Reading in a Foreign Language April 2019, Volume 31, No. 1
ISSN 1539-0578 pp. 1–18

Reading performance of Japanese high school learners following a one-year extensive reading program

Natsuki Aka
Okinawa International University
Japan

Abstract

The present study investigates the effects of extensive reading on developing the language knowledge and reading abilities of high school learners of English as a foreign language (EFL) over the course of one school year. Participants were divided into control (n = 205) and experimental groups (n = 200), with the former receiving regular grammar instruction and the latter receiving extensive reading instruction. Pre- and posttests consisting of a grammar and vocabulary section and a reading section were conducted to test the effects of the extensive reading program. The study further divided the participants into three proficiency groups. The results showed that the scores of the middle- and lower-proficiency groups increased dramatically compared to the higher-proficiency group. These findings suggest that the participants activated the knowledge they had learned in junior high school through the one-year extensive reading course, resulting in the improvement of their language knowledge and reading abilities.

Keywords: extensive reading, reading fluency, high school learners, longitudinal study, second language reading, English proficiency

An important consideration in contemporary English language education is the amount of English language input (Aka, 2015; Kanatani, 2008; Renandya, 2007; Takase, 2010). Although the Japanese Ministry of Education, Culture, Sports, Science and Technology (MEXT) (2013) emphasizes the importance of enhancing students' communication skills, English language exposure within traditional government-authorized textbooks seems inadequate for improving students' language proficiency. In fact, the number of English words students encounter in textbooks is limited; for example, junior high school textbooks consist of only approximately 5,000 words and high school textbooks only 20,000, meaning that students learning English in the Japanese public education are exposed to only 25,000 words over six years (Kanatani, 2008; Sakai, 2008; Ushiro, 2013). According to Kanatani (2008) and Sakai (2008), if that number of words were printed as a paperback novel, only 70 to 80 pages would be filled. When all the words used in the government-authorized textbooks for seventh and eighth graders were recorded, it was found that they would fill only two sheets of A3 paper (Kanatani, 2008). These government textbooks are usually the only English language matter that students are exposed to in school, although they may be exposed to other forms of English outside of school.

One possible explanation concerning why government-authorized textbooks contain insufficient English input is that fluent reading is not the main goal of English instruction. Rather, language-focused learning through intensive reading and understanding of new grammar and vocabulary items through reading passages are the primary goals of English education in Japan (Takase, 2010). Therefore, most junior and high school learners in Japan do not receive sufficient practice to ensure fluent reading. However, Nation (2013) has suggested that reading fluency development should make up approximately one quarter of reading class time through activities such as repeated reading, speed reading, and extensive reading.

Extensive reading is one approach promoted to solving the language exposure problem in Japan, by increasing the amount of second language (L2) exposure and thereby enhancing learners' English proficiency (Arakawa, 2012; Kanatani, Nagata, Kimura, & Minai, 1991; Nishizawa, Yoshioka, & Ito, 2010). Research has been conducted on the effect of extensive reading from different perspectives, including learners' motivation (Day & Bamford, 1998; Takase, 2010), reading comprehension and reading speed (Beglar & Hunt, 2014; Bell, 2001), sight vocabulary (Brown, Waring, & Donkaewbua, 2008; Pigada & Schmidt, 2006; Waring & Takaki, 2003; Webb & Chang, 2015), and grammatical knowledge (Song & Sardegna, 2014). However, researchers have pointed out that it is difficult to observe the benefits of extensive reading in a relatively short period of time (Grabe & Stoller, 2002; Kanatani et al., 1991; Yamashita, 2008; Yoshii & Lavin, 2017), although several studies on extensive reading have been implemented using more rigorous experimental approaches, including control groups (Grabe, 2007; Nakanishi, 2015). Additionally, most participants have been university students, whereas few studies have been conducted with high school and junior high school learners (Jeon & Day, 2016; Nakanishi, 2015).

Furthermore, the existing studies are not without limitations. One limitation involves the difficulty of creating a control group with many participants in an authentic educational situation due to ethical issues (i.e., one group receives the educational input whereas others do not). Therefore, the main goal of the present study is to investigate the effects of extensive reading on the increase of high school learners' language development and reading comprehension through the use of a control group. Recent curriculum changes in the school selected for this study also enabled the implementation of a one-year longitudinal study with a large sample of high school participants.

Literature Review

In extensive reading instruction, learners receive a large amount of input through reading within their linguistic abilities, but there is no consensus on how much reading should be regarded as extensive reading. However, Nation and Wang (1999) have recommended that learners practice reading at the rate of one book per week. Previous research has shown that learners who are continuously exposed to large amounts of English show improved reading skills, reading speed, comprehension, and motivation (Beglar & Hunt, 2014; McLean & Rouault, 2017), as well as improved sight vocabulary (Pigada & Schmitt, 2006; Waring & Takaki, 2003), general English abilities (Arakawa, 2012; Kanatani et al., 1991; Nishizawa, Yoshioka, & Ito, 2010), and good reading habits (Day & Bamford, 1998; Takase, 2010). For example, Beglar and Hunt (2014)

found that a group of participants who read easy materials improved their reading speed and comprehension more than a group of participants who read difficult materials, even though the two groups of participants received almost the same amount of English input. This finding indicates that reading extensively—regardless of whether or not the materials are easy—improves learners' reading speed and comprehension, but that exposure to reading materials within their reading capacity tends to be more effective than reading difficult materials in terms of improving reading skills. Nuttall (2005) has also highlighted that readers read faster and develop interest in what they read if the materials are slightly below their proficiency level. Examining vocabulary acquisition through extensive reading, Pigada and Schmitt (2006) showed that, although learning vocabulary incidentally is more difficult than learning vocabulary intentionally, the number of partially known words increases in incremental steps through encountering the same words several times. This indicates that incidental learning takes time; however, learners can progress by increasing their knowledge of sight vocabulary through extensive reading.

As indicated, there are limitations in the literature on extensive reading. Nakanishi (2015) performed a meta-analysis using the key words *extensive reading*, *pleasure reading*, and *graded readers* to investigate how many and what kind of studies had been implemented in the areas of extensive reading research, and then quantified and summarized the data. Nakanishi found that approximately 40% of the studies compared pre- and posttest results without control groups to conclude that extensive reading is effective in increasing participants' English proficiency levels. However, without a control group, it is difficult to determine whether these results were exclusively due to the effects of extensive reading instruction, even in cases where there were pre- and posttest statistical differences. Therefore, studies should include control groups. Second, even though researchers recognize that it takes a relatively long period of time to determine the effects of extensive reading, especially on the improvement of learners' English proficiency, 47% of the studies were short term; that is, less than three months, and 63% lasted less than six months, according to Nakanishi (2015). Only 15 studies had been conducted for a period between six months and one year. Because it takes time for students to experience the enjoyment of reading or to realize how well they can read, a period of less than three months seems unlikely to be sufficient to improve English proficiency or develop a love for reading. Therefore, more longitudinal research is required to better understand the effects of extensive reading on general English learning abilities. Third, while 56% of the studies employed university-level students, only 22% of the studies examined the effects of extensive reading on high school-level participants. However, if learners experience the joys of reading through extensive reading at an earlier stage of language learning, they might be motivated to study English, further contributing to improving their English proficiency. Therefore, more studies with junior high school or high school participants are necessary. In a similar manner, Jeon and Day (2016) also performed a meta-analysis to obtain a better understanding of the effects of extensive reading. They selected sample studies based solely on the definitions of extensive reading as suggested by Day and Bamford (1998), focusing in particular, on the self-selection of books. This is a major difference between the two studies. The findings showed that a greater number of studies dealt with adults while the number of studies addressing adolescents was very small, and thus, the result was the same as that of the study by Nakanishi (2015). Moreover, no statistical difference was found between the duration of the programs, meaning that extensive reading can be effective regardless of the duration. According to Jeon and Day (2016), however, since the number of long-term

studies was smaller than the number of short-term studies and the long-term studies were conducted using smaller sample sizes, we need more research on long-term studies involving a larger number of participants, especially adolescents, in order to investigate the effects of a long-term extensive reading program.

A study by Kanatani et al. (1991) is one of the few studies involving high school learners that investigated English proficiency development using a control group (n = 34) and an experimental group (n = 34), during a six-month after-school extensive reading program. Japanese EFL learners in their first year of high school participated in the study as an after-school program. There were no statistical differences between the control and experimental groups in terms of English proficiency test scores prior to the program. However, the posttest results showed statistical differences between the two groups after six months, especially on the reading section component. Kanatani et al. (1991) concluded that it is difficult to see the effects of extensive reading initially, but these gradually appear after several months have been devoted to extensive reading. However, since the extensive reading occurred in an after-school program, the participants might have had higher motivation levels than non-participants. For these reasons, it is not clear that the same conclusions from their study would apply to high school learners in regular courses, which reinforces the value of studying extensive reading programs among high school-level participants in general classes.

Nishizawa et al. (2010) conducted a study over a five-year period with engineering students, aged 17 to 22, to investigate the relationship between the number of running words the participants read and their score on the Test of English for International Communication (TOEIC). Thirty English as foreign language (EFL) engineering students participated in the study. The average TOEIC score for the participants was 518, which was above the mean of 497 for university students in Japan (Nishizawa et al, 2010). Compared to the mean of university-level participants, the mean of the engineering students engaging in extensive reading was also higher. Nishizawa et al. (2010) further divided the participants into three groups based on the number of words they encountered over five years: groups A (n = 9), B (n = 13), and C (n = 8) read approximately 310,000, 660,000, and 1,810,000 words, respectively, over four years, earning mean scores of 435, 498, and 604 points. The study showed that the more the students read, the higher their TOEIC scores were, suggesting a strong correlation between the amount of reading undertaken and language proficiency. However, there were no data on the TOEIC test scores of the participants before the study; therefore, the extent to which the participants improved their English proficiency through extensive reading is unclear. The study also divided the participants into three groups based on how much they read. However, the groups were small, which makes generalizing the results more difficult. Moreover, as there was no control group, it is uncertain to what extent the results were due to extensive reading. For these reasons, it is difficult to interpret the correlation between the number of words the students read and their TOEIC scores as an effect of extensive reading.

The present study investigates whether extensive reading is beneficial in enhancing learners' language abilities and reading comprehension while paying careful attention to the experimental design. First, both the control and experimental groups comprised a large number of participants. This large participant sample allows us to generalize the results to high school-level students. Furthermore, a larger sample made it possible to divide the participants into three different

proficiency groups based on pretest scores to identify which group showed the largest scoring gain between the pre- and posttests. This step was necessary to determine whether test scores at different proficiency levels increased. Since it was expected that an extended period would be required to see the effects of extensive reading (Jeon & Day, 2016; Nakanishi, 2015), this study was conducted over a period of one school year. Considering that few studies on extensive reading have been conducted with high school students as participants (Jeon & Day, 2016; Nakanishi, 2015), a longitudinal study with many high school participants provides helpful data. Moreover, the features of the experimental design—namely, the number of participants, inclusion of a control group, and group divisions—distinguishes the present study from prior studies on extensive reading (Arakawa, 2012; Kanatani et al., 1991; Nishizawa et al., 2010). Therefore, the results of the present study are likely to contribute to the literature on extensive reading and second language acquisition (SLA).

The main purpose of this study is to examine the effects of extensive reading over the course of one year on student language knowledge and reading ability compared to a control group receiving grammatical instruction. To investigate the effects of extensive reading on high school students, the present study posed the following research questions:

Research Question 1 (RQ1): Will a one-year extensive reading instruction program improve learners' language knowledge (grammar and vocabulary) and reading abilities?

Research Question 2 (RQ2): Are there statistically significant differences between pre- and posttest scores within groups of different starting proficiency?

Method

Participants

A total of 405 *Kosen* students aged fifteen to sixteen participated in this study in 2014 and 2015. A *Kosen* is a specialized institution for early engineering education in Japan, with a five-year educational program that is equivalent to a combination of high school and junior college. The participants in this study were all first-year students at a *Kosen*, equivalent to freshmen in high school, and were from five different majors: mechanics, electronics, information technology, chemistry, and architecture. These types of students are seldom very motivated to learn English; rather, they tend to be more focused on mathematics and science. The *Kosen* students do not have to take university entrance exams like other high school students, because the *Kosen* curriculum provides a five-year educational program. Since they do not take these exams, this can mean a further disincentive to learning English. *Kosen* students tend therefore to differ from most other high school students in terms of their English language learning motivation. As the 2014 group did not participate in an extensive reading program, it served as the control group, and as the 2015 learners took the extensive reading program, they served as the experimental group. Creating a control group is one of the challenges faced when conducting a study in an authentic classroom setting due to the ethical issue of providing students with different instructional contents. However, the present study could use a control group without encountering any ethical issues because of curriculum changes in 2015, which enabled an

experimental design with large control and experimental groups.

In 2014, the control group (n = 205) had six English classes a week, three hours of intensive reading classes, two hours of grammar classes, and one hour of listening class, while in 2015, an extensive reading program was introduced in place of grammar classes. The experimental group (n = 200) took the same number of hours of English classes, but the instruction consisted of three hours of intensive reading classes, two hours of extensive reading classes, and one hour of listening class. That is, the two groups differed solely according to whether they had 60 hours of grammar instruction or extensive reading instruction over the course of one year. The participants in the control group learned one specific grammar item each week. They learned grammatical rules and did grammar translation exercises, either from English to Japanese or from Japanese to English, during the class using a textbook called *Harvest English Grammar, Green Course* (Kirihara Editorial Department, 2009). The participants in the experimental group, on the other hand, were encouraged to choose reading materials according to their interests. They were also told to select books that were lower than their English proficiency level, that is, at the *i minus 1* level suggested by Day and Bamford (1998), from the school library. Approximately 3000 graded readers and leveled readers are available in the *Kosen* library. The participants read the following series of books: *Oxford Reading Tree*, *I Can Read*, *Foundation Reading Library*, *Cambridge English Readers*, *Cambridge Storybooks*, *Step into Reading*, *Macmillan Readers*, *Oxford Bookworms Library*, *Penguin Reading*, *Penguin Young Readers*, and *Penguin Active Readers*, among others.

English Proficiency Testing

In the present study, the Basic Assessment of Communicative English test (BACE test) was adopted to measure the participants' linguistic abilities and reading comprehension abilities. This test, developed by the English Language Proficiency Assessment (ELPA) institute, is a general English achievement test. The test consists of the following three sections: a listening section (12 items), a grammar and vocabulary section (16 items), and a reading section (12 items in eight passages). The BACE test uses a measurement method based on item response theory to estimate learners' English proficiency levels. Each item has a weighted score that differs from question to question. The score of each section is converted into 100 points. The present study used the grammar and vocabulary (G&V) and reading sections of the BACE test, excluding the listening section, as the latter was not relevant to whether the students received grammar (control group) or extensive reading instruction (experimental group) throughout the year. In addition, the primary concern in the present study was to see if reading skills would improve through extensive reading, rather than listening skills.

This study adopted the BACE test for the following reason. In extensive reading instruction, learners are instructed to read books that are below their English proficiency level. As it might have been unsuitable to check the effects of extensive reading for learners who had just graduated from junior high school using tests such as the Test of English as a Foreign Language (TOEFL) or TOEIC since the participants were studying in their first year of high school, the BACE test was deemed an appropriate measure of the students' English proficiency because the question items are all related to the junior high school level. Additionally, it takes only 50 minutes to complete, which reduces the mental burden on students. Although the participants

took different versions of the BACE test twice, before and after the treatment (in April and February), the item difficulty in the tests was controlled at the same level according to the ELPA guidelines (for more information, refer to ELPA's website (http://npo-elpa.org/bace/)). In addition to the test quality, the cost of the BACE test is 900 yen, approximately US$10 per person, which makes it more economically feasible than more expensive alternatives. This factor enabled us to obtain a larger participant sample.

Analyses

To investigate the effects of the one-year extensive reading program, a *t* test was used to compare the BACE test scores of the experimental and control groups, as well as the pre- and posttest results of each group. In order to examine whether all participants benefited from the extensive reading program, the experimental group was further divided into three groups (i.e., higher-, middle-, and lower-proficiency groups) based on their scores of each test section (i.e., G&V and reading), and not the number of participants. Grouping was done using the following formula: mean score ± 0.5 × standard deviation. Using this formula not only produced statistically different proficiency groups, but also measured the pre- and posttests within participants more accurately than simply dividing by the number of participants. One sample *t* test between the pre- and posttests of the three groups was conducted to see in which section the participants showed greater improvement (i.e., in G&V or reading), and which group (i.e., the higher-, middle-, or lower-proficiency group) benefited the most from the extensive reading program.

Results

Research Question 1

To answer the first research question, we collected and analyzed pre- and posttest data for the G&V and reading sections to measure increases in the scores of the participants in both groups under instruction. Table 1 shows the results of the pretest scores for the G&V and reading sections, including the means and standard deviations (*SD*) for the control and experimental groups. The mean for the G&V section for the control group was 55.38, while the mean for the experimental group was 56.75. For the reading section, the mean of the control group was 50.71 and the mean of the experimental group was 50.81. The mean difference between the two groups on the G&V section was 1.37 versus 0.10 for the reading section. There were no statistically significant differences between the two groups for the G&V section (t (403) = 1.24, $p < .21$) or the reading section (t (403) = .08, $p < .94$), implying that the two groups had the same proficiency level before the teaching programs began. As shown in Table 1, the effect sizes, indicated by Cohen's d for the differences between the two groups, were very small (G&V, d = .12; reading, d = .01).

Table 1. *Result of the t-test on the pretest scores of the control and experimental groups*

	Control Group (n = 205)		Experimental Group (n = 200)		t	p	*Cohen's d*
	M	*SD*	*M*	*SD*			
G&V	55.38	11.30	56.75	10.85	1.24	.21	.12
Reading	50.71	12.66	50.81	11.71	.08	.94	.01

Note. * $p < .05$. For Cohen's *d*, an effect size of 0.2 to 0.3 is considered a small effect, around 0.5 a medium effect, and 0.8 and above a large effect.

Table 2 shows the results of the posttest scores for the G&V and reading sections, including the mean scores and standard deviations (*SD*s) for the control and experimental groups. The mean for the G&V section of the control group was 54.72, while that of the experimental group was 63.85. For the reading section, the mean of the control group was 52.29 and the mean of the experimental group was 61.78. Thus, the mean differences between the two groups were 9.13 and 9.49 on the G&V and reading sections, respectively. The scores of the experimental group in both sections were significantly higher than the scores of the control group. Furthermore, the learning gains for both sections were larger than in the pretest: the control and experimental groups scored significantly differently on the G&V, t (403) = 6.89, $p < .000$, and reading sections, t (403) = 6.08, $p <.000$.

Table 2. *Result of the t-test on the posttest scores of the control and experimental groups*

	Control Group (n = 205)		Experimental Group (n = 200)		t	p	*Cohen's d*
	M	*SD*	*M*	*SD*			
G&V	54.72	12.63	63.85	14.01	6.89	.000***	.68
Reading	52.29	15.11	61.78	16.28	6.08	.000***	.60

Note. ***p < .001. For Cohen's *d*, an effect size of 0.2 to 0.3 is considered a small effect, around 0.5 a medium effect, and 0.8 and above a large effect.

Figures 1 and 2 show the differences between the pre- and posttests. The effect sizes for the differences between the two groups were medium (G&V, d = .68; reading, d = .60). Thus, for the initial research question, it was found that the one-year extensive reading program helped the students improve their grammar and vocabulary knowledge, as well as their reading proficiency.

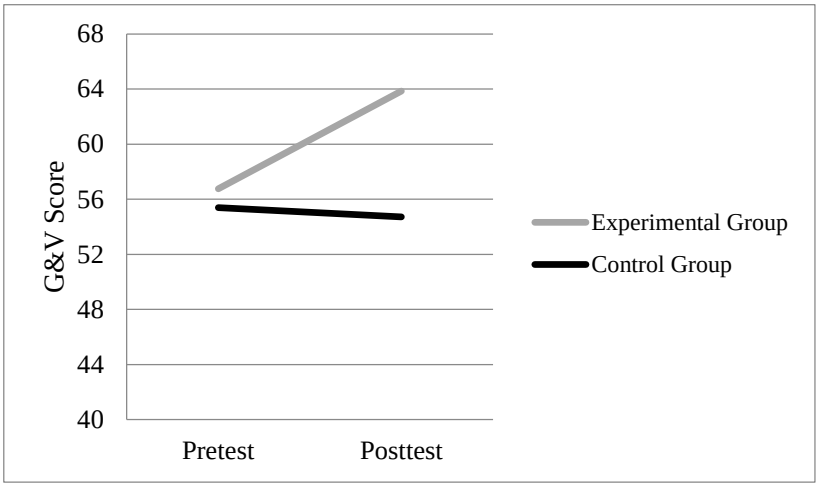

Figure 1. Pre- and posttest results of the grammar and vocabulary (G&V) section.

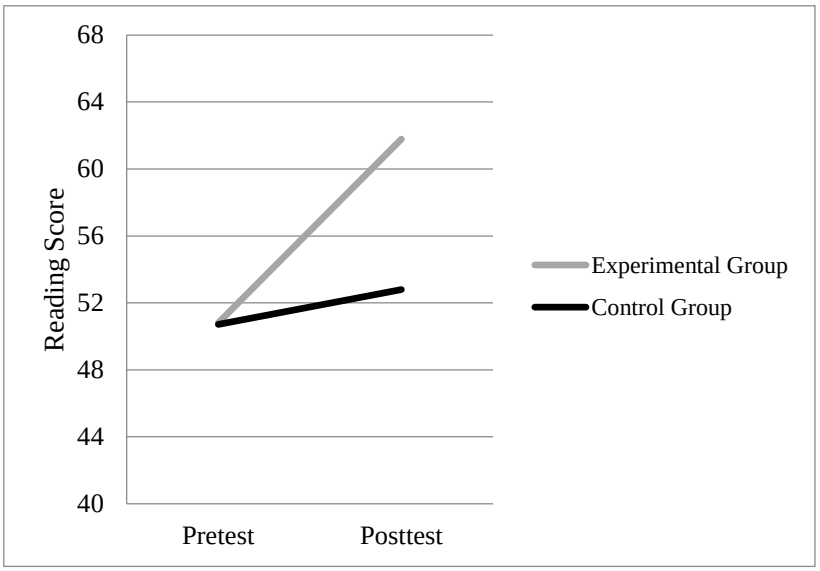

Figure 2. Pre- and posttest results of the reading section.

Research Question 2

Since a positive answer was obtained for the first research question, it was then investigated whether all students benefited from the extensive reading program. The means from the G&V and reading sections in the experimental group were used to investigate the differences in each section of the BACE test between the three proficiency groups (higher, middle, and lower), defined based on their scores on the G&V and reading sections of the pretest (Appendices A and C). Grouping was done using the following formula: mean score ± 0.5 × standard deviation. Subsequently, a one-way ANOVA was conducted to see if the three groups were statistically different. The results of the one-way ANOVA for both the G&V (Appendix B) and reading sections (Appendix D) confirmed that there were significant differences among the groups. A

post-hoc Tukey test also showed statistical differences between the three groups on both the G&V and reading sections, with a significance level of 5%. These results allowed comparison of the pre- and posttest scores of each group in the subsequent step (F (2, 198) = 339.27, $p < .000$; $F(2, 198) = 345.05$, $p < .000$), as depicted in Appendices A, B, C, and D.

First, each section of the pre- and posttests within the group of higher-proficiency learners in the experimental group was compared. The means for the G&V section, as shown in Table 3, were 66.92 and 70.64 for the pre- and posttest, respectively. For the reading section, learners scored 67.46 on the pretest and 67.80 on the posttest. The result of the t test for the higher group revealed a statistically significant difference between the pre- and posttests for the G&V section (t (82) = –1.20, $p = .007$). However, as shown in Table 3, there was no significant difference between the pre- and posttests for the reading section (t (39) = –.14, $p = .89$, *n.s.*). There was, however, a small effect size for the G&V section, but only a very small effect size for the reading section (G&V, $d = .36$; reading, $d = .03$).

Table 3. *T-test results for the pre- and posttest scores of higher-proficiency learners in the experimental group*

	Pretest		Posttest		*t*	*p*	*Cohen's d*
	M	*SD*	*M*	*SD*			
G&V	66.92	6.62	70.64	12.82	-1.20	.007**	.36
Reading	67.46	6.79	67.80	14.49	-.14	.89	.03

Note. *$p < .05$, **$p < .01$, ***$p < .001$. For Cohen's d, an effect size of 0.2 to 0.3 is considered a "small" effect, around 0.5 a "medium" effect, and 0.8 and above a "large" effect.

Next, the middle-proficiency group was compared. The means for the G&V section, as shown in Table 4, were 54.79 and 62.18 for the pre- and posttests, respectively, and 53.24 and 61.17 for the reading section pre- and posttests. There were also statistically significant differences between the pre- and posttest scores in the G&V and reading sections (t (54) = 4.03, $p = .000$; t (76) = –3.84, $p = .000$, respectively). The data in Table 4 show medium effect sizes for both sections (G&V, $d = .74$; reading, $d = .61$). Mean improvement rates for the two sections were much larger for the middle-proficiency group than for the higher-proficiency group, being 7.39 and 7.93 for the G&V and reading sections, respectively.

Table 4. *T-test results for the pre- and posttest scores of middle-proficiency learners in the experimental group*

	Pretest		Posttest		*t*	*p*	*Cohen's d*
	M	*SD*	*M*	*SD*			
G&V	54.79	2.01	62.18	13.90	-4.03	.000***	.74
Reading	53.24	2.50	61.17	18.22	-3.84	.000***	.61

Note. *$p < .05$, **$p < .01$, ***$p < .001$. For Cohen's d, an effect size of 0.2 to 0.3 is considered a small effect, around 0.5 a medium effect, and 0.8 and above a large effect.

Last, the pre- and posttest scores of the lower-proficiency group were examined. The means, as shown in Table 5, were 44.33 and 55.90 for the G&V section of the pre- and posttests, respectively, and 49.88 (pretest) and 61.15 (posttest) for the reading section. Mean improvement rates for the two sections were much larger than for the higher-proficiency group, being 11.57 and 11.27 for the G&V and reading sections, respectively. There were also statistically significant differences between the pre- and posttests in both the G&V and reading sections (t (58) = –7.96, p = .000; t (79) = –6.74, p = .000, respectively). Table 5 shows the larger effect sizes for both sections (G&V, d = 1.37; reading, d = .94) for the lower-proficiency learners. Thus, to answer RQ2, it appears that extensive reading was most effective for the lower-proficiency group. Figures 3 and 4 also demonstrate how much each group of participants improved from pre- to posttests for the G&V and reading sections of the BACE test.

Table 5. *T-test results for the pre- and posttest scores of lower-proficiency learners in the experimental group*

	Pretest		Posttest		*t*	*p*	*Cohen's d*
	M	*SD*	*M*	*SD*			
G&V	44.33	4.90	55.90	10.87	-7.96	.000***	1.37
Reading	49.88	5.19	61.15	16.16	-6.74	.000***	.94

Note. *p < .05, **p < .01, ***p < .001. For Cohen's d, an effect size of 0.2 to 0.3 is considered a small effect, around 0.5 a medium effect, and 0.8 and above a large effect.

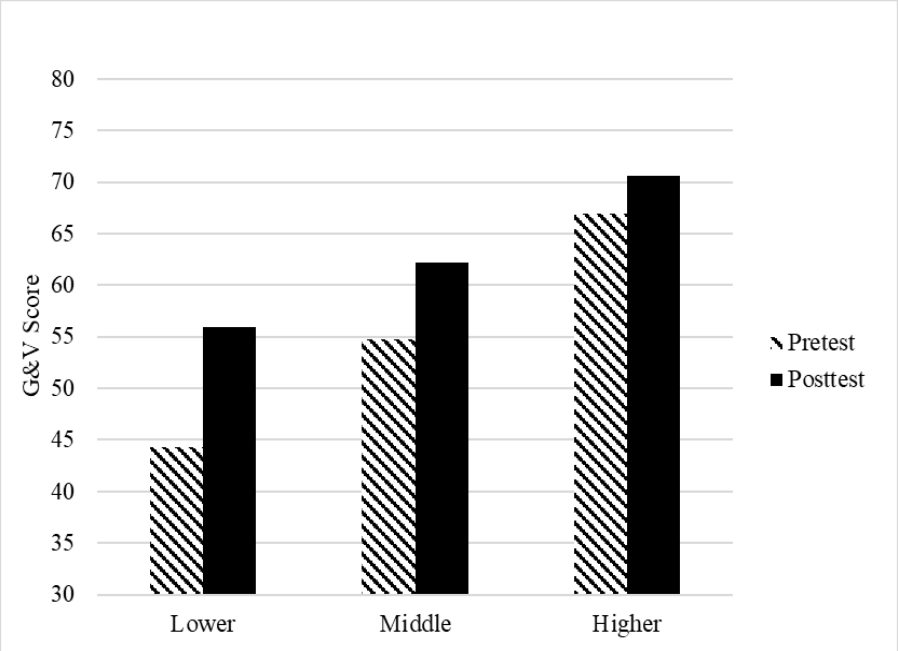

Figure 3. Comparison of the three proficiency groups on the G&V section.

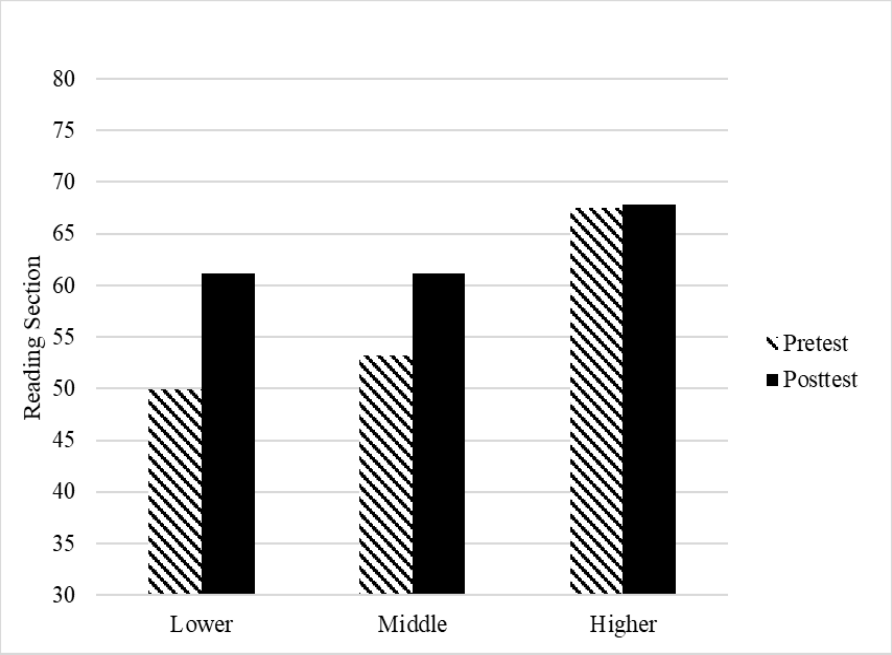

Figure 4. Comparison of the three proficiency groups on the reading section.

Discussion

The results show that the extensive reading group performed better than the control group in terms of language knowledge and reading performance. Regarding research question 1 (RQ 1), the results of the posttest support the view that extensive reading over the course of a one-year period can be effective for EFL high school learners with three years' learning experience in an EFL input-poor context. The results strongly suggest that undertaking extensive reading for one year was more effective than two hours a week of grammar instruction in building learners' language knowledge and reading skills.

One of the reasons why the extensive reading program may have been so effective is the quantity of input. The participants in the experimental group read approximately 115,000 words on average, from 49,000 in the first semester to 66,000 words in the second semester (ranging from 22,000 to 25,000 words, *SD* = 39131.66), which is equivalent to 23 government-authorized textbooks of around 5,000 words each. According to Nishizawa et al. (2010), the positive effects of extensive reading do not transfer to learners' English proficiency when reading only up to 100,000 words. The present study, however, found that the learners could improve their English proficiency level with 100,000 words read. It could be reasonably assumed that students encountered words or grammar items repeatedly in their reading, and in a variety of contexts that they had already learned in junior high school. Moreover, through the program, it is more likely that the learners grasped the meaning and use of words and grammatical concepts more accurately or deeply from reading them in context. Simply stated, the learners' partially known words might have become completely known at the level of sight vocabulary. This consideration can be supported in reference to Pigada and Schmitt (2006), who detailed the acquisition process for words through extensive reading by which students use spelling, meaning, and grammatical characteristics, showing that the participants gradually enhanced their lexical knowledge through

extensive reading. Although it remains unknown how partially known words become automatic and fluent for language users, the scores for the experimental group in the reading section improved more than those of the control group. Therefore, it seems plausible that they increased their sight vocabulary and vocabulary size, as suggested could occur by Pigada and Schmitt (2006). Furthermore, as noted, the number of words read by the participants in the second semester was 1.33 times greater than that in the first semester. Beglar, Hunt, and Kite (2012) have shown that extensive reading facilitates learners' reading speed, especially among those who read easier materials, thereby contributing to reading comprehension as well. Although no exact data on learners' reading speeds were presented, it is likely that word recognition speed improved, affecting learners' reading speeds and test scores as well. Therefore, improvements in reading speed would also be a factor enhancing learners' English proficiency.

The present study shows that a one-year extensive reading program enhanced the English proficiency level of high school EFL learners. Kanatani et al. (1991) found that the effects of extensive reading appeared after six months. As this study involved a longer period, it is not clear whether the participants in the experimental group surpassed the control group in the first six months of the process. However, this study was conducted in a regular classroom with a large sample of participants rather than as part of an after-school program, as studied by Kanatani et al. (1991). Therefore, 60 hours of extensive reading in a year appears to be beneficial for high school EFL learners. In this study, the participants undertook extensive reading over the course of a year, but it should be noted that the period under examination included approximately seven weeks of summer vacation and two weeks of winter break. Yoshii and Lavin (2017) have highlighted that learners' English proficiency levels drop during breaks between semesters. Despite occasional breaks in the programs, the effects of extensive reading could still be observed in the high school learners. Given these considerations, future studies should examine whether six-month extensive reading programs might benefit students in a regular course.

For research question 2 (RQ2), we also examined the learning gain between the pre- and posttests for each proficiency group to see which group of participants improved their English proficiency the most through extensive reading. The findings show that extensive reading was more effective for the middle- and lower-proficiency groups than for the higher-proficiency group. The learners in the middle- and lower-proficiency groups might have had more chances to activate the knowledge they had learned in junior high school than the higher group. More specifically, it appears that the learners' partial knowledge of vocabulary or grammar had gradually been enhanced by reading extensively, contributing to an improved learning outcome in both sections of the test. Although the scores on G&V slightly improved in the higher-proficiency group, the mean score gain of 3.72 was much lower compared to the other groups. Moreover, the mean of the reading section in the higher-proficiency group showed no statistical difference. As all the groups with the exception of the higher-proficiency group improved in both sections, it would appear that learners' English proficiency was measured with varying effectiveness through the BACE test, especially for the higher proficiency group. In other words, because the BACE test was created to measure the English knowledge of junior high school students, it is likely that the test measured the English proficiency levels of the middle- and lower-proficiency groups more accurately than the higher group. If a test more sensitive to differences at a higher level of proficiency had been used for the higher group, it is probable that larger learning gains between the pre- and posttests of that group would have been found.

There are some limitations to this study. First, even though the learning gain of the experimental group was higher than that of the control group on the BACE test, there were only 16 items in the G&V section and 12 items in the reading section. Furthermore, G&V were measured in the same section. These factors might have obscured the effects of extensive reading on the higher-proficiency group. There are a few English external examinations available to those who have just graduated from junior high school. The institution in this study requires all students to take a BACE test to measure their English ability every year before starting English classes in April, which enabled the author to create the experimental design of the present study. Another experimental design could have been adopted using self-created tests or another external test such as the Society for Testing English Proficiency (STEP) test. However, to minimize the mental and financial burdens on students, the author chose to use the BACE test as a measurement of the learners' English abilities. In addition, this experiment could not have been done without the curriculum changes in the institution in which the study took place.

Second, it remains unclear which skills the learners acquired through the one-year extensive reading program, even though the test scores show improvements in the learners' linguistic abilities and reading comprehension skills. One reason for this uncertainty is that individual learners were exposed to different types of reading passages at varying levels since the purpose of the extensive reading class was for the students to enjoy reading. Therefore, they were allowed to choose whatever they wanted to read according to their English level and interests, making it difficult to identify what knowledge all the learners acquired through extensive reading.

Although the results of this study demonstrate the positive effects of extensive reading, it remains necessary to investigate which skills lead to the improvement in learners' English proficiency levels. In future studies, it would be desirable to separate the G&V sections and add more questions so that more precise data on English skills acquisition will be available. Furthermore, a method needs to be developed to ascertain not only the number of words that learners encounter, but also how many times the learners encounter the same grammatical items or words during a program. More accurate information will help researchers and educators determine how well learners acquire words and grammatical items through extensive reading.

Conclusion

The present study demonstrated an improvement in English language abilities through one year of extensive reading instruction, especially among middle- and lower-proficiency learners. The findings of the study also show that reading over 115,000 running words during one academic year contributed to the development of learners' English proficiency. Adolescent EFL learners can significantly improve their reading abilities through a one-year extensive reading program, which implies they can start reading materials in a foreign language by themselves without the teachers' support if the input is relatively easy to grasp. As language-focused learning, involving for example the detailed study and understanding of grammar passages, is the preferred approach in English education in Japan, students have few opportunities to read English materials that are not demanding. However, fluency training such as extensive reading should be included in language teaching to activate what students learn in textbooks. Such training will also help students acquire word decoding skills in context and increase their sight vocabulary, leading in

turn to the development of general English proficiency. However, the number of words that participants encountered was not enough for them to develop into autonomous readers, for which one million words are assumed necessary (Sakai, 2008). Nonetheless, considering that learners became more confident in foreign language reading as they developed their reading skills during the program, extensive reading is likely to be beneficial in the early stages of language learning.

The number of studies involving high school-level learners is limited compared to those involving university-level learners. Taking account of some of the limitations of earlier studies, the present study was conducted with a large sample of learners in regular English classes. Furthermore, the low relative cost of the BACE test enabled us to conduct the study with a greater number of participants. Although more research is necessary, this study suggests that extensive reading instruction is an effective means to improve young EFL learners' reading fluency during the early stages of L2 learning.

Acknowledgments

I am grateful to Dr. Tadamitsu Kamimoto of Kumamoto Gakuen University for his cooperation with and insightful comments on this study. I would like to thank the anonymous reviewers for their feedback on my manuscript. Their advice and suggestions expanded my research knowledge and improved this paper as well. This study was supported by Grant-in-Aid for Young Scientists (B) 16K17462 and Grant-in-Aid for Young Scientists 18K13188.

References

Aka, N. (2015). Eigo gakushuu ni okeru ishiki chousa to tadoku sidou no kanousei ni tuite [Students' attitudes about English education and extensive reading]. *Bulletin of Ariake National College of Technology*, *51*, 11–17.

Arakawa, T. (2012). Tadoku Tachou wo dounyuu sita okinawa kousen no eigo kyouiku –ikki kara nanakisei made- [Extensive reading and listening in English education at Okinawa National College of Technology: From first-year enrollment to the seventh]. *Bulletin of Okinawa National College of Technology*, *6*, 27–37.

Beglar, D., & Hunt, A. (2014). Pleasure reading and reading rate gains. *Reading in a Foreign Language*, *26*, 29–48.

Beglar, D., Hunt, A., & Kite, Y. (2012). The effect of pleasure reading on Japanese university EFL learners' reading rates. *Language Learning*, *62*, 665–703. doi:10.1111/j.1467-9922.2011.00651.x

Bell, T. (2001). Extensive reading: Speed and comprehension. *The Reading Matrix*, *1*(1). Retrieved from http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.466.3793&rep=rep1&type=pdf

Brown, R., Waring, R., & Donkaewbua, R. (2008). Incidental vocabulary acquisition from reading, reading-while-listening, and listening to stories. *Reading in a Foreign Language*, *20*, 136–163.

Day, R. R., & Bamford, J. (1998). *Extensive reading in the second language classroom*. Cambridge, UK: Cambridge University Press.

Grabe, W., & Stoller, F. (2002). *Teaching and researching reading*. London, UK: Pearson Education Longman.

Jeon, E. Y., & Day, R. R. (2016). The effectiveness of extensive reading on reading proficiency in ESL/ EFL settings: A meta-analysis. *Reading in a Foreign Language*, *28*, 246–265.

Kanatani, K. (2008). *Eigo kyouiku netu kanetu sinri wo jousiki de samasu* [English education fever: Removing overheating rumors by indubitable truth]. Tokyo, Japan: Kenkyusha.

Kanatani, K., Nagata, M., Kimura, T., & Minai, Y. (1991). Koukou ni okeru tadoku puroguramu [Extensive reading program in senior high school]. *KATE Bulletin*, *5*, 19–26.

Kirihara Editorial Department. (2009). *Harvest English grammar, green course*. Tokyo, Japan: Kirihara Press.

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

Ministry of Education, Culture, Sports, Science and Technology (MEXT). (2013). *Guroubaruka ni taiou sita eigo kyouiku jissi keikaku* [English education plan based on globalization]. Retrieved from http://www.mext.go.jp/a_menu/kokusai/gaikokugo/__icsFiles/afieldfile/2014/01/31/1343704_01.pdf#search='グローバル化に対応した英語教育実施案'

Nakanishi, T. (2015). A meta-analysis of extensive reading research. *TESOL Quarterly*, *49*, 6–37. doi:10.1002/tesq.157.

Nation, P. (2013). *Learning vocabulary in another language*. Cambridge, UK: Cambridge University Press.

Nation, P., & Wang, K. M. (1999). Graded readers and vocabulary. *Reading in a Foreign Language*, *12*, 355–380.

Nishizawa, H., Yoshioka, T., & Itoh, K. (2010). Kougyoukei gakusei no nigate isiki wo kokuhuku si jiritu gakushuu e michibiku eigo tadoku jugyou [An extensive reading program that changes reluctant engineering students into autonomous learners of English]. *Journal of JSEE*, *58*(3), 12–17. https://doi.org/10.1541/ieejfms.126.556

Nuttall, C. (2005). *Teaching reading skills in a foreign language*. Oxford, UK: Macmillan Education.

Pigada, M., & Schmitt, N. (2006). Vocabulary acquisition from extensive reading: A case study. *Reading in a Foreign Language*, *18*, 1–28.

Renandya, W. A. (2007). The power of extensive reading. *RELC Journal*, *38*, 133–149. https://doi.org/10.1177/0033688207079578

Sakai, K. (2008). *Sayonara eibunpou! Tadoku ga sodateru eigo ryoku* [Good-bye, English grammar! The improvement of English ability by extensive reading]. Tokyo, Japan: Chikuma Gakugei Bunko.

Song, J., & Sardegna, G. V. (2014). EFL learners' incidental acquisition of English prepositions through enhanced extensive reading instruction. *RELC Journal*, *45*, 67–84. doi: 10.1177/0033688214522623.

Takase, A. (2007). Japanese high school students' motivation for extensive L2 reading. *Reading in a Foreign Language*, *19*, 1–18.

Takase, A. (2010). *Eigo tadoku tachou sidou manyuaru* [Manual of instruction in English extensive reading and extensive listening]. Tokyo, Japan: Taishuukan.

Ushiro, Y. (2013). Tokushuu "situ ryou juujitu jidai no riidingu sidou wo dousuru? Eigo riidingu no situ to ryou: seito wa hontou ni eigo wo yondeirunoka" [What can we do for improving

students reading performance both quality and quantity? Do students really read English in their English classes?]. *Eigo Kyouiku: The English Teacher's Magazine*, *September*, 10–12.
Waring, R., & Takaki, M. (2003). At what rate do learners learn and retain new vocabulary from reading a graded reader? *Reading in a Foreign Language*, *15*, 130–163.
Webb, S., & Chang, A. C.-S. (2015). How does prior word knowledge affect vocabulary learning progress in an extensive reading program? *Studies in Second Language Acquisition*, *37*, 651–675. doi:10.1017/S0272263114000606.
Yoshii, M., & Lavin, R. (2017). A new extensive reading and listening class: How effective is it? *Kyushu Academic Society of English Language Education*, *45*, 29–38.

Appendix A

Descriptive Statistics for the G&V Section for the Experimental Group (N = 200)

	M	*SD*	*Min*	*Max*
Higher group (*n* = 84)	66.92	6.62	62	90
Middle group (*n* = 56)	54.79	2.01	53	57
Lower group (*n* = 60)	44.33	4.90	27	49

Appendix B

Results of the One-way ANOVA for the G&V Section

	Sum of Squares	*df*	Mean Square	*F*	*p*
Between groups	18148.82	2	9074.41	339.27	.000***
Within groups	5269.18	198	26.75		
Total	23418.00	200			

Note. **p* < .05, ***p* < .01, ****p* < .001

Appendix C

Descriptive Statistics for the Reading Section for the Experimental Group

	M	*SD*	*Min*	*Max*
Higher group (*n* = 41)	67.46	6.79	62	89
Middle group (*n* = 78)	53.24	2.50	51	56
Lower group (*n* = 81)	40.04	6.83	9	45

Appendix D

Results of the One-way ANOVA for the Reading Section

	Sum of Squares	*df*	Mean Square	*F*	*p*
Between groups	21233.32	2	10616.66	345.05	.000***
Within groups	6061.46	198	30.77		
Total	27294.78	200			

Note. *$p < .05$, **$p < .01$, ***$p < .001$

About the Author

Natsuki Aka is a lecturer at Okinawa International University. Her areas of interest include second language reading, especially reading fluency development. Email: natsuki@okiu.ac.jp