

## Reading rate gains during a one-semester extensive reading course

Jeffrey Huffman  
St. Luke's International University  
Japan

### Abstract

Extensive reading (ER) is an effective way to provide large amounts of comprehensible input to foreign language learners, but many teachers and administrators remain unconvinced, and it has been argued that there is still insufficient evidence to support the claims that have been made regarding its benefits. Few studies have looked at ER's effect on reading fluency. This article reports on an investigation of the reading rate gains of Japanese nursing college freshmen during a one-semester ER course, with students in an intensive reading (IR) course serving as the comparison group ( $N = 66$ ). The ER group achieved significantly higher reading rate gains (20.73 wpm) than the IR group (-.62 wpm), without sacrificing comprehension. These results add to a growing body of empirical evidence of the effectiveness of ER.

**Keywords:** reading; extensive reading, reading fluency, reading rate, reading speed, EFL/ESL reading, second language reading

Reading is perhaps the most important language skill for students learning English as a foreign language (EFL) or second language (ESL) in academic contexts (Grabe, 1991). Given its importance, it is unfortunate that in so many EFL settings around the world, teachers and administrators remain so heavily invested in grammar-translation and other methods that involve using reading as a route to form-focused grammar and vocabulary learning, to the exclusion or near-exclusion of developing fluent reading skills. As shown below, reading comprehension, fluency development, and enjoyment and confidence in reading go hand in hand and must be built up together through practice over time. The development of strong reading skills requires, for many low-proficiency learners in EFL contexts, the breaking of a downward spiral in which low fluency and comprehension lead to lack of enjoyment, which leads to less reading, which ensures that comprehension and fluency will remain low (Nuttall, 2005; Stanovich, 2000). For precisely these reasons, extensive reading has received attention from growing numbers of teachers and researchers, particularly since the 1990s (Grabe, 2009). Day and Bamford (1998) proposed that “students’ initial successful experiences in extensive reading result in the discovery that they can read in the second language (L2) and that is rewarding and pleasurable. This stimulates the development of positive attitudes toward reading, ...and these positive beginning experiences then feed back into subsequent extensive reading experiences...” (p. 30).

Why is there still resistance to implementing extensive reading programs? It might be that, as

Grabe (2009) has suggested, “teachers are not prepared to radically rethink how reading should be taught and learned,” or that “administrators and teachers are uncomfortable with teachers not teaching, and students not preparing, for high-stakes exams” (p. 312). However, another likely reason is that reading instruction methods that emphasize fluency, such as extensive reading, are still largely viewed as new, untested approaches. Teachers may also worry that reading rate and fluency increases may not be accompanied by improvement in reading comprehension. To remedy that situation, further research to clarify and provide empirical evidence of the benefits of extensive reading is needed.

Reading fluency refers to the ability to read words and process text rapidly and accurately and with good expression and prosody (Adams, 1994; Grabe, 2009; Nathan & Stanovich, 1991; Pressley, 2006). However, this definition assumes that oral reading fluency is an exact indicator of silent reading fluency, and this may not be true for EFL learners. Such learners may be able to read silently with a certain degree of fluency and comprehend the text but still be unable to perform well on oral reading fluency measures due to an inability to recode the words orally (Lems, 2006; Jeon, 2012). For this reason, the current study uses silent reading rate (with adequate comprehension) as a proxy for silent reading fluency.

Rather than simply being a sign of comprehension, research shows that improved fluency promotes improved comprehension (Breznitz, 1988; De Soto & De Soto, 1983; Jeon, 2012; Lems, 2006; Nathan & Stanovich, 1991). It should therefore be seen as a primary goal of any comprehensive reading education program. Extensive reading is a practical way to implement this goal, so it is imperative that the effect of extensive reading on reading fluency development be clarified and empirically demonstrated.

## **Literature Review**

### *Extensive Reading*

In their book on the subject, Day and Bamford (1998) laid out a set of characteristics of extensive reading. These include having students self-select from a wide variety of enjoyable reading material which is written well within their ability level, having them read extensively but individually and orienting them to the goals of extensive reading, including an emphasis on reading speed, and encouraging them to read for pleasure or information rather than for vocabulary and grammar learning purposes.

A great deal of research on extensive reading has been conducted in the past three decades, with tentatively positive results in the areas of incidental vocabulary acquisition (Day, Omura, & Hiramatsu, 1991; Dupuy & Krashen, 1993; Hayashi, 1999; Horst, 2005; Mason & Krashen, 1997; Nagy, Herman, & Anderson, 1985; Pitts, White, & Krashen, 1989; Saragi, Nation, & Meister, 1978; Waring & Takaki, 2003) and affect (Mori, 2004; Nishino, 2007; Robb & Susser, 1989; Stoeckel, Reagan, & Hann, 2012; Yang, 2001) as well as general reading proficiency (Elley, 1991; Elley & Mangubhai, 1983; Mason & Krashen, 1997; Tudor & Hafiz, 1989). Another likely benefit of extensive reading is in the area of improved reading fluency. Fluency can be thought of as the ability to read rapidly while accurately comprehending the text. However, while many

studies on the benefits of extensive reading have reported increased comprehension ability (Elley, 1991; Homan, Klesius, & Hite, 1993; Pichette, 2005; Tanaka & Stapleton, 2007), there have been surprisingly few well-designed empirical studies which measure and report reading rate changes as an indicator of fluency development (Beglar, Hunt, & Kite, 2012; Grabe, 2009). Examples of this small but growing body of research are reviewed in a later section of this paper.

### *Reading Fluency*

There is consensus, from Sir Edmond Huey in 1908 to the seminal modern work by LaBerge and Samuels (1974) and nearly every expert on the subject since, for the notion that reading fluency consists of component subskills, such as decoding, word recognition, phonological representation, and syntactic and semantic parsing or chunking, which are gradually automatized and unitized so that the reader's attentional resources can focus on the higher level processes of comprehension, analysis, and interpretation (Adams, 1994; Grabe, 2009; Kuhn & Stahl, 2003; LaBerge & Samuels, 1974; Nathan & Stanovich, 1991; Pikulski & Chard, 2005; Wolf & Katzir-Cohen, 2001). LaBerge and Samuels (1974) argued that, "if each component process requires attention, performance of the complex skill will be impossible, because the capacity of attention will be exceeded. But if enough of the components and their coordinations can be processed automatically, then the load on attention will be within tolerable limits and the skill can be successfully performed" (p. 293). Nathan and Stanovich (1991) described the phenomenon in this way: "When processes of word recognition take little capacity (are fluent), most of the reader's cognitive capacity can be focused on comprehending the text, criticizing it, elaborating on it, and reflecting on it—in short, doing all the things we know good readers do" (p. 176). Similarly, Adams (1994) explained that, "to the extent that you are directing that attention to the mechanics of the system, it is not available to support your understanding. Only if your ability to recognize and capture the meanings of the words on a page is rapid, effortless, and automatic will you have available the cognitive energy and resources upon which skillful comprehension depends" (p. 5).

Concerning the specific component subskills of which fluent reading consists, there is more variability among researchers. Wolf and Katzir-Cohen (2001) referred to "multiple components" including sublexical decoding processes, phonological representation, lexical access and retrieval, prosodic knowledge, and connected-text comprehension. Kuhn and Stahl (2003) settled on three components: decoding, word recognition, and prosody. Grabe (2009) listed four sub-processes essential to reading fluency: automaticity (defined as not requiring attentional resources, unconscious, and not subject to interference or suppression), accuracy, rapid overall rate, and recognition of prosodic phrasing or chunking.

The picture that emerges is that reading fluency development is much more than simply increasing the speed or "smoothness" of reading. It is a complex, multifaceted phenomenon involving the automatization of lower level processes and skills to free limited attentional resources so that these resources can be employed toward comprehending the global message of the text. It therefore follows that improved fluency should go hand in hand with improved comprehension. Breznitz (1988) conducted research showing that "reading at a faster pace increased comprehension and reduced errors" (p. 47). Her results indicated that the improved comprehension was likely due to reduced distractibility, because reading slowly allows too

much opportunity for distracting stimuli to be introduced during the “empty time” between words or sentences, while a reading rate that pushes the limits of readers’ attentional resources leaves them with no spare processing capacity with which to register irrelevant stimuli. The notion that fluency gains lead to comprehension gains has been confirmed in other studies as well, such as De Soto and De Soto (1983) and Nathan and Stanovich (1991) with first language (L1) learners and Jeon (2012) and Lems (2006) with ESL or EFL learners.

For the purposes of empirical research, fluency is often operationalized in terms of reading rate, and in the L1 setting this is normally done by having participants engage in oral reading. However, L1 children can read with great speed and accuracy and yet be unable to demonstrate any real understanding of what they have read (Pressley, 2006). This is a serious drawback because, as Pressley pointed out, “nobody should be interested in or promoting fast reading with low comprehension” (p. 209). In the L2 context, on the other hand, it is unknown whether oral reading is an appropriate way to measure reading fluency, given the extreme constraints of pronunciation and intelligibility (Grabe, 2009; Lems, 2006). It may very well be that L2 readers can read faster silently than orally. Given these considerations, fluency was measured in this current study as silent reading rate, and comprehension of the text was also tested to ensure that rate increases did not come at the expense of comprehension. This method is consistent with that employed in one of the few well-designed empirical studies of the effect of extensive reading on fluency in an EFL context to date, that of Beglar, Hunt, and Kite (2012).

If reading fluency improvement is to gain credence as an important educational goal in EFL contexts, it is important to consider what sort of targets might be seen as ideal for readers. Carver (1982) found that the most efficient *rauding* rate for college students reading L1 college-level material was around 300 wpm; *rauding* refers to the fastest speed at which a reader can read easy material at an adequate level of comprehension. Nuttall (2005) also reported that the average L1 reader reads at around 300 wpm, although there is a considerably wide range around that average figure. Higgins and Wallace (1989) noted that 180 wpm is “generally agreed to be close to the minimum at which reading becomes a pleasure” (p. 394) and that it may be “a threshold between immature and mature reading” (p. 392). Beglar, Hunt, and Kite (2012) suggested that this minimum L1 rate could be used as a reasonable goal for L2 readers in many contexts. Nuttall (2005) noted that secondary school students in ESL countries (presumably meaning countries where English is not the L1 but is used as the primary or exclusive language of education in all subjects, as opposed to countries where subjects other than English are taught in the L1) read at around 120 to 150 wpm before training.

### *Previous Studies on L2 Reading Rate Gains through Extensive Reading*

As previously mentioned, few studies have conclusively and empirically demonstrated the effectiveness of extensive reading in improving reading fluency in L2 settings. Some of the previous related studies are briefly summarized in this section, along with a discussion of their strengths and weaknesses.

Robb and Susser’s (1989) investigation of extensive reading vs. intensive reading in intact year-long university classes in Japan included a pre- and post-treatment measurement of reading speed for both groups. Although the two measurements were not equivalent, the extensive

reading group (ER group) read significantly faster than the intensive reading group (IR group) post-treatment (86.55 vs. 76.75 wpm), while there was no significant difference between the two groups pre-treatment. The ER group also scored significantly higher than the IR group on comprehension measures post-treatment, where there were no significant differences pre-treatment. These results constitute a fairly strong case that extensive reading helped improve reading fluency to a greater extent than did intensive reading. Drawbacks in this study are typical of those found in extensive reading studies using intact classes. Time-on-task was nearly double for the ER group, so the reading rate gains may be due simply to increased time spent reading rather than the pedagogical approach itself. Another drawback related to the use of intact classes is that the students were taking a total of six English courses concurrently; therefore, their fluency gains may not have come from the extensive reading course alone. A final drawback is that the students in the ER group were reading from a library of authentic fiction written for L1 teenagers rather than graded readers written for L2 learners, so it is likely that they did not meet the conditions required for fluency development to occur (see Nation, 2009).

Bell (2001) looked at extensive reading vs. intensive reading methodology with young adult EFL learners in Yemen, reporting changes in reading speed and reading comprehension over one year. Time-on-task was equivalent for the two groups, and the reported results are quite impressive indeed. The IR group's rate increased from 78.45 wpm to 92.54 wpm, while the ER group increased from 68.10 wpm to 127.53 wpm, both of these gains being statistically significant and the post-treatment difference between the two groups also being significant. The ER group also made higher comprehension improvement than the IR group. Drawbacks of this study include the small number of participants ( $N = 26$ ) and the fact that the amount of text actually read by the students in the study was not reported. A much bigger drawback, however, is the fact that student comprehension of the text used for the reading rate measure itself was apparently not checked. This means that the impressive reading rate gains seen by both groups in this experiment may indeed have come at the expense of comprehension, a danger the author himself warns about multiple times in the article.

Taguchi, Takayasu-Maass, and Gorsuch (2004) researched the effects of assisted repeated reading versus extensive reading during a one-semester course on reading rate and comprehension of Japanese university students ( $N = 20$ ). They reported a decrease in reading rate for both groups after the treatment, but it should be noted that the ER group read only 205 pages during the semester, the pre-treatment and post-treatment measures of reading rate were not equivalent, and that the books were likely well-above the level required for fluency development (comprehension scores were 1.90/16 on the pretest and 4.50/16 on the posttest). For these reasons, this study may have been limited in its ability to reveal reading fluency gains that may result from semester- and year-long extensive reading programs which involve a greater amount of reading and texts that are more closely matched to students' reading ability levels.

Beglar, Hunt, and Kite (2012) recently conducted what is perhaps the strongest study to date investigating the effect of reading extensively on fluency development. The participants were first-year Japanese university students ( $N = 97$ ), consisting of an IR group and three treatment groups engaging in various amounts of pleasure reading during a one-year program. The IR group made negligible reading rate gains, while the pleasure reading groups (PR groups) ranged from gains of 8.02 wpm (89.71 wpm pre-treatment to 97.73 wpm post-treatment) to 16.85 wpm

(103.9 wpm to 119.93 wpm). These results were in line with the amount of reading done by the different PR groups, and the gains came without any accompanying decrease in comprehension of the passages. The reason this program was termed “pleasure reading” rather than extensive reading is that students were provided with and permitted to read authentic (written for native speakers) novels in addition to graded readers (written for L2 readers). However, the authors provide a clear analysis of the correlation between authentic novel reading and graded reader reading and reading rate gains, finding that reading the L2-targeted graded readers resulted in greater rate gains. A minor limitation of this study, identified as such by the authors themselves, is that reading rate was measured in a slightly inaccurate way, by writing the elapsed time in 10-second increments on a whiteboard and having students look up and record their time when they finished reading the text. Overall, however, this study represents a strong improvement on previous studies and will hopefully be used as a model and a springboard for further studies on fluency improvement through extensive reading in a variety of EFL settings.

The current body of research remains severely lacking. There is still a dearth of empirical support for the claim that extensive reading results in fluency improvement. The reason for this gap in the research may be that the connection between extensive reading and reading fluency seems more obvious than with other areas such as vocabulary or comprehension, perhaps because the connection between practice and automaticity is already well-established in L1 reading research (Kuhn & Stahl, 2003; LaBerge & Samuels, 1974; Nathan & Stanovich, 1991). However, given the fact that teachers and institutions in EFL contexts continue to show a great deal of resistance toward implementing strong extensive reading programs (Grabe, 2009), there is a salient need for further empirical studies in this area.

### **Purpose and Hypotheses**

The current study is intended to build on the small body of existing research while avoiding some of the pitfalls encountered in previous studies. Given the likelihood that long-term extensive reading improves reading fluency and the lack of strong empirical support for this claim, the purpose of this study is to assess the effect of extensive reading vis-à-vis intensive reading on the reading fluency improvement of first-year college students in Japan. The a priori hypotheses to be investigated are as follows:

1. Reading rate gains will be significantly greater for students in a one-semester college extensive reading course than those in an intensive reading course.
2. Greater amounts of reading will yield significantly greater reading rate gains for students in the extensive reading course.

Both of these hypotheses, if confirmed, would support the assumptions evident in the literature on extensive reading and reading fluency outlined in the previous sections, and they would also provide additional support for the results found in the few empirical studies that have been conducted, particularly the finding in Beglar, Hunt, and Kite (2012) that reading fluency gains are greater with extensive reading than intensive reading and that there is a positive relationship between amount read and fluency gains.

## Method

### *Participants*

The participants were 66 female first-year students (19–20 years old) at a private, 4-year nursing college in Japan. They formally studied English approximately 5–6 hours per week in an EFL environment during their six years of secondary education. The study was carried out during the students' first semester of college, during which they were enrolled in two required English courses, each meeting for 90 minutes per week for 15 weeks. The treatment group consisted of 34 students who were enrolled in an extensive reading course and a writing course (the ER group). The comparison group consisted of 32 students who were enrolled in an intensive reading course and an oral communication course (the IR group). On a pre-course questionnaire, eight (24%) of the students in the ER group and six (19%) in the IR group reported previous experience in an extensive reading course or program (defined on the questionnaire as “reading texts or books that are at an easy level for you, but reading a lot of them over a long period of time, such as 1 book per week for a semester or a year.”) A standardized measure of the participants' overall English proficiency was unavailable, but the ER group had a mean vocabulary size of 3,312 words as measured by Nation and Beglar's (2007) Vocabulary Size Test (These data were not obtained from the IR group).

### *Study Design*

These data were gathered as part of the regular coursework of intact classes, so the design of this study is quasi-experimental. However, assignment to the classes was alphabetical and therefore effectively random, and the courses were planned and the data gathered with this study in mind, so the results should be generalizable to similar contexts. The duration of the study was one 15-week semester. The ER group (treatment group) consisted of students in two extensive reading classes (17 and 17), and the IR group (comparison group) consisted of students in two intensive reading classes (16 and 16). All students completed a pre- and post-course reading rate and comprehension test, and the ER group also completed the Vocabulary Size Test (Nation & Beglar, 2007) at the beginning of the semester. The number of words read by each group during the semester was calculated using “standard words” (Carver, 1982), in which each 6-character unit is counted as a word, and spaces and punctuation are also counted as characters. The length of different books and the number of words on a page varies significantly, and this variation is of particular concern when books written at different levels are being compared, and even more so when comparing authentic texts with texts written at an easy level for L2 learners. For this reason, standard words are a more accurate unit of measurement. The use of standard words also allows for more accurate comparisons of results across studies (Beglar, Hunt, & Kite, 2012).

The research project was explained to the students, orally in English and in written format in Japanese. Students were given the opportunity to anonymously request that their data not be used in the analysis, and the names of the students in the data set were replaced with random numerical codes by a third party researcher before the analysis was conducted. Authorization for this study was obtained from the research ethics committee of the college.

### *Intensive Reading Group*

The 32 students in the IR group read a selection of stories from *Chicken Soup for the Nurse's Soul* (Canfield, Hansen, Mitchel-Autio, & Thieman, 2001). All students in the IR group read the same amount of text, estimated at 9,682 standard words, or 32.6 pages. To arrive at this number, three randomly selected full pages were counted and averaged, and this was multiplied by the total number of full pages read, counted at the quarter-page sensitivity level. The total number was then doubled to reflect the fact that the IR students read each assigned excerpt twice, once for homework and once while going over the text during class. There is no way to guarantee that all of the students actually read all of the assigned text every week. However, the course instructor attested that the students participated actively in the class discussions of the meaning of the text (described below), and that this participation demonstrated that they were familiar with the content. Furthermore, the students knew that they would have a final exam covering the content of the text in detail, and that failure on the exam would result in having to retake the course. Thus the students had a high degree of motivation to actually read the text as assigned. Text readability was at Flesch-Kincaid Grade Level 8.0 (U.S. eighth grade reading level), as estimated by randomly sampling 5 pages out of the 33 pages read. This is a far higher level than both the texts read by the ER group (ranging from Flesch-Kincaid Grade 1.1 through 3.9) and the texts used in the reading rate tests (ranging from Flesch-Kincaid Grade 3.6 through 4.4).

The IR group's course was taught by a veteran Japanese English professor. Each week, all the students in the IR group were assigned to read a passage for homework, while one small group (3–4 students) was assigned to translate that passage into Japanese thoroughly. On a post-course questionnaire, students in this course reported that they had spent an average of 2.44 hours ( $SD = 1.38$ ) per week on their reading. In class, the small group presented their translation to the class orally and led a class discussion about the meaning of the text as well as any translation difficulties they had encountered. The instructor participated in this discussion and offered corrections and additional comments where the students' explanations and translations were lacking. There was also a 15-minute vocabulary quiz each week.

### *Extensive Reading Group*

The 34 students in the ER group read from an in-class library of 237 graded readers from the Macmillan Readers series (75 titles, Levels 2-6), the Oxford Bookworms series (89 titles, Levels 1-6), and the Cambridge English Readers series (73 titles, Levels 1-6). According to the publishers, these books are written using a controlled set of vocabulary ranging from 400 to 3,800 headwords. According to Beglar, Hunt, and Kite (2012), Flesch-Kincaid readability for the Oxford readers ranges from Grade Level 1.1 through 3.9.

In order to establish exactly how much reading the students in the ER group did, estimated standard word counts were calculated. The mean amount read by the ER group was 80,201.74 ( $SD = 29,747.75$ ) standard words. The students read an average of 545.85 pages, or 10.97 books. The number of standard words was calculated by first estimating the number of standard words per full page of a book from each level of each publisher. Three full pages were randomly selected from each of these books, and standard words were counted and averaged. Then, the number of full pages in each book that each student read was counted at the quarter-page

sensitivity level. Then, for each book that each student read, the estimated full-page standard word count for that level/publisher was multiplied by the actual number of full pages in that book.

Students were encouraged to borrow and return books freely throughout the semester. The goals and benefits of extensive reading were clearly explained, and students were encouraged to read books at or below their current reading level, so that there were few words in the text which they did not already know. As a general guide, they were encouraged to consider choosing an easier level if there were more than five or so words they did not know per page. The students were also strongly encouraged to find books that interested them, using the cover illustration/photo, title, back-cover blurb, and reading a page or two before choosing each book. Students were also told to stop reading a book and try a different one if they felt it was boring or too difficult. Most students started in the lower three levels, and some moved up to Levels 4 and 5 later in the semester while others stayed at the lower levels. Students were instructed to avoid using dictionaries except when absolutely necessary to understand a word that seemed important for understanding the story. They demonstrated their knowledge of the philosophy, goals, and benefits of extensive reading on an end-of-term test.

Students in the ER group spent class time engaging in book familiarization activities such as a poster contest and communicative activities designed to get them to talk to each other about the books they had read. They also had a 30-minute journal activity for most sessions, during which they got into pairs, summarized a book for each other orally, and then produced a written summary of their partner's book. They also did a series of six timed readings in addition to the pre- and post-course reading rate measures. Finally, they engaged in at least 30 minutes of sustained silent reading at the end of each session, during which they were allowed to choose and return books, read silently, or write book reports. On a post-course questionnaire, the students reported that they spent an average of 3.59 hours ( $SD = 1.79$ ) per week reading. It is presumed that approximately 30 minutes of this was during the in-class silent reading time.

At the beginning of the semester, students were told that they would be evaluated primarily on the number of pages they read, and that they needed to submit a book report to show that they had read each book. The amount read was evaluated on a sliding scale from 400 pages (passing) up to 800 or more pages (highest possible grade). The book report form required students to provide a variety of information about the story, most notably a 50–100 word summary of the plot including all the important points of the story, as well as a few sentences describing their thoughts and feelings about the story. The summary was evaluated carefully by the instructor, and summaries that did not include most of the main points of the story or that were not written in comprehensible English or that included major inaccuracies were returned to the students as “rewrites.” For rewrites, credit for having read the book was withheld until the student rewrote the report to the instructor's satisfaction.

### *Instruments*

*Reading Rate Test.* A selection of three texts from *Speed Reading* (Quinn & Nation, 1974) was administered as a pretest to all participants during the second class session, and a selection of three different texts from the same book was administered as a posttest during the second to last class session. Different texts were used for the pretest and posttest because reading fluency is

likely to be particularly sensitive to the practice effect. Having already read a particular text is likely to make it easier to read the second time around, even after a 12-week period. Each text was 550 words long (total length in standard words: 1,444.17 pretest; 1,489.34 posttest) and the texts were written using a vocabulary of 660 of the most frequent words in English, according to the authors. On the back of each text were ten comprehension questions, the purpose of which was to ensure that students were actually reading and comprehending the passages and, more importantly, to ensure that reading speed gains did not come at the cost of comprehension. The Rasch item reliability estimate for the comprehension questions was .76.

In order to mitigate the internal validity issue raised by using differing pretests and posttests, the texts used in the pretests and posttests were carefully selected so that the readability was nearly identical. Flesch-Kincaid grade levels for the pretest texts were 3.6, 3.9, and 4.4, and for the posttest were 3.7, 3.9, and 4.4. In terms of readability, this places the reading rate texts near the top of the range of the graded readers, but the high comprehension scores of the students indicate that the texts were well within their capability. The texts were selected in part because they were primarily non-fiction expository texts, rather than narrative fiction, as in the graded readers. If narrative fiction texts had been used, it could have been argued that the reading rate gains were only applicable to that genre, whereas using a different genre increases the likelihood that reading rate gains represent true gains in the students' core reading rate, because they transfer to a different genre.

A practice reading rate test consisting of one text was administered one week before the actual test in order to familiarize students with the procedure. The actual test consisted of a cover sheet outlining the procedure, then a double-sided sheet with the first text on the front and ten multiple-choice comprehension questions on the back, a blank sheet of paper, then the second text and questions, another blank sheet, and finally the third text and questions. Students were instructed to start reading upon the instructor's signal, read as quickly as possible while maintaining comprehension, then record their time at the bottom of the sheet. A large stopwatch was displayed at the front of the classroom. Upon recording their time, students turned the sheet over and answered the comprehension questions without referring to the text. Upon finishing the questions, students put their pens down and waited, not proceeding to the next text. The instructor monitored the students carefully to ensure that they did not refer back to the text nor proceed to the next text, and the blank page ensured that they could not see the next text. When all students were finished, the instructor signaled them to start the next text, and the procedure was repeated for the second and third texts. The posttest was administered in the same fashion, without the warm-up test.

*Vocabulary Size Test.* A 50-item version of the Vocabulary Size Test (Nation & Beglar, 2007) was administered to the ER group during the third class session in order to confirm that the students had sufficient vocabulary knowledge to read the graded readers and timed reading passages. In order to reduce the time required to administer the test, a version of the test covering only the 1st through 5th 1,000 word families was administered. The test consisted of 10 words per each 1,000 word frequency level, and it had a Rasch item reliability estimate of .87.

*Pre- and Post-Course Questionnaires.* Pre- and post-course questionnaires were administered to gain insight into the students' experiences with and feelings toward reading in English. Because

this study dealt solely with fluency improvement and did not delve into the area of attitude and motivation, for the most part the results of these questionnaires are not reported here. However, one item on the post-course questionnaire asked students to estimate from memory the number of hours per week they spent reading in English during the semester, and this data is reported as a self-reported measure of time-on-task. Additionally, the information on students' previous experience with extensive reading, as reported in the Participants section above, was obtained from the pre-course questionnaire.

## Results

Descriptive statistics for the results of the Vocabulary Size Test (extensive reading group only) were inspected. The mean vocabulary size of the ER group was 3,312.12 ( $SD = 429.94$ ), indicating that the students likely had a sufficient vocabulary to read the graded readers (400–3,800 headwords) and the reading rate passages (660 headwords).

As a measure of time-on-task, descriptive statistics for the amount of time spent reading, as self-reported on the post-course questionnaire, were inspected. The mean number of hours per week spent reading was 3.59 ( $SD = 1.79$ ) for the ER group and 2.44 ( $SD = 1.38$ ) for the IR group.

### *Results for Hypothesis 1*

Hypothesis 1 stated that reading rate gains will be significantly greater for students in the ER group than those in the IR group. In order to evaluate this hypothesis, the descriptive statistics were inspected and an independent samples  $t$  test was conducted with group (ER group and IR group) as the independent variable and reading rate gain scores as the dependent variable. Table 1 shows the pre- and post-course reading rates for the ER and IR groups, as well as the reading rate gains. The ER group had a mean reading rate increase of 20.73 standard words per minute ( $SD = 15.22$ ), while the IR group had a mean decrease of .62 standard words per minute ( $SD = 13.72$ ). Because two separate  $t$  tests were conducted in this study, a Bonferroni correction was made to ensure that the cumulative Type I error rate was below .05. The criterion for statistical significance was therefore set at .025. The independent samples  $t$  test was significant,  $t(64) = 5.97, p = .000$ . The eta squared index indicated that 36% of the variance of the reading rate gain variable was accounted for by whether a student was in the ER or IR group. These results unequivocally support the hypothesis.

Table 1. *Pretest and posttest reading rates, ER and IR groups (standard words per minute)*

	ER pretest	ER posttest	ER Gain	IR pretest	IR posttest	IR Gain
<i>M</i>	110.59	131.33	20.73	103.76	103.14	-0.62
<i>SE</i>	3.93	4.32	2.61	2.96	3.10	2.43
95% CI lower	102.60	122.54	15.42	97.73	96.82	-5.57
95% CI upper	118.58	140.12	26.04	109.79	109.46	4.33
<i>SD</i>	22.90	25.17	15.22	16.73	17.54	13.72

To confirm that any reading rate increase was not accompanied by a decrease in comprehension,

the descriptive statistics of the comprehension scores were inspected and paired samples *t* tests were conducted on the pre and posttest comprehension scores of each group separately. Table 2 shows the pre and posttest comprehension scores on the reading rate test (the average score of a set of three 10-item tests) for the ER and IR groups. The ER group had a mean comprehension score decrease of .30 ( $SD = .92$ ), while the IR group also had a mean decrease of .27 ( $SD = .76$ ). A paired samples *t* test was conducted on the change in comprehension scores for the ER group,  $t(33) = 1.90, p = .067$ , and the IR group,  $t(31) = 2.00, p = .054$ . These results show that the comprehension score changes for both groups were not significant.

Table 2. *Pretest and posttest comprehension scores, ER and IR groups (average score of three 10-item tests)*

	ER pretest	ER posttest	ER change	IR pre	IR post	IR change
<i>M</i>	8.63	8.33	-0.30	9.01	8.74	-0.27
<i>SE</i>	0.14	0.21	0.16	0.11	0.15	0.14
95% CI lower	8.36	7.90	-0.62	8.79	8.43	-0.55
95% CI upper	8.91	8.76	0.02	9.23	9.05	0
<i>SD</i>	0.79	1.24	0.92	0.62	0.85	0.76

### *Results for Hypothesis 2*

Hypothesis 2 stated that greater amounts of reading will yield significantly greater reading rate gains for students in the ER group. In order to test this hypothesis a Pearson correlation coefficient between amount read and reading rate gain for the ER group only was calculated. The correlation was a very low .057 and was non-significant.

## **Discussion**

This study provides solid empirical data supporting the effectiveness of extensive reading over intensive reading for reading fluency development. The reading rate gain achieved by the ER group in this one-semester course (20.73 wpm) was similar to that achieved by the highest-performing pleasure reading group (16.85 wpm) in the year-long study by Beglar, Hunt, and Kite (2012). One possible reason for this difference is that the students in this study read only easy graded readers written for L2 learners, while those in the Beglar, Hunt, and Kite study read a combination of L2-targeted and authentic material. Another, perhaps even more influential factor accounting for the more rapid gains in this study is the fact that students in this study engaged in timed reading activities during class time throughout the semester and were regularly encouraged to work on increasing their reading speed. It is likely that these timed readings resulted in a stronger reading rate increase than would be seen in an extensive reading course without such activities.

The students in the ER group in this study were able to improve their reading speed from 110.59 wpm to 131.33 wpm in one semester. If one accepts Higgins and Wallace's (1989) 180 wpm threshold as a reasonable goal for tertiary reading instruction in an EFL setting, this represents a substantial increase. By extrapolation, the threshold would be surpassed by the average student

after four semesters of extensive reading.

Unfortunately, the hypothesis that students within the ER group who read more would experience higher rate gains was not supported by the data in this study, a finding that is disappointing but consistent with previous research (see Beglar, Hunt, & Kite, 2012). This may be due to individual differences, in that some students require less input to increase their reading rate than others. It might be that some high-performing readers have already hit a reading rate ceiling before entering the course and that low-performing readers gain more from the extensive reading approach, but further research is needed to clarify this. Whatever the reason, these results seem to indicate that while individual differences in the amount read do not have a great effect on reading rate improvement, the instructional approach of extensive reading does have a large and significant effect when compared with intensive reading.

### *Limitations*

The results of this study should be viewed in light of its limitations, the most obvious of which is the fact that intact classes were used rather than setting up truly experimental conditions. Also, the students in the two groups were simultaneously enrolled in two different additional courses, writing and oral communication. The instruction received in these courses may have affected the results of this study.

A second, and again rather serious, limitation is that the pre-course and post-course reading rate measurements were conducted using different texts. This was done because reading fluency is likely to be particularly sensitive to the effect of having read the same passage previously. However, this has the adverse effect of making the comparison of pre-course to post-course reading rate highly suspect. It should be noted however, that the texts were part of the same timed reading series which used a pre-decided set of headwords and strict simplification rules, and they were matched for readability levels.

A third limitation is that while the comprehension scores on the timed readings used in this study fulfilled their purpose in showing that comprehension was not sacrificed for reading rate gains, they were quite high overall and did not show any increase in comprehension either. It is possible that there was a ceiling effect, and that slightly more difficult questions would have allowed the ER group to demonstrate comprehension gains in line with their rate gains. This would provide further support for the idea that comprehension improves with fluency (Breznitz, 1988; De Soto & De Soto, 1983; Nathan & Stanovich, 1991).

A final limitation is that the students in the ER group spent considerably more time reading during the semester than the intensive reading students, based on their self-reported data. It is possible and even likely, therefore, that the reading rate gains achieved by the ER group are due not only to the difference between the extensive and intensive reading approaches themselves, but also to the additional time the ER group students spent reading during the semester. From an experimental standpoint it would be ideal to control time on task, but from a pedagogical standpoint it can be argued that this difference in time spent reading is in itself an argument in favor of the effectiveness of extensive reading. It is also difficult to justify placing artificial limits on the time students spend reading for the purpose of an experiment.

## Conclusion

This study provides strong empirical support for the claim that extensive reading yields substantial improvements in reading fluency in first-year students at a nursing college in Japan. There have been few previous studies providing convincing evidence for the effect of the extensive reading approach on fluency development, so the results of this study represent a substantial finding in the field of L2 reading. In terms of methodology as well, this study advances this area of research by employing Carver's (1982) standardized method of quantifying text, by using exact time measurements when calculating reading rate, and by using texts and tests that were matched to students' reading ability level. It is hoped that the body of research in this area will continue to develop so that these findings can be confirmed and so that a more precise picture of how reading fluency develops in learners of EFL can be obtained.

### *Future Research*

As mentioned previously, reading fluency development through extensive reading in EFL or ESL contexts is still an area of inquiry that is seriously lacking, so this study raises a number of questions that would benefit from both quantitative and qualitative investigation:

- (a) Which factors within an extensive reading course or program are more or less responsible for the improvement in reading fluency? Possible candidates would be amount read, the absence or presence of possibly demotivating activities such as book reports or tests, control of level placement by the teacher, and engaging in timed reading or other activities designed to push students to increase their reading speed.
- (b) At what point or points within a semester- or year-long course do the students experience the greatest gains in their reading rate? This could be explored by measuring reading rate periodically rather than just at the beginning and end of the course.
- (c) Is measuring reading rate while ensuring maintenance of comprehension the best way to measure fluency development? It is possible that some students achieve rate gains without gaining true fluency in terms of chunking and automatization of reading processes and subskills, so qualitative methods need to be developed in order to get a more complete picture of fluency development.
- (d) To what degree do reading fluency gains from reading easy, L2-targeted novels in an extensive reading course transfer to other genres and particularly to the reading of authentic texts?
- (e) Studies focusing on fluency development through different reading approaches need to include measures of both reading rate and reading comprehension in order to get a more comprehensive look at the development of reading ability that occurs.

## Acknowledgments

I would like to thank David Beglar of Temple University for his helpful advice and encouragement throughout this project, as well as for reading, editing, and commenting on an

earlier version of the manuscript. I also thank the two anonymous reviewers for their valuable comments. I wish to thank Haruko Hishida of St. Luke's College of Nursing for her assistance in collecting the data used in this study as well as for her support and encouragement. I would also like to thank Kumiko Ohashi of St. Luke's College of Nursing (now St. Luke's International University, College of Nursing) for her assistance. Finally, I thank the students who allowed me to use their data in this study.

## References

- Adams, M. (1994). *Beginning to read: Thinking and learning about print*. Cambridge, MA: MIT Press.
- 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), 1–13.
- Breznitz, Z. (1988). Reading performance of first graders: The effects of pictorial distractors. *The Journal of Educational Research, 82*, 47–52.
- Canfield, J., Hansen, M. V., Mitchell-Autio, N., & Thieman, L. (2001). *Chicken soup for the nurse's soul*. Deerfield Beach, FL: Health Communications.
- Carver, R. P. (1982). Optimal rate of reading prose. *Reading Research Quarterly, 18*, 56–88. doi: 10.2307/747538
- Day, R. R., & Bamford, J. (1998). *Extensive reading in the second language classroom*. New York, NY: Cambridge University Press.
- Day, R. R., Omura, C., & Hiramatsu, M. (1991). Incidental EFL vocabulary learning and reading. *Reading in a Foreign Language, 7*, 541–551.
- De Soto, J. L., & De Soto, C. B. (1983). Relationship of reading achievement to verbal processing abilities. *Journal of Educational Psychology, 75*, 116–127. doi: 10.1037/0022-0663.75.1.116
- Dupuy, B., & Krashen, S. (1993). Incidental vocabulary acquisition in French as a foreign language. *Applied Language Learning, 4*, 55–64.
- Elley, W. B. (1991). Acquiring literacy in a second language: The effect of book-based programs. *Language Learning, 41*, 375–411. doi: 10.1111/j.1467-1770.1991.tb00611.x
- Elley, W. B., & Mangubhai, F. (1983). The impact of reading on second language learning. *Reading Research Quarterly, 19*, 53–67. doi: 10.2307/747337
- Grabe, W. (1991). Current developments in second-language reading research. *TESOL Quarterly, 25*, 375–406. doi: 10.2307/3586977
- Grabe, W. (2009). *Reading in a second language: Moving from theory to practice*. Cambridge: Cambridge University Press.
- Hayashi, K. (1999). Reading strategies and extensive reading in EFL classes. *RELC Journal, 30*, 114–132. doi: 10.1177/003368829903000207
- Higgins, J., & Wallace, R. (1989). Hopalong: A computer reading pacer. *System, 17*, 389–399.
- Homan, S. P., Klesius, J. P., & Hite, C. (1993). Effects of repeated readings and nonrepetitive strategies on students' fluency and comprehension. *Journal of Educational Research, 87*, 94–99.
- Horst, M. (2005). Learning L2 vocabulary through extensive reading: A measurement study.

- Canadian Modern Language Review*, 61, 355–382.
- Jeon, E. H. (2012). Oral reading fluency in second language reading. *Reading in a Foreign Language*, 24, 186–208.
- Kuhn, M. R., & Stahl, S. A. (2003). Fluency: A review of developmental and remedial practices. *Journal of Educational Psychology*, 95, 3–21. doi: 10.1037/0022-0663.95.1.3
- LaBerge, D., & Samuels, S. J. (1974). Toward a theory of automatic information processing in reading. *Cognitive Psychology*, 6, 293–323. doi: 10.1016/0010-0285(74)90015-2
- Lems, K. (2006). Reading fluency and comprehension in adult English language learners. In T. Rasinski, C. Blachowicz, & K. Lems, (Eds.), *Fluency instruction: Research-based best practices* (pp. 231–252). New York, NY: Guilford.
- Mason, B., & Krashen, S. (1997). Extensive reading in English as a foreign language. *System*, 25, 91–102. doi: 10.1016/S0346-251X(96)00063-2
- Mori, S. (2004). Significant motivational predictors of the amount of reading by EFL learners in Japan. *RELC Journal*, 35, 63–81. doi: 10.1177/003368820403500106
- Nagy, W. E., Herman, P., & Anderson, R. C. (1985). Learning words from context. *Reading Research Quarterly*, 20, 233–253. doi: 10.2307/747758
- Nathan, R. G., & Stanovich, K. E. (1991). The causes and consequences of differences in reading fluency. *Theory into Practice*, 30, 176–184.
- Nation, I. S. P. (2009). *Teaching ESL/EFL reading and writing*. New York, NY: Routledge.
- Nation, I. S. P., & Beglar, D. (2007). A vocabulary size test. *The Language Teacher*, 31(7), 9–13.
- Nishino, T. (2007). Beginning to read extensively: A case study with Mako and Fumi. *Reading in a Foreign Language*, 19, 76–105.
- Nuttall, C. (2005). *Teaching reading skills in a foreign language*. Oxford, England: Macmillan.
- Pichette, F. (2005). Time spent reading and reading comprehension in second language learning. *Canadian Modern Language Review*, 62, 243–262.
- Pikulski, J. J., & Chard, D. J. (2005). Fluency: Bridge between decoding and reading comprehension. *The Reading Teacher*, 58, 510–519.
- Pitts, M., White, H., & Krashen, S. (1989). Acquiring second language vocabulary through reading: A replication of the Clockwork Orange study using second language acquirers. *Reading in a Foreign Language*, 5, 271–275.
- Pressley, M. (2006). *Reading instruction that works* (3<sup>rd</sup> ed.). New York, NY: Guilford Press.
- Quinn, E., & Nation, I. S. P. (1974). *Speed reading*. Kuala Lumpur, Malaysia: Oxford University Press.
- Robb, T. N., & Susser, B. (1989). Extensive reading vs. skills building in an EFL context. *Reading in a Foreign Language*, 5, 239–251.
- Saragi, T., Nation, I. S. P., & Meister, G. F. (1978). Vocabulary learning and reading. *System*, 6, 72–78. doi: 10.1016/0346-251X(78)90027-1
- Stanovich, K. E. (2000). *Progress in understanding reading: Scientific foundations and new frontiers*. New York, NY: The Guildford Press.
- Stoeckel, T., Reagan, N., & Hann, F. (2012). Extensive reading quizzes and reading attitudes. *TESOL Quarterly*, 46, 187–198. doi: 10.1002/tesq.10
- Taguchi, E., Takayasu-Maass, M., & Gorsuch, G. (2004). Developing reading fluency in EFL: How assisted repeated reading and extensive reading affect fluency development. *Reading in a Foreign Language*, 16, 70–96.
- Tanaka, H., & Stapleton, P. (2007). Increasing reading input in Japanese high school EFL classrooms: An empirical study exploring the efficacy of extensive reading. *The Reading*

- Matrix*, 7, 115–131.
- Tudor, I., & Hafiz, F. (1989). Extensive reading as a means of input to L2 learning. *Journal of Research in Reading*, 12, 164–178. doi: 10.1111/j.1467-9817.1989.tb00164.x
- 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.
- Wolf, M., & Katzir-Cohen, T. (2001). Reading fluency and its intervention. *Scientific Studies of Reading*, 5, 211–239. doi: 10.1207/S1532799XSSR0503\_2
- Yang, A. (2001). Reading and the non-academic learner: A mystery solved. *System*, 29, 451–466. doi: 10.1016/S0346-251X(01)00038-0

### **About the Author**

Jeffrey Huffman is an assistant professor of English at St. Luke's International University, College of Nursing in Tokyo, Japan and a part-time lecturer at Waseda University's School of Political Science and Economics. He holds an M.S.Ed. in TESOL and is currently working on a Ph.D. in Education (Applied Linguistics), both from Temple University. His research interests are EFL reading and listening, second language testing, vocabulary acquisition, and English for specific purposes. E-mail: jeffhuffman@slcn.ac.jp