



Analysis of Students' Core Skills at Grade XI of SMK N 1 Angkola Timur Towards British Council EnglishScore

Satria SB Pohan¹, Rahmat Huda², Shoufi Nisma Dewi³, Happy Sri Rezeki Purba⁴, Devita Sari Harahap⁵, Siti Rohmadhona⁶, Dyah Supraba Lastari⁷, Nasibah Awae⁸

^{1,2,3,4,5}English Education Study Program, Universitas Muhammadiyah Tapanuli Selatan, Indonesia.

⁷Education, Universiti Muhammadiyah Malaysia, Malaysia.

⁸Masjid Taluban School, Sai Buri, Pattani, Thailand.

Article Info

Article history:

Received: 14 July 2025

Revised: 30 August 2025

Accepted: 29 September 2025

Keywords:

Grammar

Vocabulary

Reading

Listening

British Council EnglishScore



Abstract:

This quantitative study sought to evaluate the English competence of grade XI software engineering students at SMK N 1 Angkola Timur and to compare their school-assessed grades with an internationally recognized norm. The research was motivated by the rising need for English proficiency in the global workforce and the possible subjectivity of internal school assessments. The main objectives were to assess the students' proficiency levels according to the Common European Framework of Reference for Languages (CEFR), analyze their distribution across these levels, and investigate the correlation between their report card grades and their British Council EnglishScore (BCES) results. The research comprised 16 students who participated in the BCES Core Skills assessment, evaluating grammar, vocabulary, reading, and listening abilities. The analysis employed descriptive statistics. The results indicated a substantial disparity: All students achieved scores at the foundational Pre-A1 (37.5%) and A1 (62.5%) levels, resulting in a mean BCES score of 122.06. Conversely, their average report card grade was 86.56, with some children attaining grades as high as 90 when assessed at the lowest Pre-A1 level. Significantly, no student attained the requisite minimum score to access the writing and speaking test modules, highlighting a profound inadequacy in fundamental skills. The report suggests that the school's internal grades are an unreliable measure of actual English competence and advocates for an immediate curriculum reform.

Corresponding Author: Huda, rahmat.huda@um-tapsel.ac.id

1. Introduction

In the age of globalization and the Fourth Industrial Revolution, the expectations for the quality of human resources are escalating. Effective communication in English is becoming an essential, obligatory skill (Husna & Nisa', 2021). For vocational high school graduates entering the industry, English proficiency is a crucial advantage for competitiveness in both national and international labor markets. English functions as a lingua franca across numerous industrial, technological, and business domains; thus, strong competence facilitates greater access to information, employment prospects, and professional collaboration (Abimanto et al., 2023). A lingua franca is a language employed

for communication among individuals from diverse linguistic backgrounds (Iriance, 2018).

Corresponding to this increasing requirement, techniques for assessing and appraising English language competency have likewise progressed. Assessments have evolved from old paper-based assessments to digital platforms that are more efficient, accessible, and globally standardized. The British Council EnglishScore is a program that is currently receiving significant recognition. This assessment employs a contemporary approach to evaluate fundamental competencies, encompassing grammar, vocabulary, reading, and listening, while also offering distinct evaluations for productive abilities like speaking and writing. Utilizing the Common European Framework of Reference (CEFR), EnglishScore delineates an individual's competence level in accordance with international norms.

Despite the evident demand for English proficiency and the existence of contemporary assessment techniques, there is a deficiency in data concerning the real preparation and ability profiles of vocational high school students, particularly those in rural areas. SMK Negeri 1 Angkola Timur, a vocational education institution in South Tapanuli Regency, confronts this difficulty directly. As of now, no targeted research has delineated the proficiency levels of pupils at the school utilizing internationally standardized assessment instruments such as EnglishScore.

A score of 90 on the report card indicates the student's proficiency in English. According to the Capaian Pembelajaran (CP) for English phase F (grades XI and XII of vocational high schools) as outlined by the Kurikulum Merdeka and analyzed using Google Gemini, this score corresponds to a minimum of B2, while certain components are beginning to manifest at the B1 level. The curriculum specifies that upon completion of this phase, students should 'utilize English to communicate on various familiar topics in academic and professional contexts' and 'generate a range of texts, both oral and written, that exhibit an understanding of purpose and audience. This description strongly indicates a targeted proficiency level of at least B1 (Intermediate) on the Common European Framework of Reference for Languages (CEFR), wherein a user can manage most situations likely to occur while traveling in a region where the language is spoken and can generate simple connected text on familiar or personally relevant topics.

The perspective, significant issues emerge: what is the student's true capability when evaluated against international benchmarks? Does a score of 90 correspond to a B2 (proficient) level or only an A2 (basic) level according to worldwide standards? Atip Latipulhayat, Deputy Minister of Primary and Secondary Education of Indonesia 2024, stated, "Report card grades tend to be subjective" (Tempo.co, 2025). In light of this gap, the researcher seeks to ascertain the answers to those inquiries. Consequently, an evaluation utilizing international criteria is necessary.

This study is to conduct a comprehensive analysis of the proficiency in core abilities among XI-grade students at SMK Negeri 1 Angkola Timur, utilizing the British Council EnglishScore test and examining its correlation with report card grades. This approach aims to yield valid empirical data on student ability profiles and offer constructive input

for educators and institutions. This information can serve as a foundation for assessing the efficacy of current teaching methods, pinpointing areas for enhancement, and formulating more focused learning strategies to equip students for global competition.

2. Literature Review

British Council EnglishScore, or BCES, is a digital tool intended to assess users' English language proficiency swiftly and precisely (British Council, 2024). This program, created by the British Council, is extensively utilized by people, educational institutions, and corporations to evaluate English proficiency. EnglishScore is the sole four-skills assessment administered by mobile phone, rendering it a handy and accessible recognized English examination.

EnglishScore utilizes a synthesis of artificial intelligence, adaptive assessment, and human knowledge to evaluate English proficiency. It employs artificial intelligence for evaluating speech and writing, along with adaptive technology to modify the test's complexity according to the user's performance. It additionally depends on human specialists for the creation and adjustment of the scoring model, as well as for delivering customized feedback. (British Council, 2024).

The test enables test takers to assess their English proficiency in listening, speaking, reading, and writing, as well as their knowledge of grammar and vocabulary (Ira Irzawati et al., 2025). This study concentrated on evaluating students' core skills (listening, reading, grammar, and vocabulary) and initially intended to include an assessment of their writing and speaking abilities as well. Nevertheless, it could not be evaluated due to a significant constraint in the architecture of the British Council EnglishScore (BCES) program.

The BCES application mandates that users attain a minimum score of 200 on the Core Skills test to access the speaking and writing test modules. During the data collection for this research, a significant result was that no student participant met the minimum criteria. As a result, it was technologically unfeasible to conduct the Speaking Test. The evaluation of speaking and writing abilities was excluded from this study not due to the researcher's preference, but because the students' foundational Core abilities performance fell short of the necessary standard to engage with that segment of the assessment.

The strength of the EnglishScore test is rooted in its advanced technological framework. The platform utilizes a synergistic integration of Artificial Intelligence (AI), adaptive testing algorithms, and human expertise to guarantee an accurate and trustworthy evaluation of English competence. It utilizes AI for the automated evaluation of productive skills, including speaking and writing, by examining factors such as pronunciation, fluency, grammar, and lexical range. The Core Skills test employs adaptive technology that modifies the question difficulty in accordance with the user's immediate performance. This guarantees that the assessment is customized to each person's proficiency, yielding a more accurate evaluation. The comprehensive evaluation system is devised and refined by professionals in language assessment, who also assist in generating customized feedback for learners (British Council, 2024).

The assessment aims to deliver a comprehensive evaluation, enabling participants to assess both their overarching English competencies (listening, speaking, reading, and writing) and the foundational elements that support them (grammar and vocabulary) (Ira Irzawati et al., 2025). This study concentrated on evaluating students' core skills: listening, reading, grammar, and vocabulary. The initial idea to evaluate writing and speaking skills could not be implemented due to a significant constraint in the design of the BCES application.

The application requires users to attain a minimum score of 200 on the Core Skills test to access the speaking and writing courses. During the data collection phase, it was found that no student participant met the minimum criterion. This technological limitation rendered the administration of the speaking and writing assessments unfeasible. The exclusion of these skills from the study was not a methodological decision but rather a direct result of the students' basic proficiency failing to achieve the necessary standards to engage with those exam components.

The primary score is reported on a point-based scale from 0 to 599 for the Core Skills test. This scale is divided into 100-point bands, each correlating directly to a CEFR level. For example, a score between 200 and 299 corresponds to CEFR Level A2, while a score of 300-399 corresponds to B1. This granular point system allows test-takers and institutions to understand proficiency not just as a discrete level but as a position along a continuum within that level.

Beyond the overall score and CEFR level, the EnglishScore certificate provides a detailed breakdown of performance. For the Core Skills test, this includes individual scores for Grammar, Vocabulary, Reading, and Listening. For the productive skills tests, such as Speaking and Writing, the certificate provides sub-scores for specific competencies like Language (accuracy and range), Organisation (cohesion and structure), and Communication (clarity and task achievement). This diagnostic detail is a key feature of the assessment, enabling learners and educators to pinpoint specific areas of strength and weakness.

Table 1. British Council EnglishScore Test Suite Overview

Component	Duration	Skills Assessed	Task Format	Scoring Mechanism
Core Skills Test	< 40 minutes	Grammar, Vocabulary, Reading, Listening	Multiple-choice, gap-filling, sentence rearrangement, comprehension questions.	Computer-adaptive; automated AI scoring; results are immediate.
Writing Test	30 minutes	Language accuracy (spelling, grammar),	Dictation, short-answer questions, image	Hybrid model: AI for language and organization; peer rating for

		lexical range, organization, comprehensibility	description, simulated chat, peer grading.	communication and task achievement.
Speaking Test	~20 minutes	Pronunciation, fluency, communication skills	Read aloud, listen and repeat, answer questions, peer grading. ⁹	Hybrid model: AI for pronunciation and fluency; human rating for communication skills.

EnglishScore tests the knowledge of English grammar and vocabulary, as well as you're reading and listening skills. The test mainly uses multiple-choice questions with three possible answers. Questions are on topics that can be from daily life, like going to the supermarket, or from the workplace, such as conversations you may have in the office.

3. Method

Creswell (2014) noted that Research is the process of developing assertions and then refining or abandoning some of them for other claims more strongly warranted. Purwanza (2022) in the other hand, believes that Research is an activity with the objective of finding, recording, analyzing and preparing reports of outcomes. In general, research is an endeavor or attempt to answer questions and solve current problems.

This study uses a quantitative method, which tries to measure and interpret data numerically. According to Creswell (2014) "Quantitative research is an approach for testing objective theories by examining the relationship among variables. These variables, in turn, can be quantified normally on instruments so that numbered data can be examined using statistical procedures." Aliaga and Gunderson (2002) describe quantitative research as "explaining phenomena by collecting numerical data that are analyzed using mathematically based methods (in particular statistics)." This method was chosen because it can provide an objective picture of the level of mastery of Core Skills of grade XI Software Engineering students at SMK Negeri 1 Angkola Timur based on the scores acquired from the British Council EnglishScore application.

Data collection was conducted through the implementation of tests using the British Council EnglishScore application. Students who are the research sample will be asked to download and install the application on their devices.

The students will take a Core Skills test sections. This test is conducted individually in a class. To avoid distractions during the process only eight students were tested in the classroom, and the other eight waited their turn outside the classroom. The Core Skills test have security test to prevent students from opening another apps or websites. After students complete the test, the scores obtained will be displayed on the EnglishScore application. The results of this test will be collected by researchers through screenshots or reports that can be shared from the application. The collected data will be recorded and

processed for further analysis to determine the level of students' English proficiency in the Core Skills.

The data obtained from the British Council EnglishScore test results will be thoroughly analysed using quantitative descriptive statistical techniques. Descriptive statistics summarize and organize characteristics of a data set. A data set is a collection of responses or observations from a sample or entire population. (Pritha Bhandari, 2020). This method was chosen because it is highly effective for summarizing, organizing, and describing the fundamental characteristics of the data set, which in turn provides a clear and comprehensive overview of the students' English proficiency levels. The analysis will go beyond a simple presentation of scores and will include the calculation of several key statistical measures to reveal deeper insights into the data.

The scores obtained from the students will be interpreted based on the assessment scale used in the EnglishScore application. This scale refers to the CEFR standard, which categorizes test results into certain levels, such as A1 (Beginners), A2 (Pre-Intermediate), B1 (Intermediate), B2 (Upper-Intermediate), C1 (Advanced), and C2 (Proficiency). Thus, this study can provide a clearer picture of the level of students' English ability based on the results obtained.

To answer the third research question, regarding the comparison of student report card grades with the international standard CEFR scores. A direct data comparison method will also be employed. This process involves two key steps. First, two distinct types of data will be collected for each student: their previous semester's English report card grade, which serves as the internal data, and their British Council EnglishScore (BCES) test results, which serve as the external data. The BCES test provides a numeric score (e.g., 0-599) and its corresponding Common European Framework of Reference for Languages (CEFR) level (e.g., A1, A2, B1). Second, these two data sets will be organized into a single comparison table for analysis. This master table will juxtapose each student's report card grade directly against their CEFR level, allowing for a clear and systematic analysis of the correlation, or lack thereof, between school-based performance and internationally standardized English proficiency.

4. Results

The data collection focused on the BCES Core Skills test and the students' most recent English report card grades. A notable and critical finding emerged during the BCES testing: not a single student achieved the minimum score of 200 required to unlock the Writing Test. Consequently, data for writing proficiency could not be collected. This limitation is a central part of the research results.

The first and second research questions sought to identify the English proficiency level of the students and their distribution across the CEFR scale. The analysis shows a proficiency. The data is unequivocal: 100% of the students in the sample scored at the A1 level or below. A majority (62.5%) are at the A1 (Elementary) level, while a significant portion (37.5%) are at the Pre-A1 (Foundation) level. Not a single student reached the A2 (Pre-Intermediate) level, which is generally considered the minimum functional level for

basic, independent communication. This finding points to a severe deficit in the foundational English skills of grammar, vocabulary, reading, and listening within this student group.

The research plan initially included a detailed analysis of writing skills. However, the inability of any student to achieve the prerequisite Core Skills score of 200 to unlock the Writing Test is a critical finding. This indicates that the students' foundational English skills (grammar, vocabulary, reading, listening) are not yet developed enough to produce even basic written responses that can be evaluated by the BCES system. Therefore, a direct mapping of writing proficiency was not possible.

The analysis reveals a profound disconnect between the two assessment systems. Students receiving the highest possible grades (90) from the school are testing at the lowest levels of international proficiency (Pre-A1). For example, Hapsa Siregar and Rismawarni Harahap both have report card grades of 90, yet their EnglishScore results place them at the Pre-A1 level. This means that while they are considered top performers in their class, their actual ability is limited to understanding and using only the most basic, familiar expressions.

This disparity holds true across the entire sample. There is no observable correlation between a higher school grade and a higher CEFR level within this group. The data strongly supports the research hypothesis that school grades are not a reliable indicator of real-world English proficiency as measured by a global standard.

5. Discussion

The research findings are striking and unequivocal. The English proficiency level of the XI-grade Software Engineering students at SMK N 1 Angkola Timur is at a foundational stage (Pre-A1 to A1), significantly below the expected level for students preparing to enter a technical workforce that increasingly requires English. Second finding of this study is the statistically non-significant negative association between students' report card grades and their CEFR scores. The results from both Pearson ($r = -.148$, $p = .585$) and Spearman ($\rho = -.153$, $p = .570$) studies indisputably confirm the absence of a meaningful statistical link between the two variables.

This outcome is quite informative. Not only does it fail to show a positive relationship (where higher grades would correspond to higher proficiency), but it also statistically confirms that a student's report card grade is a poor and unreliable predictor of their actual, real-world English language ability as measured by a global standard. The high grades awarded by the school appear to be utterly detached from the practical language skills assessed by the BCES. Another notable finding is the universal inability to unlock the writing test, which implies a systemic skill deficit in the essential building blocks of the language. However, the most concerning finding is the huge difference between internal and external ratings. The average report card grade of 87.19 conveys an image of success and great achievement. In truth, the BCES results show that 100% of these students are at a beginner's level (A1 or lower).

The revelation that all student participants scored at the A1 and Pre-A1 levels is particularly troubling when contrasted with the national curriculum's goals. As indicated in Chapter I, the Learning Outcomes (Capaian Pembelajaran) for Phase F of the Kurikulum Merdeka implicitly seek a competency of at least a B1 (Intermediate) level. The enormous disparity between this curricular goal and the measured reality at SMK N 1 Angkola Timur suggests a potential systemic failure in obtaining the expected educational outcomes. It means that either the curriculum is not being implemented successfully, the teaching techniques are not aligned with communicative goals, or the assessment practices are not appropriately monitoring progress toward these goals. This research throws into doubt the readiness of pupils under the current educational model to satisfy the criteria set by the national curriculum, let alone the demands of the global economy.

6. Conclusion

Based on the extensive examination of the British Council EnglishScore test results and school report card data from the XI-grade Software Engineering students at SMK N 1 Angkola Timur, the following conclusions are drawn:

1) The English language competency of XI-grade students majoring in Software Engineering at SMK N 1 Angkola Timur is at the Pre-A1 and A1 levels. No student scored an A2 (Pre-Intermediate) level or higher, suggesting a basic skill base that is critically underdeveloped for their educational stage; 2) A direct examination of writing proficiency was not possible. The students' Core Skills were insufficient to satisfy the minimum threshold necessary to access the Writing Test component of the BCES. This is an important finding, revealing that productive abilities cannot be established upon the current poor basis; 3) The results indicate a large and quantitative discrepancy between internal school assessments and standardized, internationally benchmarked competence examinations. With students averaging a high score of 87.19 on report cards while simultaneously placing at the lowest Pre-A1 and A1 tiers of the CEFR standard, it is argued that the school's grading system does not reflect students' actual ability to use English in a real-world setting; and 4) Failure to Meet National Curriculum Goals. The evaluated competence levels fell woefully short of the national curriculum's requirements. The Kurikulum Merdeka for this grade level aims a competency of at least B1, however the kids are functionally at the A1 level. This enormous difference reveals a systemic challenge in achieving the expected educational goals and preparing students for future success.

References

- Abimanto, D., Aprillina, A., & Oktavia, A. (2023). Bahasa Inggris Sebagai Lingua Franca Dalam Dunia Transportasi. *Sinar Dunia: Jurnal Riset Sosial Humaniora Dan Ilmu Pendidikan*, 2(1), 240–250. <https://doi.org/10.58192/sidu.v2i1.947>
- Alderson, J.C. 2000. *Assessing Reading*. Cambridge University Press, Cambridge.
- Aliaga, M., & Gunderson, B. 2002. *Interactive Statistics*. Thousand Oaks, CA: Sage.
- Anton, O., & Yodi. 2021. Predicting English language learners' proficiency level using EnglishScore android application. *Journal of Digital Ecosystem for Natural Sustainability (JoDENS)* 1 (2): hal. 106-110
- Aqil, D, F. 2021. The Importance Of English For Informatics Engineering Student. Institut Teknologi Sepuluh Nopember (ITS), Surabaya. Article in *English in Education*.
- Bhandari, P. 2020. Descriptive Statistics | Definitions, Types, Examples. (<https://www.scribbr.com/statistics/descriptive-statistics/>, diakses pada 30 juli 2025)
- Bhat, Adi. n.d. Data Collection Methods: Types & Examples. (<https://www.questionpro.com/blog/data-collection-methods/>, Diakses pada 11 Maret 2025).
- Birajdar M. Laxmi. 2020. Application of CEFR (Common European Framework of Reference) in Testing the English Language Proficiency Level of Polytechnic Students, 62 (1): hal. 7-11
- British Council. 2025. English Score Core Skills Test Guide. (<https://www.englishscore.com/resources/test-preparation>, Diakses pada 11 Februari 2025)
- British Council. 2025. Understand Your English Level. (<https://learnenglish.britishcouncil.org/english-levels/understand-your-english-level>, Diakses pada 30 April 2025)
- Cambridge. 2025. International language standards. (www.cambridgeenglish.org/exams-and-tests/cefr/, Diakses pada 30 April 2025)

- Creswell, J. W. 2014. *Research Design: Qualitative, Quantitative and Mixed Methods Approaches* (4th ed.). Thousand Oaks, CA: Sage.
- Creswell, J. W. 2012. *Educational research: Planning, conducting, and evaluating quantitative and qualitative research* (4th ed.). Boston, MA: Pearson.
- Grabe and Stoller. 2013. *Teaching and Researching Reading*. New York:Routledge.
- Harmer, J. 2001. *The Practice of English Language Teaching*. London:Longman Group Ltd.
- Husna, E. U., & Nisa', K. (2021). Strategi Pembelajaran Program Bahasa Arab dan Bahasa Inggris Untuk Meningkatkan Daya Saing Alumni Di Era Revolusi Industri 4.0 Di Pondok Pesantren Mamba'us Sholihin Suci Manyar Gresik. *JoEMS (Journal of Education and Management Studies)*, 4(1), 51–58. <https://doi.org/10.32764/joems.v4i1.355>
- Hyland, K. 2003. *Second Language Writing*. Cambridge: Cambridge University Press.
- Ira Irzawati, Agnes Felisya Unamo, & Sopian Sopian. 2025. Scoring English Proficiency Through Englishscore. *International Journal of Education, Language, Literature, Arts, Culture, and Social Humanities*, 3(1), 123–129. <https://doi.org/10.59024/ijellacush.v3i1.1276>
- Iriance. 2018. Bahasa Inggris Sebagai Bahasa Lingua Franca dan Posisi Kemampuan Bahasa Inggris Masyarakat Indonesia Diantara Anggota MEA. *Prosiding Industrial Research Workshop and National Seminar*, 9(0), 776–783. <https://jurnal.polban.ac.id/index.php/proceeding/article/view/1149/944>
- L. Vandergrift & C. M. Goh 2012. *Teaching and Learning Second Language Listening: Metacognition in Action*. New York: Routledge.
- Mardianti, D. L. 2025. Kementerian Pendidikan Tegaskan Tak Ada Pungutan Biaya dalam TKA. *Tempo.co*. (<https://www.tempo.co/politik/kementerian-pendidikan-tegaskan-tak-ada-pungutan-biaya-dalam-tka-2058087>, diakses pada 12 Agustus 2025).
- Marpurdianto, K., & Dharma, T. E. Y. 2023. English Ability Students Faculty of Economics and Business, Wijaya Kusuma University in The Competition of The World of Work.

- Neuman, S. B., & Dwyer, J. 2009. Missing in action: Vocabulary instruction in pre-k The Reading Teacher, 62(5), 194-392.
- Nunan, D. 2003. Practical English Language Teaching. International Edition, McGraw-Hill, Singapore, 88.
- Nurzaman, M. A. W. 2021. The Importance of English for the Future of Engineering Students. Sepuluh Nopember Institute of Technology.
- Pang, Elizabeth S et al. 2003. Teaching Reading. IBE, Publications Unit, P.O. Box 199, 1211 Geneva 20, Switzerland.
- Ur, P. 1996. A course in Language Teaching: Practice and Theory. Cambridge University Press, Cambridge.
- Purwanza, Sena Wahyu et al. 2022. Metodologi Penelitian Kuantitatif, Kualitatif, dan Kombinasi. Bandung, Jawa Barat: Media Sains Indonesia.
- Rayner, Keith et al. 2012. Psychology of Reading. New York: Psychology Press.
- Rost, Michael. 2011. Teaching and Researching Listening. London, United Kingdom: Pearson Education Limited.
- Schmitt, Norbert. 2000. Vocabulary in Language Teaching. United Kingdom: The Press Syndicate of The University of Cambridge.
- Thornbury, S. 1999. How to Teach Grammar. Pearson Education Ltd.