

The Impact of the AI-based Immersive Reader on Korean High School Students' English Reading Comprehension and Perceptions

Yujin Moon (Pusan National University)

Received: 14 July 2025
Revised: 1 November 2025
Accepted: 12 December 2025

Moon, Yujin. (2025). The impact of the AI-based Immersive Reader on Korean high school students' English reading comprehension and perceptions. *Modern English Education*, 26, 477-497.

Keywords

Immersive Reader, artificial intelligence, English reading comprehension, perception, high school English
몰입형리더, 인공지능, 영어 독해력, 인식, 고등학교 영어

Yujin Moon

PhD Candidate
Department of English Language Education
Pusan National University
munyujin@pusan.ac.kr

*This paper is a revised and supplemented version of the author's master's thesis.

Abstract

This study examined the impact of the AI-based Immersive Reader on Korean high school students' English reading comprehension and their perceptions. Over eight weeks, the experimental group ($n = 21$) used Immersive Reader to read L2 texts, while the control group ($n = 19$) participated in traditional reading instruction. Data from pre- and post-tests, as well as surveys, were analyzed quantitatively. The findings reveal several key points: First, the experimental group demonstrated a statistically significant improvement in reading comprehension. Second, this enhancement was particularly evident among intermediate-proficiency students, who showed the most substantial gains. Third, survey results indicated that using Immersive Reader led to more positive attitudes toward English learning and reading, boosting students' interest, confidence, and motivation while decreasing anxiety. Overall, these results suggest that Immersive Reader is an effective AI-based tool for enhancing both cognitive and affective aspects of L2 learning. The study underscores the potential of incorporating Immersive Reader into English education to foster self-directed learning.

INTRODUCTION

In the era of digital transformation, the integration of technology into education has become essential. In particular, the field of second and foreign language education has seen the emergence of technology-assisted learning tools as promising solutions to enhance language acquisition. Reflecting this global trend toward technology-enhanced language learning (TELL) (Ghanizadeh et al., 2015; Iberahim et al., 2023; Shadiev & Yang, 2020), the Korean Ministry of Education (2021) emphasized the importance of incorporating artificial intelligence (AI) into education as part of the 2022 revised national curriculum. This curriculum aims to foster digital and AI-enhanced learning environments and to promote personalized learning through a more flexible and adaptive approach.

In Korea, English has been a compulsory subject throughout primary and secondary education and remains a core

component of the College Scholastic Ability Test (CSAT). Despite spending considerable time learning English, many Korean high school students still have difficulty with reading comprehension, which is crucial for language acquisition (Adapon & Mangila, 2020). Several factors may contribute to this challenge, including reading classes that primarily focus on test-oriented strategies (Kim & Lee, 2014), a lack of effective reading strategies (Lim, 2009), and the psychological burden associated with language learning (Choi & Kim, 2022).

Research suggests that incorporating technology into reading instruction can help address comprehension gaps (Hidayat, 2024; Jamshidifarsani et al., 2019). In line with this, Immersive Reader (Microsoft, n.d.) can serve as a learning tool that offers a potentially effective approach to enhancing students' reading comprehension. It can support learners of diverse ages and abilities by providing useful features such as text-to-speech, grammar tools, text-customizing tools, line focus, translation, and picture dictionary. However, empirical research evaluating the effectiveness of Immersive Reader in authentic classroom contexts remains limited—particularly in secondary-level English as a Foreign/Second Language (EFL/ESL) settings.

Accordingly, this study investigates the impact of Immersive Reader on Korean high school students' English reading comprehension and perceptions. This research contributes to the growing body of literature on educational AI and provides valuable insights into refining AI-driven tools to facilitate self-directed learning.

LITERATURE REVIEW

Artificial Intelligence in Language Learning and Teaching

The integration of AI into language education has garnered increasing scholarly interest, with numerous studies underscoring its transformative potential in language acquisition. AI technologies create adaptive, personalized, and interactive learning environments, fundamentally reshaping how languages are taught and learned (Almelhes, 2023; Mushthoza et al., 2023). Existing research has explored the application of AI across a wide range of language skills. For example, Huang et al. (2023) investigated the effectiveness of AI-based instruction for reading, while Liu et al. (2022) and Sun (2023) examined the effects of AI on speaking proficiency. AI-supported tools have been shown to significantly improve EFL learners' writing performance (Gayed et al., 2022) and listening skills (Xiao, 2025). Collectively, these studies indicate that AI plays a meaningful role in developing learners' linguistic competence by offering tailored support across key language domains: listening, speaking, reading, and writing.

In addition to linguistic outcomes, research highlights the positive influence of AI on affective variables such as learner motivation, anxiety reduction, and willingness to communicate (AITwijri & Alghizzi, 2024). Chatbots and AI-powered platforms create emotionally supportive environments that sustain learners' interest and confidence (Alshammari, 2025; Parsakia, 2023; Yang & Zhao, 2024). AI-enhanced instruction also plays a critical role in fostering self-regulated learning. As noted by Wei (2023), AI tools that provide real-time feedback and adaptive learning pathways enable learners to set goals, monitor their progress, and reflect on their learning strategies. This ultimately fosters metacognitive awareness and strategic planning, both of which are essential for effective language learning (Qiao & Zhao, 2023).

Meta-analytical findings further affirm the effectiveness of AI in language learning. Wu (2024), in a synthesis of 49 empirical studies, found that AI tools are especially effective in improving listening and speaking skills, with moderate gains in writing and vocabulary. Psychological benefits, such as increased motivation, reduced anxiety, and greater willingness to communicate, were also reported. The findings noted that AI interventions of moderate duration (typically six weeks to six months) tend to yield the most positive outcomes.

As demonstrated in previous studies, language learners benefit from AI-based scaffolding, which “has been adapted to mean structured guidance through digital tools, particularly AI-driven systems” (Fan, 2025, p. 144). This concept aligns with Wood et al. (1976), who conceptualized scaffolding as methods and tools that help learners bridge the gap between their current abilities and potential performance. In technology-enhanced learning contexts, AI-based scaffolding can create new possibilities for the advancement of personalized learning by offering immediate feedback, systematic guidance, and detailed instruction throughout the learning process (Akram et al., 2025). For example, adaptive scaffolding that modifies text difficulty and offers glossaries tailored to learners' performance can enhance cognitive efficiency and facilitate comprehension (Wu et al., 2025).

On the other hand, researchers have cautioned that AI could negatively impact learners' cognitive growth and higher-order thinking skills, as heavy dependence on AI tools could weaken their ability to learn independently and solve problems (Darwin et al., 2024; Liu & Wang, 2024). Because these tools frequently provide instant answers or hints for linguistic

difficulties, students risk becoming passive receivers of information rather than active language users who engage in language construction (Feng, 2024).

In summary, while AI demonstrates considerable potential to enhance L2 learning by fostering linguistic competence, affective engagement, and self-regulated learning, it also presents notable challenges regarding learners' overreliance on AI technologies, which may hinder the development of independent learning, problem-solving, and higher-order thinking skills. As AI technologies continue to evolve, ongoing empirical inquiry is essential to deepen our understanding of its pedagogical affordances and limitations, ensuring that its integration supports rather than undermines learners' linguistic and affective development.

Microsoft Immersive Reader

Microsoft Immersive Reader (hereafter IR) is a digital tool originally developed to enhance literacy for students with learning difficulties, particularly those with dyslexia and attention deficit disorders (Osman, 2024). Over time, its application has expanded beyond special education contexts to include broader educational settings, especially in language learning (Santos, 2025; Sarwat et al., 2024). IR is widely integrated into Microsoft platforms such as Word, Outlook, Microsoft Edge, and Microsoft Teams as an assistive technology to support language learning.

The tool offers a suite of features—text-to-speech (TTS), adjustable font size and spacing, grammar highlighting (e.g., parts of speech), line focus, built-in translation, and picture dictionary—that are designed to reduce the cognitive load of reading and promote multimodal engagement with text. These features are aligned with contemporary understandings of how learners process and comprehend written language through multimodal texts (Januarty & Nima, 2018; Unsworth, 2014).

Among these features, the TTS feature (Figure 1) is particularly noteworthy. It allows learners to hear written text read aloud at adjustable speeds while simultaneously highlighting each word in real time with a rectangular box. This combination of auditory and visual input enhances phonological awareness and decoding skills, which are essential for beginning and intermediate L2 learners (De Araújo Gomes et al., 2018). Additionally, repeated auditory exposure via TTS supports improvements in word pronunciation (Uchihara et al., 2023).

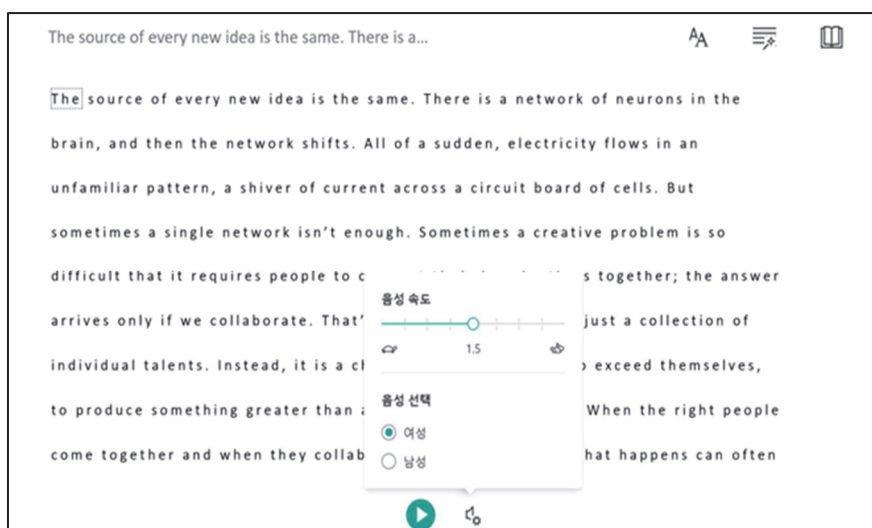


FIGURE 1
Text-to-Speech

Another core feature is the line focus (Figure 2), which allows readers to isolate one, three, or five lines at a time. According to cognitive load theory (Sweller, 1988), minimizing extraneous load helps learners allocate more resources to intrinsic processing, thereby facilitating deeper comprehension of meaning (Sweller & Chandler, 1994). By reducing visual distractions and supporting sustained attention, the line focus aligns with this principle. It is particularly beneficial for learners with limited working memory or those prone to cognitive overload when processing dense texts.

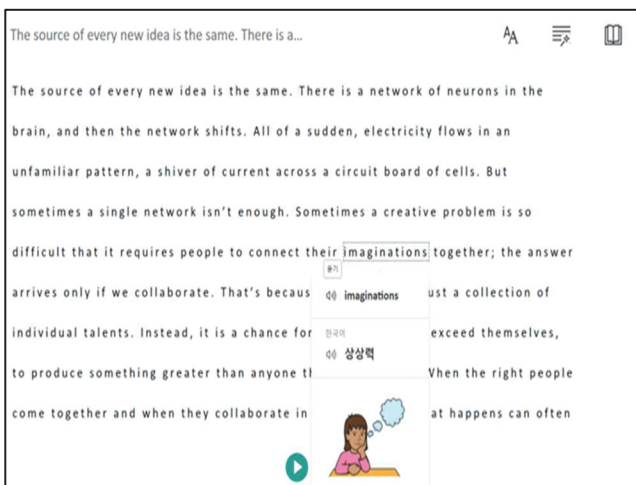


FIGURE 4
Picture Dictionary

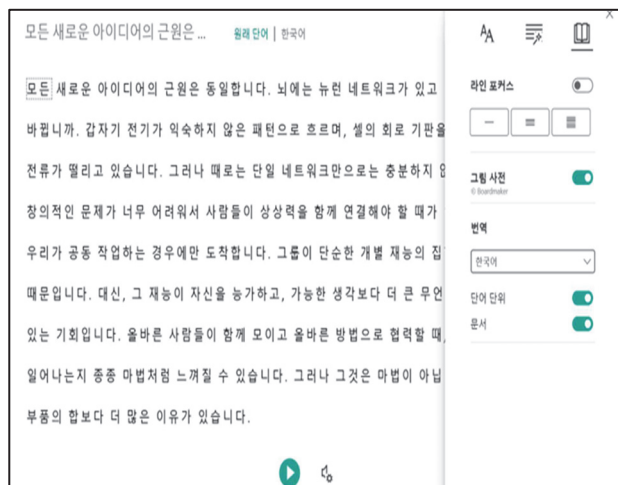


FIGURE 5
Translation Tool

Lastly, the customizable display options such as font type/size, spacing, and background color (Figure 6) serve as accessibility features for optimizing the reading environment to suit learner preferences and to make reading more comfortable. These features align with the principles of learner-centered design, supporting diverse reading needs and preferences.

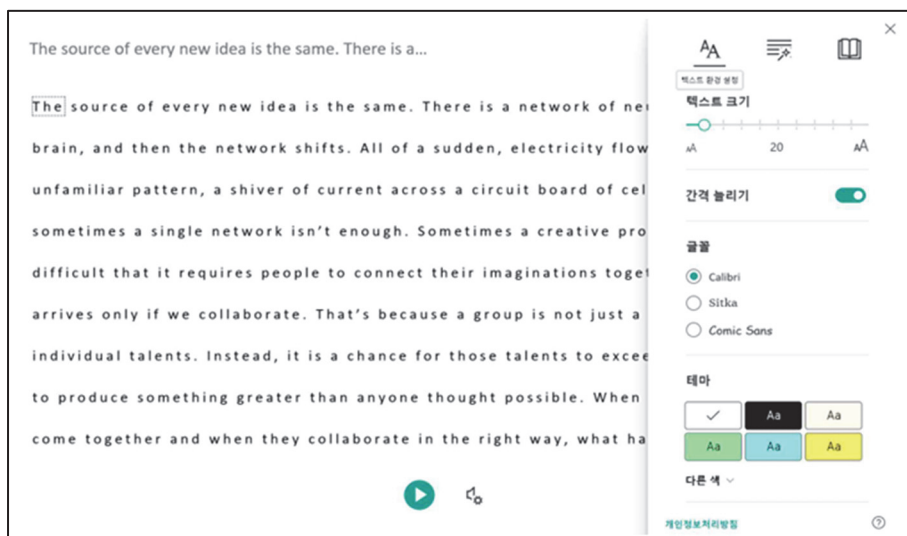


FIGURE 6
Text-Customizing Tool

Previous Studies on Microsoft Immersive Reader

Recent research on IR has demonstrated its positive impact on L2 learning. For instance, KhuramShahzad et al. (2021) conducted a quantitative study to examine the effect of IR-based computer-mediated instruction on secondary school EFL learners' reading comprehension in Pakistan. In their study, the control group ($n = 20$) received traditional reading instruction, while the experimental group ($n = 20$) engaged in reading activities using IR. The findings showed that the experimental group significantly outperformed the control group on reading comprehension tests. In the same vein, Bairaqqdar (2025)

investigated the effectiveness of IR in enhancing adult learners' listening comprehension. The study showed that students who used IR for 12-15 hours per week improved both comprehension and retention, handling listening difficulties more effectively.

Moreover, Dewi et al. (2021) demonstrated that IR has a positive effect on learner motivation and engagement in EFL reading. They conducted a qualitative study to explore the effect of IR on learner motivation among Indonesian EFL university students. Drawing on data from interviews, classroom observations, and field notes, the researchers found that both students and the instructor perceived IR as enhancing student motivation and the overall learning quality. Students reported that IR made reading more engaging, improved their understanding of texts, and encouraged more active participation in reading activities. The study also indicated that IR empowers teachers to design reading instruction that is more dynamic and creative, while fostering learner autonomy through self-directed reading.

While certain studies highlight the benefits of IR, others report less favorable findings. Santos (2025) examined the effect of IR on English reading comprehension among 22 Grade 4 pupils at an elementary school in the Philippines. The experimental group used IR to enhance their reading comprehension, and the test and survey results between the two groups were analyzed. Although the post-test mean score of the experimental group was higher than that of the control group, the difference was not statistically significant. Nevertheless, the experimental group expressed positive perceptions of using IR to support their English reading development.

These findings suggest that IR not only facilitates reading comprehension as a form of AI-based scaffolding but also supports learners' affective development in L2 settings. Its combination of features—TTS, visual aids, customizable display options, and embedded language support—aligns with pedagogical principles that emphasize personalized learning. As such, IR can be regarded as a valuable resource for promoting inclusive and accessible literacy instruction in diverse classrooms, where learners' linguistic and affective needs vary widely.

The Present Study

Despite growing interest in IR, few studies have investigated its effects on students of varying proficiency levels and its implications for secondary education. Therefore, the present study aims to examine whether the use of IR in the English reading classroom supports reading comprehension and whether its effectiveness differs across students' English proficiency levels. Furthermore, the study explores potential changes in students' perceptions of English learning after using IR. The study is guided by the following research questions:

Research Question 1. Does the use of Immersive Reader affect Korean high school students' English reading comprehension?

Research Question 2. Does the use of Immersive Reader affect English reading comprehension across different proficiency levels?

Research Question 3. Does the use of Immersive Reader affect Korean high school students' perceptions of English learning?

METHOD

Participants

A total of 40 Korean high school freshmen participated in this eight-week study. None of the participants had previously lived in English-speaking countries or used IR. The English class was held twice a week and aimed to improve students' reading comprehension. Two classes with comparable instructional hours were selected as the experimental group ($n = 21$) and the control group ($n = 19$). There was no statistically significant difference between the pre-test scores of the two groups, confirming homogeneity. The participants in the experimental group were further divided into three subgroups based on proficiency levels (high, intermediate, and low), as determined by their pre-test scores. The information about the participants is presented in Table 1.

TABLE 1
Participants

Group	Number of participants	Proficiency level	Number of participants
Experimental	21 (14 females, 7 males)	High	6
		Intermediate	8
		Low	7
Control	19 (13 females, 6 males)		

Immersive Reader-based Reading Instruction

Experimental Procedure

The experimental procedure (Table 2) was designed to evaluate the effectiveness of IR in English reading. During the first week, baseline data were collected through the administration and analysis of the pre-test and pre-survey. Over the following eight weeks, the experimental group participated in IR-based reading instruction, while the control group engaged in traditional reading instruction. The final three days were allocated for the post-test and post-survey administration to collect final data.

TABLE 2
Experimental Procedure

Period	Procedure
1 week	<ul style="list-style-type: none"> • Administration of pre-test and pre-survey • Analysis of pre-test and pre-survey results
8 weeks	<ul style="list-style-type: none"> • Experimental group: participating in Immersive Reader-based English reading instruction • Control group: engaging in traditional English reading instruction
3 days	<ul style="list-style-type: none"> • Administration of post-test and post-survey

Lesson Plan

The experimental group received a demonstration on how to use IR and was instructed to read English texts with it. Each participant was provided with a tablet and guided to scan the texts using the Microsoft Lens application, convert them into digital format, and manipulate various IR features. In contrast, the control group was instructed to listen to audio files as a whole class and read the texts individually, which reflects a typical reading instruction commonly adopted in many Korean secondary schools. Table 3 outlines the lesson plans for both the experimental and control groups, highlighting the commonalities and key differences in the lessons. Appendix A presents a sample lesson plan incorporating IR in L2 reading.

The lesson was structured into three stages: pre-reading, while-reading, and post-reading. The pre-reading stage, lasting five minutes, served as a common experience for both groups. During this phase, students predicted the content from the pictures or photos in the text or activated background knowledge by watching short videos or listening to storytelling. According to Shin et al. (2019), when learners lack relevant background knowledge, the instructor should provide it to facilitate understanding. If learners already possess it, guidance should be offered to activate it. Ultimately, providing or activating background knowledge increases familiarity with the text and supports comprehension.

The while-reading stage, lasting 20 minutes, constituted the core of the instructional treatment and marked the primary procedural difference between the two groups. The experimental group engaged with L2 texts using IR: listening to the text through the TTS function and manipulating other features such as font size, line spacing, background color, syllable separation, parts of speech, line focus, reading speed, translation, and picture dictionary. In contrast, the control group carried out their reading activities in a conventional way: they listened to the text as a whole class using audio files and individually read the text with a word list provided by the teacher. Following this, both groups engaged in the same 20-minute activities: taking a true/false quiz on the text, sharing their answers, receiving verbal feedback from the teacher, and listening to the teacher's further explanation of the text.

In the post-reading stage, the lesson concluded with a final five-minute activity in which both groups summarized the text using a graphic organizer that visually presents the relationship between the main idea and supporting details. Employing a graphic organizer has been shown to enhance reading comprehension (Hernández-Chérrez et al., 2020) and facilitate the long-term internalization of textual content.

Reading Materials

One text from the Visang high school English I textbook (Hong et al., 2017) and nine passages from EBS Olympus English Reading Basic I (EBS, 2022) were selected as the reading materials for this study. The Visang (Hong) English I textbook has been widely recognized as a representative English textbook in high school English education. EBS supplementary materials have also commonly been used because they are nationally authorized and provide reliable content aligned with the high school English curriculum and the CSAT. Therefore, the use of both Visang and EBS supplementary materials in high school can be considered pedagogically valid, as they reflect the curricular standards that Korean high school students are required to meet.

TABLE 3
Lesson Plan

Stage	Treatment (Time)	The experimental group	The control group
Pre-reading	Common (5 min)	<ul style="list-style-type: none"> • predicting the content from pictures or photos in the text • activating background knowledge by watching a short video or listening to storytelling 	
While-reading		<ul style="list-style-type: none"> • listening to the text using IR individually <ul style="list-style-type: none"> - 'Text-to-Speech' feature to select the speed and voice type 	<ul style="list-style-type: none"> • listening to the text using audio files as a whole class
	Differentiated (20 min)	<ul style="list-style-type: none"> • reading the text individually by using IR features <ul style="list-style-type: none"> - 'highlight parts of speech' feature - 'adjust font type, font size, line space' feature - 'choose background color' feature - 'split words into syllables', 'show parts of speech', and 'use line focus' features - 'translate content in real-time' feature - 'display pictures for common words' feature 	<ul style="list-style-type: none"> • reading the text individually - learning vocabulary from a word list
	Common (20 min)	<ul style="list-style-type: none"> • taking a true/false quiz and sharing answers • getting verbal feedback on answers from the teacher • listening to the teacher's further explanation of the text 	
Post-reading	Common (5 min)	<ul style="list-style-type: none"> • summarizing the text using a graphic organizer 	

Specific information on the reading materials is presented in Table 4. The text from Chapter 2 of the Visang (Hong) English I textbook is about teenagers' advice about fashion, with the theme of self-understanding through fashion styles, which is considered an appealing topic for first-year high school students. This chapter consists of 648 words across 38 sentences, with an average of 17.1 words per sentence. The Flesch Reading Ease Score (FRES)—a readability measure that indicates how easy or difficult a text is to understand based on sentence length and word complexity—is 62.6. The Flesch Kincaid Grade Level (FKGL)—a readability test that estimates the U.S. school grade level required to understand a given text—is 8.8. This indicates that the text is suitable for 8th-9th-grade students and is written in plain English. Appendix D presents reading difficulty and school level by Flesch Reading Ease Score.

TABLE 4
Reading Materials

Textbook	Chapter	Passage	Domain	Word count	Sentence count	Avg. words per sentence	FRES	FKGL	Reading grade level
Visang (Hong)	2	-	Fashion	648	38	17.1	62.6	8.8	8th-9th
		1	Social studies	144	10	14.4	65.3	7.7	8th-9th
EBS	10	2	Psychology	141	5	28.2	68.2	10.7	8th-9th
		3	Science	152	11	13.8	65.9	7.5	8th-9th

Note. FRES = Flesch Reading Ease Score, FKGL = Flesch Kincaid Grade Level

The EBS passages come from three chapters: Chapters 10, 12, and 15. The three passages from Chapter 10 focus on the topics in social studies, psychology, and science. The three passages in Chapter 12 address the topics in the domains of history, philosophy, and science. Chapter 15 consists of three passages, which cover the topics of psychology, philosophy, and economics. Chapter 10 focuses on completing blanks with long phrases, Chapter 12 deals with identifying irrelevant sentences, and Chapter 15 covers summary completion. Since all three chapters represent major reading comprehension question types included in CSAT, they were considered appropriate for this study. The number of sentences, word count, and reading grade level of each text/passage were analyzed and found to be appropriate for first-year high school students. For example, passage 1 in Chapter 10 consists of 144 words across 10 sentences, with an average of 14.4 words per sentence, and a reading grade level of 8th-9th grade. Passage 2 contains 141 words in 5 sentences, with an average of 28.2 words per sentence, and a reading grade level of 8th-9th grade. Passage 3 comprises 152 words in 11 sentences, with an average of 13.8 words per sentence, and a reading grade level of 8th-9th grade.

The information about the other passages from Chapters 12 and 15 is presented in Appendix E. The passages in the EBS supplementary materials show a wide range of reading grade levels, from 7th grade to college graduate level; however, since their difficulty is similar to that of the actual CSAT reading passages, they are considered appropriate for high school students' English learning. In addition, because the EBS passages are not limited to a specific level of difficulty or topic, they were regarded as suitable materials for objectively analyzing the educational effects of IR.

Data Collection

Pre- and Post-Test

The study investigated the impact of using IR on high school students' English reading comprehension and perceptions. The pre- and post-test items were collected from the National Unified Academic Achievement Test (NUAAT), administered by the Seoul Metropolitan Office of Education. The test items from the March 2019 test (Appendix B) and the March 2020 test (Appendix C) were used as the pre- and post-test items in this study. The objectivity of the pre- and post-tests can be ensured because every March NUAAT is administered by the Seoul Metropolitan Office of Education and maintains a consistent level of difficulty and a stable set of question types, reflecting those of the CSAT. It was also unlikely that the students had been exposed to both tests because the school, designated as an Educational Capacity Enhancement School by the Metropolitan Office of Education, primarily serves students with low academic achievement and limited access to private tutoring, largely due to financial constraints, low motivation, and other related factors. This background information suggests minimal prior exposure to the test items, which helped reduce test-related bias and supported the validity of the test results.

The pre- and post-tests comprised a total of 28 multiple-choice questions, each worth 2-3 points, yielding a maximum score of 63. The reading comprehension questions are divided into six main categories. Main idea test items evaluate the ability to understand the overall meaning of a text. Detail comprehension test items require careful reading to match specific information with given options. Grammar and vocabulary judgment test items assess the proper use of grammar and expressions within the context of the passage. Blank completion test items are often difficult because they require students to fill in a blank based on the logical relationships and overall content of the text. Indirect writing test items evaluate the ability to identify the flow and logical relationships in a text by finding an irrelevant sentence or correctly ordering sentences. Finally, complex-type test items combine several reading skills within a single long passage with multiple sub-questions.

Pre- and Post-Survey

Surveys were conducted to investigate the effects of using IR on students' perceptions of English learning. The pre- and post-survey items were adapted and reconstructed from the survey items developed by Oh et al. (2016), which examined middle school students' affective attitudes toward English learning. All survey items employed a four-point Likert scale (4: strongly agree, 3: agree, 2: disagree, 1: strongly disagree), excluding the ambiguous midpoint option of "neutral." This design was intended to prevent students from indiscriminately choosing a neutral response without careful judgment. The pre- and post-surveys consisted of 22 items measuring students' overall perceptions of English learning, and the items were identical in both surveys. As shown in Table 5, the surveys included five items on attitudes toward English learning, five on English reading strategies, and three each on interest, confidence, motivation, and anxiety related to English learning. The complete set of the survey items is provided in Appendix F.

Data Analysis

Pre- and Post-Test

TABLE 5
Constructs of Survey Items

Survey domain	Item number	Number of items	Response type
Attitude toward English learning	1, 2, 3, 4, 5	5	
English reading strategy	6, 7, 8, 9, 10	5	
Interest	11, 12, 13	3	a four-point
Confidence	14, 15, 16	3	Likert scale
Motivation	17, 18, 19	3	
Anxiety	20, 21, 22	3	

The results collected from the pre- and post-tests were analyzed using IBM SPSS Statistics 28.0 with a significance level set at $p < .05$. The Shapiro-Wilk test was employed to assess the normality of the pre- and post-test results in each group with fewer than 30 participants. The Mann-Whitney test was used to compare the pre-test results between the experimental and control groups. The independent t -test was applied to compare the post-test results between the two groups. Effect size (Cohen's d) was calculated to measure the magnitude of the observed effect. The Wilcoxon signed-rank test was used to compare the pre- and post-test results within each group. Additionally, the Wilcoxon signed-rank test was conducted to examine whether reading comprehension differed across English proficiency levels.

Pre- and Post-Survey

The results collected from the pre- and post-surveys were analyzed using IBM SPSS Statistics 28.0 with a significance level set at $p < .05$. A reliability analysis was conducted to determine the internal consistency of the survey items, and a t -test was performed to examine the difference in the experimental group's perceptions before and after using IR.

FINDINGS

Comparison of Pre- and Post-test Results between Two Groups

To assess the normality of the pre- and post-test results for each group, the Shapiro-Wilk test was conducted. The p -values for the pre-test result were 0.005 for the experimental group and 0.043 for the control group. Since both significance values are below the 0.05 significance level, the null hypothesis was rejected at the 95% confidence level, indicating that the data violated the assumption of normality. Because the pre-test results of the groups follow a non-normal distribution, the Mann-Whitney test was used to assess homogeneity between the groups. As shown in Table 6, the Mann-Whitney U value was

185.000, the Z value was -0.393, and the p -value was 0.708. Since this p -value is greater than the 0.05 threshold, there was no statistically significant difference, confirming homogeneity between the groups.

TABLE 6
Comparison of Pre-test Results between Two Groups

Group	Number	Average rank	Sum of ranks	Mann-Whitney U	Z	p
Experimental	21	21.19	445.00	185.000	-0.393	0.708
Control	19	19.74	375.00			

Note. * $p < .05$

To examine whether the use of IR led to a change in English reading comprehension, the post-test results of the groups were compared. A normality test of the post-test results showed that the p -value for the experimental group was 0.194 and that for the control group was 0.578. Since both groups had p -values greater than 0.05, the null hypothesis was not rejected at a 95% confidence level, indicating that the assumption of normality was satisfied. Therefore, the post-test results of the two groups followed a normal distribution, and an independent t -test was conducted.

Table 7 presents the results of the independent t -test for the post-test results between the two groups. The average post-test score of the experimental group was 37.71 points, while that of the control group was 29.37 points, resulting in a mean difference of 8.34 points between the two groups. According to Levene's test for equality of variances, the p -value was 0.660, which is greater than the 0.05 significance level, indicating that the assumption of equal variances was satisfied. Under this assumption, the independent t -test yielded a p -value of 0.045, indicating a statistically significant difference at the 0.05 significance level.

In order to determine the effect size, which shows how much the means of the two groups actually differ, Cohen's d was calculated. Cohen's d value was 0.657, indicating a moderate difference in post-test scores between the two groups. Since this value falls between the thresholds for a medium effect size (0.5) and a large effect size (0.8), it can be concluded that IR has a moderately strong effect on learners' English reading comprehension.

TABLE 7
Comparison of Post-test Results between Two Groups

Group	Number	Mean	Standard deviation	t	p	Effect size (Cohen's d)
Experimental	21	37.71	13.22	2.073	0.045*	0.657
Control	19	29.37	12.44			

Note. * $p < .05$

Figure 7 highlights the changes in median scores from pre-test to post-test for both groups. While both groups demonstrated an increase in their post-test medians, this improvement was substantially more pronounced in the experimental group. The decision to focus on the median, which represents the 50th percentile or the exact midpoint of the data, is methodologically appropriate. This is because, unlike the mean, the median is not skewed by statistical outliers, such as exceptionally high-performing or low-performing students, which makes it a more robust and stable indicator of typical participants' performance. Therefore, a pronounced upward shift in the experimental group's median scores provides strong evidence that IR had a significant positive effect, confirming its effectiveness in improving overall performance in reading comprehension, rather than merely affecting a small number of individuals.

Table 8 presents the results of the Wilcoxon signed-rank test conducted to compare pre- and post-test scores for each group. The p -value for the experimental group was 0.002, indicating a statistically significant difference between the pre- and post-test results. This significant difference is further supported by the sum of the ranks. The positive rank sum refers to the sum of the ranks of samples that resulted in a positive difference when the pre-test scores were subtracted from the post-test scores. Conversely, the negative rank sum represents the sum of the ranks of samples that resulted in a negative difference when the same subtraction was performed. For the experimental group, the positive rank sum is 203.00, and the negative rank sum is 28.00, showing a large difference between the two sums. These figures indicate that reading comprehension of the experimental group significantly improved by using IR, demonstrating that IR is an effective tool for enhancing reading comprehension. In contrast, the Wilcoxon signed-rank test for the control group shows a p -value

of 0.983 at the 0.05 significance level, indicating no significant difference between the pre- and post-test results. This suggests that the traditional reading instruction did not lead to a statistically significant improvement in students' reading comprehension.

The comparison of the post-test results between the two groups, together with the pre- and post-test comparison within the experimental group, indicates that the improvement in reading comprehension was significantly greater in the experimental group than in the control group. This suggests that IR-based reading has a more positive effect on students' English reading comprehension than traditional reading instruction.

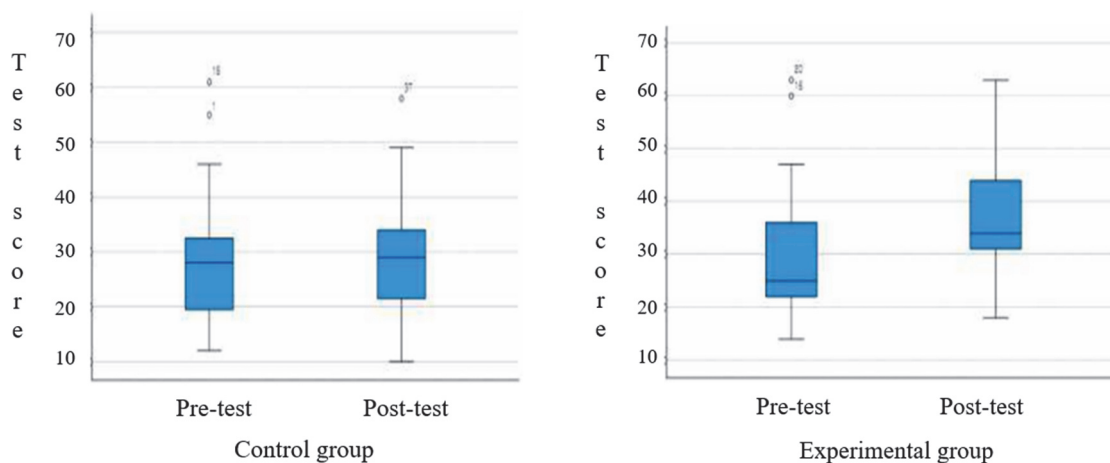


FIGURE 7
Comparison of Median Score Changes from Pre- to Post-Tests for Each Group

TABLE 8
Comparison of Pre- and Post-test Results for Each Group

Group	N	Pre-test	Post-test	Negative rank		Positive rank		Tied rank	Wilcoxon signed-rank test	
		M (SD)	M (SD)	N	MR (SR)	N	MR (SR)	N	Z	p
Experimental	21	30.14 (13.39)	37.81 (13.31)	6	4.67 (28.00)	15	13.53 (203.00)	-	-3.043	0.002*
Control	19	29.10 (13.85)	29.37 (12.45)	9	9.56 (86.00)	9	9.44 (85.00)	1	-0.022	0.983

Note. * $p < .05$

Comparison of Pre- and Post-test Results by Proficiency Level

The study investigated whether the use of IR affects learners' reading comprehension across proficiency levels. Based on the pre-test scores, the experimental group was divided into three proficiency-based subgroups: high-proficiency group ($n = 6$), intermediate-proficiency group ($n = 8$), and low-proficiency group ($n = 7$). Because each subgroup included fewer than 10 students, the Wilcoxon signed-rank test, a non-parametric method, was employed to compare the pre- and post-test results within each group. Table 9 shows the results of the Wilcoxon signed-rank test for the pre- and post-test results across proficiency levels.

For the high-proficiency group, the post-test mean score increased by 4.67 points from the pre-test mean score. The sum of ranks for participants whose scores improved (15.00) was higher than that for those whose scores decreased (6.00). However, the p-value (0.344), which is above the 0.05 significance level, indicates that the improvement was not statistically significant. This suggests that high-proficiency students, who already possessed strong reading comprehension skills, may not have been substantially affected by the experimental treatment using IR for English reading.

TABLE 9
Comparison of Pre- and Post-test Results by Proficiency Level

Group	N	Pre-test	Post-test	Negative rank		Positive rank		Tied rank	Wilcoxon signed-rank test	
		M (SD)	M (SD)	N	MR (SR)	N	MR (SR)	N	Z	p
High	6	48.00 (11.22)	52.67 (11.98)	3	2.00 (6.00)	3	5.00 (15.00)	-	0.946	0.344
Inter- mediate	8	26.38 (2.26)	34.63 (6.50)	1	1.00 (1.00)	7	5.00 (35.00)	-	-2.383	0.017*
Low	7	19.14 (2.91)	28.71 (9.12)	2	1.50 (3.00)	5	5.00 (25.00)	-	-1.866	0.062

Note. * $p < .05$

Regarding the intermediate-proficiency group, the post-test mean score increased by 8.25 points from the pre-test mean score. The sum of ranks for score increases (35.00) was substantially higher than that for score decreases (1.00). The p-value (0.017) is statistically significant, indicating that the use of IR had a positive effect on intermediate-proficiency students. These learners may achieve higher levels of reading comprehension when provided with appropriate IR-based support. For this group, IR functioned as AI-based scaffolding that enhances L2 reading comprehension, demonstrating its educational effectiveness.

For the low-proficiency group, the post-test mean score increased by 9.57 points from the pre-test mean score. Although the sum of ranks for score increases (25.00) exceeded that for score decreases (3.00), the p-value (0.062), which is slightly above the 0.05 significance level, indicates that the improvement did not reach statistical significance. This suggests that while the use of IR led to a positive trend for low-proficiency students, the effect was not strong enough to be considered statistically conclusive. For this group, IR did not function effectively as AI-based scaffolding to support learners in overcoming comprehension challenges and achieving the intended learning outcomes.

Survey Results on Students' Perceptions

Pre- and post-surveys were conducted with the experimental group to examine students' perceptions of English learning, and the results were analyzed. The survey consisted of 22 items across six categories: attitudes toward English learning, English reading strategies, interest, confidence, motivation, and anxiety. Responses were measured using a 4-point Likert scale. Reliability analysis for all 22 items revealed Cronbach's alpha values of 0.858 for the pre-survey and 0.910 for the post-survey. Since a Cronbach's alpha above 0.70 is generally considered acceptable for internal consistency, the reliability of both surveys was considered high.

Table 10 presents a comparison of the pre- and post-survey results for the experimental group. The group showed statistically significant differences between the pre- and post-survey results, with particularly low p-values for all survey items. These results suggest that the use of IR helped students develop more positive attitudes toward English learning, strengthen their understanding of reading strategies, increase their interest, confidence, and motivation in learning English, and reduce their anxiety about it. This finding aligns with the studies by Dewi et al. (2021) and Santos (2025), which reported that the use of IR enhanced learners' motivation, encouraged active participation in reading classes, and promoted independent use of IR by students.

CONCLUSION

The purpose of this study was to investigate the effects of Microsoft Immersive Reader (IR) on Korean high school students' reading comprehension and perceptions. Specifically, the study examined whether the use of IR improves L2 reading comprehension and whether its effectiveness differs across students' English proficiency levels. Surveys were also administered to identify changes in students' perceptions of English learning, reading strategies, interest, confidence, motivation, and anxiety after using IR. The study was conducted over eight weeks with 40 high school freshmen, who were assigned to the experimental group ($n = 21$) and the control group ($n = 19$). The experimental group engaged in L2 reading

TABLE 10*Comparison of Pre- and Post-survey Results*

Category	Pre-/Post-test	Mean	Standard deviation	<i>t</i>	<i>p</i>	Effect size (cohen's <i>d</i>)
Attitude toward English learning	pre	3.06	0.85	-2.07	0.04*	0.80
	post	3.29	0.74			
Reading strategies	pre	3.05	0.70	-4.59	0.001*	0.65
	post	3.46	0.59			
Interest	pre	2.29	0.85	-3.45	0.001*	0.90
	post	2.84	0.95			
Confidence	pre	2.11	0.84	-3.01	0.003*	0.89
	post	2.59	0.93			
Motivation	pre	2.83	1.04	-2.32	0.02*	0.96
	post	3.22	0.87			
Anxiety	pre	2.84	0.81	-2.46	0.015*	0.76
	post	3.17	0.71			

Note. **p* < .05

using IR features such as text-to-speech, text customization, highlighting, grammar tools, focus mode, translation, and picture dictionary. On the other hand, the control group participated in traditional reading instruction, where they listened to audio files and read the texts individually. Pre- and post-tests and surveys were used to analyze the effects of IR on L2 reading comprehension and perceptions.

The findings of the study are as follows. First, the experimental group showed a statistically significant improvement between pre- and post-tests, whereas the control group did not. This indicates that IR-based L2 reading instruction is more effective than traditional methods for enhancing reading comprehension. Second, an analysis of reading comprehension gains across proficiency levels demonstrated that only the intermediate-proficiency group exhibited a statistically significant improvement. This suggests that IR functions as effective AI-based scaffolding for learners at this level by narrowing the gap between their independent performance and potential achievements. Third, survey results revealed positive perceptual changes in English learning, reading strategies, interest, confidence, motivation, and anxiety among students who used IR.

These results are consistent with previous research reporting that IR-supported language learning facilitates comprehension (Bairaqqdar, 2025; KhuramShahzad et al., 2021) and promotes positive learner perceptions (Dewi et al., 2021; Santos, 2025). In line with KhuramShahzad et al.'s (2021) findings with Pakistani secondary learners, this study confirms that IR provides meaningful scaffolding for Korean secondary students as well. In addition, this study extends prior evidence by demonstrating IR's effectiveness specifically in secondary-school contexts, where empirical studies remain limited compared to elementary (Santos, 2025) or tertiary settings (Bairaqqdar, 2025; Dewi et al., 2021). Moreover, while the overall trend of enhanced comprehension and positive perceptions observed in this study aligns with previous research, the differentiated outcomes across English proficiency levels refine existing conclusions and underscore the importance of learner characteristics—particularly learners' proficiency levels—when implementing IR. Taken together, the findings demonstrate that IR can create a dynamic and responsive L2 reading environment that surpasses the limitations of traditional instruction and supports self-directed learning. Ultimately, as AI-based scaffolding, IR enables learners to navigate challenging texts more independently and strategically.

Several recommendations can be made for future research. First, the small sample size required the use of non-parametric analyses, which may have reduced statistical power and generalizability. Studies with larger samples would provide more robust evidence of IR's effectiveness across proficiency levels. Second, the eight-week duration limits understanding of long-term effects of IR. Extending the intervention could allow learners to engage more deeply with the tool, internalize reading strategies, and achieve more sustained improvements in both linguistic and affective outcomes. Future research should therefore consider larger participant pools and longitudinal designs to further investigate the impact of IR on L2 reading comprehension and learner perceptions. It is hoped that the findings of this study will support broader applications of IR in diverse educational contexts and foster learners' self-directed L2 learning.

References

- Adapon, M. T., & Mangila, B. B. (2020). Helping struggling readers to read: The impact of the care for the non-readers (CRN) program on Filipino pupils' reading proficiency. *ETERNAL (English, Teaching, Learning, and Research Journal)*, 6(2), 171-187. <https://doi.org/10.24252/eternal.v6i2.2020.a2Albano>
- Akram, A., Kiren, A., Sabir, S., & Aslam, S. (2025). The effect of AI based scaffolding on problem solving and metacognitive awareness in learners. *The Critical Review of Social Sciences Studies*, 3(3), 1074-1089. <https://doi.org/10.59075/ja7dvy78>
- Alhumsy, M. H., & Shabdin, A. A. (2016). The relationship between phonemic segmentation skill and EFL word recognition-A review of literature. *International Journal of Linguistics*, 8(2), 31-46. <https://doi.org/10.5296/ijl.v8i2.9097>
- Almelhes, S. A. (2023). A review of artificial intelligence adoption in second-language learning. *Theory and Practice in Language Studies*, 13(5), 1259-1269. <https://doi.org/10.17507/tpls.1305.21>
- Alshammari, M. E. (2025). Promoting emotional and psychological well-being in students through a self-help ChatGPT intervention: Artificial intelligence powered support. *International Journal of Instruction*, 18(3), 511-530. <https://doi.org/10.29333/iji.2025.18326a>
- AITwijri, L., & Alghizzi, T. M. (2024). Investigating the integration of artificial intelligence in English as foreign language classes for enhancing learners' affective factors: A systematic review. *Heliyon*, 10, 1-12. <https://doi.org/10.1016/j.heliyon.2024.e31053>
- Bairaqdar, M. (2025). Using Microsoft's Immersive Reader tool to enhance students' listening comprehension. *Dialog on Language Instruction*, 35(1), 78-82.
- Choi, Eun-Kyung, & Kim, Sung-Yeon. (2022). Korean high school students' anxiety and affective strategies when taking English listening and reading tests. *Korean Journal of English Language and Linguistics*, 22, 978-997. <https://doi.org/10.15738/kjell.22..202210.978>
- Darwin, K., D. Rusdin, N. Mukminatien, N. Suryati, E. D. Laksmi, & Marzuki. (2024). Critical thinking in the AI era: An exploration of EFL students' perceptions, benefits, and limitations. *Cogent Education*, 11(1), 1-18. <https://doi.org/10.1080/2331186X.2023.2290342>
- De Araújo Gomes, A. A., Cardoso, W., & De Lucena, R. M. (2018). Can TTS help L2 learners develop their phonological awareness? In P. Taalas, J. Jalkanen, L. Bradley & S. Thoušny (Eds.), *Future-proof CALL: language learning as exploration and encounters—short papers from EUROCALL 2018* (pp. 29-34). Research-publishing.net. <https://doi.org/10.14705/rpnet.2018.26.808>
- Dewi, S. R., Fatmawati, S., & Widodo, A. (2021). Integrating Microsoft Immersive Reader in SMAR model to promote students' reading comprehension. *Reflection on Education: Journal of Educational Sciences*, 11(2), 190-196. <https://doi.org/10.24176/re.v11i2.5354>
- EBS Editorial Department. (2022). *Olympus English Reading Basic I*. Educational Broadcasting System.
- Fan, K. (2025). Enhancing cognitive reading strategy use for L2 learners with lower working memory through AI-based scaffolding. *Journal of English Language Teaching and Applied Linguistics*, 7(3), 143-151. <https://doi.org/10.32996/jeltal.2025.7.3.14>
- Feng, L. (2024). Investigating the effects of artificial intelligence-assisted language learning strategies on cognitive load and learning outcomes: A comparative study. *Journal of Educational Computing Research*, 62(8), 1741-1774. <https://doi.org/10.1177/07356331241268349>
- Gayed, J. M., Carlon, M. K. J., Oriola, A. M., & Cross, J. S. (2022). Exploring an AI-based writing assistant's impact on English language learners. *Computers and Education: Artificial Intelligence*, 3, 100055. <https://doi.org/10.1016/j.caeai.2022.100055>
- Ghanizadeh, A., Razavi, A., & Jahedizadeh, S. (2015). Technology-enhanced language learning (TELL): A review of resources and upshots. *International Letters of Chemistry, Physics and Astronomy*, 54, 73-87. <https://doi.org/10.18052/www.scipress.com/ILCPA.54.73>
- Hayikaleng, N. (2019). The effects of using dual-coding theory in teaching English reading comprehension among vocational students at Narathiwat Technical College. *International Journal of Arts Humanities and Social Sciences Studies*, 4(6), 7-12.
- Hernández-Chérrez, E., Hidalgo-Camacho, C. S., & Carrera-Martínez, V. (2020). Graphic organizers for the enhancement of EFL reading comprehension. *Chakíñan Journal of Social Sciences and Humanities*, 12, 106-117. <https://doi.org/10.1590/scielopreprints.885>
- Hidayat, M. T. (2024). Effectiveness of AI-based personalised reading platforms in enhancing reading comprehension. *Journal of Learning for Development*, 11(1), 115-125. <https://doi.org/10.56059/jl4d.v11i1.955>
- Hong, Min-Pyo, Ahn, Hyun-Ki, Kim, Jung-Tae, Park, Yeon-Mi, Richard Pak, Jang, Hyun-Ok, Shin, Jeong-Seop, & Cho, Geum-Hee. (2017). *English I*. Visang Education.
- Huang, B., Dou, J., & Zhao, H. (2023). Reading bots: the implication of deep learning on guided reading. *Frontiers in Psychology*, 14, 980523. <https://doi.org/10.3389/fpsyg.2023.980523>
- Iberahim, A., Yunus, M. M., & Sulaiman, N. A. (2023). A review on technology enhanced language learning (TELL). *International Journal of Academic Research in Business and Social Sciences*, 13(2), 1509-1519. <https://dx.doi.org/10.6007/IJARBS/v13-i2/16496>
- Jamshidifarsani, H., Garbaya, S., Lim, T., Blazevic, P., & Ritchie, J. M. (2019). Technology-based reading intervention programs for elementary grades: An analytical review. *Computers & Education*, 128, 427-451. <https://doi.org/10.1016/j.compedu.2018.10.003>

- Januarty, R., & Nima, H. N. A. (2018). Energizing students' reading comprehension through multimodal texts. *International Journal of Language Education*, 2(2), 14-22. <https://doi.org/10.26858/ijole.v2i2.4347>
- KhuramShahzad, S., Panwar, A. H., & Ansari, S. (2021). Impact of computer-mediated instruction on reading proficiency skills of ESL learners at secondary level. *Elementary Education Online*, 20(5), 468-475. <https://doi.org/10.17051/ilkonline.2021.05.50>
- Kim, Da-Hee, & Lee, Jin-Hwa. (2014). University students' reflective perception on their experience of private English education during high school years. *Modern English Education*, 15(3), 242-262.
- Korean Ministry of Education. (2021). 2022 Revised national curriculum for primary, secondary and special schools announced. <https://english.moe.go.kr/boardCnts/viewRenewal.do?boardID=265&boardSeq=93810&lev=0&searchType=null&statusYN=W&page=1&s=english&m=0201&opType=N>
- Lim, Hye-Jung. (2009). Effects of the reading task difficulty and students' achievement on the reading strategy use by Korean high school students. *Modern English Education*, 10(3), 242-266.
- Liu, J., Liu, X., & Yang, C. (2022). A study of college students' perceptions of utilizing automatic speech recognition technology to assist English oral proficiency. *Frontiers in Psychology*, 13, 1-9. <https://doi.org/10.3389/fpsyg.2022.1049139>
- Liu, W., & Y. Wang. (2024). The effects of using AI tools on critical thinking in English literature classes among EFL learners: An intervention study. *European Journal of Education*, 59(4), e12804. <https://doi.org/10.1111/ejed.12804>
- Microsoft. (n.d.). *Immersive Reader*. Microsoft Learn. <https://learn.microsoft.com/en-us/training/educator-center/product-guides/immersive-reader/>
- Mushthoza, D. A., Syariatun, N., Tahalele, O., Telussa, S. I., Rasmitha, R., & Mokodenseho, S. (2023). Analyzing the impact of artificial intelligence on the future of English language teaching and learning. *Journal on Education*, 6(1), 1549-1557. <https://doi.org/10.31004/joe.v6i1.3115>
- Nurhayati, D. (2014). Using picture series to inspire reading comprehension for the second semester students of English department of Iain Tulungagung. *Dinamika Ilmu*, 14(2), 176-189. <https://doi.org/10.21093/di.v14i2.14>
- Oh, Ro-Sa, Park, Young-Min, & Lee, Hee-Kyung. (2016). The effects of syntactically visualized text reading on English reading comprehension of middle school students in Korea. *Language Research*, 52(3), 581-610.
- Osman, F. M. (2024). Innovative teaching: A proposed Immersive Reader training program for special needs teachers. *Journal of ICSAR*, 8(1), 1-5. <https://dx.doi.org/10.17977/um005v8i1p1>
- Paivio, A. (1971). *Imagery and verbal process*. New York: Holt, Rinehart & Winston.
- Park, Young-Min, Xu, Y., Collins, P., Farkas, G., & Warschauer, M. (2019). Scaffolding learning of language structures with visual-syntactic text formatting. *British Journal of Educational Technology*, 50(4), 1896-1912. <https://doi.org/10.1111/bjet.12689>
- Parsakia, K. (2023). The effect of chatbots and AI on the self-efficacy, self-esteem, problem-solving and critical thinking of students. *Health Nexus*, 1(1), 71-76. <https://doi.org/10.61838/hn.1.1.14>
- Qiao, H., & Zhao, A. (2023). Artificial intelligence-based language learning: Illuminating the impact on speaking skills and self-regulation in Chinese EFL context. *Frontiers in Psychology*, 14, 1255594. <https://doi.org/10.3389/fpsyg.2023.1255594>
- Roslina, R. (2017). The effect of picture story books on students' reading comprehension. *Advances in Language and Literary Studies*, 8(2), 213-221. <https://doi.org/10.7575/aiac.all.v.8n.2p.213>
- Sadoski, M., Paivio, A., & Goetz, E. T. (1991). Commentary: A critique of schema theory in reading and a dual coding alternative. *Reading Research Quarterly*, 26, 463-484. <https://doi.org/10.2307/747898>
- Sadoski, M., & Paivio, A. (1994). A dual coding view of imagery and verbal processes in reading comprehension. In R. B. Ruddell, M. R. Ruddell, & H. Singer (Eds.), *Theoretical models and processes of reading* (4th ed.) (pp. 582-601). Newark, DE: International Reading Association.
- Santos, C. (2025). Improving reading comprehension skills in English for grade 4 pupils through Microsoft Immersive Reader. *Asia Pacific Higher Education Research Journal*, 12(1), 1-17. <https://doi.org/10.56278/apherj.v12i1.3198>
- Sari, D. (2017). Using text-related picture to increase students' reading comprehension of descriptive text. *Inovish Journal*, 2(2), 116-132. <https://doi.org/10.35314/inovish.v2i2.238>
- Sarwat, S., Manzoor, A., Shakoor, A., Tufail, N., & Kokab, N. (2024). Investigating the perceptions and attitudes of ESL learners towards the use of Immersive Reader technology in enhancing reading comprehension at the secondary school level. *Spry Contemporary Educational Practices*, 3(1), 212-228. <https://doi.org/10.62681/sprypublishers.scep/3/1/12>
- Shadiev, R., & Yang, M. (2020). Review of studies on technology-enhanced language learning and teaching. *Sustainability*, 12(2), 524. <https://doi.org/10.3390/su12020524>
- Shin, Ji-Hye., Dronjic, V., & Park, Boon-Joo. (2019). The interplay between working memory and background knowledge in L2 reading comprehension. *TESOL Quarterly*, 53(2), 320-347. <https://doi.org/10.1002/tesq.482>
- Sun, W. (2023). The impact of automatic speech recognition technology on second language pronunciation and speaking skills of EFL learners: A mixed methods investigation. *Frontiers in Psychology*, 14. <https://doi.org/10.3389/fpsyg.2023.1210187>
- Sweller, J. (1988). Cognitive load during problem solving. *Cognitive Science*, 12, 257-285.
- Sweller, J., & Chandler, P. (1994). Why some material is difficult to learn. *Cognition and Instruction*, 12(3), 185-233. https://doi.org/10.1207/s1532690xc1203_1
- Uchihara, T., Webb, S., Saito, K., & Trofimovich, P. (2023). Frequency of exposure influences accentness and comprehensibility in learners' pronunciation of second language words. *Language Learning*, 73(1), 84-125. <https://doi.org/10.1111/lang.12517>
- Unsworth, L. (2014). Multimodal reading comprehension: Curriculum expectations and large-scale literacy testing practices.

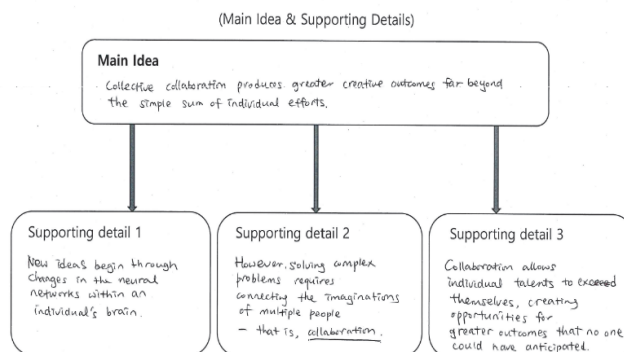
- Pedagogies: An International Journal*, 9(1), 26-44. <https://doi.org/10.1080/1554480x.2014.878968>
- Wei, L. (2023). Artificial intelligence in language instruction: Impact on English learning achievement, L2 motivation, and self-regulated learning. *Frontiers in Psychology*, 14, 1261955. <https://doi.org/10.3389/fpsyg.2023.1261955>
- Wood, D., Bruner, J. S., & Ross, G. (1976). The role of tutoring in problem solving. *Journal of Child Psychology and Psychiatry*, 17(2), 89-100.
- Wu, J., Wang, Y., Chen, F., Yin, X., & He, Y. (2025). Impact of AI-powered adaptive learning platforms on English reading proficiency: Evidence from structural equation modeling. *IEEE Access*, 88230-88242. <https://doi.org/10.1109/access.2025.3571055>
- Wu, X. (2024). AI in L2 learning: A meta-analysis of contextual, instructional, and social-emotional moderators. *System*, 126, 103498. <https://doi.org/10.1016/j.system.2024.103498>.
- Xiao, Y. (2025). The impact of AI-driven speech recognition on EFL listening comprehension, flow experience, and anxiety: A randomized controlled trial. *Humanities and Social Sciences Communications*, 12(1), 1-14. <https://doi.org/10.1057/s41599-025-04672-8>
- Yang, L., & Zhao, S. (2024). AI-induced emotions in L2 education: Exploring EFL students' perceived emotions and regulation strategies. *Computers in Human Behavior*, 159, 108337. <https://doi.org/10.1016/j.chb.2024.108337>

Appendix A

Sample lesson plan for the experimental group

Stage	Time (min)	Students' activity
Pre-reading	5	<ul style="list-style-type: none"> • watching a short video related to creativity or teamwork (e.g., the activity of neurons in the brain or a scene of successful team collaboration) • looking at an image of the brain along with a photo of people collaborating, and discussing how brain activity might be connected to teamwork
	20	<ul style="list-style-type: none"> • listening to the passage individually using IR's TTS feature -selecting the speed and voice type (male/female) <div style="border: 1px solid black; padding: 10px; margin: 10px 0;"> <p>Sample passage</p> <p>The source of every new idea is the same. There is a network of neurons in the brain, and then the network shifts. All of a sudden, electricity flows in an unfamiliar pattern, a shiver of current across a circuit board of cells. But sometimes a single network isn't enough. Sometimes a creative problem is so difficult that it requires people to connect their imaginations together: the answer arrives only if we collaborate. That's why a group is not just a collection of individual talents. Instead, it is a chance for those talents to exceed themselves, to produce something greater than anyone thought possible. When the right people come together and when they collaborate in the right way, what happens can often feel like magic. But it's not magic. There is a reason why some groups are more than the sum of their parts.</p> </div>
While-reading		<ul style="list-style-type: none"> • using IR features -choosing font type/size, line space, and background color -highlighting parts of speech: emphasizing key verbs in the text (e.g., shifts, flows, collaborate, exceed) or important nouns (e.g., network, electricity, imagination, talents) to understand sentence structure -splitting words into syllables: breaking down difficult vocabulary (e.g., un-fam-i-liar, col-lab-o-rate, in-di-vid-u-al) and practicing their pronunciation. -translating complex phrases in real-time (e.g., 'a shiver of current across a circuit board of cells') -using a picture dictionary to immediately check visual materials for words (e.g., neurons, brain, electricity, magic) to clarify meaning in context
	20	<ul style="list-style-type: none"> • completing a T/F quiz and checking answers e.g., 1. A single network of neurons is always enough for a creative problem. (T/F) 2. Collaboration allows a group to produce something greater than was thought possible. (T/F) • discussing the concept of 'synergy' and the key expression (e.g., 'more than the sum of their parts') • listening to the teacher's further explanation of what students discussed
Post-reading	5	<ul style="list-style-type: none"> • summarizing the text using a graphic organizer

Graphic Organizer for Summarizing & Visualizing



Appendix B

Sample test items from the pre-test (the March 2019 test)

34. 다음 빈칸에 들어갈 말로 가장 적절한 것을 고르시오. [3 점]

It is difficult to know how to determine whether one culture is better than another. What is the cultural rank order of rock, jazz, and classical music? When it comes to public opinion polls about whether cultural changes are for the better or the worse, looking forward would lead to one answer and looking backward would lead to a very different answer. Our children would be horrified if they were told they had to go back to the culture of their grandparents. Our parents would be horrified if they were told they had to participate in the culture of their grandchildren. Humans tend to _____. After a certain age, anxieties arise when sudden cultural changes are coming. Our culture is part of who we are and where we stand, and we don't like to think that who we are and where we stand are short-lived. [3 점]

- ① seek cooperation between generations
- ② be forgetful of what they experienced
- ③ adjust quickly to the new environment
- ④ make efforts to remember what their ancestors did
- ⑤ like what they have grown up in and gotten used to

35. 다음 글에서 전체 흐름과 관계 없는 문장은?

Public speaking is audience centered because speakers “listen” to their audiences during speeches. They monitor audience feedback, the verbal and nonverbal signals an audience gives a speaker. ① Audience feedback often indicates whether listeners understand, have interest in, and are ready to accept the speaker's ideas. ② This feedback assists the speaker in many ways. ③ It helps the speaker know when to slow down, explain something more carefully, or even tell the audience that she or he will return to an issue in a question-and-answer session at the close of the speech. ④ It is important for the speaker to memorize his or her script to reduce on-stage anxiety. ⑤ Audience feedback assists the speaker in creating a respectful connection with the audience.

* verbal: 언어적인

Appendix C

Sample test items from the post-test (the March 2020 test)

34. 다음 빈칸에 들어갈 말로 가장 적절한 것을 고르시오.

Say you normally go to a park to walk or work out. Maybe today you should choose a different park. Why? Well, who knows? Maybe it's because you need the connection to the different energy in the other park. Maybe you'll run into people there that you've never met before. You could make a new best friend simply by visiting a different park. You never know what great things will happen to you until you step outside the zone where you feel comfortable. If you're staying in your comfort zone and you're not pushing yourself past that same old energy, then you're not going to move forward on your path. By forcing yourself to do something different, you're awakening yourself on a spiritual level and you're forcing yourself to do something that will benefit you in the long run. As they say, _____. [3 점]

- ① variety is the spice of life ② fantasy is the mirror of reality ③ failure teaches more than success
- ④ laziness is the mother of invention ⑤ conflict strengthens the relationship

39. 다음 글에서 전체 흐름과 관계 없는 문장은?

Paying attention to some people and not others doesn't mean you're being dismissive or arrogant. ① It just reflects a hard fact: there are limits on the number of people we can possibly pay attention to or develop a relationship with. ② Some scientists even believe that the number of people with whom we can continue stable social relationships might be limited naturally by our brains. ③ The more people you know of different backgrounds, the more colorful your life becomes. ④ Professor Robin Dunbar has explained that our minds are only really capable of forming meaningful relationships with a maximum of about a hundred and fifty people. ⑤ Whether that's true or not, it's safe to assume that we can't be real friends with everyone.

* dismissive: 무시하는 ** arrogant: 거만한

Appendix D

Reading difficulty and school level by the Flesch Reading Ease Score

Flesch Reading Ease Score	Level of difficulty	School level (US)
90 to 100	very easy to read	5th grade
80 to 90	easy to read	6th grade
70 to 80	fairly easy to read	7th grade
60 to 70	plain English	8th & 9th grade
50 to 60	fairly difficult to read	10th to 12th grade
30 to 50	difficult to read	college
10 to 30	very difficult to read	college graduate
0 to 10	extremely difficult to read	professional

Appendix E

Information on the passages from chapters 12 and 15 in EBS Olympus English Reading Basic I

Chapter	Passage	Domain	Word count	Sentence count	Avg. words per sentence	FRES	FKGL	Reading grade level
12	1	History	111	5	22.2	23.6	15.5	college graduate
	2	Philosophy	121	7	17.3	70.8	7.7	7th
	3	Science	114	6	19	43.7	11.9	college
15	1	Psychology	168	8	21	50.2	11.5	10th-12th
	2	Philosophy	136	8	17	45.8	11.1	college
	3	Economics	118	5	23.6	39.1	13.7	college

Note. FRES = Flesch Reading Ease Score, FKGL = Flesch Kincaid Grade Level

Appendix F

Pre- and Post-survey items

* Please read the following questions and mark the number that applies to you.

(4: Strongly agree, 3: Agree, 2: Disagree, 1: Strongly disagree)

Survey item		4	3	2	1
1	I think studying English is important.	4	3	2	1
2	I think my school English classes help improve my English reading skills.	4	3	2	1
3	I concentrate on reading English during my school's English reading time.	4	3	2	1
4	I try to actively participate in my school's English classes.	4	3	2	1
5	I find and read other books in English on my own, in addition to textbooks.	4	3	2	1
6	I focus on bold or colored words when I read.	4	3	2	1

	Survey item	4	3	2	1
7	I read carefully and slowly to have an accurate understanding of the text.	4	3	2	1
8	I adjust my reading speed according to the text.	4	3	2	1
9	I use a dictionary when I read.	4	3	2	1
10	I try to guess the meaning in Korean when I read a text.	4	3	2	1
11	I find studying English fun.	4	3	2	1
12	I enjoy my English class time.	4	3	2	1
13	I enjoy reading writings or books written in English.	4	3	2	1
14	I think I'm good at English.	4	3	2	1
15	I have a certain level of confidence in my English reading skills.	4	3	2	1
16	I can read and understand most English texts on my own.	4	3	2	1
17	I want to be good at reading and understanding English.	4	3	2	1
18	I plan time in my daily schedule to study English reading.	4	3	2	1
19	I think my English reading skills will gradually improve in the future.	4	3	2	1
20	I don't feel nervous or anxious when it's time for English class.	4	3	2	1
21	I don't feel nervous or confused when I read books or texts written in English.	4	3	2	1
22	When I feel nervous or confused while reading books or texts written in English, I can control those feelings to some extent.	4	3	2	1