

# *The effect of a teacher's guidance on Japanese University EFL learners' voluntary reading outside class*

Kiyomi Yoshizawa Atsuko Takase Kyoko Otsuki

多読における教師の役割の一つは、学習者が自分の英語力にあった教材を選び、読み進めていくように指導することである。また、教師は日頃学習者がどのような読みをしているのかを観察し、問題のある読み方をしている学習者を指導する。更に、学習者の多読記録から学習者の読みのパターンや興味を読み取るなども教師の役割として挙げられる。本研究は授業外での多読において、学習者の本の選択に関する教師の指導がL2リーディングにどのように影響するかを検証した。多読クラス5つのうち、4クラスを実験群、1クラスを統制群とし、130名の大学生英語学習者の一年間の読書量、読み能力の伸びを観察した。研究結果は、学習者がL2リーディングのどの発達段階にいるのかを教師が的確に判断し、学習者の英語力、興味などを鑑みながら本の選択をアドバイスすることの重要性を示している。

## Key words

多読 (extensive reading)、teacher's roles (教師の役割)、the amount of reading (読書量)

## INTRODUCTION

Reading researchers have recognized that the insufficient exposure to print is one of the major causes of reading difficulty (Stanovich, West, Cunningham, Cipelewski, & Siddiqui, 1996). Extensive Reading (henceforth, ER) is one of the effective teaching methods to expose learners to print. Grabe (2009) reviewed the research studies on extensive reading both in L1 and L2 and wrote that “a greater amount of reading and extensive reading” will have beneficial effects on learners' reading abilities: “the research is persuasive that greater amount of reading and extensive reading, when carried out consistently and appropriately over an extended period of instructional time, will significantly improve students' reading abilities” (p. 322).

Day and Bamford (1998, 2002) listed ten principles for teaching extensive reading successfully. One of the principles is that “students read as much as possible” (1998, p. 7) both in and

out of class. For learners to learn to read, it has the utmost importance that the amount of time spent on reading is secured. Day and Bamford argue that even though most reading teachers would not disagree on this point, students may not be given the “opportunity or incentive” (2002, p. 138) to read extensively.

Day and Bamford (1998) stress the importance of continuing guidance and counseling in an ER program. In an ER class, students are encouraged to select their reading materials on their own and read at their own pace. However, this does not mean that a teacher plays a minor role in an ER classroom. Two of Day and Bamford’s principles refer to ER teachers. The first principle is that teachers introduce the goals and methodology of extensive reading to their students, and further they guide their students throughout the ER program: “Teachers orient students to the goals of the program, explain the methodology, keep track of what each student reads, and guide students in getting the most out of the program” (p. 8). After the introduction of the goals and methodology of the ER program, teachers can advise students what to read at the beginning of the program, and they keep track of their students’ reading amounts and reactions to the books they read. Based on the information the teachers have about their students’ reading behaviors, they can further guide their students to the reading materials suitable for them. The second principle referring to ER teachers is that “the teacher is a role model of a reader for students” (p. 8). In effective ER classes, teachers themselves are readers and they show their students “the attitudes and behaviors of a reader” (Day & Bamford, 2002, pp. 139–140).

Similar characteristics of successful ER classes are repeated in Takase (2010). She mentions that one of the key factors which make an extensive reading class successful is to have Sustained Silent Reading (henceforth, SSR) in ER classes. Further, Takase stresses that advice a teacher gives to his or her students is of great importance.

Takase (2010) describes three major roles a teacher plays in extensive reading classes. The first role is to advise or guide students to read books which are appropriate for their English proficiency. Students do not know what books are good for them, especially at the beginning stage of an ER class. If a student asks the teacher which books to read, he or she should be able to advise considering the students’ proficiency and interests.

The second role a teacher plays in ER classes is to observe how his or her students read in class and guide them to engage themselves in reading as independent readers. Takase (2010) describes different types of problematic readers in an ER class. The first type of problematic readers is those who seem to have difficulty in concentrating in class during SSR. Many of those students have chosen books which are too difficult for their English proficiency level. The

teacher can advise them to read easier books. Another type of problematic readers is those who may translate an L2 text sentence by sentence into L1 in order to understand what they read. Even though L1 is important as a resource for EFL learners (Grabe, 2009; Kern, 1994), L2 learners may depend on L1 translation so much while reading in L2 that L1 translation may have become a major means for obtaining a text model of understanding (Grabe, 2009). On the other hand, some readers seem to be reading too fast to create a text model of understanding, or they may depend on the pictures or drawings in text and infer the content of the book they are reading. Also, some readers pretend to read, i.e., reading across text, but they do not understand what they are reading. When the teacher finds those problematic readers, it is important that the teacher gives proper advice which addresses each type of reading problems his or her students face in a manner that the teacher does not demotivate them.

The third role a teacher can play in an ER class is to check reading records which students submit. Reading records provide information concerning the levels of the books the students read and the amount of reading they engage themselves in. Also, the ER teacher can find a particular student's interest by observing the kinds of books he or she reads. This information is useful when the teacher recommends the student books to read.

There are many empirical studies reporting the effectiveness of ER on reading achievement (e.g., Elley, 1991, 2000; Elley & Mangubhai, 1983; Mason & Krashen, 1997; Robb & Susser, 1989). Also, previous studies have shown that the participants' experience with ER changed their attitudes to reading positively (e.g., Camiciottoli, 2001; Robb & Susser, 1989; Takase, 2007). Further, ER has positive effects on the development of language skills and fluency (e.g., Beglar, Hunt, & Kite, 2012; Elley, 1991; Hafiz & Tudor, 1989; Maruhashi, 2011). However, there is a lack of empirical research about the effect of a teacher's guidance on reading achievement in ER. The present study aims to investigate the effect of teachers' guidance on the growth in L2 reading. Specifically, we investigate the following research question: To what extent does a teacher's guidance on book selection affect L2 readers' growth in reading ability? We focus on the teacher's guidance on book selection for voluntary reading outside class (i.e., out-of-class reading).

## **METHOD**

### **Participants**

A total of 130 EFL students participated in the present study. The students were in five ER classes at a four-year university located in the western part of Japan. The ER classes lasted for

two semesters. Class met once a week, 15 weeks each semester. In this study, the first four classes were treated as an experimental group; the fifth class was treated as a control group. The students in the experimental group were second-year students and their majors were commerce, economics, law, and letters. The students in the control group were first-year students and they majored in economics. The experimental group was taught by one of the authors and the control group was taught by another. The number of the participants in the experimental group was 90 at the onset of the program. However, ten of them did not take the second semester due to several reasons such as their course change, study abroad, or earning the credit by taking a standardized English test. Another six students were deleted because of the absence from the posttest or no submission of the reading data. Therefore, the final number of the participants was 74. Similarly, the number of the participants in the control group was 40 at the onset of the program. However, four students did not take the posttest.

### Reading Materials

In regular sessions, the experimental group used the textbook *Reading Advantage 2* and *3* (Thomson Heinle) in the first half of the class. The article in each chapter from the textbook was utilized for speed reading practice. After that, working in pairs, students read the article aloud to his/her pair for one minute from the beginning. Taking turns, they repeated the same parts three times. The latter half of the class, which is approximately 45 minutes, students were engaged in Sustained Silent Reading. The participants were required to check out books from the library and bring them for in-class reading. The reading materials were from graded readers: Foundations Reading Library (levels 1-7), Cambridge English Readers (levels 1-4), Macmillan Readers (levels 1-5), Penguin Readers (levels 0-5), and Oxford Bookworms (levels 0-4) and from leveled readers: Oxford Reading Trees (levels 2-9), Longman Literacy Land Story Street (levels 1-12), Scholastic Readers (level 3), Capstone Series, and Rookie Reader Series.

On the other hand, the control group read SRA Reading Laboratory materials (SRA/McGraw-Hill). These are reading kits designed to provide individualized instruction for learning to read in English. The reading selections are divided into levels which are coded in different colors. Those colors correspond to “approximate reading levels” of elementary school students in the US (Parker, 2004). *SRA Reading Laboratory 2a* kit contains six colors: they correspond to reading levels of the second to seventh graders. Similarly, *SRA Reading Laboratory 2b* kit contains six colors: they correspond to reading levels of grade 2 or 3 to grade 8. In the spring semester, the students read *SRA Reading Laboratory 2a*; in the fall semester, they read *SRA*

*Reading Laboratory 2b.*

*SRA reading Laboratory* mainly consists of two types of cards: Rate-Builder and Power-Builder. Rate-Builder cards are designed to improve learners' reading fluency. The time for Rate-Builder practice is three minutes; however, the length and difficulty of passages increase as the student progresses from one level to another. Within three minutes, the student has to read a passage and answer five to eight comprehension questions. After each three-minute exercise, the student checks his or her answers with a key so that the student can have immediate feedback on his or her reading efficiency. This ensures "that the gain in reading speed does not result in a loss of comprehension" (Parker, 1985). Each class period starts with Rate-Builder exercise and each student read three Rate-Builder cards in one class period. The instructor checked the students' progress each week. When a student's record showed that his or her comprehension scores were 80 to 100 percent constantly, the instructor asked him or her to move to a next higher level.

Power-Builder cards are designed to improve learners' reading comprehension. Each Power-Builder card consists of three parts: a reading selection, five to ten reading comprehension questions, and thirty-three vocabulary questions. Concerning the vocabulary questions, the students were asked to answer the first ten vocabulary questions. In each class, the students were asked to select their Power-Builder cards, which matched their reading levels. After they finished reading and comprehension and vocabulary exercises of one card, they checked their answers with a key and start to work on another card. When students continue to score more than 80 percent of comprehension questions consistently, they were asked to consult with the instructor about moving to a next higher reading level.

The control group started with three Rate-Builder cards for timed reading each week. Then, they read the Power-Builder cards and check their answers with answer keys. The class reading, i.e., reading Power-Builder cards, time was about 45 minutes. During the Power-Builder reading, the instructor observed the students. Sometimes, students asked questions on the directions of comprehension questions and they asked the instructor to help with the directions. Other times, students had difficulty with the answers for some questions and they came to the instructor for explanation. The students' Rate-Builder and Power-Builder records were checked after class every week to see whether students were reading the cards appropriate for their reading ability. In some cases, the instructor wrote her comments on the students' record books concerning their reading performance.

Further, the students in the control group were required to read graded and leveled readers outside class: the outside reading was carried out throughout the year, including the summer

recess. The students kept the reading log and recorded the information of a book they finished reading: the title, the publisher, the estimated number of words and the reading time in minute. The reading log also had a space for students to write a short comment on each book after they finished reading it. The students were asked to submit their reading logs in the middle and end of each semester. Also, students were asked to submit two book reports per semester. Concerning the first book report, the students were asked to make a group of four or five and to introduce their book to the other group members.

### **Individual counseling**

The second difference between the experimental and control groups was individual counseling about selection of leveled and graded readers for out-of-class reading. The experimental group read leveled and graded readers in and outside class. As mentioned in the previous section, the instructor had short counseling sessions with individual students in class. When she did not have enough class time for giving guidance directly, she wrote her comments on the students' reading logs. The topics of individual counseling were mainly about the comprehension of the book a student finished reading, his or her reading rate, and the level of a book for each student. The instructor gave advice based on the pretest result and a student's reading log.

The instructor of the control group also had individual counseling with students in class. However, the counseling was mainly about the SRA materials. When students seemed to have difficulty with reading Power-Builder cards assigned to them, the instructor wrote in their reading record books and asked to have individual counseling in the following class. Also, individual counseling was held when a student was ready to move to a next higher level card. However, the instructor did not examine the students' reading logs for outside reading on a weekly basis to such an extent that she could guide students for selecting books.

### **Instrument**

The Edinburgh Project on Extensive Reading Tests (henceforth, EPER) were developed by the team of Edinburgh Project on Extensive Reading (EPER, 1992). The EPER tests are semi-fixed ratio open-ended cloze tests, consisting of 10 to 12 short passages taken from different levels of obsolete graded readers and arranged in order of increasing difficulty. Examinees are required to use information from clause level to text level to fill in the gaps (Yoshizawa, Takase, & Otsuki, 2012).

The EPER Placement/Progress Test Form A was used in the present study. The EPER Form A consists of 12 short passages. The readability of the passages is 92.7 according to Flesch Reading Ease and 2.3 according to Flesch-Kincaid Grade Level. There were 141 deletions and the number of words between deletions ranges from 4 to 12.

All the students' responses to the EPER Test Form A were first marked by assigning numbers ranging from 0 to 4: 0 for an incorrect answer and 4 for a correct answer. The alternative scoring method was used in the study: a full mark was given to both the original answers and alternative answers. Two points were subtracted when a correct tense was not used. One point was subtracted when a number was wrong. Also, one point was subtracted for a misspelled word. First, the same sets of answer sheets were marked by all the authors independently. The differences in marking were discussed and assignments of points were decided by consensus. While marking, correct, partially correct, and incorrect answers were listed. Later, different sets of answer sheets were marked independently, using the answer list. The list was revised constantly as marking progressed.

## **Procedure**

The participants in the experimental and control groups were in two-semester reading program: the class met once a week for 15 weeks each semester. The EPER test was administered at the beginning and end of the program as the pre- and the posttests to measure the participants' overall reading proficiency.

In the first week of the first semester, the participants in the experimental group took the EPER cloze test. Also, the first session was used for orientating the students in the experimental students to extensive reading. The last two sessions were used for administering the post-EPER test, semester final test, questionnaire and reviewing the class, leaving 12 sessions for class work including extensive reading.

The instructor of the experimental group determined the initial reading levels of her students based on the results of the pretest. Initially, the instructor assigned two or three levels lower than what the results of the EPER suggested. This was because most of Japanese students were not accustomed to reading in English without translation and reading the books lower than their reading ability could encourage students to understand the text without using translation. The other reason was that students could read at an appropriate speed. The instructor encouraged the students to read 50 to 100 easy books first and they started to read the books appropriate for their reading levels.

The students in the experimental group were required to keep their reading log of what they

read. In the reading log the students' initial reading levels were written based on the pretest results as a suggestion or advice to the students together with the student's reading history. The instructor read their reading logs every week or every other week (half of the class each week) and gave them advice by writing comments on reading logs or talking to them in class. As there were 43 to 50 students in each class, there was not enough time to talk to all the students in one class period. Walking around the classroom and examining the reading log of each student, the instructor talked to those who appeared to have trouble reading smoothly or who had chosen inappropriate level of books.

Example advice and comments made by the instructor of the experimental group are as follows: Raise the level of books to YL1.2-1.5<sup>1)</sup>; Read five more books from the same series or the same level. If you feel comfortable reading them, then raise one level; Your reading speed is very slow. Aren't you translating?; You should slow down. Nobody can read at the speed of 300 wpm at this level; Did you understand the story? If the book is too difficult, choose an easier one; Try some books in the lower level. To those who seem to have no problem in reading, words of encouragement were written: Good speed, keep reading; Congratulations! 100,000 words!; A little more to go to 100 books!!; Oh, you like history, aren't you? If you like to read more advanced history books, let me know. I'll lend you mine; If you like non-fiction, why don't you try this series; Enjoy reading! In order to share the interest in the books the instructor often expressed her own opinions on students' comments: I like this story, too; I agree with your opinion; I don't think so.

Similarly, the control group took the EPER cloze test in the first week of the first semester. Also, the first session was used to orient the students to the class and materials. A brief explanation was given to the students about extensive reading, including the locations where they could check out the books for reading outside class. Further, the students in the control group took a starting level guide check at the beginning of each semester so that they could read the SRA Rate Builder and Power Builder cards appropriate for their reading proficiency. Also, the instructor gave specific instructions for completing a *SRA Student Record Book* at the beginning of the first semester.

Further, the reading logs of the students in the experimental group were examined for obtaining the information about the number of books the students read and the number of words included in those books. Also, the teacher's comments written on the students' reading logs were examined. Similarly, the *SRA Record Books* of the students in the control group were examined to obtain information about the number of words in the Power-BUILDER cards the students read in class. The words in the Rate-BUILDER cards were not included. Also, the infor-

mation about the starting and final reading levels of the students was collected. Further, the reading logs of voluntary reading outside class were examined for calculating the number of words read by the students in the control group<sup>2)</sup>.

## RESULT

### The amount of reading

Figure 1 shows the number of words the students in the experimental groups read. The experimental group read 125.4 ( $SD=42.97$ ) books on average in one year. The average and median number of words read by the students was 197,391 ( $SD=141,374$ ) and 163,416, respectively. The highest number of words read was 943,582; the smallest number of words read was 44,686. Figure 1 indicates that 26.7% of the students in the experimental group read more than 200,000; 47.7% of the students read between 100,000 to 200,000 words; and 25.6% read less than 100,000 words.

Figure 2 shows the distribution of the starting levels of SRA Rate- and Power-Builder cards. The starting level of the SRA materials in the spring semester was Brown, which corresponds approximately to the reading level of the second graders. 37.5% of the students started with

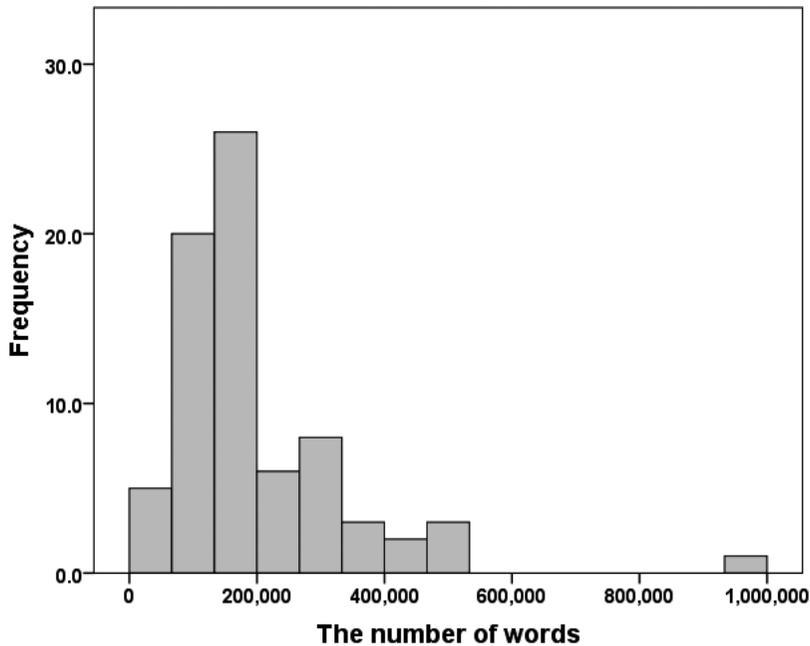


Figure 1. The number of words read by the experimental group

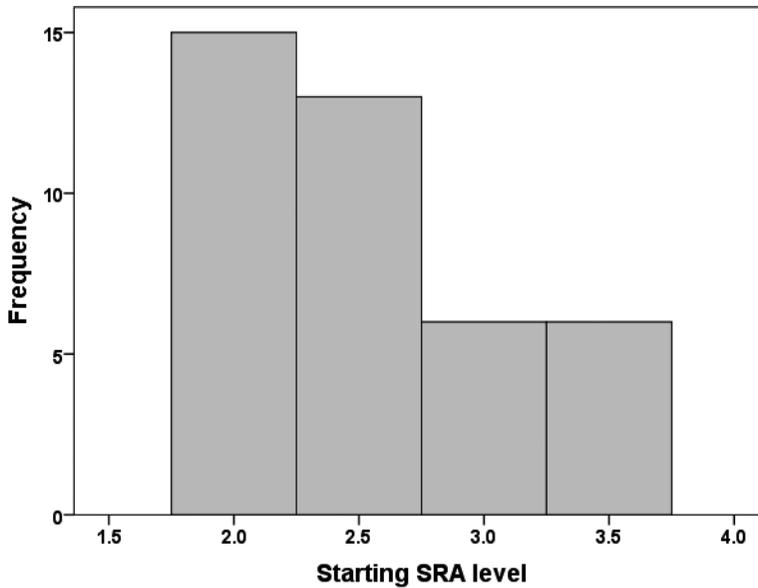


Figure 2. The distribution of the starting levels of SRA cards

Brown. 32.5% of the students started with Lime, which corresponds approximately to the reading level between the second and third graders. 15% of the students started with Aqua, which corresponds approximately to the reading level of the third graders, and another 15% started with Blue, which corresponds approximately to the reading level between the third and fourth graders. At the beginning of the program, the students reading levels ranged approximately from the second graders and that of between the third and fourth graders. 70 percent of the students started with the reading levels approximately of the second graders or close to the third graders, and 30 percent started with the reading level approximately of the third graders or the reading level close to the fourth graders.

Figure 3 shows the distribution of the final levels of SRA Rate- and Power-Builder cards. Compared with the starting levels, the reading levels of the control group spread, ranging from the level between second and third graders to the sixth graders.

Table 1 shows the number of words the students in the control group read in class, outside class, and the total of the two. The average number of words read in-class reading for two semesters was 14,577 ( $SD = 4272$ ). The smallest number of words read in class was 4,579; the largest number of words read in class was 23,797. Concerning the outside reading, the students' reading logs were available only for the second semester. The median number of words the control group read in the second semester was 49,524. The smallest number of words read was

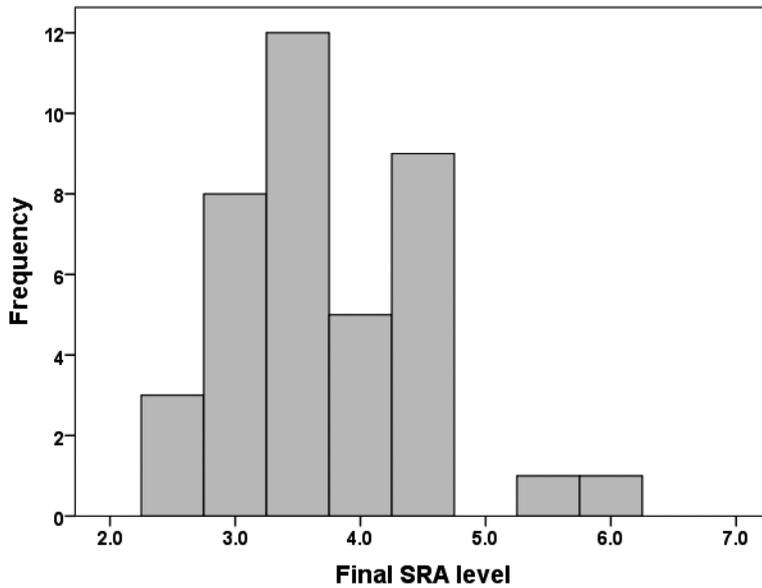


Figure 3. The distribution of the final levels of SRA cards

Table 1 *The descriptive statistics of the reading amount of the control group*

	Experimental Group	Control Group		
		In-class reading	Outside class reading	Total
Mean	197,391	14,577	83,919	98,496
Median	163,416	14,270	49,524	66,809
SD	141,374	4,272	106,147	106,420
Minimum	44,686	4,579	3,202	14,317
Maximum	943,582	23,797	630,255	643,815

3,202 and the largest number of words read was 630,255. Only about 8% of the students in the control group read more than 200,000 words, and about 24% of the students read between 100,000 and 200,000 words. On the other hand, about 70% of the students read less than 100,000 words. When we combine the number of words for in-class reading and outside reading, the median number of words read was 66,809; the minimum was 14,317 and the maximum was 643,815.

We cannot compare the number of words read by the experimental and the control groups directly since the amount of voluntary reading by the control group in the first semester was not available. However, if we assume that the control group had read approximately the same amount of reading outside class in the first semester, we can estimate that the total number of

words the control group read for a year would probably amount to between 100,000 and 110,000.

### The reading proficiency

The pretest and posttest data were merged and stacked together and they were analyzed using Rasch Unidimensional Measurement Model software (RUMM Laboratory). The main purpose of conducting Rasch analysis was to convert the data into an interval scale. The dichotomous model was used<sup>3)</sup>. In order to confirm that the data fit the dichotomous model, the following aspects were examined prior to obtaining the participants' ability scores (i.e., pretest and posttest scores) in logits: the item-total statistics, fit of items and persons to the model, the assumption of the local item independence, unidimensionality, targeting of the scale, and reliability.

By the end of the Rasch analyses, 12 items and two persons were deleted as misfitting items or persons. Table 2 shows the result of the final analysis. When the data fits the model, the mean of the fit residuals becomes close to zero and the standard deviation becomes close to one. The results show a fairly good fit to the model. However, the means of the items and persons indicate that item difficulty has a higher logit than the person logit, meaning that items were rather difficult for the participants. Also, the person separation index was .933 and Cronbach alpha was .932.

Figure 4 shows the person-item threshold distribution. The upper graph shows the distribution of the persons; the lower graph shows the distribution of items. Although the distribution of the items is greater than that of the persons, indicating that there are easier and more difficult items than the abilities of the participants, the items are well targeted.

After confirming that the data fit the model, the person abilities resulted from the final Rasch analysis were converted to standard scores ranging from 0 to 100: 0 was assigned to the smallest logit and 100 was assigned to the largest logit. All the further analyses were conducted, using IBM SPSS Statistics version 20. Then, the mixed-design ANOVA analysis was conducted. Further, Pearson's correlation coefficient  $r$  was calculated to show effects sizes. Following

Table 2 *The Summary statistics of the Rasch analysis*

	Items		Persons	
	Location	Fit residuals	location	Fit residuals
Mean	0.000	-0.112	-0.497	-0.211
SD	2.293	1.274	0.944	0.672

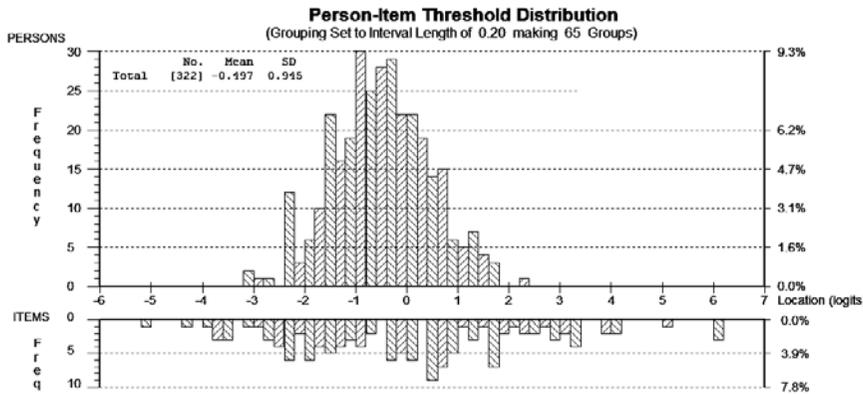


Figure 4. The distribution of the persons and items

Table 3 The descriptive statistics for the pretest and posttest (standard scores)

	Pretest		Posttest	
	Control	Experimental	Control	Experimental
Mean	40.27	41.63	45.42	58.37
SD	15.21	16.40	15.27	17.04
Min	17.48	0.00	16.20	14.88
Max	84.01	80.47	87.69	99.97

Note. The standard scores range from 0 to 100. 0 is assigned to the smallest logit and 100 is assigned to the largest logit.

Cohen's guideline (1988, 1992) about the size of an effect, effect sizes are defined as follows: small ( $r = .10$ ), medium ( $r = .30$ ) and large ( $r = .50$ ) (Field, 2009).

Prior to the mixed-design ANOVA analysis, the data was analyzed to confirm whether they would meet the assumptions for conducting a parametric test: normality and homogeneity of variance. To examine the former assumption, Kolmogorov-Smirnov test and Shapiro-Wilk test were conducted. The pretest and posttest of the control group and the experimental group indicate that the data were normally distributed. Further, to confirm the homogeneity of variance assumption, Levene's test was conducted. The results show that the variances are homogeneous for the control and experimental groups:  $F(1, 110) = .358, p > .05$  for the pretest;  $F(1, 110) = 1.211, p > .05$  for the posttest.

Table 3 shows the descriptive statistics for the standard pretest and posttest scores. The mean pretest scores of the control and experimental groups were 40.27 and 41.63, respectively. The two groups have a similar reading ability at the onset of the program<sup>4</sup>. On the other hand, the reading ability gain scores for the control group was 5.15 and that for the experimental

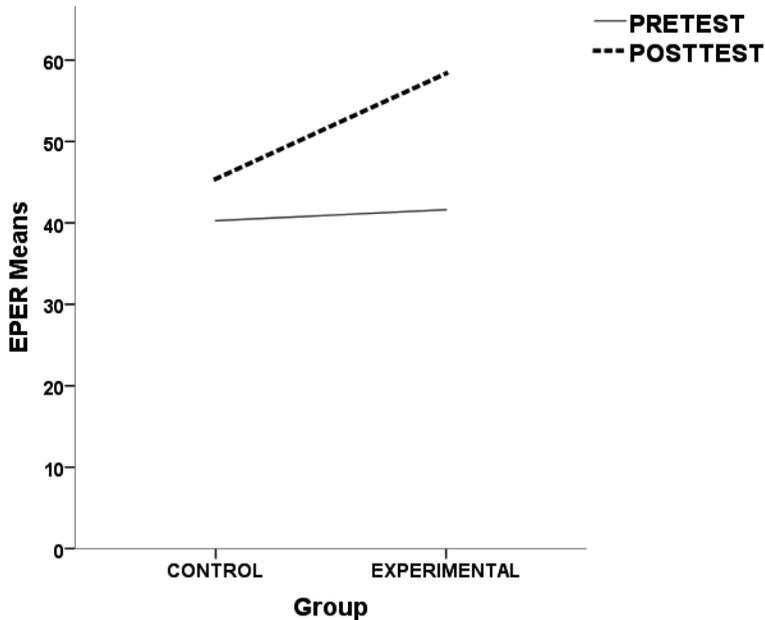


Figure 5. The EPER pretest and posttest results

group was 16.74, which was more than three times larger than that of the control group. Figure 5 shows the mean scores of the pretest and posttest of the experimental and control groups. The straight line shows the means of the EPER pretest and the dotted line shows the means of the EPER posttest.

The results of the mixed-t-design ANOVA showed a significant main effect of EPER:  $F(1, 111) = 121.25, p < .001, r = .72$ . The result shows a large effect size. Also, there was a significant main effect of class on the exam scores:  $F(1, 111) = 5.22, p < .001, r = .21$ . Further, there was a significant interaction between EPER and class:  $F(1, 111) = 34.06, p < .001, r = .48$ . That is, the reading ability did not develop uniformly. The effect size of the interaction between EDER results and class was large.

## DISCUSSION

Our research question asked “To what extent does a teacher’s guidance on book selection affect L2 readers’ growth in reading ability? The results show that the experimental group had significantly larger gain scores than the control group. We conclude that the differences in the amount of reading the two groups did resulted in the differences in the two groups’ gain scores.

As explained in the result section, the experimental group read 125.4 ( $SD=42.97$ ) books on average in one year. In terms of words, the average number of words read by the students in the experimental group was 197,391, and the median was 163,416. The control group, on the other hand, read 14,577 words on average in class. They read about 50,000 words outside class for one semester. We have estimated that the control group read approximately between 100,000 and 110,000 words for a year. This is about 50,000 to 60,000 words less than what the experimental group read.

The instructor's counseling on the selection of books for voluntary reading helped the learners to select books appropriate for their reading levels. Throughout the program, the instructor of the experimental group continued to give guidance to students so that they would read books appropriate for their reading abilities. On the other hand, the instructor of the control group encouraged the students to read outside class; however, she did not make any specific guidance about which books to read outside class. Further, she did not check the students' logs for voluntary reading to confirm that they were reading books appropriate for their reading abilities outside class.

The instructor in the experimental group provided an effective scaffolding for her students to read both in class and outside class. On the other hand, the instructor of the control group could not provide an effective scaffolding for her students' voluntary reading outside class. Scaffolding is generally defined as the support given to a younger learner by a more experienced adult (Wood, Bruner, & Ross, 1976). "A more experienced adult" is the instructor of the experimental group in this study. The instructor gradually removed her scaffolding so that her students could select books on their own and become independent readers. She first asked her students to read 100 easy books which are below their reading abilities in the first and second month of the instruction (Takase, 2010). The aim of this instruction was to eliminate the reading habits her students had before they took her class, including L1 translation, syntactic analysis, and the use of a dictionary for every single unknown word. After students finished reading books easier than their reading abilities, the instructor started to advise what kind of books the students should read. Further, she monitored the students' reading logs on a regular basis so that her students would read the books which were appropriate for their reading abilities. Also, she paid close attention to the students who seemed to go back to their old reading habits. Sometimes, she measured the reading time of the students who seemed to be reading too slowly, and gave proper guidance. Thus, the instructor continued to guide the students so that they could become independent readers.

In order to provide an effective scaffolding for students to read both in and outside class, a

teacher needs to analyze at which developmental stage each student is. In ER classes, teachers need to diagnose their students' developmental stages in learning to read in L2 and give them appropriate advice about which books to read or unlearn what they have learned in their previous learning experiences. This requires teachers to observe their students on a regular basis and utilize students' initial reading levels, the rate of their growth, and students' interests. A teacher plays diverse and significant roles in ER classes, which contribute to successful extensive reading practice.

### Limitation

The present study is not without limitations. First, the experimental group and the control group used different reading materials for sustained silent reading in class. The former used mainly leveled readers and graded readers in and outside class; the control group read the SRA Reading Laboratory cards in class, and read leveled readers and graded readers outside class. There may be a possibility that students in the control group did not take outside reading so seriously as they did in-class reading. A second limitation of the study is related to the EPER test as a measure of reading ability. The EPER test is originally designed to place students to proper reading levels using graded readers and examine their reading progresses. However, no validation study has been reported. The third limitation is that the reading logs of the control group's voluntary reading in the first semester was not available.

Although the present study has the above-mentioned limitations, the findings support the point that an ER teacher plays a crucial role in encouraging students to read voluntarily outside class. A replication study will be helpful to further clarify how a teacher can help learners to become independent readers through extensive reading.

### Notes

- 1) YL is the abbreviation of a Japanese phrase, *yomiyasusa reberu* (ease of reading). It is a readability index widely used among ER practitioners in Japan. YL was established subjectively, taking such factors as illustrations, the size of fonts, different text styles, genres, and the backgrounds of the books, familiarity with the content into consideration.
- 2) The number of words the participants in the control group read in the first semester was not available at the time of data collection.
- 3) In order to fit the data to the dichotomous model, the students' responses were converted to dichotomous data. This was due to a small sample size to analyze data using the Rasch measurement model.
- 4) An independent *t*-test was conducted to confirm that the experimental and control groups were equivalent. The result showed that there was no significant difference between the two groups ( $t =$

.047,  $df=128$ ,  $p=.963$ ).

## ACKNOWLEDGEMENT

This research was supported by Kansai University's Overseas Research Program for Kiyomi Yoshizawa in the academic year of 2009. Also, this work was supported by JSPS KAKENHI Grant Number 24520678.

## REFERENCE

- Begler, D., Hunt, A., & Kite, Y. (2012). The effect of pleasure reading on Japanese university EFL learners' reading rates. *Language Learning*, 62 (3), 665-703.
- Cohen, J. (1988). *Statistical Power Analysis for the Behavioural Sciences* (2<sup>nd</sup> ed.). New York: Academic Press.
- Day, R., & Bamford, J. (1998). *Extensive reading in the second language*. New York: Cambridge University Press.
- Day, R., & Bamford, J. (2002). Top ten principles for teaching extensive reading. *Reading in a Foreign Language*, 14 (2), 2002.
- Elley, W. (1991). Acquiring literacy in a second language: The effect of book-based program. *Language Learning*, 41, 375-411.
- Elley, W. (2000). The potential of book floods for raising literacy levels. *International Journal of Education*, 46, 233-255.
- Elley, W., & Mangubhai, F. (1983). The impact of reading on second language learning. *Reading Research Quarterly*, 19, 53-67.
- Field, A. (2009). *Discovering Statistics Using SPSS*. Los Angeles: Sage.
- Grabe, W. (2009). *Reading in a second language: Moving from theory to practice*. Cambridge: Cambridge University Press.
- Hafiz, F., & Tudor, I. (1989). Extensive reading and the development of language skills. *English Language Teaching Journal*, 43 (1), 4-13.
- Kern, R. (1994). The role of mental translation in L2 reading. *Studies in Second Language Acquisition*, 16, 441-61.
- Maruhashi, K. (2011, March). *The effect of extensive reading on Japanese EFL learners' grammatical competence*. Paper presented at the American Association for Applied Linguistics Annual Conference, Chicago, Illinois.
- Mason, B., & Krashen, S. (1997). Extensive reading in English as a foreign language. *System*, 24, 91-102.
- Parker, D. H. (1985). *SRA Reading Laboratory 3a Teacher's Handbook*. Chicago: Science Research Associates, Inc.
- Parker, D. H. (2004). *SRA Reading Laboratory 2a Teacher's Handbook*. Columbus, Ohio: McGraw-Hill.
- Robb, T., & Susser, B. (1989). Extensive reading vs. skills building in an EFL context. *Reading in a*

*Foreign Language*, 5 (2), 239-251.

Stanovich, K., West, R., Cunningham, A., Cipelewski, J., & Siddiqui, S. (1996). The role of inadequate print exposure as a determinant of reading comprehension problems. In C. Cornoldi & J. Oakhill (Eds.), *Reading comprehension difficulties* (pp. 15-32). Mahwah, NJ: L. Erlbaum.

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

Takase, A. (2010). *Eigo tadoku tacho shido manuaru* [Manual for instructing extensive reading and listening]. Tokyo: Taishukan.

Wood, D., Bruner, J. S., & Ross, G. (1976). The role of tutoring in problem solving. *Journal of Child Psychology and Psychiatry*, 25, 45-62.

Yoshizawa, K., Takase, A., & Otsuki, K. (2012, October). *Can we treat the EPER Form A and Form E as alternate forms?* Paper presented at the 16<sup>th</sup> Annual Conference of the Japan Language Testing Association, Senshu University, Kanagawa.