

## Test Review

### EF Standard English Test (EF SET)

Pongsiri Kamkankaew\*

Test Information	
<b>Test Name:</b>	EF Standard English Test (EF SET)
<b>Price:</b>	Free
<b>Author(s) and Publisher:</b>	EF Education First
<b>Purpose:</b>	To enable non-native English speakers to accurately assess their own English language proficiency.
<b>Contact Information:</b>	Headquarters Address: Selnaustrasse 308001 Zürich Tel: 043 430 41 18 Email: <a href="mailto:info@efset.org">info@efset.org</a> . Website: <a href="https://www.efset.org/">https://www.efset.org/</a>

---

\* Faculty of Business Administration, North-Chiang Mai University, Chiang Mai, Thailand

Reviewer: Kpongsiri85@gmail.com

## Background

The EF Standard English Test (EF SET) was initially created by EF Education First (EF) to serve as an internal tool for assessing language learners within their programs. The goal was to develop a high-quality, free assessment for adults (16 years and older) that could evaluate English skills, specifically listening and reading, and be comparable to expensive, high-stakes tests such as IELTS and TOEFL. Development began with input from language assessment experts, including Lyle Bachman, whose work provided foundational principles (Bachman & Palmer, 2010), and Mari Pearlman, who helped translate theory into practice. The process followed a structured design, starting with an Assessment Use Argument (AUA) and a Design Statement to define the test's rationale and intended use. An extensive 15-month trial period involving thousands of students across EF schools began in May 2012. These trials allowed for data collection and refinement of the test tasks and design through detailed psychometric analysis, including Rasch modeling. This iterative process led to the final test specifications and the decision to use computer-adaptive multi-stage testing (ca-MST) for delivery. The official EF SET was launched publicly in September 2014.

The EF SET is recognized as a standardized and objectively scored test focusing on English listening and reading abilities. Its core design purpose is to classify a test taker's language performance into one of the six distinct levels defined by the Common European Framework of Reference for Languages (CEFR), ranging from A1 (Beginner) to C2 (Proficient) (EF Education First, 2014a). The test uses a Computer Adaptive Multi-Stage Testing (ca-MST) model. This adaptive technology adjusts the test's difficulty in real-time based on the test taker's responses, making the assessment efficient and accurate across a wide range of proficiency levels. The EF SET is freely available online, allowing broad accessibility. Its results are used in large-scale studies, such as the EF English Proficiency Index (EPI), which analyzed data from over 2.1 million test-takers in 2023 to rank global English skills (EF Education First, 2024). Furthermore, studies conducted shortly after its launch indicated that the EF SET PLUS (a longer version of the test) demonstrates high reliability and correlates positively with established assessments such as TOEFL iBT and IELTS (EF Education First, 2014b).

## Description of the test

The EF Standard English Test (EF SET) is an assessment used to measure English skills. This section explains how the test is developed. It also describes the types of questions and

tasks included in the test content. Finally, it details how the test performance is scored and how results are reported.

### **Development of the EF Standard English Test (EF SET)**

The journey to create the EF Standard English Test (EF SET) began with the goal of developing a free, high-quality English assessment tool for adults. From the outset, EF envisioned a comprehensive test covering all four language skills: listening, reading, writing, and speaking. However, the initial design prioritized the receptive skills of reading and listening, as these were considered the most feasible to automate reliably in the early development stages. The primary purpose was to create an internal assessment tool for placement, advancement, and certification within EF's programs. This internal tool was also intended to become a publicly available standardized test that could rival expensive, high-stakes exams such as IELTS and TOEFL in quality.

The foundational principles for the test's development were heavily influenced by the work of Lyle Bachman, particularly his book *Language Assessment in Practice* (Bachman & Palmer, 2010). During the trial phase, Mari Pearlman facilitated the practical application of these theories by focusing on test development and delivery processes. EF employed a formal and highly structured design approach, beginning with the creation of an Assessment Use Argument (AUA) and a Design Statement. These documents guided the development of the initial test blueprint, which specifically targeted reading and listening skills.

A significant period of exploration and testing followed, starting in May 2012 and lasting approximately 15 months through nine formal trials (EF Education First, 2014a). During this time, numerous task types and response formats were tested within EF's International Language Schools to ensure the sample was representative of the intended population. Data gathered from each trial underwent detailed psychometric analysis using both classical test theory and Rasch model analyses.

Following the September 2012 trial, the development process entered a data-driven refinement phase. The initial blueprint was revised using assessment engineering principles, resulting in more detailed Task Model Grammars for reading and listening. These grammars were continuously improved as new data emerged, aiming to enhance the quality and consistency of the automated tasks. Assessment engineering influenced not only task design but also the selection of the delivery model. Key decisions made during this phase included finalizing task types in May 2013, adopting Rasch model analyses in January 2013, calibrating

the adaptive item pool in April 2013, and choosing computer-adaptive multi-stage testing (ca-MST) as the delivery approach in April 2013. Expert guidance from Richard M. Luecht and John Willse was instrumental in these developments.

In parallel with refining the reading and listening components, EF began to develop separate writing and speaking modules to provide a fuller picture of learners' communicative abilities. These additional sections were designed to complement the receptive skills assessments and support EF's goal of measuring all four skills comprehensively.

A critical step in finalizing the EF SET was linking it to the Common European Framework of Reference (CEFR). This involved formal Standard Setting Studies where experts established cut scores on the EF SET scale corresponding to CEFR levels. An initial study in June 2013 set operational definitions for CEFR levels on the beta version. After further analysis, a second Standard Setting Study in July 2014 defined the final CEFR levels and cut scores for the operational EF SET, launched in September 2014. These cut scores guide how results are reported. Since its launch, EF has continued to monitor item performance, the adaptive testing system, and the online delivery platform, while further developing the writing and speaking sections to extend the test's coverage of productive skills and ensure ongoing validation of content and standards (EF Education First, 2014a).

### **Content of EF Standard English Test (EF SET)**

The EF Standard English Test (EF SET) evaluates two core English language skills: reading and listening in each of the two versions (EF Education First, 2025). Writing and speaking are also evaluated in the second, longer version. The first version takes 50 minutes to complete, while the second, longer version takes 90 minutes as shown in Figure 1.

#### *50 Minutes for Two Skills*

The EF Standard English Test (EF SET) includes a section specifically designed to assess two key language skills within a limited timeframe. This test takes a total of 50 minutes to complete. The time is divided equally between the two skills: reading comprehension and listening comprehension. Test takers are allocated 25 minutes for the reading section and another 25 minutes for the listening section. This structure ensures that both receptive skills are evaluated within this focused 50-minute period.

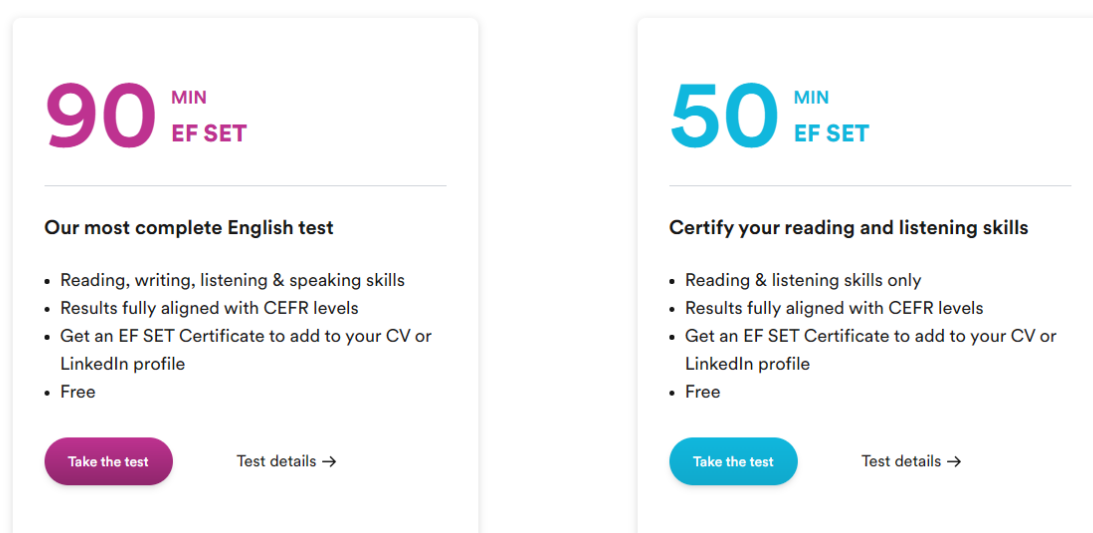
#### *90 Minutes for Four Skills*

The EF Standard English Test (EF SET) assesses four key language skills within a total time frame of 90 minutes. The test allocates specific durations for each skill section. Test-

takers spend 20 minutes on the reading section and another 20 minutes on the listening section. The writing section requires 35 minutes to complete. Finally, the speaking section is allotted 15 minutes. It is important to note that once the test begins, the timer cannot be paused, and the test cannot be restarted. While brief breaks between sections are permitted, these breaks are also timed. This structure ensures a comprehensive evaluation of reading, listening, writing, and speaking abilities within a fixed 90-minute duration.

**Figure 1**

*Content of EF Standard English Test (EF SET)*



Note. Source from <https://www.efset.org/>

### Scoring of EF Standard English Test (EF SET)

The EF Standard English Test (EF SET) utilizes a score scale ranging from 1 to 100. This scale is designed to measure English proficiency across all levels defined by the Common European Framework of Reference for Languages (CEFR), starting from A1 Beginner up to C2 Proficient as shown in Table 1. The test employs a computer-based multi-stage adaptive design, meaning that the difficulty of questions changes in real-time based on the test taker's performance. This adaptive approach helps ensure that the scores are reliable across the entire proficiency range. For instance, scores between 41 and 50 correspond to a B1 Intermediate level, while scores from 71 to 100 indicate a C2 Proficient level (Luecht, 2014a, 2014b).

**Table 1**

*Scoring of EF Standard English Test (EF SET)*

Scoring of EF Standard English Test	Common European Framework of Reference for Languages (CEFR)
1 - 30	A1 Beginner
31 - 40	A2 Elementary
41 - 50	B1 Intermediate
51 - 60	B2 Upper Intermediate
61 - 70	C1 Advanced
71 - 100	C2 Proficient

*Note.* Sources from Luecht (2014a, 2014b)

### **Analysis of the EF Standard English Test (EF SET)**

This section analyzes the EF Standard English Test (EF SET). It first looks at whether EF SET scores are as dependable as scores from TOEFL and IELTS. The analysis will also discuss the advantages and disadvantages of using the EF SET.

#### **The EF Standard English Test (EF SET) as reliable as TOEFL and IELTS scores**

The EF Standard English Test (EF SET) is an English proficiency test designed to measure reading and listening skills. It utilizes an adaptive testing framework, specifically computerized adaptive multistage testing (ca-MST), to tailor the test difficulty to the individual examinee, aiming for efficient and precise assessment. The scores are aligned with the Common European Framework of Reference for Languages (CEFR) levels A1 to C2. As a relatively new assessment compared to established tests such as the Test of English as a Foreign Language (TOEFL) and the International English Language Testing System (IELTS), evaluating its external validity and reliability is crucial. Studies were conducted to explore the statistical relationship between EF SET scores and those from TOEFL iBT and IELTS to provide evidence for its validity and reliability.

A key finding from the validation studies concerns the reliability of the EF SET scores. Reliability refers to the consistency and precision of test scores. Due to the adaptive nature of the EF SET, which adjusts difficulty based on performance, traditional reliability calculations need adaptation. Using marginal reliability coefficients, which are comparable to standard

measures such as Cronbach's alpha for adaptive tests, the studies found the EF SET to be highly reliable. The reliability coefficients for EF SET reading and listening were reported to be very high, specifically 0.949 for reading and 0.944 for listening based on large samples (Luecht, 2014a). This level of reliability was found to be comparable to the reliability of TOEFL iBT total scores and generally highly similar to IELTS scores, indicating that EF SET provides consistent measurements of reading and listening skills across different proficiency levels (Luecht, 2014a; Luecht, 2014b). The test design aims for uniform precision across the score scale, minimizing measurement error effectively (Luecht, 2014b).

The studies also investigated the correlation between EF SET scores and those from IELTS and TOEFL iBT as shown in Table 2. Correlation analysis measures the degree of linear association between scores from different tests. The research found reasonably strong, positive correlations between EF SET reading and listening scores and the corresponding sections on both IELTS and TOEFL iBT (Luecht, 2014a; Luecht, 2014b). Specifically, the observed correlation between EF SET reading and IELTS reading was 0.52, and the correlation between EF SET listening and IELTS listening was also 0.52 (Luecht, 2014b). For TOEFL iBT, after correcting for measurement error (disattenuated correlations), the correlation was 0.70 between EF SET reading and TOEFL reading and 0.77 between EF SET listening and TOEFL listening (Luecht, 2014a). These moderate to strong positive correlations suggest convergent validity, meaning that EF SET measures language skills with a similar degree of accuracy as the two more established tests (Luecht, 2014a; Luecht, 2014b).

In conclusion, the research indicates that the EF SET demonstrates high reliability for its reading and listening scores, comparable to the reliability found in established assessments such as TOEFL iBT and IELTS (Luecht, 2014a; Luecht, 2014b). Furthermore, there is solid evidence of positive correlations between EF SET scores and those from TOEFL and IELTS, supporting the argument that EF SET measures related English proficiency constructs. While the correlations are significant, they are not perfect, suggesting that EF SET may also assess slightly different aspects of language ability or that study limitations such as score range restrictions might affect the observed correlations (Luecht, 2014a; Luecht, 2014b). Consequently, while EF SET is shown to be a reliable assessment tool, the researcher advises against creating direct score concordance tables between EF SET and TOEFL/IELTS at this stage due to the moderate nature of the correlations (Luecht, 2014a; Luecht, 2014b).

Table 2

*The comparison of EF Standard English Test (EF SET) Scoring of TOEFL and IELTS scores*

Scoring of EF Standard English Test	Common European Framework of Reference for Languages (CEFR)	TOEFL	IELTS
1 - 30	A1 Beginner	N/A	N/A
31 - 40	A2 Elementary	N/A	N/A
41 - 50	B1 Intermediate	42 - 71	4.0 - 5.0
51 - 60	B2 Upper Intermediate	72 - 94	5.5 - 6.5
61 - 70	C1 Advanced	95 - 120	7.0 - 8.0
71 - 100	C2 Proficient	N/A	8.5 - 9.0

*Note.* Sources from Luecht (2014a, 2014b)

### EF English Proficiency Index

The EF Standard English Test, often called the EF SET, is a tool used to measure English skills. It is a free online test that anyone can take to check their English reading and listening abilities. The results from this test are used to create the EF English Proficiency Index, or EF EPI (EF Education First, 2024). The EF EPI is a large report that ranks countries and regions based on the average English skills of adults who took the EF SET. The main goal of the EF EPI is to show how English proficiency varies around the world and to track changes over time (EF Education First, 2024). This information can be useful for governments, schools, and companies looking at English language learning.

Creating the EF English Proficiency Index involves several steps based on the EF SET results. The EF SET itself is designed to be an adaptive test, meaning that the difficulty changes based on the test taker's answers, focusing on reading and listening skills. Millions of adults take this test each year, but it is important to know that these test-takers choose to take the test themselves, often because they are interested in learning English. This means that the sample might not perfectly represent everyone in a country. To calculate a country's score, the average EF SET score for that country over the past three years is used. When calculating scores for larger regions or the world, the scores of individual countries are weighted based on their population size.



The EF English Proficiency Index groups countries into five different levels, called proficiency bands: Very High, High, Moderate, Low, and Very Low. These bands help compare the general English ability across different places. For example, someone in a country with the Very High proficiency might be able to use complex language in social situations or negotiate a contract in English. In contrast, someone in a country with the Low proficiency might be able to handle simple tourist situations or understand basic work emails (EF Education First, 2024). These bands provide a simple way to understand what the average score in a country means in terms of practical English skills (EF Education First, 2024). It is a tool for comparison between large groups, not a measure of every individual's specific skill.

The EF English Proficiency Index in 2024 provides insights into global English skills trends. This report notes a decline in worldwide English proficiency for the fourth consecutive year, suggesting less focus on improving English skills in many areas. The report also highlights ongoing differences, such as varying skill levels between genders in different regions and age groups. For instance, while women's scores remained steady, men's scores declined slightly, narrowing the overall gap, though significant gaps favoring men persist in many countries. The index also shows correlations between national English proficiency and factors such as economic development, innovation, and social indicators, suggesting that English skills play a role in a country's global engagement and development. By tracking these patterns, the EF EPI aims to offer valuable information for discussions about English language education worldwide.

### **Advantages and Disadvantages of the EF Standard English Test (EF SET)**

The EF Standard English Test (EF SET) represents a significant development in the landscape of English language proficiency assessment. The test utilizes modern assessment technologies and aligns itself with internationally recognized standards, offering a distinct alternative to established high-stakes tests such as TOEFL and IELTS. However, as in the case of any assessment tool, the EF SET possesses both strengths and weaknesses that warrant careful consideration by potential users, educators, and institutions.

#### *Advantages of the EF Standard English Test (EF SET)*

One of the most significant advantages of the EF SET is its accessibility and cost-effectiveness. Unlike many traditional standardized English proficiency tests that carry substantial fees, the basic EF SET is available online completely free of charge to anyone with an internet connection. This removes a major financial barrier, democratizing access to

standardized English assessment for a global audience. This free access allows individuals worldwide to gauge their English skills against a recognized framework without the financial burden often associated with tests such as TOEFL or IELTS. Online delivery further enhances accessibility, allowing users to take the test from various locations, provided they have a stable internet connection.

In addition, the standardization and alignment of the EF SET with the Common European Framework of Reference for Languages (CEFR) are considered reliable. The test is designed as a standardized, objectively scored instrument specifically focused on assessing reading and listening abilities. Its development process involved rigorous design, trialing, and psychometric analysis to ensure quality comparable to high-stakes tests. The scores are explicitly mapped onto the six CEFR levels (A1-C2), providing users with results that are interpretable within a widely accepted international standard for language proficiency. Furthermore, correlation studies have demonstrated positive associations between EF SET scores and those from TOEFL iBT and IELTS, suggesting that while potentially measuring slightly different aspects, the tests converge in assessing common language proficiency traits.

Lastly, it involves the use of advanced adaptive testing technology. The EF SET employs a Computer Adaptive Multi-Stage Testing (ca-MST) design. This means the test's difficulty dynamically adjusts based on the individual test-taker's performance in real time. As a user answers questions, the system routes them to subsequent sections (modules) that are statistically optimized for their apparent proficiency level. This adaptive approach enhances measurement precision across a wide spectrum of abilities, from beginner (A1) to proficient (C2), and makes the testing process more efficient compared to traditional fixed-form tests.

#### *Disadvantages of the EF Standard English Test (EF SET)*

Firstly, the widely accessible, free version of the EF SET has limited scope in the skills it assesses. It primarily focuses on receptive skills, namely reading and listening comprehension. While EF Education First also offers tests covering writing and speaking, the core free assessment does not evaluate productive skills. This contrasts sharply with comprehensive exams such as TOEFL iBT and IELTS, which are designed to assess all four language modalities, including reading, listening, speaking, and writing, providing a more holistic view of a candidate's communicative competence. Users needing certified proof of all four skills may find the standard EF SET insufficient for their needs.

Secondly, the test results may be influenced by sampling bias due to its self-selected test-taker population. As the EF SET is free and online, participation is voluntary, attracting

individuals who are specifically interested in learning English or assessing their current skill level. This self-selection process means the sample of test-takers is not guaranteed to be representative of the broader population within a country or globally. The resulting scores, therefore, might be skewed either higher or lower than the average proficiency of the general population, potentially limiting the generalizability of findings based solely on EF SET data, such as the EF English Proficiency Index.

Thirdly, a disadvantage is its dependence on technology and internet connectivity. Being an exclusively online test, the EF SET requires users to have access to a device (computer or mobile) and, crucially, a stable internet connection throughout the test duration. Test-takers in areas with poor or unreliable internet infrastructure may be unable to complete the test, as partial tests are not saved if the connection drops. This technological requirement inherently excludes individuals without internet access, potentially introducing a socioeconomic bias towards more educated or affluent populations, particularly in regions where internet penetration is low.

Finally, the test administration procedures impose certain limitations. Test-takers are permitted only a single attempt at the EF SET; it cannot be repeated for practice or to achieve a better score. Once a test session begins, the timer cannot be paused, even for breaks between sections (which are also timed), requiring users to allocate a specific, uninterrupted block of time. Furthermore, navigation is restricted; users cannot return to previous questions or sections once they have submitted an exercise. These constraints differ from the flexibility offered by some other testing environments and require careful planning from the test-taker.

## Conclusion

In conclusion, the EF Standard English Test, or EF SET, is an important tool for checking English skills. It was carefully created using assessment theories and expert advice to be a high-quality test. One of its best features is that the main test is free and available online, making it easy for many people around the world to use. The test uses smart computer technology that changes the difficulty based on the test-taker's answers, which helps give accurate scores. Studies show that the EF SET scores are reliable and consistent, similar to well-known tests such as TOEFL and IELTS. The scores match the common European levels (CEFR), which helps people understand their English level based on an international standard. However, the test also has some drawbacks. The free version mostly checks reading and listening skills, not speaking or writing. Because people choose to take the test themselves,

the results might not represent everyone in a country. Also, people need a good internet connection to take it, which might be a problem for some. Finally, there are rules such as only one try allowed and no pausing, which requires careful planning. Overall, the EF SET is a valuable, accessible, and reliable test for reading and listening, but its limitations, such as the skills it covers and the need for internet access, should be considered.

### The Reviewer

**Pongsiri Kamkankaew** is a lecturer at the Faculty of Business Administration, North-Chiang Mai University. He received a B.A. in German (Second Class Honor), an M.B.A., and a D.B.A.

### AI Declaration Statement

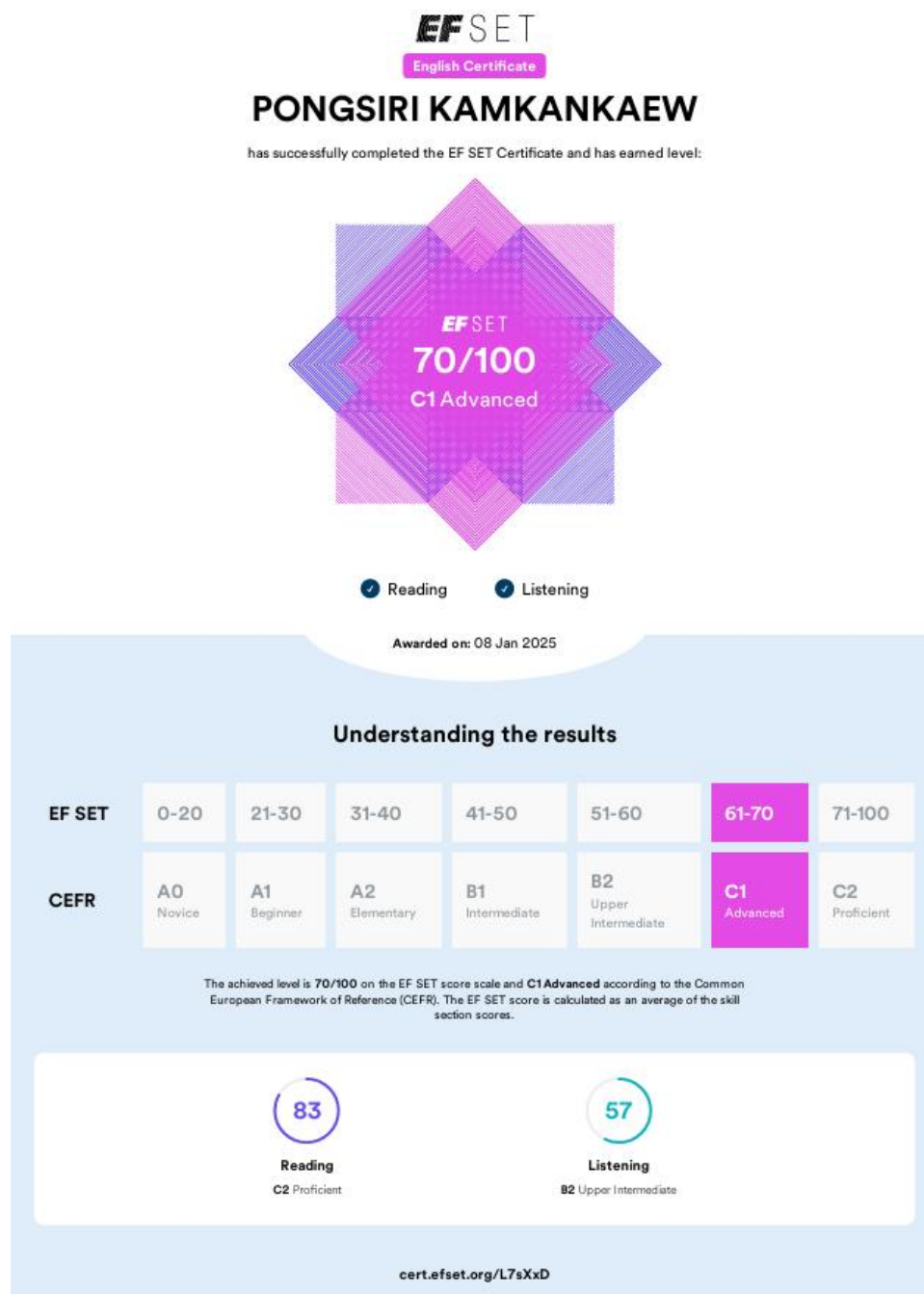
During the preparation of this work, the author used ChatGPT in order to enhance readability and improve language. After using this tool/service, the author reviewed and edited the content as needed and takes full responsibility for the content of the publication.

### References

- Bachman, L. F., & Palmer, A. S. (2010). *Language assessment in practice: Developing language assessments and justifying their use in the real world*. Oxford University Press.
- EF Education First. (2014a). *EF SET academic & technical development report*. EF Education First. <https://a.storyblok.com/f/79503/x/6c8302dc10/ef-set-academic-development-report.pdf>
- EF Education First. (2014b). *EF SET PLUS correlation study executive summary*. EF Education First. <https://a.storyblok.com/f/79503/x/016fca1e06/ef-set-correlation-study-summary.pdf>
- EF Education First. (2024). *EF English proficiency index: A ranking of 116 countries and regions by English Skills*. Signum International AG. <https://www.ef.com/assetscdn/WIBlwq6RdJvcD9bc8RMd/cefcom-epi-site/reports/2024/ef-epi-2024-english.pdf>
- EF Education First. (2025). *EF Standard English Test (EF SET)*. <https://www.efset.org>
- Luecht, R. M. (2014a). *EF SET PLUS & TOEFL iBT correlation study report*. EF Education First. <https://a.storyblok.com/f/79503/x/8f28ee8ebc/ef-set-toefl-correlation-study-report.pdf>
- Luecht, R. M. (2014b). *EF SET PLUS & IELTS correlation study report*. EF Education First. <https://a.storyblok.com/f/79503/x/3044206eea/ef-set-ielts-correlation-study-report.pdf>

## Appendix 1

An Example EF SET Certificate for 2 Skills



## Appendix 2

An Example EF SET Certificate for 4 Skills

