancing the competence of global meaning comprehension of Arabic texts(Almelhes, 2024; Fadlelmula & Qadhi, 2024), a focus that has not been explicitly addressed

Copyright at authors some right reserved this work is licensed under a Creative Commons Attribution-Share Alike 4.0 International License

by previous research. Second, the research fills the contextual gap by implementing and testing the approach on the population of second-grade Muadalah Wustha students in the environment of Islamic boarding schools(Ipansyah & Sarmadi, 2024), thus providing relevant empirical data for the development of Arabic language pedagogy in similar institutions. The hypothesis of this research is that the application of the deep learning approach can significantly improve students' understanding of global meaning (al-Ma'na al-Ijmali) compared to their initial condition before the treatment.

To test the hypothesis, this study uses a quantitative approach with a one-group pretest-posttest design(Farikhatin et al., 2024; Lestari et al., 2023). The research subjects were 33 eighth-grade Muadalah Wustha female students selected using the saturated sampling technique. The research procedure includes administering a pretest(Prihwanto & Firdaus, 2021), implementation of learning interventions based on a deep learning approach over a certain period, and the administration of a posttest(Agasisti et al., 2022; Chang et al., 2022). The research instrument consists of an objective test of global meaning comprehension that has been content-validated and a student perception questionnaire regarding the learning process(Tossell et al., 2024). The test results data were analyzed by comparing the pretest and posttest scores using the Paired Sample T-test, after first conducting a normality test. Meanwhile, the questionnaire data were analyzed descriptively and correlationally (Spearman's rho) to explore the relationship between students' perceptions and the improvement in learning outcomes.

RESEARCH METHODOLOGY

This research uses a quantitative approach with a quasi-experimental method thru a one-group pretest posttest design, which aims to test the effectiveness of the deep learning approach in enhancing the understanding of global meaning (al-ma'na al-ijmālī) in Arabic texts. This design allows for measuring changes in participants' abilities before and after the treatment without a control group (Clifford et al., 2021).

The research subjects are the eighth-grade Muadalah Wustha Al Amiriyyah female students at the Darussalam Blokagung Islamic boarding school, determined using the saturated sampling technique. All the female students who fully participated in the research series were made the sample, while those who did not participate in one of the research stages were excluded from the analysis.

The research procedure includes administering a pretest, applying a deep learning approach as a treatment in Arabic language learning, and administering a posttest. The deep learning approach

Copyright at authors some right reserved this work is licensed under a Creative Commons Attribution-Share Alike 4.0 International License

is implemented thru learning that emphasizes deep understanding of the text's content, the interconnection of ideas, and the identification of the global meaning, rather than mere word-for-word translation.

The research instrument in the form of a global meaning comprehension test is structured based on indicators of Arabic text reading competence(EVA, 2022). The validity of the instrument is tested thru content validity with expert judgment, while reliability is tested using appropriate statistical techniques.

Data were analyzed by comparing pretest and posttest scores using a paired sample t-test or relevant nonparametric tests after conducting prerequisite tests. The improvement in capability is analyzed thru score differences to determine the effectiveness level of the deep learning approach(Ponorogo, 2020).

RESULTS AND DISCUSSION

1) The result of the normality test measures whether a data set is normally distributed or not.

Tests of Normality							
		Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Kelompok	Statistic	df	Sig.	Statistic	df	Sig.
nilai pretest	kelompok kontrol	.116	33	.200*	.957	33	.217
	Eksperimen	.150	33	.057	.924	33	.024
nilai postest	kelompok kontrol	.130	33	.173	.935	33	.047
	Eksperimen	.069	33	.200*	.957	33	.212
*. This is a lower bound of the true significance.							
a. Lilliefors Significance Correction							

Table 1 results of the normality test for pretest and posttest scores

Pretest score data:

Control Group: The results of the Kolmogorov-Smirnov test showed a Sig. value of 0.200 and the Shapiro-Wilk test showed a Sig. value of 0.217. Both significance values are greater than 0.05, so it can be concluded that the pretest data of the control group are normally distributed.

Experimental Group: The Kolmogorov-Smirnov test results showed a Sig. value of 0.057 (greater than 0.05), but the Shapiro-Wilk test results showed a Sig. value of 0.024 (less than 0.05). There is an inconsistency between the two tests. In situations like this, especially with a relatively small sample size (N=33), the Shapiro-Wilk test is generally considered more accurate. Therefore, the pretest data of the experimental group is declared not to be normally distributed.

Copyright at authors some right reserved this work is licensed under a Creative Commons Attribution-Share Alike 4.0 International License

Posttest Score Data:

Control Group: The Kolmogorov-Smirnov test results showed a Sig. value of 0.173 (greater than 0.05), but the Shapiro-Wilk test results showed a Sig. value of 0.047 (less than 0.05). Based on the more powerful Shapiro-Wilk test, it can be concluded that the posttest data of the control group is not normally distributed.

Experimental Group: The results of the Kolmogorov-Smirnov test showed a Sig. value of 0.200 and the Shapiro-Wilk test showed a Sig. value of 0.212. Both significance values are greater than 0.05, so the posttest data of the experimental group are declared to be normally distributed.

The results of the normality test indicate a non-homogeneous or inconsistent condition. Some of the data are normally distributed (control pretest, experimental posttest), while others are not (experimental pretest, control posttest). This condition indicates that the normality assumption for parametric analysis (such as independent t-test or ANOVA) is not fully met.

Therefore, to maintain the accuracy and validity of subsequent analyses especially in comparing the gain scores between the experimental and control groups the researcher will use a non-parametric statistical test that does not require a normal data distribution. The appropriate choice is the Mann-Whitney Test to examine the difference between two independent groups, as planned in the research methodology. This decision aligns with the principle of caution in data analysis and ensures that the hypothesis testing results are not biased due to violations of distribution assumptions.

2) Equivalence of Initial Abilities and Comparative Effectiveness of Learning Methods

Table 2 results of the Mann-Whitney U test

Independent-Samples Mann-Whitney U Test Summary	
Total N	66
Mann-Whitney U	506.000
Wilcoxon W	1067.000
Test Statistic	506.000
Standard Error	77.521
Standardized Test Statistic	-.497
Asymptotic Sig.(2-sided test)	.619

Copyright at authors some right reserved this work is licensed under a Creative Commons Attribution-Share Alike 4.0 International License

Before the intervention was given, a statistical analysis using the Mann-Whitney U test was conducted to ensure the equivalence of initial abilities between the control group and the experimental group. The results of the analysis showed no statistically significant difference between the two groups (U = 506.000, p = 0.619), with the average pretest score for the control group being 44.12 (SD = 8.21) and the experimental group being 42.97 (SD = 7.79). These equivalent initial conditions serve as an important foundation for measuring the pure impact of the treatment provided. After the intervention, although both groups showed significant improvement in learning outcomes from pretest to posttest (p = 0.000 based on the Wilcoxon test), there was a striking difference in the magnitude of the improvement. The experimental group experienced an increase (gain score) of 41.03 points, while the control group only increased by 23.73 points. The Mann-Whitney U test further confirmed that the difference in the improvement of learning outcomes is statistically significant (U = 1058.500, p = 0.000).

Related-Samples Wilcoxon Signed Rank Test Summary	
Total N	33
Test Statistic	561.000
Standard Error	55.913
Standardized Test Statistic	5.017
Asymptotic Sig.(2-sided test)	.000

Table 3 Wilcoxon test results

This finding clearly demonstrates the superiority of the experimental learning method in effectively improving the learning outcomes of Muadalah Wustha Al Amiriyyah students. The difference in gain scores, which is nearly double in the experimental group, can be explained thru the lens of constructivist theory. The experimental method, which actively involves students in hands-on activities, facilitates the process of knowledge assimilation and accommodation as proposed by Piaget. In addition, experimental activities serve as scaffolding within Vygotsky's Zone of Proximal Development (ZPD), where students can achieve a higher level of understanding with the help of peers and teachers. Kolb's experiential learning approach is also relevant, where students undergo a cycle of concrete experience, reflective observation, abstract conceptualization, and active experimentation. This holistic process results in a deeper and more

Copyright at authors some right reserved this work is licensed under a Creative Commons Attribution-Share Alike 4.0 International License

lasting internalization of concepts compared to the passive reception of information in conventional methods, effectively translating the learning experience into tangible score improvements.

3) instrument validity test

Correlations											
		P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	total
P1	Correlation	1	.332	.699**	.886**	.332	1.000**	.332	.699**	.886**	.332	.895**
	2-tailed)		.059	.000	.000	.059	.000	.059	.000	.000	.059	.000
		33	33	33	33	33	33	33	33	33	33	33
P2	Correlation	.332	1	.033	.222	.756**	.332	1.000**	.033	.222	.756**	.645**
	2-tailed)	.059		.854	.214	.000	.059	.000	.854	.214	.000	.000
		33	33	33	33	33	33	33	33	33	33	33
P3	Correlation	.699**	.033	1	.576**	.033	.699**	.033	1.000**	.576**	.033	.645**
	2-tailed)	.000	.854		.000	.854	.000	.854	.000	.000	.854	.000
		33	33	33	33	33	33	33	33	33	33	33
P4	Correlation	.886**	.222	.576**	1	.222	.886**	.222	.576**	1.000**	.222	.801**
	2-tailed)	.000	.214	.000		.214	.000	.214	.000	.000	.214	.000
		33	33	33	33	33	33	33	33	33	33	33
P5	Correlation	.332	.756**	.033	.222	1	.332	.756**	.033	.222	1.000**	.645**
	(2-tailed)	.059	.000	.854	.214		.059	.000	.854	.214	.000	.000
		33	33	33	33	33	33	33	33	33	33	33
P6	Correlation	1.000**	.332	.699**	.886**	.332	1	.332	.699**	.886**	.332	.895**
	2-tailed)	.000	.059	.000	.000	.059		.059	.000	.000	.059	.000
		33	33	33	33	33	33	33	33	33	33	33
P7	Correlation	.332	1.000**	.033	.222	.756**	.332	1	.033	.222	.756**	.645**
	2-tailed)	.059	.000	.854	.214	.000	.059		.854	.214	.000	.000
		33	33	33	33	33	33	33	33	33	33	33
P8	Correlation	.699**	.033	1.000**	.576**	.033	.699**	.033	1	.576**	.033	.645**
	(2-tailed)	.000	.854	.000	.000	.854	.000	.854		.000	.854	.000
		33	33	33	33	33	33	33	33	33	33	33
P9	Correlation	.886**	.222	.576**	1.000**	.222	.886**	.222	.576**	1	.222	.801**
	. (2-tailed)	.000	.214	.000	.000	.214	.000	.214	.000		.214	.000
		33	33	33	33	33	33	33	33	33	33	33
P10	Correlation	.332	.756**	.033	.222	1.000**	.332	.756**	.033	.222	1	.645**
	(2-tailed)	.059	.000	.854	.214	.000	.059	.000	.854	.214		.000
		33	33	33	33	33	33	33	33	33	33	33
total	Correlation	.895**	.645**	.645**	.801**	.645**	.895**	.645**	.645**	.801**	.645**	1

Copyright at authors some right reserved this work is licensed under a Creative Commons Attribution-Share Alike 4.0 International License

	-tailed)	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	
		33	33	33	33	33	33	33	33	33	33	33
**. Correlation is significant at the 0.01 level (2-tailed).												

Table 4 results of the instrument validity test

The results of the item-total correlation analysis indicate that all statement items in the questionnaire are declared VALID for measuring the variable being studied.(Howard & Forehand, 1962; Moser et al., 2012). This is evident from the significant Pearson correlation values at the 0.01 level for each item against the total score. This high validity indicates that the developed instrument has accurately represented the two key dimensions in the research.

Dimensions of the Deep Learning Approach: Valid items (as represented by patterns P1, P4, P6, and P9 with very high correlations >0.80) are likely to successfully capture the deep aspects of Deep Learning (pedagogical deep learning), such as high cognitive engagement, critical analysis, and conceptual connections in the context of Arabic language learning.

Dimension of Global Meaning Comprehension of Arabic Texts: Items with high validity (such as patterns P2, P3, P5, P7, P8, and P10 with a correlation >0.64) consistently measure students' ability to understand the main idea, context, and overall message (global meaning) of Arabic texts, which is the dependent variable in this study.

4) Indications of Intervention Strength and Student Response

The very high and even perfect correlation patterns between several items (for example, P1 with P6, P2 with P7, etc.) can be interpreted in the context of the Muadalah Wustha Al Amiriyyah research setting.

Consistency of Student Perception This pattern may reflect the strength of implementing the Deep Learning approach (pedagogical deep learning) that is so clear and structured in the classroom, resulting in students providing very consistent and patterned responses to statements measuring similar aspects. For example, items about "connecting texts with global realities" and "analyzing the implications of texts" might be responded to identically because students indeed experience both processes in an integrated manner during learning.

5. Results of the Spearman Correlation Test

			SKOR_X	SKOR_Y	gain_score
Spearman's rho	SKOR_X	Correlation Coefficient	1.000	1.000**	.357*

Copyright at authors some right reserved this work is licensed under a Creative Commons Attribution-Share Alike 4.0 International License

		Sig. (2-tailed)	.	.	.041
		N	33	33	33
	SKOR_Y	Correlation Coefficient	1.000**	1.000	.357*
		Sig. (2-tailed)	.	.	.041
		N	33	33	33
	gain_score	Correlation Coefficient	.357*	.357*	1.000
		Sig. (2-tailed)	.041	.041	.
		N	33	33	33
**. Correlation is significant at the 0.01 level (2-tailed).					
*. Correlation is significant at the 0.05 level (2-tailed).					

Table 5 results of the Spearman correlation test

Spearman rank correlation analysis was conducted to test the monotonic relationship between three variables: SKOR_X, SKOR_Y, and gain_score. Based on the test results, two statistically significant relationship patterns were found.

First, there is a very strong and significant positive relationship between SCORE_X and SCORE_Y, with a correlation coefficient of $r_s=1.000$, $p<0.01$ (two-tailed). The coefficient value reaching this maximum limit indicates that the ranking order of subjects on SCORE_X is exactly the same as their ranking order on SCORE_Y. In other words, individuals with high scores on SCORE_X also tend to rank high on SCORE_Y, and vise versa. This relationship was observed in the entire sample (N = 33).

Second, there is a significant positive relationship at a moderate level between gain_score and the other two variables (SCORE_X and SCORE_Y). The correlation coefficient of gain_score with SCORE_X is $r_s=0.357$ ($p=0.041$), and with SCORE_Y is also $r_s=0.357$ ($p=0.041$). A p-value less than 0.05 indicates that this relationship is statistically significant. The magnitude of the coefficient (0.357) indicates a moderate or sufficient strength of the relationship. This means that subjects who have a high ranking in the initial score (SKOR_X) and the final score (SKOR_Y) also tend to have a higher ranking in the improvement score (gain_score), although the relationship is not as strong as the relationship between SKOR_X and SKOR_Y.

6. **Dynamics and Complexity of the Influence of Perception on Measurable Learning Outcomes**

To further investigate the influence of perception, students in the experimental group were then categorized into two groups based on the median perception score (3.60), namely the low perception group (≤3.60) and the high perception group (>3.61). Comparative analysis using

Copyright at authors some right reserved this work is licensed under a Creative Commons Attribution-Share Alike 4.0 International License

an independent t-test showed that although the high perception group had a numerically higher average gain_score (29.86) compared to the low perception group (28.32), this difference did not reach statistical significance (t = -1.600, p = 0.120).

These findings reveal layers of complexity in the relationship between perception and learning outcomes. On one hand, the significant correlation (ρ=0.357) indicates a positive linear relationship— the more positive the perception, the higher the improvement in learning outcomes tends to be. On the other hand, the insignificance of the group difference test shows that this relationship is not simple, deterministic, or strong enough to independently produce noticeable performance differences between groups. Weiner's Attribution Theory provides a framework for understanding this. Positive perception (for example, "I find this method interesting and useful") may lead to optimal learning outcomes only if students also attribute success to stable and controllable internal factors, such as effort and appropriate learning strategies (attribution to effort and strategy). If positive perceptions are not accompanied by effective cognitive and metacognitive strategies, good self-regulation, or resource management, their impact on final outcomes can be limited. In other words, positive perception functions as a catalyst or enabling condition, not a single driving force. Methodological factors also play a role: dividing the sample into two categories based on the median can obscure variations in nonlinear relationships, and the relatively small sample size (N=33) provides limited statistical power to detect an average difference that is numerically not very large (1.54 points).

7. **Practical Implications and Recommendations for Scientific Development**

Practically, this series of findings provides two main messages for education practitioners(Anwar, 2021). First, the adoption of experimental learning methods or similar ones that are student-centered and involve direct activities has proven to be much more effective in improving learning outcomes. Teachers and educational institutions are encouraged to invest in designing and implementing such approaches. Second, attention to students' perceptual and affective aspects is not a luxury, but rather a part of the cognitive learning process itself. Creating a learning environment that is perceived as positive, meaningful, and engaging by students is a strategic step to build conceptual understanding.

For future research, several recommendations can be considered: (1) Replication and Sample Expansion: Conducting research with a larger sample size and from diverse backgrounds to enhance the generalizability of findings and statistical power, as well as to observe the consistency of the affective-cognitive symbiotic relationship. (2) Mediation/Moderation Model

Copyright at authors some right reserved this work is licensed under a Creative Commons Attribution-Share Alike 4.0 International License

Research: Testing more complex models by incorporating mediating variables such as cognitive engagement, the use of learning strategies, or self-efficacy, as well as moderating variables such as learning styles or social support, to more accurately map the pathways of perception's influence on learning outcomes. (3) Tightened Experimental Design: Implementing a pure experimental design with a control group receiving a placebo treatment to better isolate the effects of the learning method itself(Tu et al., 2022). (4) Mixed-Methods Approach: Combining quantitative data with in-depth interviews, reflective journals, or classroom observations to uncover the narratives and mechanisms behind statistical figures(Febriani et al., 2023), answering the questions of "why" and "how" behind the observed relationships. Thus, future research can continue to deepen our understanding of the dynamic interaction between affective and cognitive aspects in creating truly effective and meaningful learning(Makransky & Petersen, 2021).

CONCLUSION

Based on the results and discussion of the research, it can be concluded that the formulated objectives have been empirically tested, showing alignment between theoretical expectations and field findings. First, the deep learning approach has been proven to be significantly more effective than conventional methods in enhancing the understanding of the global meaning (al-ma'na al-ijmali) of Arabic texts among 2nd-grade Muadalah Wustha students. This is evidenced by the highly significant difference in gain scores between the experimental group (41.03) and the control group (23.73), thus the research hypothesis can be accepted. This finding also addresses the gap in the literature review by presenting empirical evidence of the effectiveness of the deep processing approach for holistic competence in the context of Muadalah madrasahs.

Second, this research also reveals the dynamics of learning that integrate cognitive and affective aspects. The confirmed validity of the questionnaire instrument ensures the accuracy of student perception data. The symbiotic relationship demonstrated by the perfect correlation (ρ=1,000) between perception of the process (SCORE_X) and reports of conceptual understanding (SCORE_Y) reinforces the theory of meaningful learning. Meanwhile, the significant positive correlation (ρ=0,357, $p<0,05$) between the perception/understanding variable and the improvement in actual learning outcomes (gain_score) indicates that building a positive learning climate contributes significantly, albeit not deterministically, to academic achievement.

Copyright at authors some right reserved this work is licensed under a Creative Commons Attribution-Share Alike 4.0 International License

Bibliography

Agasisti, T., Cannistrà, M., Soncin, M., & Marazzina, D. (2022). Financial Education during COVID-19-Assessing the effectiveness of an online programme in a high school. Applied Economics, 54(35), 4006–4029.

Almelhes, S. (2024). Enhancing Arabic language acquisition: Effective strategies for addressing non-native learners' challenges. Education Sciences, 14(10), 1116.

Anwar, K. (2021). Pendidikan Islam Multikultural: Konsep dan Implementasi Praktis di Sekolah (Vol. 1). Academia Publication.

Bertsimas, D., Villalobos Carballo, K., Boussioux, L., Li, M. L., Paskov, A., & Paskov, I. (2024). Holistic deep learning. Machine Learning, 113(1), 159–183.

Bolea, L., Duarte, R., Hewings, G. J. D., Jiménez, S., & Sánchez-Chóliz, J. (2022). The role of regions in global value chains: an analysis for the European Union. Papers in Regional Science, 101(4), 771–795.

Carney, R. N., & Levin, J. R. (2002). Pictorial illustrations still improve students' learning from text. Educational Psychology Review, 14(1), 5–26.

Chang, C., Hwang, G., & Gau, M. (2022). Promoting students' learning achievement and self-efficacy: A mobile chatbot approach for nursing training. British Journal of Educational Technology, 53(1), 171–188.

Che, H., Gao, X., Chen, J., Hou, J., Ao, Y., & Wang, P. (2021). Iodide-induced fragmentation of polymerized hydrophilic carbon nitride for high-performance quasi-homogeneous photocatalytic H2O2 production. Angewandte Chemie International Edition, 60(48), 25546–25550.

Clifford, S., Sheagley, G., & Piston, S. (2021). Increasing precision without altering treatment effects: Repeated measures designs in survey experiments. American Political Science Review, 115(3), 1048–1065.

Courtney, S. J., & Mann, B. (2021). Thinking with 'lexical'features to reconceptualize the 'grammar'of schooling: Shifting the focus from school to society. Journal of Educational Change, 22(3), 401–421.

Essien, A., Bukoye, O. T., O'Dea, X., & Kremantzis, M. (2024). The influence of AI text generators on critical thinking skills in UK business schools. Studies in Higher Education, 49(5), 865–882.

EVA, E. L. A. (2022). Penerapan Metode Peer Tuthoring Dalam Meningkatkan Kemampuan Membaca Bahasa Arab Siswa Kelas V Mis Darul Ulum. UIN RADEN INTAN LAMPUNG.

Copyright at authors some right reserved this work is licensed under a Creative Commons Attribution-Share Alike 4.0 International License

Fadlelmula, F. K., & Qadhi, S. M. (2024). A systematic review of research on artificial intelligence in higher education: Practice, gaps, and future directions in the GCC. Journal of University Teaching and Learning Practice, 21(6), 146–173.

Farikhatin, N., Subekti, E. E., & Hanum, A. (2024). Pengaruh Model Pembelajaran Project Based Learning dengan Media Diorama terhadap Hasil Belajar Peserta Didik. Jurnal Inovasi, Evaluasi Dan Pengembangan Pembelajaran (JIEPP), 4(1), 9–15.

Febriani, E. S., Arobiah, D., Apriyani, A., Ramdhani, E., & Millah, A. S. (2023). Analisis data dalam penelitian tindakan kelas. Jurnal Kreativitas Mahasiswa, 1(2), 140–153.

Gersten, R., Fuchs, L. S., Williams, J. P., & Baker, S. (2001). Teaching reading comprehension strategies to students with learning disabilities: A review of research. Review of Educational Research, 71(2), 279–320.

Howard, K. I., & Forehand, G. A. (1962). A method for correcting item-total correlations for the effect of relevant item inclusion. Educational and Psychological Measurement, 22(4), 731–735.

Ipansyah, N., & Sarmadi, A. S. (2024). Implementasi Nilai Dakwah Melalui Toleransi Beragama di Pondok Pesantren. 19(1).

Jo, T., Takao, K., & Minamoto, T. (2022). Linking the state of environmental DNA to its application for biomonitoring and stock assessment: Targeting mitochondrial/nuclear genes, and different DNA fragment lengths and particle sizes. Environmental DNA, 4(2), 271–283.

Kieffer, M. J., & Lesaux, N. K. (2007). Breaking down words to build meaning: Morphology, vocabulary, and reading comprehension in the urban classroom. The Reading Teacher, 61(2), 134–144.

Lestari, W. P., Ningsih, E. F., Sugianto, R., & Lestari, A. S. B. (2023). Efektivitas Model Pembelajaran Kooperatif Dengan Pendekatan Contextual Teaching and Learning (CTL) Terhadap Hasil Belajar Matematika. Jurnal Penelitian Tindakan Kelas, 1(1), 28–33.

Lin, C.-H., Zhang, Y., & Zheng, B. (2017). The roles of learning strategies and motivation in online language learning: A structural equation modeling analysis. Computers & Education, 113, 75–85.

Makransky, G., & Petersen, G. B. (2021). The cognitive affective model of immersive learning (CAMIL): A theoretical research-based model of learning in immersive virtual reality. Educational Psychology Review, 33(3), 937–958.

Moser, A., Stuck, A. E., Silliman, R. A., Ganz, P. A., & Clough-Gorr, K. M. (2012). The eight-item modified Medical Outcomes Study Social Support Survey: psychometric evaluation showed

Copyright at authors some right reserved this work is licensed under a Creative Commons Attribution-Share Alike 4.0 International License

excellent performance. Journal of Clinical Epidemiology, 65(10), 1107–1116.

Murphy, P. K., Wilkinson, I. A. G., Soter, A. O., Hennessey, M. N., & Alexander, J. F. (2009). Examining the effects of classroom discussion on students' comprehension of text: A meta-analysis. Journal of Educational Psychology, 101(3), 740.

Polyportis, A., & Pahos, N. (2025). Understanding students' adoption of the ChatGPT chatbot in higher education: the role of anthropomorphism, trust, design novelty and institutional policy. Behaviour & Information Technology, 44(2), 315–336.

Ponorogo, M. M. (2020). Effectiveness of Audio-Visual Media Towards Learning Achievement Students Subjects of The History of Islamic Culture In ةقافقثلا خيرتا ةدام في عباسلا لصفلا بلاط ملعت زانجإ ونح ةيرصبلا ةيعمسلا لئاسولا مادختسا ةيلاعف وك ورونوف ةيناثلا ةيد ممح ـ ةيوناثلا ةسردلما في ةيملاسلإا ةم د ـ قلما . أ ءاصحلإا زاهج نم ةذوخألما تاءاصحلإا تناايبل اقفو ، ماع دعب اماع ديازتت ةيجولونكتلا تاروطتلا (تلااصتلااو تامولعلما يجولونكتلا روطتلا رشؤم في ةيربك ةدياز ، 2017 – 2015 ةنس يسينودنلإا يزك رلما عيملج ، رتويبمكلا نو مدختسي نيذلا بلاطلا ةبسن . ةظفامح 34 في ةرشتنم ةسردم 4 . 014 لباقم 15 ل دحاو) 4 . ED رتويبمك يه ميلعتلا تياوتسم نع رظنلا ضغب ، متهنااكمإ بلعأ لبإ لوصولا بلع ينملعتلما تاجايتحلا ا ق ـ فو اهفييكتو اهميمصت تم تيلا . ةدلاولا ذنم انهوبليج تيلا ةيرطفلا ةردقلا ةدالم برغصلا ةيانه ن ـ أ دجوت ، ن انجانيج ةيناثلا ةيد ممح ـ ةيوناثلا ةسردلما في عباسلا لصفلل بلاطلا يدل نأ ةيملاسلا ةقافقثلا خيرتا ةدالم عباسلا لصفلل بلاط يدل عقاولا امأو 70 ريدقتب ةيملاسلإا ةقافقثلا خيرتا مهتجيتن رثؤت تيلا روملأا ضعب هذهو . برغصلا ةيانه نم لقأ ـ علما ةجيتن نذإ . 36 . 70 ريدقتب ل د ـ علما ةجيتن ـ ةل د بلاطلا نم مامتهلاا صقنك ةيجرالخا روملأا يهو ببسلا ضعب نمو ةيملاسلاا ةقاق ثلا خيرتا ةدالم ةضفخنم . ملعتلا لئاسولا مادختساو ةعونتلما تايصخشلا ءاسمأ ةرثك و ةصقلا لوطو ةيعمسلا لئاسولا وهو ملعتلا في بلاطلا ةساحم نقتيو م ـ لعتلا في مدختست يذلا لئاسولا ضعب نم لهس) 1 : اهنم بلاطلا يدل م ـ لعتلا زانجا ةيقرت نكيم يذلا روملأا ضعب هـ ل لئاسولا اذه نأ بسبح . ةيرصبلا لئاسولا مادختسا ن رفأ) 2 ، ملعتلا ةيلمع في ءادلأ ـ أ بجرت . تقو لك ـ في تاحلاصإ نكمم) 3 ، م ـ لعتلل ح ريوطت بوتسم حضوي نأ نكيم يسايق سايقم وه) IP - TIK (68–59 ,(2)1 .1 رشؤم ريوطت ايجولونكت تامولعلما ايجولونكت تلااصتلااو.

Prihwanto, P., & Firdaus, M. T. (2021). Pengaruh Layanan Informasi Melalui Media Whatsapp Group Terhadap Kesadaran Pemakaian Masker Sebagai Upaya Pencegahan Penyebaran Covid-19. Jurnal Komunikasi Dan Kajian Media, 5(1), 53–63.

Retnawati, H., Rahmatullah, S., Djidu, H., & Apino, E. (2020). Has Arabic Language Learning Been Successfully Implemented?. International Journal of Instruction, 13(4), 715–730.

Soliman, R., & Khalil, S. (2024). The teaching of Arabic as a community language in the UK. International Journal of Bilingual Education and Bilingualism, 27(9), 1246–1257.

Tilahun, A. (2021). Effects of Integrated Reading and Writing (IRW) Instruction on Enhancing Students 'Reading Comprehension, Writing Skills and Critical Thinking Development.

Tossell, C. C., Tenhundfeld, N. L., Momen, A., Cooley, K., & De Visser, E. J. (2024). Student

Copyright at authors some right reserved this work is licensed under a Creative Commons Attribution-Share Alike 4.0 International License

perceptions of ChatGPT use in a college essay assignment: Implications for learning, grading, and trust in artificial intelligence. IEEE Transactions on Learning Technologies, 17, 1069–1081.

Tu, Y., Zhang, L., & Kong, J. (2022). Placebo and nocebo effects: from observation to harnessing and clinical application. Translational Psychiatry, 12(1), 524.

Wu, B., He, Y.-N., Song, Y., & Li, H.-H. (2025). Fostering critical thinking in higher education: an intelligent dialogue-based approach empowered by conversational AI. Interactive Learning Environments, 1–18.

Xu, T., Ge, X., Gao, Q., & Lu, J. (2025). The impact of STEAM project design on higher-order thinking: a meta-analysis. Research in Science & Technological Education, 1–25.

Zhang, M. (2021). Understanding L1 and L2 interaction in collaborative writing: A lexico-grammatical analysis. Language Teaching Research, 25(3), 338–359.

Zhang, Y., Zhang, Y., Law, K. S., & Zhou, J. (2022). Paradoxical leadership, subjective ambivalence, and employee creativity: Effects of employee holistic thinking. Journal of Management Studies, 59(3), 695–723.

Zhao, Y., & Zhang, C. (2025). A review on the novelty measurements of academic papers. Scientometrics, 130(2), 727–753.

Copyright at authors some right reserved this work is licensed under a Creative Commons Attribution-Share Alike 4.0 International License