

THE EFFECTIVENESS OF ENGLISH FOR SPECIFIC PURPOSES (ESP) VOCABULARY TEACHING TO ENHANCE VOCABULARY MASTERY: A STUDY OF ELEVENTH - GRADE STUDENTS AT MA NURUL IMAN MONTONG BAIK

Arifudin Sultan^{1✉}, Liana Opinia², M. Hilmi Syauqani³, Wasilatul Afiah⁴

(1) English Education Department, Universitas Hamzanwadi

(2) English Education Department, Universitas Hamzanwadi

(3) English Education Department, Universitas Hamzanwadi

(4) English Education Department, Universitas Hamzanwadi

✉ sultan120796@gmail.com

Abstrak

Rendahnya penguasaan kosakata bahasa Inggris pada siswa sekolah kejuruan/madrasah seringkali disebabkan oleh ketidakrelevanan materi General English dengan kebutuhan akademik siswa. Penelitian ini bertujuan untuk mengukur efektivitas penerapan pendekatan English for Specific Purposes (ESP) dalam meningkatkan penguasaan kosakata siswa Kelas XI jurusan Bisnis di MA Nurul Iman Montong Baik. Penelitian ini menggunakan metode kuantitatif dengan desain Pre-Experimental (One-Group Pretest-Posttest Design). Subjek penelitian dipilih menggunakan teknik sampling jenuh, yang terdiri dari 26 siswa Kelas XI. Instrumen penelitian berupa tes kosakata yang telah divalidasi ahli, dan data dianalisis menggunakan Paired Sample T-Test serta Normalized Gain (N-Gain) dengan bantuan SPSS 27. Hasil analisis menunjukkan adanya peningkatan skor rata-rata yang signifikan dari 45.00 pada pre-test menjadi 75.38 pada post-test. Uji hipotesis menunjukkan nilai signifikansi (*Sig. 2 – tailed*) sebesar 0.000 ($p < 0.05$), yang berarti hipotesis alternatif diterima. Selain itu, perhitungan N-Gain menghasilkan skor rata-rata 0.56 yang masuk dalam kategori efektivitas “Sedang”. Penelitian ini menyimpulkan bahwa pengajaran kosakata berbasis ESP efektif dalam meningkatkan kompetensi siswa karena materi yang disajikan relevan dengan jurusan mereka, sehingga mempermudah proses retensi memori dan motivasi belajar.

Kata Kunci: *English for Specific Purposes* (ESP), Penguasaan Kosakata, *Business English*, Siswa Madrasah Aliyah, Efektivitas Pembelajaran.

Abstract

The low English vocabulary mastery among vocational/Islamic high school students is often caused by the irrelevance of *General English* materials to the students' academic needs. This research aims to evaluate the effectiveness of the *English for Specific Purposes* (ESP) approach in improving the vocabulary mastery of eleventh-grade Business students at MA Nurul Iman Montong Baik. This study employed a quantitative method with a Pre-Experimental design (*One-Group Pretest-Posttest Design*). The subjects were selected using saturated sampling

Article History

Received: December 2025

Reviewed: December 2025

Published: December 2025

Plagiarism Checker No
234.872.737

Prefix DOI : Prefix DOI :
10.8734/argopuro.v1i2.365

Copyright : Author

Publish by : Argopuro



This work is licensed under
a [Creative Commons
Attribution-NonCommercial
4.0 International License](https://creativecommons.org/licenses/by-nc/4.0/)

technique, consisting of 26 eleventh-grade students. The instrument used was a content-validated vocabulary test, and the data were analyzed using *Paired Sample T-Test* and *Normalized Gain* (N-Gain) with SPSS 27. The results showed a significant increase in the mean score from 45.00 in the pre-test to 75.38 in the post-test. The hypothesis test revealed a significance value (*Sig. 2 – tailed*) of 0.000 ($p < 0.05$), indicating that the alternative hypothesis was accepted. Furthermore, the N-Gain calculation resulted in an average score of 0.56, which falls into the “Medium” effectiveness category. This study concludes that ESP-based vocabulary teaching is effective in enhancing students’ competence because the materials presented are relevant to their major, thereby facilitating memory retention and learning motivation.

Keyword: *English for Specific Purposes* (ESP), Vocabulary Mastery, Business English, Islamic Senior High School Students, Learning Effectiveness.

INTRODUCTION

English has established itself as the pivotal medium for global communication and a vital instrument in professional, scientific, and commercial domains. In the context of the global economy, proficiency in this language has become a core competency required for secondary education graduates to remain competitive. Furthermore, mastery of English is no longer merely an academic requirement but a prerequisite for accessing information and technology in the digital era (Kumar, 2023). Among all language components, vocabulary constitutes the fundamental building block that underlies successful communication. Without an adequate lexical repertoire, students’ productive skills (speaking and writing) and receptive skills (listening and reading) will be fundamentally hindered. Low vocabulary mastery directly limits students’ ability to comprehend instructions and participate actively in the learning process (Machfudi & Afidah, 2022).

In alignment with current educational demands, the Indonesian Government has implemented the *Kurikulum Merdeka* (Merdeka Curriculum), which emphasizes relevant, in-depth, and contextual learning (Kartika, 2024). This curriculum demands that students not only understand language theory but also apply it according to their interests and vocational fields. However, the idealism of this curriculum often conflicts with the reality of students’ basic abilities in the field. This discrepancy is clearly evident at MA Nurul Iman Montong Baik. Based on preliminary observations, it was found that the vocabulary mastery level of Grade XI students falls into the very low or minimal category. This condition serves as a primary barrier in the teaching and learning process, where standard English curriculum materials are difficult to absorb because students do not understand the basic vocabulary used in texts or teacher instructions.

The lack of vocabulary mastery creates a serious psychological impact in the form of learning demotivation. Students tend to feel “hopeless” as they perceive English as a subject that is too difficult and abstract (Triastuti, 2021). The failure to understand simple texts renders them passive in class, ultimately forming a recurring cycle of failure. Therefore, a new approach is required to not only teach the language but also restore students’ confidence with materials they perceive as “manageable” and relevant to their major’s identity (Tomlinson, 2023). This issue is exacerbated by limited human resources; the school only has one English teacher who must cover all grade levels. This high workload results in the absence of additional

programs such as extracurricular activities (English Club), which should serve as a venue for students to enrich their vocabulary outside formal lesson hours. Consequently, students' exposure to English is strictly limited to classroom hours.

For a single teacher with a maximum workload, teaching "General English," which has a vast scope, presents a distinct challenge as learning targets often become blurred and overly broad (Christison & Murray, 2021). In this situation of limitations, a more tactical and focused teaching strategy is necessary. By limiting the teaching scope to specific vocabulary relevant to the major, the teacher can allocate limited energy and time to ensure students' deep understanding, rather than pursuing extensive material that is only understood superficially. The condition in Grade XI also has specific characteristics; this class is a Business major (Business Economics) consisting of 26 students. Low mastery of General English makes students feel that English lessons are irrelevant and tedious. Therefore, a teaching strategy is needed to build a vocabulary foundation while simultaneously increasing student motivation by connecting English materials to their major.

Furthermore, as a religious-based educational institution (*Madrasah Aliyah*), students at MA Nurul Iman Montong Baik face a denser curriculum compared to general high schools due to intensive religious subjects. Consequently, the time available for learning English becomes extremely valuable. Conventional approaches that require long periods to memorize general vocabulary are often ineffective in this environment. An efficient method is needed where every vocabulary item learned has immediate value for their vocational studies, thereby optimizing learning time. As a solution, the researcher proposes the *English for Specific Purposes* (ESP) approach. ESP is a teaching approach where materials and methods are tailored to the specific needs of the learner's field of study.

Various prior studies have examined the implementation of ESP. For instance, Sari and Wirza (2021), investigated the specific target needs of Office Authorization and Governance students, identifying a disconnect between General English instruction and the specific speaking and writing skills required for future job markets. Similarly, Hajar and Triastuti (2021), analyzed the needs of Agriculture students, advocating for the blending of English materials with agricultural topics like plant cultivation to enhance communication skills. However, their research focused primarily on macro-level needs analysis and industrial material suitability for purely vocational schools (SMK). Consequently, less attention has been paid to the cognitive mechanisms of vocabulary acquisition for students with low basic competence outside of the vocational school environment. Cognitively, teaching vocabulary through an ESP approach facilitates students in the retention process because the taught vocabulary resides within the same "semantic field," namely business. When students learn words such as profit, demand, supply, and transaction within a single economic story context, their brains form stronger connections compared to memorizing random words. This is predicted to accelerate the improvement of Grade XI students' vocabulary mastery compared to conventional methods.

Other research by Thi Ngu et al. (2021), examined the use of English to teach economics at the university level through complex case studies, such as analyzing stock prices and financial market determinants. Although this method helps expand discussion via open-mind methods, it presumes a level of cognitive maturity and linguistic proficiency that creates a gap if applied directly to Madrasah Aliyah students who are still struggling with basic terminology. On the other hand, recent studies by Sadikova (2025), same as Rachmawati and Irawan (2024) focused on enhancing ESP vocabulary through advanced digital ecosystems, including AI-based platforms and Learning Management Systems (LMS), specifically within economic contexts. While effective in improving learner engagement and retention, these studies highlight a reliance on digital literacy and appropriate resource selection. This presents a technological gap, as high-technology implementation is difficult to execute in schools with limited facilities, such as the

research location. Based on the review above, an empirical gap exists. There is a scarcity of research testing the effectiveness of ESP vocabulary teaching specifically on Madrasah Aliyah (MA) Business students with the conditions of a single teacher resource and limited facilities, utilizing direct face-to-face methods without reliance on paid digital applications.

The selection of Grade XI students as research subjects is based on academic urgency. Grade XI represents a crucial transition phase before students face final exams in Grade XII. If vocabulary reinforcement intervention is not conducted at this stage, it is feared that students will enter the final level without sufficient language foundation, which will hinder their graduation and competitiveness post-graduation. Therefore, this research is not merely an academic experiment but an effort to salvage students' basic competence. The novelty of this research lies in the application of a simplified ESP vocabulary teaching intervention focused on the Business Economics syllabus to build a basic foundation. This study aims to provide a practical solution for teachers in schools with similar limitations so that learning remains effective in accordance with the *Kurikulum Merdeka*.

Given the limited population size, this research utilizes a *Quasi-Experimental One-Group Pretest-Posttest* design on 26 Grade XI students. This method was chosen to accurately measure whether there is a gain in vocabulary mastery before and after the implementation of the ESP method. Based on the urgent background of the problem and the proposed solution, the research problem is formulated as: "How effective is *English for Specific Purposes* (ESP) Vocabulary Teaching in improving the vocabulary mastery of Grade XI students at MA Nurul Iman Montong Baik?". Consequently, the objective of this research is to measure and analyze the effectiveness of vocabulary mastery improvement among Grade XI students at MA Nurul Iman Montong Baik after the implementation of ESP Vocabulary Teaching, thereby serving as a recommended learning model for schools with similar characteristics.

METHODS

A. Research Design

This study employed a quantitative approach utilizing a pre-experimental research design. Specifically, the *One-Group Pretest-Posttest Design* was selected for this investigation. The selection of this design aligns with the research objective to assess the effectiveness of an intervention on a single subject group without a control group, where assessments are conducted both prior to and following the treatment (Santika, 2024). The design pattern is illustrated as follows:

$$O_1 \ X \ O_2 \quad \quad \quad (\text{MSDM, n.d., Sugiyono (2019)})$$

Description:

- O_1 : Pre-test (Initial vocabulary assessment before treatment).
- X : Treatment (Implementation of ESP vocabulary teaching focusing on Business Economics)
- O_2 : Post-test (Final vocabulary assessment after treatment).

B. Participants and Setting

The research was conducted at MA Nurul Iman Montong Baik, East Lombok, spanning from October to November 2025. The population comprised all eleventh-grade students majoring in Business (Business Economics).

Table 1 Timelines of the Research

Activities	OCTOBER				NOVEMBER			
Week	1	2	3	4	1	2	3	4
A. PREPARATION								
1. Title Proposal & School Permission	X							
2. Developing Lesson Plan & ESP Module		X						
3. Constructing Vocabulary Test Items		X						
4. Instrument Validation (Expert)			X					
5. Revision & Duplication			X					
B. IMPLEMENTATION								
6. PRE-TEST				X				
7. Treatment 1: <i>Basic Economic Terms</i>					X	X		
8. Treatment 2: <i>Business Activities</i>					X	X		
9. Treatment 3: <i>Job & Profession</i>					X	X		
10. Treatment 4: <i>Review & Games</i>					X	X	X	
11. POST-TEST							X	
C. REPORTING								
12. Data Analysis (Scoring & SPSS)								X
13. Article Writing								X
14. Finalization & Submission								X

The implementation of activities, focusing on the *English for Specific Purposes* (ESP) teaching intervention, was rigorously planned over a two-month period, specifically October and November. The entire process was divided into three main phases. The Preparation Phase commenced in October (Weeks 1-3), beginning with securing school permission and title submission. This was followed by the development of the Lesson Plan (RPP), ESP module, and vocabulary test instruments (Week 2). These instruments subsequently underwent expert validation and revision (Week 3) before duplication.

Upon the completion of preparation, the Implementation Phase began in late October (Week 4) with the administration of the PRE-TEST to measure the subjects' baseline abilities. Subsequently, four Treatment sessions were conducted intensively during Weeks 1 to 3 of November. The treatment covered specific topics such as Basic Economic Terms, Business Activities, and Job & Profession, concluding with a Review & Games session. To finalize the intervention, the POST-TEST was administered in Week 3 of November to measure the resulting improvement. Finally, the Reporting Phase was

conducted in Week 4 of November, encompassing comprehensive data analysis (including scoring and SPSS processing), drafting the report or article, and finalization.

Table 2 Total Student of XI Grade MA Nurul Iman Montong Baik

XI Grade Students	
Gender	Total
Male	14
Female	12
Total	26

The subjects of this study included all eleventh-grade students enrolled at MA Nurul Iman Montong Baik. Based on the population data, the total class size consisted of 26 students, comprising 14 males and 12 females, as detailed in Table 2. Given the relatively small population size, the sampling technique employed was Saturated Sampling. According to Parsaeian et al. (2021), this technique is utilized when the entire population serves as the research sample, which is generally considered effective for populations fewer than 30 individuals. Consequently, the total number of participants in this study amounted to 26 students. The decision to utilize this technique aimed to maximize data representation, ensuring that every individual within the population was included in the research process..

C. Research Instruments

To ensure the validity of the intervention and data acquisition, this study employed two specially developed primary instruments:

1. Treatment Instrument

The instrument used to guide the intervention was a teaching module titled “*Business English for Senior High School*”. This module was developed using a dual-approach framework, combining *English for Specific Purposes* (ESP) and *English for General Purposes* (EGP). This integration was selected to establish a robust foundation in business English while simultaneously addressing the general language proficiency required at the secondary level.

The materials within the module were allocated for six sessions (2 × 45 *minutes*), covering fundamental economic topics such as *Introduction to Business English* (Unit 1), *Marketing and Advertising* (Unit 3), and *Business Ethics* (Unit 6). It incorporates specific target vocabulary highly relevant to the students’ major, including terms like *revenue*, *profit*, *entrepreneur*, and *stakeholder*. Furthermore, the module features structured lexical reinforcement activities, such as the *Business Word Map*, to facilitate vocabulary retention. Prior to implementation, the module underwent a content validity process by a practitioner expert, specifically the English Master Teacher at MA Nurul Iman Montong Baik, to ensure the material’s alignment with the curriculum and the students’ cognitive levels (Pingil, 2022). Based on the validator’s assessment, the module was declared valid and feasible for use as an intervention guide.

2. Test Instrument

The data collection instrument was a *Vocabulary Test* comprising 20 items, designed to comprehensively measure students’ business vocabulary mastery. The test

structure is divided into two main sections. Part A consists of 10 Multiple Choice Questions designed to assess students' understanding of vocabulary usage within sentence contexts (*contextual meaning*), aligning with Schmitt and Schmitt, (2020) assertion that vocabulary knowledge encompasses understanding meaning in context. Part B contains 10 Matching items (*Word Map*) to evaluate students' ability to associate business terms with their definitions. This format was directly adopted from the lexical activities within the Teaching Module, ensuring that the evaluation aligns with the learning process (Hristov et al., 2023).

To guarantee instrument quality, the researcher conducted validation and reliability checks through *Expert Judgment* by the Assistant Teacher Mentor. In this phase, the validator assessed the suitability of test items against the material indicators to ensure content validity and construct feasibility (El, 2023). This adheres to the principles proposed by Beck (2020), stating that content validity is crucial to ensure the test truly represents the taught material (*content representativeness*). Based on the expert's evaluation, the instrument was declared to have met validity and reliability requirements, making it eligible for research use without requiring further empirical statistical testing.

D. Data Collection Technique

The data collection procedure was executed through three systematic stages:

1. Pre-test (O_1)

This was administered at the first meeting prior to the provision of any treatment. Students completed a vocabulary test consisting of 20 items to obtain baseline data regarding their initial business vocabulary proficiency.

2. Treatment (X)

The researcher implemented the teaching intervention utilizing the "*Business English for Senior High School*" module over the course of 4 to 6 face-to-face sessions. The instructional method was conducted explicitly and contextually, where students were engaged to comprehend vocabulary through real business sentence contexts and interactive activities, such as job interview role-plays or business meeting simulations.

3. Post-test (O_2)

This assessment was conducted after the entire sequence of learning units within the module had been taught. Students completed a test equivalent to the pre-test to measure significant changes in their vocabulary mastery following the treatment.

E. Data Analysis Technique

The quantitative data obtained from pre-test and post-test scores were analyzed statistically using SPSS 27 software. The analysis process was conducted through three primary steps as follows:

1. Normality Test

The normality test was conducted to determine whether the data were normally distributed. Given that the sample size ($N = 26$) was below 50, the Shapiro-Wilk test was employed because it possesses better power for small samples compared to the Kolmogorov-Smirnov test (de Souza et al., 2023). Data is categorized as normally distributed if the significance value (Sig.) is greater than 0.05 ($Sig. > 0.05$).

2. Hypothesis Testing (Paired Sample T-Test)

To test the hypothesis regarding the effectiveness of the teaching, the *Paired Sample T-Test* was utilized. This test compares the mean scores of the 26 students before and after the treatment. The formula used is as follows (Sugiyono, 2019, as cited in Afifah et al., 2022).

$$t = \frac{\bar{D}}{\frac{s_D}{\sqrt{N}}}$$

Where $N = 26$, $\bar{D}$ represents the mean difference, and s_D represents the standard deviation of the difference. The alternative hypothesis (H_a) is accepted if the significance value (*Sig. 2 – tailed*) is less than 0.05.

3. N-Gain Score (Effectiveness Level)

To measure the effectiveness level of the improvement, the *Normalized Gain* (N-Gain) was calculated using the normalized gain formula (Hake, 1999, as cited in Rahman et al., 2024):

$$g = \frac{S_{post} - S_{pre}}{S_{max} - S_{pre}}$$

The calculation result g is interpreted into the following categories: $g > 0.7$ (High), $0.3 \leq g \leq 0.7$ (Medium), and $g < 0.3$ (Low).

RESULT AND DISCUSSION

This section presents the empirical findings obtained from the quantitative analysis conducted on 26 eleventh-grade Business students at MA Nurul Iman Montong Baik. The analysis commences by elucidating the initial state of the students' vocabulary proficiency (Pre-test) to establish a baseline of competence prior to the implementation of the *English for Specific Purposes* (ESP) method.

Pre-test Analysis: Initial Student Competency

Before implementing the learning intervention utilizing the ESP module, a pre-test consisting of 20 vocabulary items was administered. This assessment aimed to gauge the extent of the students' business vocabulary mastery under conventional learning conditions (*General English*). The results of the descriptive statistical calculations for the pre-test scores are summarized in Table 3.

Table 3. Descriptive Statistics of Pre-Test Score

Statistics	
Pre - Test Score	
N	Valid
	Missing
Mean	
Median	
Mode	
Std. Deviation	
Variance	
Range	
Minimum	
Maximum	

The data in Table 3 illustrates a concerning overview of the students' basic competence. The class mean score reached only 45.00. This figure is significantly below the mastery standard, with a minimum score of 30 and a maximum of only 60. This low achievement corroborates Machfudi and Afidah (2022) theory, which asserts that low vocabulary mastery fundamentally limits students' abilities, particularly in comprehending instructions and texts. In the context of MA Nurul Iman students, this low score indicates that without an adequate lexical repertoire, students experience cognitive blocking when confronting English questions, preventing their receptive skills from functioning optimally. Furthermore, the distribution of student scores can be observed through the frequency distribution in Table 4.

Table 4. Frequency Distribution of Pre-Test Score

Pre - Test Score					
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	30	1	3,8	3,8	3,8
	35	3	11,5	11,5	15,4
	40	5	19,2	19,2	34,6
	45	8	30,8	30,8	65,4
	50	5	19,2	19,2	84,6
	55	3	11,5	11,5	96,2
	60	1	3,8	3,8	100,0
Total		26	100,0	100,0	

Table 4 reveals that the majority of students (8 students or 30.8%) obtained a score of 45, which also represents the mode of the data. Cumulatively, 65.4% of students achieved scores of ≤ 45 . This phenomenon can be explained by examining the learning approach previously received by the students. As stated by Christison and Murray (2021), English proficiency is often viewed merely as a general academic requirement (*General English*) with a scope that is too broad and abstract for vocational students.

Students struggled to answer the test because the tested vocabulary (e.g., *profit*, *demand*, *revenue*) had never been learned within a cohesive "semantic field". Previously, their brains did not form strong inter-word connections because materials were taught randomly and irrelevant to their Business major. Consequently, as mentioned in the background, students perceived English as a difficult and "abstract" subject, resulting in these low pre-test scores. This empirical finding underscores the necessity of an ESP intervention which, according to Kartika (2024), must be based on specific learner needs to ensure the material is relevant and easily absorbed.

Post-test Analysis: Final Student Competency

After establishing the students' baseline abilities through the pre-test, the researcher implemented an *English for Specific Purposes* (ESP) vocabulary learning intervention over four sessions. The instruction focused on materials relevant to the students' major, such as *Business Activities* and *Job & Profession*, utilizing the developed Teaching Module. At the conclusion of the learning cycle, a post-test was administered to measure changes in students' vocabulary mastery. The summary of descriptive statistics from the post-test results is presented in Table 5.

Table 5. Descriptive Statistics of Post-Test Score

Statistics	
Post - Test Score	
N	Valid
	Missing
Mean	
Median	
Mode	
Std. Deviation	
Variance	
Range	
Minimum	
Maximum	
Sum	

The data in Table 5 demonstrates a statistically significant surge in achievement. The students' mean score increased drastically to 75.38, compared to the pre-test mean of only 45.00. Another compelling fact is observed in the shift of the minimum score; the lowest score in the post-test was 60, a figure that previously represented the *highest* score in the pre-test. Meanwhile, the maximum score reached 90. These increases in average and minimum scores indicate that the ESP intervention successfully elevated the performance of all students inclusively, including those categorized as low achievers. To observe the score distribution pattern in greater detail, the frequency distribution of the post-test scores is presented in Table 6.

Table 6. Frequency Distribution of Post-Test Score

Post - Test Score					
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	60	1	3,8	3,8	3,8
	65	3	11,5	11,5	15,4
	70	4	15,4	15,4	30,8
	75	8	30,8	30,8	61,5
	80	6	23,1	23,1	84,6
	85	3	11,5	11,5	96,2
	90	1	3,8	3,8	100,0
	Total	26	100,0	100,0	

Based on Table 6, the frequency distribution indicates a strong positive trend. The mode of the data lies at the score of 75 (30.8%), and cumulatively, 69.2% of students achieved scores above 70.

From a discussion perspective, this significant improvement validates the effectiveness of the ESP approach in addressing the issue of low vocabulary mastery. This success can be elucidated through the theory of Kartika (2024), who emphasize that language learning is effective when the taught material aligns with the target needs of the learners. In this study, economic vocabulary (such as *revenue* and *transaction*) was perceived as “meaningful” by students

because it was relevant to their major's identity, in contrast to abstract *General English* materials.

In addition to the relevance factor, the enhancement of student memory retention was also supported by contextual teaching methods. As proposed by Schmitt and Schmitt (2020), vocabulary is best learned through sentence context rather than as isolated words. The teaching module employed facilitated students in learning words within a single business "semantic field," thereby enabling their brains to form associations between concepts more easily. This aligns with the findings of Sari and Wirza (2021) as well as Hajar and Triastuti (2021), who stated that specific English materials for vocational students can increase student motivation and engagement in learning, which ultimately culminates in improved learning outcomes as evidenced by this post-test data.

Prerequisite Analysis (Normality Test Analysis)

Prior to proceeding with the primary hypothesis testing regarding the effectiveness of the ESP method, it was imperative to verify that the sample data met the fundamental statistical assumption of normal distribution. This step is critical in determining the appropriate statistical test; if the data is normally distributed, parametric statistics (*Paired Sample T-Test*) are applicable. Conversely, if the distribution is not normal, non-parametric statistics must be employed. Given that the number of participants in this study is classified as small ($N = 26 < 50$), the normality test was conducted using the *Shapiro-Wilk* technique. This technique was selected due to its superior sensitivity (*power of test*) in detecting normality within small samples compared to the Kolmogorov-Smirnov test (de Souza et al., 2023). The results of the normality test using SPSS 27 are presented in Table 7.

Table 7. Normality Test Result on Pre and Post Test (Shapiro-Wilk)

	Tests of Normality					
	Kolmogorov-Smirnov ^a			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
Pre - Test Score	0,154	26	0,116	0,962	26	0,425
Post - Test Score	0,171	26	0,049	0,959	26	0,371

a. Lilliefors Significance Correction

Based on Table 7, the analysis focuses on the significance value (*Sig.*) in the Shapiro-Wilk column. The obtained *Sig.* value for the Pre-test score is 0.425, and for the Post-test score is 0.371. According to statistical rules, data is considered normally distributed if the p -value > 0.05 . Since both significance values are well above 0.05 ($0.425 > 0.05$ and $0.371 > 0.05$), it can be statistically concluded that the students' vocabulary ability data, both before and after the treatment, are normally distributed. To corroborate the numerical test results, a visual inspection was also performed using *Normal Q-Q Plot* graphs. These graphs map the distribution of actual data against the diagonal normality line.

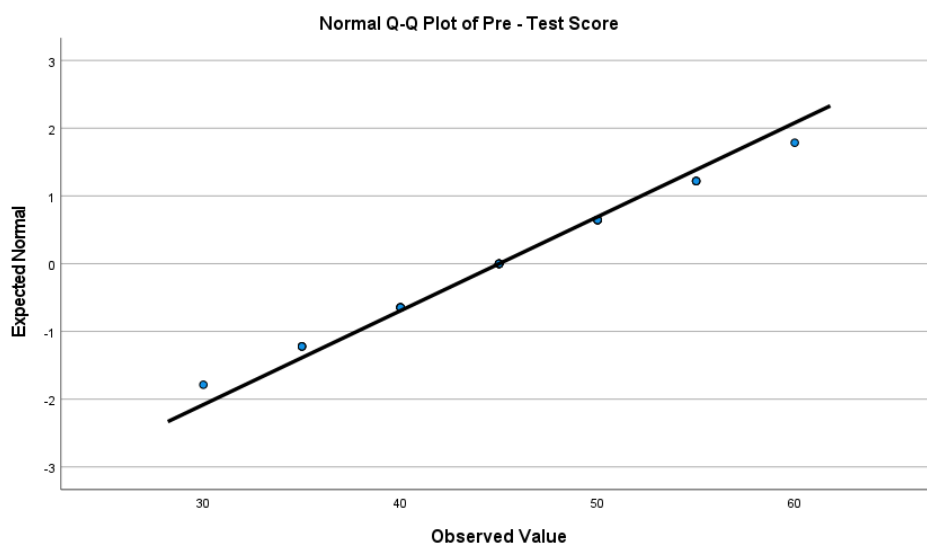


Figure 1. Normal Q-Q Plot of Pre-test Score

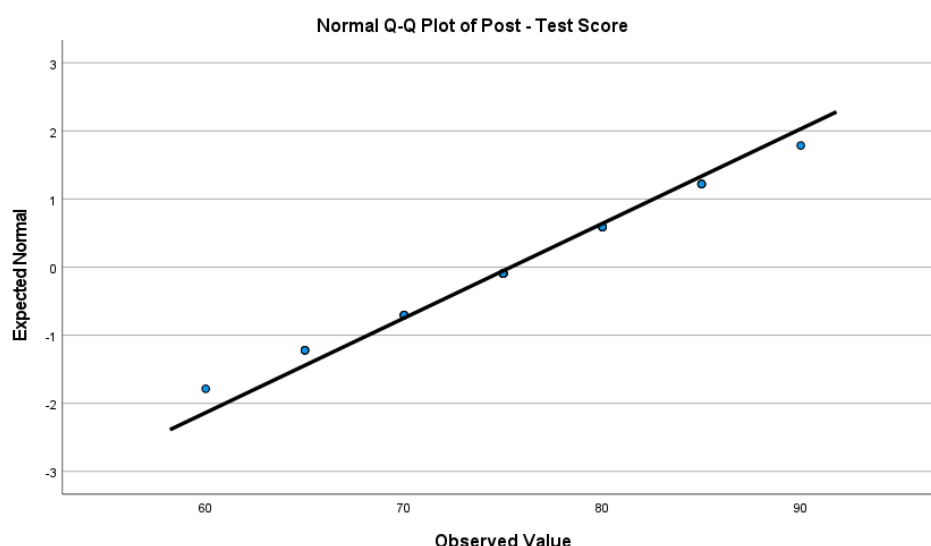


Figure 2. Normal Q-Q Plot of Post-test Score

As observed in Figure 1 and Figure 2, the distribution pattern of data points (*dots*) tends to approximate and follow the straight diagonal line. There are no extreme deviation patterns or significant outliers diverting from the reference line. The consistency between the Shapiro-Wilk statistical calculation results and the Q-Q Plot visual confirmation provides strong empirical evidence that the normality assumption has been met. The implication of this finding is that the analysis of the mean difference can safely proceed using the parametric statistical test, specifically the *Paired Sample T-Test*. Normal data ensures that the subsequent T-test results are unbiased and scientifically justifiable for drawing conclusions regarding the effectiveness of ESP learning.

Uji Hipotesis (Hypothesis Testing Analysis)

Following the confirmation that the data was normally distributed, the subsequent step involved validating the hypothesis utilizing the parametric *Paired Sample T-Test*. This test was conducted to address the research problem regarding whether the *English for Specific Purposes*

(ESP) approach significantly enhances students' vocabulary mastery. The statistical hypotheses proposed are as follows:

H_0 (Null Hypothesis): There is no significant difference in student vocabulary scores before and after ESP implementation ($Sig. > 0.05$).

H_a (Alternative Hypothesis): There is a significant difference in student vocabulary scores before and after ESP implementation ($Sig. < 0.05$).

The results of the statistical calculation using SPSS 27 are summarized in Table 8.

Table 8. Paired Sample T-Test on Pre and Post Test Result

Paired Samples Test									
		Paired Differences							
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference		t	df	Sig. (2-tailed)
					Lower	Upper			
Pair 1	PreTest Score - PostTest Score	-30,385	1,359	0,266	-30,933	-29,836	-114,027	25	0,000

Based on Table 8, the conclusion is drawn from the analysis of three primary statistical indicators. The first indicator is the Significance Value ($Sig.$), where a $Sig. (2 - tailed)$ value of 0.000 was obtained. This value is substantially lower than the significance level $\alpha = 0.05$ ($0.000 < 0.05$). In accordance with statistical decision rules, since the $Sig. < 0.05$, the Null Hypothesis (H_0) is rejected and the Alternative Hypothesis (H_a) is accepted. This decision is supported by the second indicator, the Mean Difference, which shows a value of -30.385. The negative sign indicates that post-test scores were consistently higher than pre-test scores, confirming a real improvement in student capability by an average of 30 points following the treatment.

The third indicator is observed in the comparison between the t_{count} and t_{table} . The analysis yielded a t_{count} of -114.027 (absolute value 114.027). When compared to the t_{table} at 25 degrees of freedom (df) with a 5% significance level (approximately 2.060), the t_{count} far exceeds the t_{table} ($114.027 > 2.060$). The magnitude of this $t - value$ is influenced by the high correlation between pre-test and post-test scores ($r = 0.982$), indicating that the pattern of competence improvement occurred very consistently and uniformly across every individual student, rather than being limited to a select few.

Table 9. Effect Size Sample

Paired Samples Effect Sizes

				Point Estimate	95% Confidence Interval	
Standardizer ^a					Lower	Upper
Pair 1	PreTest Score - PostTest Score	Cohen's d	1,359	-22,362	-28,525	-16,187
		Hedges' correction	1,380	-22,025	-28,095	-15,943

a. The denominator used in estimating the effect sizes.

Cohen's d uses the sample standard deviation of the mean difference.

Hedges' correction uses the sample standard deviation of the mean difference, plus a correction factor.

Overall, these statistical results prove that the implementation of the ESP method delivers a highly significant impact on improving the vocabulary mastery of Grade XI students at MA Nurul Iman Montong Baik. The rejection of H_0 confirms that the score changes were not coincidental but the direct result of the learning intervention. This effectiveness aligns with the findings of Sadikova (2025) as well as Rachmawati and Irawan (2024), who stated that the ESP approach is capable of bridging the gap between students' English proficiency and their vocational academic needs. The high t – value reflects the success of the teaching materials in accurately addressing students' cognitive needs; when students are presented with specific vocabulary relevant to their world, learning resistance decreases, and memory retention increases sharply. This finding is bolstered by the massive *Effect Size* (Cohen's d) reaching 22.362, a figure which statistically indicates a “huge effect,” demonstrating that the ESP module was proven capable of drastically transforming the students' vocabulary competence profile within a short period.

Effectiveness Level Analysis (N-Gain Score Analysis)

Although the T-test statistical analysis has proven a significant difference, this test has not yet measured the magnitude of the real effectiveness level of student ability improvement. Therefore, as the final step of data analysis, the researcher calculated the *Normalized Gain* (N-Gain) score to eliminate potential bias arising from differences in students' initial scores, ensuring the measured improvement is purely the result of the learning intervention. The summary of descriptive statistics for student N-Gain scores is presented in Table 10.

Table 10. Descriptive Statistics of N-Gain Test Score

Descriptive Statistics					
	N	Minimum	Maximum	Mean	Std. Deviation
NGain	26	0,43	0,75	0,5616	0,07662
Valid N (listwise)	26				

Based on Table 10, the calculation results show an average (*Mean*) N-Gain value of 0.5616. To interpret the meaning of this figure, the researcher utilized the effectiveness criteria developed by Hake (1999, as cited in & Halim, 2020), as presented in Table 11 below:

Table 11. N-Gain Score Interpretation Criteria

Koefisien N-Gain (g)	Kategori Efektivitas
$g > 0,7$	Tinggi
$0,3 \leq g \leq 0,7$	Sedang
$g < 0,3$	Rendah

Source: Hake (1999)

Referring to the categories in Table 11, the average value of 0.5616 falls within the range of $0.3 \leq g \leq 0.7$. Thus, the effectiveness level of using the ESP method in this study is categorized as “Medium”. This finding provides a more objective overview of the learning impact. The data indicates that the minimum N-Gain value obtained by students is 0.43. This fact is crucial as it indicates that not a single student falls into the “Low” improvement category ($g < 0.3$). All students, without exception, experienced competence improvement at a minimum of a medium level. Furthermore, there were students who successfully achieved a maximum score of 0.75, falling into the “High” effectiveness category.

From a discussion perspective, achieving this “Medium” category represents a very positive result, especially considering the relatively short intervention duration (4 meetings). As noted by (Triastuti, 2021), the main challenge of ESP teaching in vocational schools is often hindered by limited time allocation. However, this N-Gain data proves that despite limited time, the use of a focused teaching module (*focused material*) is capable of optimizing the learning process. This effectiveness occurs because the ESP approach trims irrelevant material, allowing students to focus their cognitive capacity solely on target vocabulary. The improvement confirmed by this N-Gain score underscores that the *Semantic Field* method works effectively in helping students understand business concepts holistically, rather than merely memorizing words.

CONCLUSION

Based on the data analysis and hypothesis testing, this research concludes that the implementation of English for Specific Purposes or ESP vocabulary teaching is significantly effective in improving the vocabulary mastery of the eleventh grade students at MA Nurul Iman Montong Baik. This conclusion is supported by the Paired Sample T Test result which showed a significance value of 0.000, which is lower than the significance level of 0.05. Furthermore, there was a substantial increase in the students mean score from 45.00 in the pretest to 75.38 in the posttest. This improvement confirms that the intervention successfully addressed the problem of the students low basic competence effectively.

More specifically, the effectiveness level of this learning model falls into the Medium category with an average N Gain score of 0.56. This achievement indicates that the ESP approach, which presents vocabulary materials relevant to the students major or Business Economics, is able to accelerate the memory retention process and reduce the students cognitive load compared to conventional methods. Therefore, this approach is recommended as an efficient learning model solution for schools with similar characteristics, particularly those with limited time and resources, to enhance students English competence according to their vocational needs.

However, despite the significant positive results, this study has several limitations that should be acknowledged. First, this research employed a pre-experimental design without a control group, which means other external factors beyond the treatment might have influenced the outcomes. Second, the relatively small sample size ($N = 26$) and the short duration of the

intervention (four meetings) limit the generalizability of these findings to a broader population. Therefore, future research is recommended to use a true experimental design with a control group and a larger sample size, as well as to extend the duration of the study to examine the consistency of the ESP method's effectiveness in the long term.

REFERENCES

- Afifah, S., Mudzakir, A., & Nandiyanto, A. B. D. (2022). How to calculate paired sample t-test using SPSS software: From step-by-step processing for users to the practical examples in the analysis of the effect of application anti-fire bamboo teaching materials on student learning outcomes. *Indonesian Journal of Teaching in Science*, 2(1), 81-92.
- Beck, K. (2020). Ensuring content validity of psychological and educational tests-the role of experts. *Frontline Learning Research*, 8(6), 1-37.
- Christison, M., & Murray, D. E. (2021). *What English language teachers need to know Volume III: Designing curriculum*. Routledge.
- de Souza, R. R., Toebe, M., Mello, A. C., & Bittencourt, K. C. (2023). Sample size and Shapiro-Wilk test: An analysis for soybean grain yield. *European Journal of Agronomy*, 142, 126666.
- El, A. A. M. M. A. (2023). The role of educational initiatives in EFL teacher professional development: a study of teacher mentors' perspectives. *Heliyon*, 9(2).
- Hajar, I., & Triastuti, A. (2021). A Need Analysis of ESP Materials for Agriculture Students at Indonesian Vocational School in West Sulawesi. *Indonesian Journal of English Language Teaching and Applied Linguistics*, 6(1), 13-26.
- Hristov, S., Nakov, D., & Miočinić, J. (2023). Constructive alignment between objectives, teaching and learning activities, student competencies and assessment methods in higher education. *Journal of Agriculture and Plant Sciences*, 21(2), 21-36.
- Junina, I., & Halim, A. (2020). The effect of discovery learning-based worksheet on students' metacognition skill and learning outcomes. *Journal of Physics: Conference Series*, 1460(1), 12100.
- Kartika, A. (2024). Unveiling the Merdeka Curriculum in Indonesia: Insights from Educators and Policymakers on Its Effectiveness and Implications. *INJURY: Journal of Interdisciplinary Studies*, 3(11), 802-819.
- Kumar, R. (2023). English language and digital literacy: Navigating the information age. *Journal of International English Research Studies (JIERS)*, ISSN: 3048-5231, 1(2), 25-31.
- Machfudi, M. I., & Afidah, A. (2022). Students' difficulties in vocabulary mastery. *Critical Review of English-Arabic World Journal*, 1(1), 1-12.
- MSDM, I. I. (n.d.). Sugiyono.(2019). Metode Penelitian Kuantitatif, Kualitatif, dan R&D. Bandung: CV. Alfabeta. Sianipar, R., & Salim, V.(2019). Faktor Etos Kerja Dan Lingkungan Kerja Dalam Membentuk 'Loyalitas Kerja'Pegawai Pada Pt Timur Raya Alam Damai. Anuar, S. 2019. Th. *E-Jurnal Manajemen Universitas Udayana*, 8(6), 3674.
- Parsaeian, M., Mahdavi, M., Saadati, M., Mehdipour, P., Sheidaei, A., Khatibzadeh, S., Farzadfar, F., & Shahraz, S. (2021). Introducing an efficient sampling method for national surveys with limited sample sizes: application to a national study to determine quality and cost of healthcare. *BMC Public Health*, 21(1), 1414.
- Pingil, N. P. (2022). Development and validation of instructional modules in enhancing the study and thinking skills of learners. *Psychology and Education: A Multidisciplinary Journal*, 3(9),

1.

- Rachmawati, D. L., & Irawan, N. (2024). Integrating digital technologies in ESP classroom: A comprehensive overview of current practices and pedagogical implications. *International Journal of Pedagogical Language, Literature, and Cultural Studies (i-Plural)*, 1(3), 27-35.
- Rahman, Y. A., Yeni, F., Apriyanti, F., & Habibah, F. A. F. (2024). Measuring the effectiveness of process writing approach to improve EFL students' writing proficiency via normalized gain. *Scope: Journal of English Language Teaching*, 9(1), 566-573.
- Sadikova, K. (2025). TEACHING ENGLISH VOCABULARY THROUGH DIGITAL TECHNOLOGY. *Педагогика и Психология в Современном Мире: Теоретические и Практические Исследования*, 4(9), 108-109.
- Santika, A. A. (2024). *THE EFFECTIVENESS OF ANAGRAM TECHNIQUE IN TEACHING VOCABULARY MASTERY (A Pre-Experimental Research to the Eighth Grade Students of SMPN 5 Sungai Kunyit in the Academic Year of 2022/2023)*. IKIP PGRI PONTIANAK.
- Sari, F. P., & Wirza, Y. (2021). Needs analysis of English for specific purposes in vocational high school level. *Thirteenth Conference on Applied Linguistics (CONAPLIN 2020)*, 188-194.
- Schmitt, N., & Schmitt, D. (2020). *Vocabulary in language teaching*. Cambridge university press.
- Thi Ngu, D., Huong, D. T., Huy, D. T. N., Thanh, P. T., & Dongul, E. S. (2021). Language teaching application to English students at master's grade levels on history and macroeconomic-banking management courses in universities and colleges. *Journal of Language and Linguistic Studies*, 17(3), 1457-1468.
- Tomlinson, B. (2023). *Developing materials for language teaching*. Bloomsbury Publishing.
- Triastuti, A. (2021). Voicing the challenges of ESP teaching: Lessons from ESP in non-English departments. *Studies in English Language and Education*.