

## **Evaluating Extensive Reading Speed and Words Read with IELTS Reading Scores**

Hanon Junn  
Musashi University  
Japan

### **Abstract**

This evaluation report examines extensive reading (ER) speed and words read by students at a Japanese university using a digital ER platform called Xreading and compares them to standardized test reading performance using the International English Language Testing System (IELTS). The study focuses on first- and second-year students who participated in ER for one year in an academic reading and writing class. The research combines quantitative data on reading speed and words read with their IELTS reading performance over two testing points during the academic year. The findings indicate significant reading progress in reading speed and words read. Total words read showed a moderate correlation with IELTS reading scores; however, no significant correlation was found between increased reading speed and IELTS reading score changes across semesters. Based on the findings, the study concludes with future implications to improve and streamline the ER program for better results and participation.

**Keywords:** Extensive Reading, Extensive Reading and standardized test scores, Extensive Reading and IELTS, L2 reading program evaluation, Extensive Reading theoretical framework evolution, L2 reading speed, L2 words read

At the university under evaluation, the principal aim of the reading curriculum is to maximize intensive reading (IR) and extensive reading (ER) for students to become successful readers. The university defines successful readers as readers who are comfortable with academic reading for comprehension and composition with guidance and feedback. Secondly, successful readers should naturally improve their vocabulary, reading speed, and the number of words read through independent, enjoyable reading. These characteristics are inspired by Nuttall's (1996) virtuous circle of the good reader, where the interdependence of reading more and faster leads to better understanding and enjoyment.

In academic reading classes, the importance of IR comes from direct instructor guidance during the reading process including a deep understanding of lexical and syntactic structures as well as development of inferencing, deducing word meanings, vocabulary building, and text construction (Tagane et al., 2018). With heavy emphasis on analysis, accuracy, and

understanding, IR limits the number of words and genres that English as foreign language (EFL) students need to increase their reading ability and fluency. Incorporation of ER can alleviate these limitations and can be an important component of reading instruction as benefits include “vocabulary recycling, sight word reading gains, reading speed gains, reading practice, and a habit of being a successful reader in the [second language]” (Grabe & Stoller, 2011, p. 122). Nation and Waring (2020) comment that through ER, reading knowledge and proficiency is gained through faster understanding of what they read, learning new words in combination of strengthening their existing vocabulary knowledge through context, and development of unfamiliar grammatical features due to the sentence structure of each graded reader.

Standardized tests, like the academic International English Language Testing System (IELTS), are one way to gauge students' language knowledge and proficiency. Iwashita et al. (2021) accentuate how the Japanese Ministry of Education, Culture, Sports, Science, and Technology (MEXT) have actively emphasized English proficiency in the four skills: listening, reading, speaking, and writing, in English education and university entrance examinations. Emphasizing the four skills combats the low levels of English ability of Japanese learners, recognizes the importance of cross-cultural communication, and increases global opportunities like study abroad programs. This has placed the IELTS at the forefront of English proficiency testing in Japan and abroad, due to its rigorous four-skills testing structure and high international applicability, with accreditation by over 12,000 institutions in 140 countries (IDP IELTS, n.d.). Thus, evaluating changes in IELTS scores provides a measure of English proficiency aligned with institutional, national and international standards.

This report aims to quantify if a correlation exists between being a successful reader using ER reading speed and words read with English proficiency using IELTS reading scores. Reading speed is used as a measure of increased rate of progress while words read measures cumulative volume and effort over the entire year. IELTS scores, specifically academic reading scores, are significant for evaluating student reading proficiency. ER reading speed and words read improve comprehension, vocabulary, and grammar skills which are critical for success in IELTS academic reading. If any correlation is found between ER and IELTS, it would further strengthen the inclusion and continued investment of ER in a reading curriculum, as IELTS scores provide a widely-accepted scoring system for institutions implementing ER.

The main research questions that will be investigated include:

- RQ1: What measurable ER progress in reading speed and words read have students made after one year?
- RQ2: Have student IELTS scores positively changed over the same one-year period at two testing points after participating in ER?
- RQ3: To what extent do ER reading speed and words read correlate with IELTS reading scores after one year?

These questions focus on one university but addresses an important research gap in several ways. First, it aims to find correlations between two ER factors with IELTS reading scores as there are no research studies related to these measures available. Secondly, by using IELTS reading scores, proficiency in standardized tests will be applicable to other institutions. Finally, this

Junn: Evaluating extensive reading speed and words read with IELTS reading scores

report will investigate if one year is sufficient to notice ER progress and its correlation to IELTS scores.

## **Literature Review**

### *Evolving Extensive Reading Principles*

Renandya et al. (2021) have attempted to refine the ten ER principles formulated by Day and Bamford (1998, 2002) into five core principles that may be broader in scope in ER implementation: (a) The reading materials are interesting and comprehensible, (b) students read as much as possible, (c) reading speed is faster rather than slower, (d) students choose what they want to read, where possible, and (e) the purpose of reading is mostly for enjoyment and general information.

Waring and McLean (2015) rationalize the broad scope of the five core ER principles by advocating core and variable dimensions. Core dimensions are fluent comprehension, high reading speed, reading large amounts of text, and focus on meaning of text which accomplishes the goal of “fast, fluent comprehension sustained over long periods with minimal distractions” to maintain the core characteristics of ER (Waring & McLean, 2015, p. 164). Variable dimensions, such as ER in or outside of the classroom, the length of texts, freedom to choose texts, whether the reading is monitored, and the presence or absence of follow-up activities offer more selection from each individual ER setting “in varying degrees depending on pedagogical aims, research questions or what is practical” (Waring & McLean, 2015, p. 164). Therefore, each ER program is unique to its specific university but abiding by the five core ER principles provides a relatable reference guide for other institutions.

The theoretical framework behind the ER principles is based on the cognitive application of the schema theory where the reader connects information from written messages with previous knowledge to arrive at meaning and understanding (Day & Bamford, 1998). Schema theory suggests a mental structure that interactively organizes background knowledge with text input to help readers develop comprehension and fluency (Nuttall, 1996). Day and Bamford (1998, p. 16) believe extensive reading plays “an important role in developing the components in which fluent second language reading depends including a large sight vocabulary, a wide general vocabulary, and knowledge of the target language, the world, and text types”. Additionally, ER aligns with Krashen’s (1985, p. 15) input hypothesis where the importance of providing “a great deal of comprehensible input” within a “low-anxiety environment” can lead to language development. Alternatively, Carver (1990) proposed the rauding theory which combines *reading* and *auding* (the process of understanding what is being read). Carver (1990) argues that reading too fast or too slow can hinder comprehension so rauding is seen as efficient reading that balances reading speed with comprehension and fluency, which align with the core ER principles.

### *Measurable ER progress with Reading Speed and Words Read*

In this report, reading speed, measured in words per minute (wpm), and words read will be the most critical measurable data because it reveals students’ real-time and personal reading activity. Increasing reading speed increases fluency while helping reading proficiency. In a comparative study of 50 ER and non-ER students, 23 students participating in ER had increased reading rates

Junn: Evaluating extensive reading speed and words read with IELTS reading scores

to 130.34 wpm while 27 students not participating in ER averaged 103.05 wpm during time-reading passages (McLean & Rouault, 2017). Most importantly, the comprehension rates, based on passing quizzes at 70%, were sustained by the ER group. Therefore, McLean and Rouault (2017) believe that ER is effective in increasing reading rates while sustaining comprehension. Additionally, the increased reading speed allows for the development of a broader vocabulary and grammar through increased exposure within the context of graded readers.

Nishizawa et al. (2010) believe that reading as many words as possible through a multi-year ER program is the best indicator of standardized test performance, using the Test Of English for International Communication (TOEIC). In their four-year study, students who read less than 100,000 words did not see significant results compared to the control group but after 300,000 words, there were significant improvements in TOEIC scores. It took two to three years for the students to reach 300,000 words but saw an increase of 4 points per 100,000 words. Students who read over 1,000,000 words achieved gains in TOEIC scores comparable to students spending at least ten months studying abroad in an English-medium university (Nishizawa et al., 2010).

Nakanishi (2015) revealed that longer ER durations to increase word exposure led to better outcomes in his meta-analysis of 34 studies. Studies lasting a year or more had a larger effect size ( $d = 0.52$  vs  $d = 0.46$ ) for ER, non-ER group studies and for pre-, post-test comparisons ( $d = 0.74$  vs  $d = 0.71$ ). In a similar meta-analysis of 49 primary studies, Jeon and Day (2016) found a small to medium effect ( $d = 0.57$ ) on the effectiveness of ER on reading proficiency (comprehension, reading rate, and vocabulary) comparing ER, non-ER group studies. More strikingly, their meta-analysis reveals a larger effect size ( $d = 0.79$ ) for reading proficiency when utilizing ER in a reading curriculum with pre- and post-tests. Jeon and Day (2016) conclude that comprehension and vocabulary gained through words read, and reading speed is a proven approach for reading proficiency and can be maximized for adult groups, computer reader programs, and ER in the curriculum validating the research aims of this report.

### *ER and Standardized Test Scores*

Theoretical underpinnings of ER principles suggest three key points in the relationship between language proficiency and ER: (a) it is a long and gradual process, (b) interesting and comprehensible ER materials are a necessity, and (c) ER promotes incidental language learning to develop implicit knowledge of the language (Renandya et al., 2021). Learning through extensive reading is largely incidental as learners' attention is focused on the story not on items to learn. Since learning gains tend to be small, the quantity of comprehensible input is important so that the readers meet vocabulary and grammatical features extensively to build a network of knowledge relationships between them (Nation & Waring, 2020).

Incidental learning highlights difficulties in measuring improved reading skills over a limited period especially in standardized tests with measurable data. Similar to Nishizawa et al. (2010) study, Yamashita (2008) notes that while ER may benefit language acquisition in the long term, it is too ambitious to expect that all skills would show equal improvement over a limited period. Additionally, skills that are developed over a fixed period vary individually from student to student. Limited research investigating ER's direct impact on standardized testing performance,

specifically the IELTS reveals the difficulty in finding consistent results and correlation. Previous studies have shown mixed results when correlating ER factors with standardized test scores, with some reporting significant gains and others finding marginal or no improvement.

Sakurai (2011) correlates a relationship between the number of words read with TOEIC reading scores. She indicates that every 10,000 words, read up to 69,999 words, can result in score improvements for lower-level learners (Sakurai, 2011, p. 118). Sakurai reported a statistically significant increase of 57 points in TOEIC reading scores, leading her to conclude that there was an improvement in general English proficiency and “extensive reading has played a positive role and contributed to the gain to some extent” (Sakurai, 2011, p. 118). Mason and Krashen (2017) show performance gains from a small sample of 8 students of nearly .6 points on the TOEIC for every hour spent on ER. Adult learners logged their extensive reading over self-determined durations ranging from 45 weeks to 162 weeks, with pre- and post-ER TOEIC scores. On average, students who read 100 hours averaged a gain of 60 points and as a group, the students gained 166 points in their TOEIC scores after spending 260.5 reading hours and reading 7846 pages.

Contrarily, Yamashita (2008) notes that 31 students who participated in ER over fifteen weeks improved general reading ability but not their second language (L2) linguistic ability when taking pre- and post- ER reading tests. They were given placement progress tests (PPT) to test linguistic ability such as grammar, vocabulary and syntactic knowledge and extensive reading tests (ERT) which aims to test their general reading ability focusing on comprehension and reading speed. The difference was significant for the ERT ( $t(30) = 2.957, p < .001$ ), but not significant for PPT ( $t(30) = 1.84$ ). Yamashita (2008) surmises that reading ability increases first because reading is a familiar activity for L2 learners and their L1 reading strategies are more readily transferable than L2 linguistic ability which needs more exposure, especially for lower-level learners.

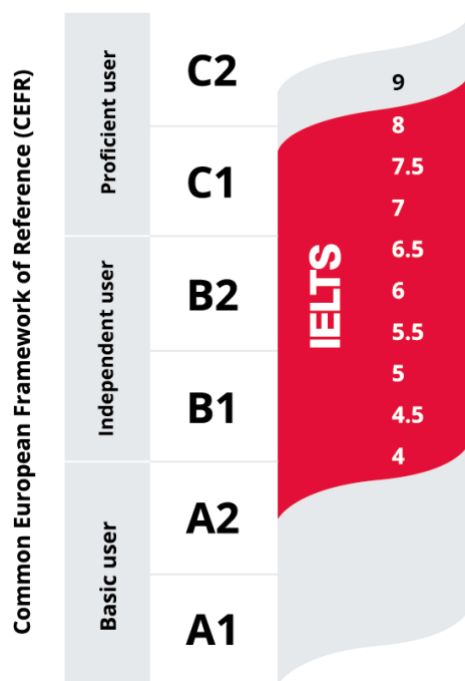
Supporting Yamashita’s (2008) mixed results, in a substantial study of 497 students over two academic years by Lyddon and Kramer (2019, p. 261) reveal that the number of words read “showed a small but consistent positive contribution” to TOEIC reading scores in post-ER TOEIC tests conducted at semesters’ end in both years ( $M = 156.71$  to  $171.39$ , in year one and  $M = 156.69$  to  $167.29$  in year two). They caution, however, there was a lack of proportional increase as the number of words read doubled from one year to the next while TOEIC reading scores remained nearly identical (Lyddon & Kramer, 2019). They believe that ER words read can be a predictor of TOEIC reading performance but general diligence of students completing assignments may be a bigger indicator of standardized test performance (Lyddon & Kramer, 2019).

Only one research study was discovered by the author analyzing the effects of ER and IELTS performance. The study conducted by Shareef (2023) measured the effects of ER on students’ IELTS speaking performance. After a ten-week ER program of 15,000 words from pre-selected graded readers, ER was supplemented by discussions and summaries of the stories. Students who participated in ER significantly increased their IELTS pre-ER, post-ER speaking performance by 20 points from 50 to 70 measured on an adjusted 100-point scale. This equates to an increase from a IELTS band score of 5.0-5.5 to 6.0-6.5.

The small sample of studies above highlights the varied results in relation to ER and standardized test performance, but the dearth of studies relating ER factors to IELTS reading allows this evaluation report to further contribute to the discussion of ER progress and gains in standardized testing.

### *IELTS Reading as a Measure of Proficiency*

The academic IELTS is a high-stakes exam covering four sections: listening, reading, writing, and speaking in approximately three hours. IELTS academic reading is a 60-minute exam with 3 academic passages and 40 questions. The question types include (a) diagram, note, or sentence completion, (b) identifying information or writer's views, and (c) matching, multiple choice, short answer questions (IELTS, 2023). All questions aim to check the test taker's ability to understand main ideas, details, logical arguments, and writer's opinions, attitudes and purpose (IELTS, 2023). Scoring is based on a band score from 1-9 where 1 represents a non-user meaning the test taker "has no ability to use the language except a few isolated words" while 9 represents expert or "fully operational command of the language" (IELTS, 2023). Figure 1 shows how IELTS scores are framed within the Common European Framework of Reference for Languages (CEFR).



**Figure 1.** *IELTS Band Score and CEFR Comparison (IELTS, 2023)*

IELTS reading has become a reliable test for proficiency because IELTS reading performance can be a predictor of academic performance (Kerstjens & Nery, 2000, p. 105). Kerstjens and Nery (2000) note that IELTS academic reading, and to a lesser extent, writing should be given special consideration for international students being selected for tertiary courses.

Dang and Dang (2023), in a mixed-methods study of Vietnamese students studying abroad in the UK, show a positive correlation between IELTS scores and academic performance using overall IELTS scores before attending the university and their GPA in the first semester abroad ( $r = .44$ ,  $p < .001$ ). Of the four sections, IELTS reading had the strongest correlation ( $r = .44$ ,  $p < .001$ ) over the other three sections. Dang and Dang (2023) argue that the higher the IELTS scores, the more likely the students were to perform well in their study abroad education. In follow-up interviews with four students, Dang and Dang (2023, p. 94) learned that IELTS academic reading preparation provided foundational knowledge like skimming and scanning for academic text and improved their reading speed but did not prepare them in “critical thinking, synthesizing sources, and comprehending long and technical texts.” For students studying abroad, IELTS scores may facilitate readiness in English-medium universities but does not guarantee success. Jang et al. (2019) caution that IELTS academic reading is reliable in measuring English proficiency but lacks in assessing higher-order thinking skills such as inference and synthesizing ideas. Therefore, institutions need to have other measures in place to ensure academic success.

Craven (2012) supports the above findings as non-English speaking students in her study were able to achieve higher IELTS scores pre- and post-undergraduate degrees, specifically in listening and reading. However, she believes that there is “no level playing field” when developing English proficiency and academic success as previous language background and “sense of agency in improving one’s English” are critical factors beyond measurable data (Craven, 2012, p. 43). IELTS can provide some positive correlations to academic success, especially at the university level but proficiency tests should not be the main factor in determining higher academic success. Hence, this report will focus on quantifiable measures while framing all results and discussions with the recognition that non-measurable ER factors including diverse English backgrounds and motivation, likely affect performance.

## Methodology

### *University and Learner Profile*

The one-year evaluation was conducted within the Global Studies (GS) department at a private university in Tokyo, Japan with 82 first- and second-year students. In this department, students are expected to complete required academic reading and writing classes, in addition to speaking and listening classes, in an English for Academic Purposes (EAP) curriculum in their first two years. This prepares them for study abroad programs in their third year and graduation research projects in their fourth year. Students are expected to have Common European Framework of Reference (CEFR) B1+ or B2 English ability based on the university entrance requirements; however, exam contents and results were not publicly available for analysis. The instructors developed the curriculum based on the above CEFR level assumption and with the belief that first-year students have never taken formal academic reading and writing classes. Students’ actual levels may vary based on previous English exposure and education before entering the GS department. Another assumption is that second-year students, as a group, should have a higher English proficiency level as they have had one extra year of EAP classes. Under these assumptions, the first two years are seen as an integral part of their academic and English development.

The students participated in ER through mandatory and randomly divided academic reading and writing classes that met twice per week for 105 minutes over a 13-week semester. The students were divided into six classes ranging from 12 to 16 students per class. Classes 1-3 were first-year students and classes 4-6 were second-year students.

The weekly schedule was one class focusing on intensive reading and one class focusing on academic writing skills related to the intensive reading text. The intensive reading focused on globally-broad topics relevant to Japanese learners of English such as Sociology, education, and technology. First-year classes read professional and academic-level articles ranging from two to three pages in length, supported by vocabulary lists to align with the Academic Word List (AWL). Second-year students read longer articles, usually between four to eight pages in length with more comprehensive vocabulary from the AWL. As the readings were intensive, there was frequent instructor interference to develop understanding of lexical and syntactic structures, text flow, construction, and vocabulary development. Group discussion activities also provided peer support for the aforementioned linguistic features. The writing aspect supported the IR articles by having them think more deeply about the topic and produce in-class academic paragraphs and out-of-class longer essays that connect the readings to their global understanding of issues through their own research.

With the classes being exclusively IR focused; the implementation of ER was conducted outside of class as homework using Xreading. Students were required to purchase Xreading subscriptions for one year and were given bi-weekly assignments to read a predetermined word count. Xreading, a digital ER platform created in 2014, was developed “to make graded readers more accessible for students and extensive reading programs easier for teachers to manage and assess” (Xreading, n.d.). As such, Xreading is a platform that can easily achieve Renandya et al.’s (2021) five core principles while offering a learning management system (LMS) for instructors to assess reading speed and words read in real time. The rationale of using Xreading lies in its easy implementation of ER with over 1,800 graded readers (as of December 2024) across various levels and genres, which students can freely choose. There are 15 different levels of books that align with CEFR pre-A1(level 1) to CEFR C2 (level 15).

For instructors, the LMS provides real-time data on student reading activity. Xreading automatically calculates how many words are on a page and records the time a reader spends on each page, effectively monitoring reading speed. By default, Xreading tracks books added, books read (passed), words read, reading time, and reading speed. Additionally, each book is accompanied by a five-question, multiple-choice comprehension quiz to ensure that the book was read and to assess readers’ general understanding of the book contents (Tabata-Sandom, 2023). The comprehension quizzes are defaulted with a passing rate of 60% but are customizable.

Nation and Waring (2020, p. 56) advocate Xreading because it “takes away the cost, time, availability, and management issues that might prevent teachers from running an extensive reading program”. Instructors can customize settings to fit the reading requirements of each ER assignment and student reading levels. Some customizable features include the type of books that can be accessed by criteria, an inactivity timer that alerts students when they are not reading, a

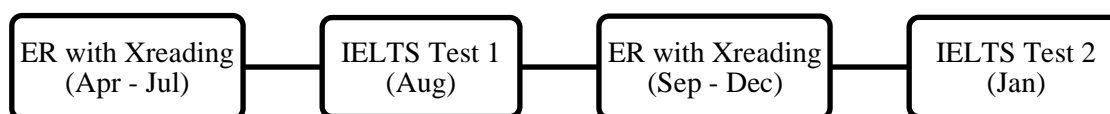
Junn: Evaluating extensive reading speed and words read with IELTS reading scores

maximum reading speed permitted to ensure that students are reading with integrity, and acceptable pass rates for comprehension quizzes. The customizable features offer a significant advantage to varied dimensions of ER implementation (McLean & Waring, 2015).

Based upon instructor preferences, the word count requirement was modified to accommodate class pace on intensive reading and writing assignments. Specifically, classes 1 and 4 had a lower word count in the fall semester, at 165,000 words compared to 234,000 words, because the instructor prioritized academic writing skills necessary to complete academic essays related to intensive reading texts (Table 1). In all classes, some class time may have been allotted for Xreading at instructor discretion after the formative English instruction and activities were completed.

At the end of each semester, students were expected to participate in IELTS tests (Figure 2). The timing of the IELTS tests allows for relevant observation of measurable ER factors to English proficiency after the ER period because it provides two equal proficiency check points providing comparable data collection for analysis.

Although IELTS is not designed to measure progress over short periods, the university in this report utilizes the scores to track student development across academic semesters via expected achievement scores. First-year students are expected to achieve an overall IELTS band score of 5.5 and second-year students are expected to reach 6.0 at the end of the academic year, which would be the equivalent of CEFR B1+ or B2 level. This approach is driven by administrative requirements to provide quantifiable measures of student performance and progress, which are used to assess academic standing for future class placements. Another benefit of using IELTS is the opportunity to study abroad as students are highly encouraged to do so for one or more semesters in their third year. IELTS academic scores can help students recognize their English proficiency for acceptance into study abroad programs that have standardized test admission requirements.



**Figure 2.** *ER and IELTS Timeline*

### *Data Collection*

Quantified data were pooled from the Xreading LMS by assignment over two semesters with specified parameters. Students were prohibited to read more than 300 wpm, as exceeding this limit would flag possible manipulation of ER. Nation and Waring (2020) comment that many non-native speakers of English read well below 300 wpm but should eventually aim to read silently at a reasonable speed of 200 wpm if they know most of the words. If a student reads more than 300 wpm, they will receive an alert from the platform to slow their reading pace. If they violate this parameter again, the book's word count would be disqualified. Another parameter was to pass each book's comprehension quiz on Xreading's platform with a 60% passing grade to receive credit for words read. If students fail to reach the passing rate, they

would not receive word count credit for that book. This ensured accountability and integrity when students read.

For RQ1, collected data specifically focused on reading speed and words read, as these two variables would best represent the broad-based core ER principles (see Renandya et al., 2021). The data was distinguished into three categories: combined, first- and second-year students. Progress was measured from Spring semester to Fall semester using a paired-samples t-test to determine statistically significant differences in reading speed. Reading speed was selected as the measurement of ER progress as the primary aim of ER is to read a lot at a good pace of around 120-150 wpm (Nation & Waring, 2020). Furthermore, reading speed best represents independent reading progress, as the number of words read depends on instructor requirements for each assignment. A paired-samples t-test was analyzed with IBM's SPSS software. Larson-Hall (2010, p. 266) advocates paired-samples t-tests for two-related measures of the same group because it takes "into account the violation of the assumption of independence to report on whether the two mean scores are considered to be different from one another."

For RQ2, 56 out of 82 students completed both IELTS tests. The remaining students completed only one test, disqualifying them from the analysis. Although five score components: (a) overall, (b) listening, (c) reading, (d) speaking and (e) writing were graded, IELTS reading scores and overall scores are highlighted. Overall scores are included in RQ2 data analysis to frame IELTS reading scores within the context of their overall performance. IELTS listening, speaking and writing scores were excluded as non-reading variability factors like speaking proficiency, listening skills, and understanding IELTS writing composition rules would particularly affect performance in these sections. Moreover, detailed analysis of the other IELTS scores would detract the primary focus on ER factors and IELTS reading score changes. The rationalization behind this choice is that in IELTS reading, Craven (2012) highlights transferable skills such as recognizing text structure, skimming, scanning, and wider vocabulary exposure contributed to improved reading scores, based on qualitative interviews in his study. Therefore, using overall scores will recognize non-ER factors and general English proficiency in their IELTS performance but the aforementioned transferable skills are particularly relevant to ER, establishing a stronger correlation to IELTS reading scores. Like their ER performance, a paired-samples t-test was conducted to determine any significance from their first testing point to their second testing point.

For RQ3, two types of correlations will be presented. The first correlation compares reading speed differences with IELTS reading scores differences from Spring to Fall to analyze any gains over the academic year. Additionally, total words read and Fall IELTS reading scores will analyze if the cumulative volume of ER activity over time affects their IELTS reading scores. Unlike R2, IELTS overall scores are excluded in this part to isolate reading activity and achievements as much as possible. Spearman correlations were extracted as it does not assume a linear relationship between reading speed and IELTS reading scores as well as words read and IELTS reading scores.

For L2 research study effect sizes and correlations, Plonsky and Oswald (2014) argue for a more localized interpretation of effect sizes as their empirical analysis of 346 primary studies and 91 meta-analyses reveals that Cohen's *d* underestimates the range of effects obtained in L2 research.

Although, they recommend effect sizes for  $d$  at 0.4 (small), 0.70 (medium), and 1.0 (large) over Cohen's  $d$  of 0.2 (small), 0.5 (medium), and 0.8 (large), this evaluation report will implement benchmarks of 0.6 (small), 1.0 (medium), and 1.4 (large) for within-group, pre-post contrasts as the same group is being analyzed at two different time periods (Plonsky & Oswald, 2014). For correlation coefficients, Plonsky and Oswald (2014) recommend,  $r$  values at .25 (small), .40 (medium) and .60 (large) for Pearson correlations for linear relationships but these benchmarks will be used for Spearman's rho ( $r_s$ ) as it similarly measures strength and direction of relationships, although  $r_s$  is used for non-linear data. Their rationale behind these sizes is that "the ability of future studies and meta-analyses, in particular, to look inward and gauge the relative magnitude in individual subdomains of L2 research, rather than using omnibus benchmarks" (Plonsky & Oswald, 2014, p. 12–13).

## Results

*RQ1: What measurable ER progress in reading speed and words read have students made after one year?*

Table 1. *Reading Performance by Class*

	Class 1	Class 2	Class 3	Class 4	Class 5	Class 6	All
<b>Spring 2023 (Apr-Jul)</b>							
Word Count Goal (words)	150,000	150,000	150,000	138,000	150,000	150,000	148,000
100%+	8	6	5	5	2	10	36
90%-99%	4	1	0	2	3	0	10
80%-89%	1	1	1	4	0	1	8
70%-79%	0	0	2	1	2	0	5
60%-69%	0	1	3	0	0	0	4
~59%	0	5	1	4	8	1	19
0%	0	0	0	0	0	0	0
Total (students)	13	14	12	16	15	12	82
<b>Fall 2023 (Sept-Dec)</b>							
Word Count Goal (words)	165,000	234,000	234,000	165,000	234,000	234,000	211,000
100%+	8	1	6	5	0	8	28
90%-99%	4	2	0	2	3	2	13
80%-89%	1	3	1	4	2	1	12
70%-79%	0	2	3	1	0	0	6
60%-69%	0	1	1	0	0	0	2
~59%	0	3	1	2	7	1	14
0%	0	2	0	2	3	0	7
Total (students)	13	14	12	16	15	12	82

Table 1 provides the reading performance of the entire group divided into six classes. On average, students were expected to read 148,000 words in Spring (approximately 12,300 words per week) and 211,000 words in Fall (approximately 16,000 words per week). In Spring, 36 out of 82 students (43.9%) were able to meet or exceed the word count requirement and 63 out of 82 students (76.8%) received a passing grade of at least 60%. In Fall, the higher word count requirements may have led to a decrease in students meeting the 100% word count goal, from 36 to 28 (34.1%) reading 100% or more and a minor decrease to 61 students (74.4%) receiving a passing grade of 60% or more.

Table 2. *Reading Speed and Word Count (Original vs Eligible)*

Group	Semester	<i>n</i>	Reading Speed (wpm)	Words Read (Avg)	Words Read (Total)
Combine	Spring	2	122.79	118,479	9,687,546
	Fall		125.06	150,189	12,249,411
	Year		123.93	134,334	21,936,957
1st-year	Spring	3	123.98	125,412	4,891,068
	Fall	9	140.25	166,700	6,501,300
	Year		132.12	146,056	11,392,368
2nd-year	Spring	4	121.71	111,546	4,796,478
	Fall	3	111.28	133,677	5,748,111
	Year		116.5	122,611	10,544,589
Eligible Combine	Spring	5	120.71	135,405	7,582,696
	Fall	6	137.2	170,889	9,569,752
	Year		128.96	153,147	17,152,448
Eligible 1st-year	Spring	3	122.94	140,780	4,504,960
	Fall	2	143.57	176,996	5,663,872
	Year		133.26	158,888	10,168,832
Eligible 2nd-year	Spring	2	117.74	128,239	3,077,736
	Fall	4	128.72	162,745	3,905,880
	Year		123.23	145,492	6,983,616

Table 2 shows the performance over the academic year as a group, separated into original and eligible, in two measurable categories collected in the Xreading LMS: words read and reading speed. Eligible students are students who participated in ER and completed IELTS tests both semesters. There were seven students who did not participate in ER in the fall semester but took both IELTS tests. There were an additional 19 students who participated in ER but did not take both IELTS tests. Therefore, these students were removed for a better understanding of correlations in RQ2 and RQ3, resulting in a final sample of 56 eligible students.

All groups naturally increased their word count due to the increased word count requirement from Spring to Fall. The combined performance of the original group reveals that they were able to read 118,479 words or approximately 80% of the required word total of 148,000 words in spring. In fall, the students read 150,189 or approximately 71% of the required word count of 211,000 words.

The most significant data show reading speed and word count average of eligible students. When comparing these 56 students with the original ER group, the eligible group had a lower reading speed than the original group (120.71 vs. 122.79 wpm) in Spring but had a higher word count average (135,405 vs. 118,479 words). In Fall, the eligible group had a higher reading speed (137.20 vs. 125.06 wpm) and a higher word count (170,889 vs. 150,189 words). First-year eligible students had comparable reading speeds (122.94 vs. 123.98 wpm in Spring and 143.57 vs. 140.25 wpm in Fall) but achieved higher word count averages (140,780 vs. 125,412 words in Spring and 176,889 vs. 166,700 words in Fall). The biggest disparity can be seen in second-year students. Eligible students had a lower reading speed in Spring (117.74 vs. 121.71 wpm) but a dramatically higher reading speed in Fall (128.72 vs. 111.28 wpm) compared with the original group of second-year students. The word count averages were also much higher in both semesters (128,239 vs. 111,546 words in Spring and 162,745 vs. 133,677 words in Fall). Comparing the eligible group of students, first-year students outperformed the second-year students in reading speed and total words read in both semesters emphasizing ER diligence and commitment from first-year students.

Table 3. Paired-samples *t*-test for Reading Speed (Eligible Students)

Group	Measure	<i>n</i>	<i>M</i> ( <i>SD</i> )	<i>MD</i> (Spring - Fall)	<i>t</i>	<i>df</i>	<i>p</i>	Cohen's <i>d</i>
Combined	Spring Reading Speed (wpm)	56	120.71 (38.07)					
	Fall Reading Speed (wpm)	56	137.20 (39.96)	-16.49	-3.675	55	<.001	-0.491
1st-year	Spring Reading Speed (wpm)	32	122.94 (43.77)					
	Fall Reading Speed (wpm)	32	143.57 (45.47)	-20.62	-2.929	31	.003	-0.518
2nd-year	Spring Reading Speed (wpm)	24	117.74 (29.46)					
	Fall Reading Speed (wpm)	24	128.72 (30.02)	-10.98	-2.402	23	.012	-0.49

To test how significantly reading speed increased over the academic year, a paired-samples *t*-test between Spring and Fall was conducted for eligible students in both semesters (Table 3). Overall, there was a statistical significant increase in reading speed ( $t(55) = -3.675, p < .001$ ); however, the effect size (Cohen's  $d = -0.491$ ) fell below Plonsky and Oswald's small effect size benchmark ( $d = 0.6$  for within-group effect sizes). For first-year students reading speed showed similar statistical significance ( $t(31) = -2.929, p = .003$ ) but a larger effect size (Cohen's  $d = -0.518$ ) that closely approaches the small effect size benchmark. For second-year students, reading speed increased numerically from 117.74 to 128.72 wpm and the increase had statistical significance ( $t(23) = -2.402, p = .012$ ) and a similar effect size to the combined group (Cohen's  $d = -0.490$ ).

For all three groups, the increase in reading speed was statistically significant but using within-group benchmarks, the practical effect may be limited.

*RQ2: Have student IELTS scores positively changed over the same one-year period at two testing points after participating in ER?*

Table 4. Paired Samples t-tests between IELTS Scores in Spring and Fall

Combined Group	<i>n</i>	<i>M</i> Spring ( <i>SD</i> )	<i>M</i> Fall ( <i>SD</i> )	<i>MD</i> (Spring - Fall)	<i>t</i>	<i>df</i>	<i>p</i>	Cohen's <i>d</i>
IELTS Overall	56	5.73 (0.72)	5.82 (0.68)	-0.09	-1.299	55	.199	-0.174
IELTS Reading	56	5.84 (0.83)	5.91 (0.92)	-0.07	-0.746	55	.459	-0.1
1st-year students								
IELTS Overall	32	5.50 (0.79)	5.70 (0.81)	-0.2	-1.981	31	.057	-0.35
IELTS Reading	32	5.63 (0.88)	5.83 (1.07)	-0.2	-1.603	31	.119	-0.283
2nd-year students								
IELTS Overall	24	6.04 (0.46)	5.98 (0.43)	0.06	0.827	23	.417	0.169
IELTS Reading	24	6.13 (0.66)	6.02 (0.67)	0.1	0.738	23	.468	0.151

Table 4 shows mean differences and paired-samples t-tests for IELTS overall and reading scores. The combined group of students showed a slight increase in their IELTS overall scores from Spring ( $M = 5.73$ ,  $SD = 0.72$ ) to Fall ( $M = 5.82$ ,  $SD = 0.68$ ). The mean difference of -0.09 was not statistically significant ( $t(55) = -1.299$ ,  $p = .199$ ) and was well below the small effect size benchmark (Cohen's  $d = -0.174$ ). This suggests that the slight improvement in IELTS overall scores was minimal and not statistically meaningful. Similarly, the IELTS reading scores increased slightly from Spring ( $M = 5.84$ ,  $SD = 0.83$ ) to Fall ( $M = 5.91$ ,  $SD = 0.92$ ). The mean difference of -0.07 was also statistically non-significant ( $t(55) = -0.746$ ,  $p = .459$ ) with an even smaller effect size (Cohen's  $d = -0.10$ ). This also indicates that the change in IELTS reading scores for the combined group was minor and did not reach statistical significance.

First-year students demonstrated a more notable increase in IELTS overall scores, rising from  $M = 5.50$ ,  $SD = 0.79$  in Spring to  $M = 5.70$ ,  $SD = 0.81$  in Fall. The mean difference of -0.20 approached statistical significance ( $t(31) = -1.981$ ,  $p = .057$ ) with a larger, though still below benchmark, effect size (Cohen's  $d = -0.35$ ). The improvement in IELTS overall scores for first-year students was more substantial, though it narrowly missed being statistically significant. For IELTS reading, first-year students' scores increased from Spring ( $M = 5.63$ ,  $SD = 0.88$ ) to Fall ( $M = 5.83$ ,  $SD = 1.07$ ). The mean difference of -0.20 did not reach statistical significance ( $t(31) = -1.603$ ,  $p = .119$ ) with limited effect size (Cohen's  $d = -0.283$ ). This indicates some improvement in IELTS Reading scores for first-year students, but the change was not statistically significant.

Conversely, second-year students exhibited a slight decrease in their IELTS overall scores from Spring ( $M = 6.04$ ,  $SD = 0.46$ ) to Fall ( $M = 5.98$ ,  $SD = 0.43$ ), with a mean difference of 0.06. This decrease was not statistically significant ( $t(23) = 0.827$ ,  $p = .417$ ), and the effect size was less than the small effect size benchmark (Cohen's  $d = 0.169$ ). This indicates that second-year

students' overall performance slightly declined, but the change was not significant. Similarly, their IELTS reading scores showed a slight decrease from Spring ( $M = 6.13$ ,  $SD = 0.66$ ) to Fall ( $M = 6.02$ ,  $SD = 0.67$ ), with a mean difference of 0.11. This change was also not statistically significant ( $t(23) = 0.738$ ,  $p = .468$ ), and the effect size was even smaller (Cohen's  $d = 0.151$ ).

At all testing points, reading score means were higher than overall score means. The mean differences are comparable except for the second-year group. There were slight improvements in IELTS scores from Spring to Fall for the combined and first-year group, though these changes were not statistically significant. Second-year students saw a slight decrease in both scores, but these changes were also non-significant. Additionally, the effect sizes were small to moderate across all comparisons, indicating that differences in scores were minimal. For the t-tests, a Bonferroni correction was implemented due to multiple variables being analyzed and is recommended when multiple means need to be accounted for Type I error (Larson-Hall, 2010). Even with the Bonferroni correction at ( $\alpha = .0083$ ) accounting for six paired-samples t-tests, none of the comparisons met statistical significance.

*RQ3: To what extent do ER reading speed and words read correlate with IELTS reading scores after one year?*

A difficult measure in ER is quantifying overall reading progress such as vocabulary building, reading fluency and comprehension as ER is seen as a long-term effort that activates students' internal language learning mechanisms and the effect is delayed rather than immediate (Renandya et al., 2021). Understanding the data of students' ER factors and their IELTS scores in RQ1 and RQ2, a descriptive statistics table (Table 5) and two Spearman correlations (Tables 6-7) were conducted to extrapolate any relationship. The first correlation reveals differences between their Fall minus Spring reading speed and IELTS reading scores. The second correlation accounts for words read by volume with IELTS reading scores after Spring, and total words read over the entire academic year with Fall IELTS reading scores. Correlation effect sizes were interpreted using Plonsky and Oswald's (2014) benchmarks of  $r_s = .25$  (small),  $r_s = .40$  (medium), and  $r_s = .60$  (large).

Table 5. *Differences in Reading Speed and IELTS Scores from Fall to Spring*

Variable	Combined Group Difference (Fall-Spring)			1st-year Difference (Fall-Spring)			2nd-year Difference (Fall-Spring)		
	<i>n</i>	<i>M</i>	<i>SD</i>	<i>n</i>	<i>M</i>	<i>SD</i>	<i>n</i>	<i>M</i>	<i>SD</i>
Reading Speed Diff	56	16.49	33.57	3	20.62	39.83	2	10.98	22.4
IELTS Reading Diff	56	0.07	0.72	3	0.20	0.72	2	-0.10	0.69

*Note.* The difference was calculated using Fall data minus Spring data.

Table 5 shows an increase for the combined group of 16.49 wpm ( $SD = 33.57$ ) from Spring and an increase in IELTS reading scores of 0.07 ( $SD = 0.72$ ). First-year students saw increases in reading speed ( $M = 20.62$ ,  $SD = 39.83$ ) and IELTS reading scores ( $M = 0.20$ ,  $SD = 0.72$ ), while second-year students saw a slight decrease in reading scores ( $M = -0.10$ ,  $SD = 0.69$ ) even though their reading speed increased ( $M = 10.98$ ,  $SD = 22.4$ ).

Table 6. Spearman Correlation between IELTS Reading scores and Reading Speed

Group	<i>n</i>	<i>r<sub>s</sub></i>	<i>p</i>
Combined Group	5	.138	.309
	6		
1st-year	3	.109	.553
	2		
2nd-year	2	.240	.259
	4		

Note. IELTS reading score and reading speed difference (Fall minus Spring)

Table 6 indicates no statistically significant correlation between reading speed and IELTS reading scores across all groups. For the combined group, the Spearman correlation coefficient was  $r_s = .138$ ,  $p = .309$ . This weak positive correlation suggests a slight association between reading speed and IELTS reading scores, but it is not statistically significant and falls below Plonsky and Oswald's (2014) threshold for a small effect size. For first-year students, the correlation coefficient was  $r_s = .109$ ,  $p = .553$ . The slightly weaker correlation further shows that reading speed does not have a significant impact on IELTS reading scores. For second-year students, the correlation was  $r_s = .240$ ,  $p = .259$ . This slightly stronger correlation suggests a small trend but lacks statistical correlation. These results imply that while reading speed may be a factor in reading performance, it does not directly correlate with the ability to score higher on the IELTS reading test.

Table 7. Spearman Correlation between Words Read and IELTS Reading Scores by Semester

Group	<i>n</i>	<i>r<sub>s</sub></i> (Spring Words and Spring IELTS Reading)	<i>p</i>	<i>r<sub>s</sub></i> (Total Words and Fall IELTS Reading Score)	<i>p</i>
Combined Group	5	.172	.20	.372	.00
	6		4		5
1st-year	3	.202	.26	.282	.11
	2		8		8
2nd-year	2	.23	.27	.604	.00
	4		9		2

To explore the relationship between reading volume and IELTS performance, words read in Spring and the entire year were correlated with corresponding IELTS reading scores. For the combined group, Spring words read showed a weak and non-significant correlation with reading scores ( $r_s = .172$ ,  $p = .204$ ). Each group showed similar correlations coefficients ( $r_s = .202$ ,  $p = .268$ ;  $r_s = .23$ ,  $p = .279$ ) both falling below the small effect size threshold. However, cumulative reading data at the end of the academic year reveals a stronger and significant correlation.

The data for the cumulative academic year using their total word count and Fall IELTS reading scores show significant correlation. For the combined group, total words read had a significant positive correlation with IELTS reading scores ( $r_s = .372$ ,  $p = .005$ ). These correlations represent a small to moderate effect size. First-year students showed a small but weaker, non-significant correlation between total words read and IELTS reading scores ( $r_s = 0.282$ ,  $p = .118$ ). These values are above the threshold for a small effect size, suggesting the impact of words read on

IELTS performance is less significant but measurable. Second-year students exhibited the strongest correlation, with significant and large effect size for reading scores ( $r_s = .604, p = .002$ ). These findings highlight a stronger relationship between words read and IELTS reading performance over the longer duration of one year compared to just Spring. These results are encouraging, yet surprising, as first-year students read more words during the entire year (on average) and had higher reading speeds (Table 2) compared to the second-year group. While ER was impactful, the results suggest that the extra year of EAP classes likely contributed to second-year students' performance and stronger correlation.

## Discussion

*RQ1: What measurable ER progress in reading speed and words read have students made after one year?*

The main principles of ER emphasize that students should read as much as possible and read faster as they read more (Renandya et al., 2021). As a group, all levels were able to increase their reading volume (Table 2) as the progressive nature of the assignments from Spring to Fall was an important factor throughout the year. As the ER program was in its first year, instructor intervention was necessary to set word counts and deadlines to establish the program as an integral part of the curriculum and to create sustained ER habits. Renandya et al. (2021, p. 17) recognize that “there is little chance that students would spend time doing self-selected reading over a period of time, [therefore] teachers play a key role in motivating their students to do ER.” Using Xreading was extremely helpful to expose the students to a variety of genres and levels for them to choose books freely.

These guidelines resulted in a statistically significant increase in reading speeds for the combined eligible students ( $p < .001$ ) including first-year students ( $p = .003$ ), and second-year students ( $p = .012$ ) after applying a Bonferroni correction ( $\alpha = .017$ ). However, all three measures failed to meet the effect size benchmarks. Encouragingly, the lowest group's reading speed reached 128 wpm, which allows EFL learners to comfortably comprehend what they are reading compared to higher reading speeds of 250 wpm, which may hinder EFL learners with reading comprehension (Waring & McLean, 2015). Waring and McLean (2015) comment that 120 wpm is helpful for EFL learners to achieve comprehensibility and adds that the importance of reading speed is sustainability over longer periods rather than just short bursts. By the end of the academic year, the combined group was able to increase their reading speed to 137.20 wpm.

These results align with the meta-analysis by Jeon and Day (2016), which noticed reading fluency (reading rate) had the highest effect size ( $d = 0.83$ ), compared to comprehension ( $d = 0.54$ ) and vocabulary development ( $d = 0.47$ ). When participating in ER, increases in words read and reading speed are typically the first noticeable positive effect, with subsequent improvements in reading comprehension and vocabulary after sustained participation.

*RQ2: Have student IELTS scores positively changed over the same one-year period at two testing points after participating in ER?*

Students were able to increase their IELTS performance; however, the increase was not statistically significant, and effect sizes were small. Two factors may explain these findings: the short gap between the two testing points and increasing test scores being more difficult when the initial test scores are higher.

A study conducted in Australia with 40 international students who retook the IELTS showed some improvements with 8 students (20%) improving their score by one band, 14 (35%) improving by half a band, 12 (30%) maintaining their level, and 6 (15%) decreasing their scores (Craven, 2012). Craven's (2012) students were at a higher level with an overall mean score of 6.71 ( $SD = 0.34$ ) in their initial test and the gap between tests was between 24-36 months. This longer gap allowed a majority of students to improve by at least half a band, but a significant number of students (45%) did not increase their scores. Craven's (2012, p. 40) study emphasizes the difficulties in dramatically improving IELTS scores as students with two or three years of undergraduate study typically see only slight improvements. In comparison, the six-month gap between tests in this report may have contributed to the minimal gains. A longer gap may yield more statistically significant distinctions.

Another factor is the difficulty of improving IELTS scores at higher initial band scores for more advanced English learners (Craven, 2012; Dang & Dang, 2023). A student with a 5.0 to 6.0 band score will probably find it easier to improve than a student with a 6.0 trying to achieve a 6.5-7.0 score (Craven, 2012). This group of first-year students averaged a 5.5 overall band score (5.63 in reading) while second-year students averaged a 6.0 overall band score (6.13 in reading) in Spring, but first-year students improved by 0.2 in both scores while second-year students slightly decreased by 0.06 (-.11 in reading) in Fall. Despite the decline in reading scores, second-year students averaged higher reading scores than first-year students by 0.19.

From the university's perspective, achievement expectations were met with raw scores as first-year students achieved a 5.5 score, and second-year students achieved a 6.0 score, despite the lack of statistically significant improvements between semesters. This score achievement validates the EAP curriculum design and execution as it allows students to confidently apply for future study abroad programs. Each study abroad institution has varying sets of acceptance guidelines but the average minimum acceptance scores tend to be 5.5 overall score or B2 equivalent on the CEFR scale (Legge & Wagner, 2019, p. 47).

*RQ3: To what extent do ER reading speed and words read correlate with IELTS reading scores after one year?*

Two types of data were investigated in RQ3: semester-to-semester progress (reading speed improvements and IELTS score differences) and cumulative progress (Spring words read with Spring IELTS reading scores and total words read with Fall IELTS reading scores). A strong, significant correlation was observed between words read and IELTS scores after one year ( $r_s = .372, p = .005$ ) compared to one semester ( $r_s = .172, p = .204$ ). The correlation is even stronger for second-year students at the end of the year ( $r_s = .604, p = .002$ ). The number of words a learner is exposed to has a positive effect on English proficiency, as demonstrated by the stronger correlation to IELTS reading scores. This aligns with studies in the meta-analyses by Nakanishi (2015) and Jeon and Day (2016) where they conclude that the available research

suggests consistent participation in ER improves students' reading proficiency and it is an effective instructional component.

On the other hand, it was difficult to pinpoint if ER reading speed progress correlates or is statistically significant to IELTS reading score differences (Table 6), even if progress in their standardized test scores correlates with their reading speed progress in real values (Table 5). This counters a study conducted by McLean and Rouault (2017) where they discovered that increased reading rates by ER-participating students positively affected comprehension. One major difference between their study and this report is McLean and Rouault's (2017) concedes that their students' English levels were lower than the students in this report. Their students had an average TOEIC score of 280 (CEFR A1 level) and a starting reading speed of 77 wpm. The students in this report were expected to have at least CEFR B1+ level and their starting reading rate was at least 118 wpm. Secondly, the types of comprehension tests were vastly different. McLean and Rouault's (2017, p. 98) comprehension questions consisted of 10 multiple-choice questions that were "factual questions with answers explicitly available in the text without any negative fact questions or higher-order inference required". The academic IELTS reading test is a much more demanding and grueling test that requires cognitive and critical thinking processes including expeditious and careful reading to locate information, understanding main ideas, analytical reading, evaluation and inferencing (Jang et al., 2019).

Therefore, any correlation should be dealt with caution, as underlying factors beyond ER may have influenced the results. First-year students read more words and at a faster rate than second-year students yet there was a stronger correlation with second-year students' ER and IELTS performance. Second-year students had one extra year of university EAP classes and this is probably the biggest factor. The difference in university level English education and high school level is noticeable because they are better equipped cognitively and have better capacity to apply the factors from ER since they have more experience, background knowledge, and vocabulary (Jeon & Day, 2016; Nakanishi, 2015). The IELTS test takers in Dang and Dang's (2023) study considered university reading tasks to be more challenging than IELTS reading tasks because they were longer and had more complicated vocabulary. It is fair to assume that the extra year of EAP classes gave the second-year students more foundational knowledge to perform better on the IELTS. Additionally, familiarity of exam questions by second-year students also put them at an advantage because they took the IELTS test in their first year. It would have been their third and fourth time taking the test. Familiarity with standardized test format is another critical factor in standardized test success (Kerstjens & Nery, 2000; Lydon & Kramer, 2019)

Lastly, the mixed results support the notion that longer durations are necessary to observe significant effects of ER factors on English proficiency. Semester-to-semester data showed a lack of correlation and significance for both reading speed and volume; however, second-year students had a much higher correlation with reading speed and IELTS reading scores differences ( $r_s = .259$  vs  $r_s = .109$  for first-year students). Additionally, when looking at the effect after one academic year, word volume correlations became stronger. These results closely align with Nishizawa et al.'s (2010) conclusion that short-term ER programs may not show desirable effects and recommend at least two to three years of regular ER to accomplish significant increase in standardized test scores. This highlights that even if statistically mixed results are revealed in the short term, that does not negate the possible benefits of ER factors transferring to

standardized testing. The emphasis of practicing and developing positive habits in ER may not reveal desired results using standardized testing over a short period (Renandya et al., 2021). Grabe and Stoller (2011) believe that building ER habits requires a curriculum-wide, long-term commitment to significantly impact reading development and fluency.

## **Implications and Limitations**

The practical motivation behind this evaluation report was to examine the effectiveness of ER factors, such as reading speed and words read, with quantifiable measures of English proficiency using standardized IELTS test scores. The practicality of using the IELTS as a measure of English proficiency from the university administration is the urgency to produce noticeable year-to-year gains in students' English progress and achievements, place students in appropriate-level classes in their third year, and use the scores for acceptance into study abroad programs. The analysis in this report reveals that implementation of ER into an academic English program can lead to significant reading progress and demonstrates some correlation with total words read and IELTS reading scores. The implications for the ER program for the university are that words read have a more significant effect on IELTS reading scores than reading speed after one year. Even with the lack of correlation with reading speed and IELTS reading scores, the significant gains in reading speed were encouraging and achieving the IELTS score targets set by the university reaffirms the overall objectives of the EAP curriculum.

In post-academic meetings, the decision-makers agreed that assessing IELTS scores and ER reading progress within a holistic approach will provide a more comprehensive view of student reading and English development. Even if one does not directly correlate with the other, the overall achievements in both areas suggest a complementary relationship of the ER program to create successful readers. The inability to find a strong correlation between reading speed and standardized reading test scores necessitates further consideration on how to improve the program to reveal more discernible outcomes. Therefore, the decision-makers are committed to continuing the ER program with adjustments to further align with the core ER principles. For example, to build positive habits of reading, ER assignments will require activity a few times a week instead of completing them at the last minute to meet deadlines. The hope is that this will add better quality to their reading, along with increased quantity in the future. Additionally, the university is considering measuring ER factors with other standardized tests like the TOEIC or Duolingo English Test to offer deeper comparisons. Continuation of ER will provide a longer period for students to participate in ER and present an opportunity to assess any meaningful gains in standardized testing similar to other longitudinal studies (Nishizawa et al., 2010; Yoshizawa, 2008)

This leads to a few limitations that need to be addressed. The most apparent limitation is the number of students participating in ER ( $n = 82$ ) and those who took both IELTS progress checks ( $n = 56$ ) were quite disparate. This resulted in a smaller sample size than originally intended. The lack of a control group to compare ER progress and English proficiency has proved more difficult to reveal results with students that did not participate in ER. Nakanishi (2015) believes two or more control groups are ideal when investigating effects of some measurable treatment. This was impossible to do in this report due to the all-inclusive nature of the curriculum. All

students were required to partake in ER because it was a part of their class grade and this group of students had access to IELTS at two testing points funded by the university. Students in other departments did not have this opportunity, further limiting comparisons.

Another major limitation is the inability to quantify or report the effects of their non-ER reading education and activity, especially for second-year students. Their first-year EAP education most likely affected their higher performance results but it was impossible to quantify or account for these measures. The extra year provided a larger wealth of background English knowledge than first-year students, giving them an advantage. Second-year students also took the IELTS in their first year so familiarity with the test offered another significant advantage. With the narrowed focus on ER performance and IELTS reading scores, the impact of other English instruction along with individual English factors could not be investigated. A separate analysis, including qualitative interviews or questionnaires, could provide a more cogent analysis on how non-reading and non-measurable factors affect ER and other aspects of English proficiency.

The mixed results also make it difficult to pinpoint how duration affects English progress and proficiency. Nuttall (1996) believes that “progress is only measurable over an extensive period” and any measurable test should be in connection with classroom lessons. Second-year students showed a stronger correlation with ER factors and IELTS reading scores even with more diligence and participation from first-year students. Another year of ER investigation may lead to more consistent results due to the incidental effect of ER on linguistic knowledge. One year of ER showed reading rate progress but its relationship with measurable tests is unknown. Reading volume showed a stronger relationship to IELTS reading scores. The longer the students are involved in ER, a more discernible result may be recognized.

Another limitation is this report’s biases toward English and possible overreliance on Xreading to collect ER data. Firstly, the university under evaluation was an EAP program and the IELTS is an English proficiency test. Secondly, Xreading is an English-medium platform so other language decision makers may not find it useful to implement in their own L2 curriculum. Additionally, word count was credited after passing the inherent comprehension quizzes in Xreading. With the main goal of conducting ER in the student’s own time, the assessment of vocabulary and comprehension skills were conducted with the assumed reliance that Xreading quizzes sufficiently measured their comprehension and vocabulary understanding of each book at a 60% passing rate. More critical analysis of Xreading’s comprehension quizzes may be useful to measure the validity and quality of students’ word count.

Thus, implications from this evaluation report of one university must be viewed through the belief that each ER environment is unique in their adoption of ER principles. Day (2015) in a comprehensive review of the ten core ER principles through 44 studies, recognizes the unique approaches of each ER environment by categorizing ER in a continuum: Pure ER (implementing all ten principles), Modified ER (many ER principles), ER Light (some ER principles), and Fringe ER (ER in name only). Day (2015) concedes that there is no single approach to implement ER and the best ER practice is one that benefits the readers most. Using Day’s (2015) continuum, this research falls in between Modified ER and ER Light as there were some variances from each instructor due to individual class progress on formative instruction that can

affect ER implementation. Also, the measurable choices of Xreading and IELTS further highlight the variable dimensions of this evaluation report.

## Conclusion

With an understanding of the limitations and implications, the results from this evaluation report show that words read had a bigger correlation to IELTS reading scores than reading speed. Although this report focused on one university, the correlation to an internationally recognized test like the IELTS offers valuable insights for other decision-makers seeking to quantify ER benefits. The mixed results show difficulties in understanding the true relationship between ER factors and English proficiency but it can be safely argued that exposure to words via ER does have a positive effect in the IELTS reading test. Above all, the importance of commitment and ER curriculum design showed the increased reading fluency of the students after one year. When initiating and strengthening an ER program, the instructor's role is crucial in establishing a supportive environment that fosters positive ER habits and success. Day and Bamford (1998, p. 138) comment that a successful and effective ER program does not just happen; rather, it requires careful thought and preparation from the teacher including providing guidance, counseling, and even creation of activities to support ER. With the results from this report and understanding the successes and limitations, additional support and guidance can enable students to become more successful readers. After the second year of ER implementation, further evaluation and analysis will be presented for dissemination for those interested in or skeptical of ER effects on English proficiency measures.

## References

- Carver, R. P. (1990). *Reading rate: A review of research and theory*. Academic Press.
- Craven, E. (2012). The quest for IELTS band 7.0: Investigating English language proficiency development of international students at an Australian university. *IELTS Research Reports*, 13(1), 1–61. <https://ielts.org/researchers/our-research/research-reports/the-quest-for-ielts-band-7-0-investigating-english-language-proficiency-development-of-international-students-at-an-australian-university>
- Dang, C. N., & Dang, T. N. Y. (2023). The predictive validity of the IELTS test and contribution of IELTS preparation courses to international students' subsequent academic study: Insights from Vietnamese international students in the UK. *RELC Journal*, 54(1), 84–98. <https://doi.org/10.1177/0033688220985533>
- Day, R. R. (2015). Extending extensive reading. *Reading in a Foreign Language*, 27(2), 294–301. <http://hdl.handle.net/10125/66893>
- Day, R. R. & Bamford, J. (1998). *Extensive reading in the second language classroom*. Cambridge University Press.
- Day, R. R. & Bamford, J. (2002). Top ten principles for teaching extensive reading. *Reading in a Foreign Language*, 14(2), 142–146. <http://hdl.handle.net/10125/66761>
- Grabe, W., & Stoller, F. L. (2011). *Teaching and researching reading* (2nd ed.). Routledge.
- IDP IELTS. (n.d.). *IELTS for Organizations*. <https://ieltsjp.com/japan/about/ielts-for-organisations/en-gb>

- IELTS. (2023, July 18). *Everything you need to know about IELTS and the CEFR*. IELTS.  
<https://ielts.org/news-and-insights/everything-you-need-to-know-about-ielts-and-the-cefr>
- Iwashita, N., Sasaki, M., Stell, S., & Yucel, M. (2021). Japanese stakeholders' perceptions of IELTS writing and speaking tests and their impact on communication and achievement. *IELTS Research Reports Online Series*, 4, 1–81. <https://ielts.org/researchers/our-research/research-reports/japanese-stakeholders-perceptions-of-ielts-writing-and-speaking-tests-and-their-impact-on-communication-and-achievement>
- Jang, E. E., Kim, H., Vincett, M., Barron, C., & Russell, B. (2019). Improving IELTS reading test score interpretations and utilisation through cognitive diagnosis model-based skill profiling. *IELTS Research Reports Online Series*, 2, 1–77.  
<https://ielts.org/researchers/our-research/research-reports/improving-ielts-reading-test-score-interpretations-and-utilisation-through-cognitive-diagnosis-model-based-skill-profiling>
- Jeon, E. Y. & Day, R. R. (2016) The effectiveness of ER on reading proficiency: A meta-analysis. *Reading in a Foreign Language*, 28(2), 246–265.  
<http://hdl.handle.net/10125/66901>
- Kerstjens, M., & Nery, C. (2000). Predictive validity in the IELTS test: A study on the academic performance of students at a university in Australia. *IELTS Research Reports*, 3, 86–108.  
<https://ielts.org/researchers/our-research/research-reports/predictive-validity-in-the-ielts-test-a-study-of-the-relationship-between-ielts-scores-and-students-subsequent-academic-performance>
- Krashen, S. (1985). *The input hypothesis: Issues and implications*. Longman.
- Larson-Hall, J. (2010). *A guide to doing statistics in second language research using SPSS and R*. Routledge.
- Legge, T., & Wagner, A. (2019). Awareness as an indicator of success in a study abroad preparation program. *Journal of Humanities Research*, 11, 43–79.  
<https://cir.nii.ac.jp/crid/1050565162986611456>
- Lyddon, P. A., & Kramer, B. (2019). Connecting extensive reading to TOEIC performance. In F. Meunier, J. Van de Vyver, L. Bradley & S. Thouësny (Eds.), *CALL and complexity – Short Papers from EUROCALL 2019* (pp. 257–262).  
<https://doi.org/10.14705/rpnet.2019.38.1019>
- Mason, B., & Krashen, S. (2017). Self-selected reading and TOEIC performance: Evidence from case histories. *Shitennoji University Bulletin*, 63, 469–475.
- McLean, S., & Rouault, G. (2017). The effectiveness and efficiency of extensive reading at developing reading rates. *System*, 70, 92–106.  
<https://doi.org/10.1016/j.system.2017.09.003>
- Nakanishi, T. (2015). A meta-analysis of extensive reading research. *TESOL Quarterly*, 49(1), 6–37. <https://doi.org/10.1002/tesq.157>
- Nation, I.S.P., & Waring, R. (2020). *Teaching extensive reading in another language*. Routledge.  
<https://doi.org/10.4324/9780367809256>
- Nishizawa, H., Yoshioka, T., & Fukada, M. (2010). The impact of a 4-year extensive reading program. In A. M. Stoke (Ed.), *JALT2009 Conference Proceedings* (pp. 632–640).  
<https://jalt-publications.org/archive/proceedings/2009/E035.pdf>
- Nuttall, C. (1996). *Teaching reading skills in a foreign language*. Oxford University Press.
- Plonsky, L., & Oswald, F. L. (2014). How big is “big”? Interpreting effect sizes in L2 research. *Language Learning*, 64(4), 878–912. <https://doi.org/10.1111/lang.12079>

Junn: Evaluating extensive reading speed and words read with IELTS reading scores

- Renandya, W. A., Hidayati, M., & Ivone, F. M. (2021). Extensive reading: Top ten implementation issues. *JACET Journal*, 65, 11–21.  
[https://doi.org/10.32234/jacetjournal.65.0\\_11](https://doi.org/10.32234/jacetjournal.65.0_11)
- Sakurai, N. (2011). The relationship between the total number of words read and test scores in an extensive reading program. *Kyoto Sangyo University essays. Humanities series*, 43, 112–122.
- Shareef, L.A. (2023). Improving IELTS speaking skills of EFL learners through the extensive reading at a tertiary level. *Amazonia Investiga*, 12(63), 123–133.  
<https://doi.org/10.34069/AI/2023.63.03.11>
- Suss, M. (2023). Examining the reliability of the International English Language Testing System design. *Bulletin of Advanced English Studies*, 8(2), 72–92.  
<https://doi.org/10.31559/BAES2023.8.2.4>
- Tabata-Sandom, M. (2023). A case study of the impact of online extensive reading on the L2 reading motivation, habits, and linguistic abilities of advanced L2 English learners. *Reading in a Foreign Language*, 35(2), 160–189. <https://hdl.handle.net/10125/67446>
- Tagane, Y., Naganuma, N., & Dougherty, P. (2018). Xreading versus paper-based extensive reading: Perceptions of students. In R. Ruegg and C. Williams (Eds.), *Teaching English for Academic Purposes (EAP) in Japan: Studies from an English-medium University* (pp. 77–97).
- Waring, R., & Mclean, S. (2015). Exploration of the core and variable dimensions of extensive reading research and pedagogy. *Reading in a Foreign Language*, 27(1), 160–167.  
<http://hdl.handle.net/10125/66708>
- Xreading. (n.d.). *About Xreading*. <https://www.xreading.com/login/index.php#nav-aboutxreading>
- Yamashita, J. (2008). Extensive reading and development of different aspects of L2 proficiency. *System*, 36, 661–672. <https://doi.org/10.1016/j.system.2008.04.003>

### About the Author

Hanon Junn is a language specialist at Musashi University in Tokyo, Japan. Concurrently, he is a PhD candidate at Temple University. His research interests include L2 motivation, CALL implementation in curriculum design, and the effects of social and tutorial digital platforms in developing L2 competence. E-mail: [hanonjunn@gmail.com](mailto:hanonjunn@gmail.com); ORCID: <https://orcid.org/0000-0002-8180-8882>.