

## **Increasing reading self-efficacy and reading amount in EFL learners with word-targets**

Stuart McLean  
Osaka Jogakuin University  
Kansai University Graduate School of Foreign Language Education and Research  
Japan

Joseph Poulshock  
Senshu University  
Japan

### **Abstract**

We compared three methods for increasing reading amount and reading self-efficacy among L2 learners. (1) We required a word-target group to read at least 2,500 words a week outside class. (2) We required a sustained silent reading (SSR) group to do (a) in class SSR for 15 minutes every week, and (b) to read one book per week. (3) We required a comparison group to read one book per week. In the post-treatment period, we required all participants to read one book per week outside class, and during the post-treatment, the word-target group read significantly more, relative to a previously established baseline. We argue that learners in the word-target group internalized extrinsic motivation from the word-targets, and this led them to do more free reading and increase their reading self-efficacy more than the other groups.

**Keywords:** additive reading, Sustained Silent Reading, reading targets, MReader, extensive reading

All practitioners of extensive reading (ER) face the problem of how to motivate readers to read more. Though we would like to believe that reading motivation begins and ends where students find reading its own reward, (Krashen, 2015), we also acknowledge that motivation is complex and varies between individuals and across cultures (Dörnyei, 1990; Schmidt, Boraie & Kassabgy, 1996).

To promote and motivate ER, many first language (L1) and second language (L2) researchers and teachers recommend sustained silent reading (SSR) in class (Krashen, 2011; Matsui & Noro, 2010; McCracken, 1971; Nishizawa, Yoshioka, & Fukuda, 2010; Pilgreen, 2000; Takase & Otsuki, 2012). However, in many settings, teachers may find SSR difficult or impossible. Moreover, though SSR may be one key principle that facilitates ER, teachers need a toolbox of principles that will help motivate students to read. We argue that one of these principles may be

weekly attainable word-targets. In this study, students who had word-targets apparently became more self-efficacious because they continued to read more than comparison groups even after word-targets were removed. It may be a subtle point, but we believe this research discloses an elemental principle, which can help teachers and schools improve their ER programs; that is, *weekly word targets may help our students read more extensively and efficaciously even after the reading targets are later removed.*

## Literature Review

If we can find tangible and empirically supported ways to help learners increase their reading amount, then we may find the holy grail of extensive reading research, for "[t]he main goal of extensive reading is to read a lot. The best learner is the one who reads the most" (Nation, 2013, p. 56). However, we lack research that points to specific and actionable methods and techniques that will help students read more. For example, many studies tell us that students with intrinsic and extrinsic motivation read more (Grabe, 2009; Guthrie, Wigfield, Metsala, & Cox, 1999; Komiyama, 2013; Wang & Guthrie, 2004). However, classroom teachers may perceive intrinsic and extrinsic motivation as almost metaphysical concepts that they find hard to materialize in the real world.

Grabe (2009, p. 192) gives us a list of 24 practices for promoting motivation for reading. Some of the suggestions are somewhat tangible and doable: for example, "[g]ive students choices"; "[h]ave students share their interests"; and "[i]nvolve learners in decision-making related to reading tasks and goals." However, many of the suggestions are somewhat vague and not specifically actionable. For example, the list suggests that we "[b]uild students' self-confidence"; "[p]romote the development of group cohesiveness"; and "[p]romote effective learning strategies." Last of all, we need to "[g]enerate "flow"." This is all good advice, but teachers will need to make a lot of effort (with trial and error) to specify and apply the advice.

We want our students to read more extensively and efficaciously, and because reading is often a self-guided behavior, readers need self-efficacy as a key motivator for reading. In Bandura's (1977) social cognitive theory, self-efficacious individuals strongly expect to be able to perform a task or successfully maintain a behavior. These individuals are more likely to initiate coping behaviors, apply more effort to tasks, and overcome obstacles. In English as a Foreign Language (EFL) and English as a Second Language (ESL) settings, we can see that empirical results show a strong relationship between reading motivation and reading self-efficacy (Mori, 2002; Tremblay & Gardner, 1995). And Burrows (2013) posits that when learners have more reading self-efficacy (RSE), they read more.

Regarding self-efficacy, there are two different measures. General self-efficacy (GSE) concerns how people perceive "their ability to perform across a variety of different situations" (Judge, Erez, & Bono, 1998, p. 170). In contrast, specific self-efficacy (SSE) is defined as "beliefs in one's capabilities to mobilize the motivation, cognitive responses, and course of action needed to meet given situation demands" (Wood & Bandura, 1989, p. 408). Thus, based on the two-factor theory, we set out to investigate L2 English RSE.

Burrows (2013) created an instrument to measure a specific self-efficacy called *L2 English reading self-efficacy* or RSE, and Burrows concurs with Wood and Bandura (1989) who states that SSE instruments are good measures and predictors of human functioning in relation to foreign language education. Burrow's research is significant and relevant to this study because RSE and reading comprehension seem to constantly interact. Thus, as learners improve reading ability, they increase RSE, which leads to more reading. In his doctoral research, Burrows (2013) investigated RSE and reading comprehension related to four factors.

Specifically, Burrows (2013) looked at reading self-efficacy in relation to ER ( $n = 74$ ), grammar translation ( $n = 79$ ), reading strategies (RS) ( $n = 90$ ), and combined reading strategies with ER ( $n = 79$ ). The ER group participants significantly improved their RSE over the academic year but significantly less than the RS group and the RS+ER group. A latent growth curve indicated that when learners increased RSE, they also increased reading comprehension.

Nevertheless, Burrows (2013) did not run the same statistical analysis to investigate whether improved reading comprehension led to increased RSE. We expect that students who had not previously done ER would increase RSE after reading extensively. Therefore, we wanted to see if attainable weekly word-targets would help students read more, and that if by reading more they would also increase RSE. Hence, in addition to tracking student reading amount, we used a variation of Burrows' (2013) instrument to measure RSE.

In this paper, we address a gap in the literature regarding self-efficacy for Asian learners (Burrows, 2013; Oettingen & Zosuls, 2006), and specifically, we ask if it is possible to cultivate greater RSE in Japanese university students. While many studies deal with affective factors in foreign language achievement (Gardner & MacIntyre, 1993; MacIntyre, 1995; Saito & Samimy, 1996), our study is the only research that we know of which investigates how variations of ER influence RSE. Specifically, we asked the following questions:

1. Who increased free reading the most: the comparison group who received only verbal encouragement to read large amounts, the SSR group who engaged in in-class SSR, or the target group who received weekly reading word targets?
2. Who increased RSE the most: the comparison group who received only verbal encouragement to read large amounts, the SSR group who engaged in in-class SSR, or the target group who received weekly reading word targets?
3. Does reading amount during the treatment period correlate positively and significantly with increased free reading and RSE?

## Method

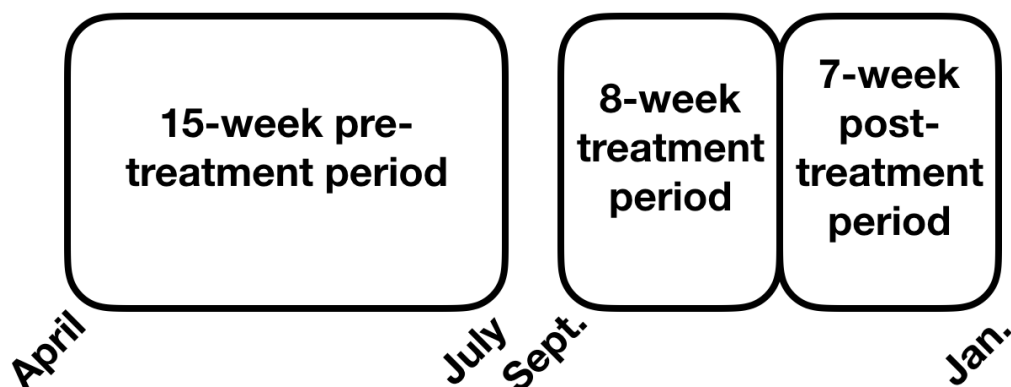
For this study, we used a quasiexperimental research design, including three intact reading classes. One class acted as a comparison group and two classes acted as treatment groups. We compared three treatments: (a) verbal encouragement, (b) SSR, and (c) word-target setting. We divided these three treatments over three reading periods: (a) the first semester 15-week, free reading observation period, which we used as a baseline (pre-treatment); (b) the 8-week reading period (treatment), and (c) the 7-week free reading observation period (post-treatment).

The pre-treatment free reading observation period lasted for the whole first semester, during which participants of all three groups did ER as homework, and they received similar amounts of verbal encouragement to read beyond the class requirement of at least one book a week. During the 15-week pre-treatment period, participants read a mean of 1,388 words ( $SD = 831$ ) per week. Students recorded the number of words that they read each week for both semesters on MReader, an online program (mreader.org) where students take short tests for the graded readers they read to demonstrate that they have read the books.

In the first half of the second semester, we conducted the eight-week treatment, during which the participants received their group-specific treatments. The SSR group participants did 15-minute SSR sessions, and they were given credit towards their final grades in return for reading at least one book a week. Students often started books during the SSR session and completed them outside of class. Though this situation is not ideal, 15-minute SSR sessions and reading one book per week are commonly practiced, and many people still refer to Nation and Wang's (1999) recommendation of reading one book per week. See also (Siyanova-Chanturia & Webb, 2016).

The word-target group did not do in-class SSR, but they were required to read at least 2,500 words per week outside of class to receive credit towards their final grade. We recommended that these participants read two books of around 1,250 words a week if they felt that books of 2,500 words were too difficult. For the third treatment, the verbal encouragement comparison group received marks towards their final grades in return for reading at least one book a week outside of class, and they did not do SSR in class. During the treatment periods, we equally encouraged all participants of each group to read as much as possible.

Right after the 8-week reading treatment period, we conducted the 7-week post-treatment free observation period. During this second free reading period, none of the participants did in-class SSR. While we encouraged all groups to read as much as possible each week, all participants were assigned marks towards their final grades for reading at least one book a week. Figure 1 shows the three periods.



**Figure 1.** Free and treatment reading periods in this study.

### Participants

The 59 participants (female = 5 and male = 54, 18–19 years of age) were first-year Japanese EFL learners attending a private university in Japan. The participants belonged to three different classes, which formed each treatment group: comparison group ( $n = 18$ ), SSR group ( $n = 22$ ), and the word-target group ( $n = 19$ ). Participants were students of the Department of Economics, and the courses for this study met once a week for a total of 30 weeks of classes over two 15-week semesters. Standardized test scores were unavailable for these participants. However, students, not included in this study, of a similar English proficiency had a mean TOEIC score of 280. Students in this study were enrolled in other English courses, but these courses did not do ER. Additionally, the participants reported no previous experience of ER in English. The treatments were randomly assigned to the intact classes, and permission from the university and participants was sought and received before the study was conducted.

When comparing groups of learners, it is important to provide evidence that prior to the treatment period they did not display significantly different reading behaviors or possess significantly different levels of RSE. Among the three groups, we found no statistically significant difference in either first semester free reading amounts ( $p = .93$ ,  $\eta^2 = .00$ ) nor pre-treatment RSE ( $p = .88$ ,  $\eta^2 = .01$ ), with as seen in Table 1. Table 2 shows mean group pre-treatment RSE estimates.

Table 1. Pre-treatment MANOVA analysis of RSE and first semester free reading amounts

Source	Multivariate			Univariate					
	F	p	$\eta^2$	RSE			Weekly reading		
				F	p	$\eta^2$	F	p	$\eta^2$
Class	.09	.99	.00	.13	.88	.01	.07	.93	.00

Note. RSE = reading self-efficacy

Table 2. Pre-treatment RSE descriptive statistics (values are in Rasch logits)

	Comparison	Word-Target	SSR
M	-.43	-.59	-.62
SE	.37	.27	.21
95% CI lower bound	-1.21	-1.16	-1.05
95% CI upper bound	.36	-.02	-.18
SD	1.58	1.19	.98

Note. SSR = sustained silent reading

Table 3 shows the number of words each group read during the pre-treatment and treatment periods. The comparison and SSR group participants slightly decreased their weekly reading amount during the treatment period relative to the first semester pre-treatment reading period. The target group read significantly more during the treatment period relative to the first semester free reading period ( $t = 3.37$ ,  $df = 18$ ,  $p = .002$ ,  $g = .81$ ). This was not the case for the comparison group ( $t = .46$ ,  $df = 17$ ,  $p = .650$ ,  $g = .10$ ) and SSR group ( $t = .63$ ,  $df = 21$ ,  $p = .534$ ,  $g = .20$ ).

Table 3. *Reading amount descriptive statistics (values in words)*

15-week pre-treatment free reading period			
	Comparison	Word-Target	SSR
<i>M</i>	1271.50	1343.50	1318.86
<i>SE</i>	147.00	185.76	58.92
<i>95% CI lower bound</i>	961.36	953.23	1196.34
<i>95% CI upper bound</i>	1581.64	1733.77	1441.39
<i>SD</i>	623.67	809.72	276.35
Eight-week treatment period			
	Comparison	Word-Target	SSR
<i>M</i>	1175.72	2852.05	1214.86
<i>SE</i>	179.04	564.65	96.28
<i>95% CI lower bound</i>	797.99	1665.77	1014.64
<i>95% CI upper bound</i>	1553.46	4038.34	1415.08
<i>SD</i>	759.59	2461.25	451.58

### *Instruments*

To measure the RSE, we used the L2 Reading Self-Efficacy Survey, a 14-item survey developed by Burrows (2013), which we administered at the start and end of the treatment period. The items are measured on a 6-point Likert scale: ranging from 1 (*I cannot do it at all*), 2 (*I cannot do it*), 3 (*I probably cannot do it*), 4 (*I can probably do it*), 5 (*I can do it*), and 6 (*I can definitely do it*). Items ranged from asking participants to judge to what degree they could "Read and understand the items on a menu written in English at a fast-food restaurant" to "Read and understand the main ideas of an article in a newspaper published in an English-speaking country that is written about a topic related to your major (economics) at university." We calculated the reliability of the data using WINSTEPS Version 3.75.0 (Linacre, 2007). We found a Rasch item reliability estimate of .99, and a Rasch person reliability of .89. Rasch person reliability estimates are more conservative than Cronbach alpha estimates but are preferable because they are based on data that conform stochastically to the Rasch model. As such, the Rasch reliability estimate indicates the repeatability of the linear measures constructed by the Rasch model.

To measure word counts, participants completed tests about the books they read using MReader. MReader was utilized to check if participants comprehended the books that they read. MReader provides an online admin page through which teachers may monitor reading word counts and books read. For students to demonstrate reading comprehension, they had to score 70% or more on comprehension questions. Nuttall (2005) recommended this criterion, and Anderson (2008, p. 67) designated it as "adequate comprehension." If a student failed a test, the word count for that graded reader did not count toward the student's reading requirement, and the student was required to read more to make up their weekly reading target. The number of tests failed was negligible. Students were closely monitored and encouraged every week to complete their reading target and catch up quickly when necessary.

### *Materials*

Participants of all three groups had access to the same ER library located in the university language center. Participants chose graded readers predominantly from the *Foundations Reading Library* series published by Heinle Cengage and the *Building Block Library* series levels five to seven, published by Scientific Education Group Co., Ltd. (SEG). The full range of *Oxford*, *Cambridge* and *Penguin* graded readers were available to readers, but only a few participants read *Oxford*, *Cambridge* and *Penguin* graded readers beyond the simplest level, which they did predominantly during the treatment period and post-treatment reading period.

Both the *Foundations Reading Library* series and levels five to seven of the *Building Blocks Library* are around the 301–400 headword level of the Extensive Reading Foundation Graded Reading Scale (2017). McLean (2014) showed that 94.5% and 97.7% of the tokens within the Foundation Reading Library are within the first 1,000 and 2,000 words of the BNC, respectively. Moreover, a number of the words outside of the first 1,000 and 2,000 words are loanwords in Japanese, like *cookie* or *pirate* that the participants are expected to have known. We therefore judged these books to be within the students' reading abilities. Students were also regularly reminded how to choose appropriate reading materials: (a) open a page and read it and understand it, (b) make sure there are no more than two unknown words per page, and (c) choose something enjoyable.

### *Interviews*

To investigate possible reasons for statistical findings, we conducted exploratory follow-up interviews with at least three students from each group. They were each informed that they could request at any time to withdraw with no penalty. These interviews were conducted in a public study area at the university, which provided a well-lit and quiet place to focus and speak. To make the students comfortable, the interviews were conducted in the language initiated by the students, either Japanese or English, although there was often a high degree of code-switching throughout. The interviewer was judged to be proficient in Japanese, having obtained the second highest-level certification on the Japanese Language Proficiency Test.

The interviews with students took about 20 minutes each, which consisted of showing the students the findings of the study in the form of tables and graphs. For example, the data from table 3 was presented in a graph along with table three itself and students were asked to provide possible explanations for the findings. The interviews were not recorded. However, notes were made of explanations provided by the students. Notes were also made from informal conversations with participants during class, and before and after class.

## **Procedures**

### *Types of Data Analysis*

In this study, we obtained raw RSE scores from Burrows' (2013) RSE survey. To increase the instrument's sensitivity to changes in students' RSEs, we used a Rasch procedure called stacking with data collected from the survey. This procedure reduces interference resulting from

instability within (a) the rating scale and (b) the item difficulty (Linacre, 2007). Stacking helps to reduce error within measurement caused by differences in the functioning of survey items and rating scales even when identical data collection protocols are used so that measurements of change will have unambiguous numerical representations and substantive meaning. See Linacre (2007) for more details on the functioning and uses of stacking.

Research Question 1 asked who increased free reading the most: the comparison group who received only verbal encouragement to read large amounts, the SSR group who engaged in in-class SSR, or the target group who received weekly reading word targets? We investigated Research Question 1 by first comparing the mean for the pre-treatment period reading amount against the mean for the post-treatment period reading amount, using paired sample *t*-tests individually for each group.

Research Question 2 asked which group increased RSE the most: the comparison group who received only verbal encouragement to read large amounts, the SSR group who engaged in in-class SSR, or the target group who received weekly reading word targets? We investigated Research Question 2 by first comparing the mean for pre-treatment and post-treatment RSEs for each group. To investigate Research Questions 1 and 2, we conducted three paired sample *t*-tests for each group. Because of the large number of statistical analyses, we set the criterion for statistical significance among paired sample *t*-tests with Bonferroni correction at  $< .008$  ( $.05/6 = .008$ ).

Besides statistical significance testing, we calculated the effect sizes of difference to investigate the efficacy of the treatments. Effect size is a simple way of quantifying the difference between groups, or the influence of a treatment on a group. Effect size emphasizes the impact of a treatment rather than confounding this with sample size (Coe, 2002, Carver, 1978; 1993, Hojat & Xu, 2004; Rosenthal, 1996). Carver (1993, p. 288) insists "that attention be paid to the size of the effect, whether it is statistically significant or not." This study adapts Plonsky and Oswald's (2014) recommendations for small, medium, and large effect sizes for mean differences (*g*) resulting from within-groups contrasts, .6, 1.0, 1.4; respectively for the correlation coefficients (*r*), .25, .4, .6.

Research Question 3 asked if reading amount during the treatment period correlated positively and significantly with increased free reading and increased RSE. We addressed Research Question 3 in two steps: (a) by calculating the Pearson's correlation coefficient between weekly treatment reading amount and changes in RSE, and (b) by calculating the Pearson's correlation coefficient between weekly treatment reading amount and changes in weekly reading amount between pre-treatment and post-treatment free reading periods. The criterion for statistical significance for the Pearson's correlation coefficients was set at  $< .025$  ( $.05/2 = .025$ ). The significance threshold of .05 was divided by two as two Pearson's correlations were conducted.

## Results

Research Question 1 asked who increased free reading the most between the pre-treatment free reading period and the post-treatment free reading period. Table 4 shows the number of words



each group read during the post-treatment. Table 5 shows that only target group participants increased the amount of reading during the post-treatment free reading period relative to pre-treatment free reading period. The increase in reading amount was not statistically significant ( $t(18)=2.153, p = .045$ ); however, the word-target group treatment yielded a small effect size,  $g .62$ .

Table 4. *Post-treatment seven-week free reading period (values in words)*

	Comparison	Word-Target	SSR
<i>M</i>	1260.00	2873.47	1211.27
<i>SE</i>	170.07	775.75	82.17
<i>95% CI lower bound</i>	901.18	1243.69	1040.39
<i>95% CI upper bound</i>	1618.82	4503.26	1382.16
<i>SD</i>	721.55	3381.41	385.42

Table 5. *Summary of paired sample t tests*

Group	<i>M</i>	<i>SD</i>	<i>SE</i> Mean	95% Confidence Interval		<i>t</i>	<i>df</i>	Sig (2-tailed)
				Lower	Upper			
Comparison	-11.50	387.89	91.43	-204.39	181.39	-.13	17	.901
Word-Target	1529.98	3097.64	710.65	36.96	3022.99	2.15	18	.045
SSR	-78.96	277.21	59.10	-201.86	43.95	-1.34	21	.196

Research Question 2 asked who increased RSE the most. Table 6 shows the statistical descriptions of RSE across groups in the post-treatment. Table 7 demonstrates that only the word-target group participants demonstrated a statistically significant increase in RSE ( $p = .001$ ), in contrast to the comparison ( $p = .250$ ) and SSR groups ( $p = .637$ ). The comparison ( $g .28$ ) and SSR ( $g .09$ ) group treatments yielded negligible effect sizes. In contrast, the word-target group treatment yielded a medium effect size ( $g 1.07$ ).

Table 6. *RSE descriptive statistics (values are in Rasch logits)*

	Comparison	Word-Target	SSR
<i>M</i>	-.02	.66	-.51
<i>SE</i>	.32	.26	.29
<i>95% CI lower bound</i>	-.71	.11	-1.11
<i>95% CI upper bound</i>	.66	1.21	.09
<i>SD</i>	1.37	1.14	1.35

Table 7. *Paired sampled t tests comparing pre-treatment and post-treatment RSE values*

Group	<i>M</i>	<i>SD</i>	<i>SE</i> Mean	95% Confidence Interval		<i>t</i>	<i>df</i>	Sig (2-tailed)
				Lower	Upper			
Comparison	.40	1.43	.34	-.31	1.11	1.19	17	.250
Word-Target	1.25	1.44	.33	.56	1.94	3.78	18	.001
SSR	.11	1.08	1.08	-.37	.59	.48	21	.637

Research Question 3 asked if reading amount during the treatment period correlated positively and significantly with increased free reading and RSE. The positive correlation between treatment reading amount and increase in free reading amount was statistically significant ( $p =$

<.001,  $r = .54$ ). However, the correlation between the number of words read during the treatment period and changes in RSE were not significant ( $p = .08$ ,  $r = .19$ ).

## Discussion

Tables 4 and 5 show that only word-target group participants increased their free reading amount. The increase in reading amount was not statistically significant ( $p = .045$ ); however, the word-target group treatment yielded a small effect size ( $g .62$ ). While not significant, the SSR and comparison groups slightly decreased their free reading amounts from the first to second semester. Some participants reported reading less because of end of term deadlines.

This supports the above cited idea of doing extended periods of SSR, which provides students with ample in-class time to read. However, interviews with participants from all groups found that there was no difference in the end of semester workloads experienced by participants from the different groups. SSR, comparison, and word-target group participants who did not increase their free reading amounts reported that they did not have time or did not make the effort to read more despite believing that ER would increase their reading ability and performance on standardized tests.

Again, this supports the provision of in-class reading time for SSR where possible. When in-class reading is not possible, this study suggests that weekly word-targets and weekly monitoring of student's reading amount may facilitate large amounts of reading—even after word-targets are removed—and that weekly word-targets encourage students to read beyond what is required to receive credit towards their final grade.

Participants from all groups who read more the second semester gave three reasons for why they read more. First, by being required to read at least 2,500 words a week, a majority of word-target group participants read longer books, and they found that longer books were more interesting. As a result, after we removed the 2,500 word-target, participants did not return to very short books, which they commonly considered as less interesting than longer and more difficult books.

However, the SSR and comparison group participants did not commonly read longer books, and so they did not commonly experience that reading longer books is more enjoyable. Books containing only about 1,300 words are written with a highly-restricted vocabulary, and writing interesting stories with such highly restricted vocabulary is problematic (McLean, 2014). Further, these stories are understandably more simplistic than the stories in longer books. Second, by reading larger amounts during the treatment period, participants stated they felt more confident and able to read relatively difficult material rather than the simplest of graded readers which they read during the first term. Finally, a few students reported simply getting used to reading books of around 3,000 words rather than 1,000 words.

Takase (2007) claims that intrinsic motivation is key to students reading more. However, statistical analysis and interviews with participants suggest that word-target group participants experienced extrinsic motivation when they were required to read roughly double what they had previously read, and this extrinsic motivation was internalized and resulted in intrinsic

motivation. This process by which extrinsic motivation is internalized and becomes intrinsic motivation is described by Deci and Ryan's (1985) self-determination theory. Participants increased L2 reading self-efficacy by reading appropriate but more difficult material during the treatment period than the pre-treatment period.

Participants from all groups were interviewed, including those who did and did not increase their reading amount and RSE. Over 90% of the interviewed students reported that the books of 1,300 words were below their reading ability. As a result, it seems that most students had the ability to read longer books than they often read during the first semester free reading period. However, it was clearly the word-target group participants who strengthened their expectations to conduct L2 reading, and who discovered that they enjoyed longer books relative to the very simple short books which were more commonly read during the first term.

These findings question the appropriateness of Nation and Wang's (1999) recommendation of reading a minimum of a book a week, or the use of the unit of books when setting reading goals with EFL students in compulsory English classes. Additionally, we found that when low proficiency participants selected books of an appropriate level, the books commonly consisted of around only 1,000 words. As a result, if low proficiency students read at least a book of an appropriate level each week, it could be only 1,000 words long, and so over a Japanese academic year, they would only read around 30,000 words. This is much less than the recommendations made by Beglar, Hunt, and Kite (2012) of around 230,000 words, and Nishizawa, Yoshioka, and Fukuda (2010) of 300,000 words.

Research Question 2 asked who increased RSE the most. As shown in Tables 6 and 7, only word-target group participants significantly increased their RSE. Following interviews with participants, we believe that this happened because word-target participants significantly increased their reading amount during the treatment period (Table 3) relative to participants of other groups. This is supported by the significant positive correlation ( $p = <.001$ ), between the reading amount and RSE. For all participants, this was their first experience of ER, and over half reported it as their first experience of reading material of an appropriate level.

Participants who read large amounts reported finding the reading process increasingly easy as the term progressed. They also reported being able to concentrate for longer periods of time while reading. This led in turn to greater confidence and more positive feelings towards reading. While participants did not use the term *L2 self-efficacy*, they described experiences of reading more difficult material than they had read in the first semester, realizing that they were able to read such L2 reading material comfortably. Thus, they continued to read when time allowed them to despite not being required to do so.

These participants may also have experienced a *virtuous cycle* of reading appropriate material (McLean, 2014) where readers understand better, and thus read faster, enjoy more, and then read more. This cycle exemplifies the extensive reading bootstrap hypothesis (Day & Bamford, 1998), where gains from reading enable readers to get greater results more quickly and efficiently. A majority of participants who increased their weekly free reading amount reported feelings of pride in being able to simply comprehend L2 material. Participants were proud to conduct reading they considered as "moto otonapoi" (i.e., *more adult like*), and not reading books which

participants commonly described as "kantan sugiru" (i.e., *very easy*), the reading of which some saw a "kodomopoi" (i.e., *childish*).

In contrast, some SSR group participants reported that they did not read outside of class at all, as they were able to read at least one book per week as required within the 15-minute SSR period. As a result, we might hypothesize that SSR and comparison group participants did not experience the feelings of achievement from reading relatively large amounts by themselves outside of class, in contrast to word-target participants. Yet the participants that were interviewed did not independently offer this explanation themselves, and the interviewer did not suggest such an explanation to the participants to avoid influencing their responses.

Research Question 3 asked if reading amount during the treatment period correlated positively and significantly with increased free reading and RSE. We found a significant ( $p = <.001$ ,  $r = .54$ ) positive correlation between treatment period reading amount and changes in free reading. This finding does not provide evidence of causality, but it supports the causal claims related to Research Questions 1 and 2 that weekly reading facilitated more reading and in turn further reading and increased RSE. We did not find a significant correlation ( $p = .08$ ) nor a meaningful effect size ( $r = .19$ ) between reading amount during the treatment period and changes in RSE. This may result from the indirect and weaker link between reading more during the treatment period and RSE.

The lack of significant correlation and meaningful effect size could be the result of four possible causes. First, reading during the treatment period and reading during the post-treatment reading period were very similar constructs. Thus, we expect that when participants read more during the treatment period, this increase will influence reading amount more than it influences RSE. That is, increased reading may influence even more reading, more than it influences RSE. Secondly, we can more easily and accurately measure reading amount than we can measure RSE. Third, we expect that we would find a statistically significant correlation with a larger effect size if we had a larger sample size of participants who read more during the treatment period.

However, we found limited variance within the data for the change of reading amount, with most participants only slightly changing the amount they read between pre-treatment and post-treatment free reading periods. Fourth, we believe that greater amounts of reading are necessary to see the increase in RSE. During the treatment period, all the participants read on average 1,728 words a week, while word-target group participants read 2,461 words a week, or 12,101 and 17,227 words respectfully over the seven-week post-treatment period. This amount is low compared with the 300,000 words that Nishizawa, Yoshioka, and Fukuda (2010) state is necessary for students to become confident readers.

### *Limitations*

We need to view the results, inferences, and the following conclusions in the context of the limitations of this research. The participants in this study were Japanese learners of compulsory English in their first year at a private university. We can only cautiously generalize this study in relation to learners of different proficiency levels, majors, or native languages. We would expect

other kinds of students to respond differently, such as L2 learners majoring in English and L2 learners with more extensive reading experience than those in this study.

This research is also limited by the quasi-experimental research design, using intact classes rather than a full experimental design in which learners are randomly assigned to different treatments. Future researchers should utilize an experimental design to explore the treatments done in this study and their relationship to RSE and reading amount.

In addition, we used only a short version of SSR, and the period of only 15 minutes a week may have been too short. However, we chose a period of 15 minutes because it is a reasonable maximum amount of time that general classes (not reading classes) might be able to devote to in-class reading. In compulsory EFL courses at Japanese universities, students seldom take classes devoted exclusively to reading, and extensive reading classes are very uncommon. We know many teachers who conduct weekly 15-minute in-class SSR sessions because they do not want to limit in-class teaching time. As a result, it was unrealistic and of limited ecological validity to conduct any more than 15 minutes of SSR.

## Conclusions

We claim that simply encouraging and requiring students to read one graded reader per week (Nation & Wang, 1999), does not suffice to facilitate large reading amounts, improve RSE, and encourage students to read beyond requirements. Thus, we recommend setting weekly word targets, not by number of books, but by word count. Furthermore, if we require weekly word-targets and not book number targets, we can increase the probability that low proficiency readers will read more than around 1,000 words a week, which is the number of words commonly found in books at the appropriate level for low L2 reading proficiency students.

To follow our recommendation, teachers may need to set weekly word-targets in line with appropriate homework loads. For example, if an hour of homework is expected of pupils per week for a class, and a teacher sets only ER as homework, we might multiply a pupil's mean reading speed per minute of the class by 60 to establish a weekly word-target. We set a weekly reading target of 2,500 words a week, which was roughly double the mean of the weekly reading amount during the first semester. However, all participants interviewed stated that they could read more than 2,500 words an hour, provided the material was at an appropriate level. Second, we found SSR to be of limited use of in-class time, for students with no intrinsic motivation used SSR to complete and not supplement reading goals.

Finally, without using limited class time, we set word-targets at 2,500 words per week. These word-target participants generally doubled their reading amount compared to when they were required to read one book per week. These readers also tended to read more than required; they increased their RSE, and that in reading longer materials, they found enjoyment. We submit, therefore, that relational evidence supports these causal claims. We therefore suggest that for non-English majors in institutional settings, we do ER with weekly word-targets while monitoring reading comprehension. Though we may verbally encourage students to read and

give them some time for SSR, we claim that weekly word-targets will more effectively motivate students to read more and develop reading self-efficacy.

### Acknowledgments

We would like to thank the anonymous reviewers for their constructive comments. Many thanks to Michael Carrol and Myles Grogan for assistance when planning and conducting this research. This work was supported by JSPS KAKENHI (grant number JP 16K16890).

### References

- Anderson, N. J. (2008). *Practical English language teaching: Reading*. New York: McGraw-Hill.
- Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavioral change. *Psychological Review*, *84*, 191–215. doi:10.1016/0146-6402(78)90009-7
- 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
- Burrows, L. (2013). *The effects of extensive reading and reading strategies on reading self-efficacy* [e-book]. US: ProQuest Information & Learning; 2013. Retrieved from: PsycINFO, Ipswich, MA.
- Carver, R. P. (1978). The case against statistical significance testing. *Harvard Educational Review*, *48*, 278–399.
- Carver, R. P. (1993). The case against statistical significance testing, revisited. *Journal of Experimental Education*, *61*, 287–292.
- Day, R. R., & Bamford, J. (1998). *Extensive reading in the second language classroom*. Cambridge: Cambridge University Press.
- Deci, E. L., & Ryan, R. M. (1985). *Intrinsic motivation and self-determination in human behavior*. New York: Plenum.
- Dörnyei, Z. (1990). Conceptualizing motivation in foreign-language learning. *Language Learning*, *40*, 45–78. doi:10.1111/j.1467-1770.1990.tb00954.x
- Extensive Reading Foundation Graded Reader Scale. (2017). Retrieved from <http://erfoundation.org/wordpress/graded-readers/erf-graded-reader-scale/>
- Gardner, R. C., & MacIntyre, P. D. (1993). A student's contributions to second language learning. Part II: Affective variables. *Language Teaching*, *26*, 218–233. doi:10.1017/S0261444800000045
- Grabe, W. (2009). *Reading in a second language: Moving from theory to practice*. New York, NY: Cambridge University Press.
- Guthrie, J. T., Wigfield, A., Metsala, J. L., & Cox, K. E. (1999). Motivational and cognitive predictors of text comprehension and reading amount. *Scientific Studies of Reading*, *3*, 231–256. doi:10.1207/s1532799xssr0303\_3

- Hojat, M., & Xu, G. (2004). A visitor's guide to effect sizes: Statistical significance versus practical (clinical) importance of research findings. *Advanced Health Science Education Theory Practice, 9*, 241–249. doi:10.1023/B:AHSE.0000038173.00909.f6
- Judge, T. A., Erez, A., & Bono, J. A. (1998). The power of being positive: The relation between positive self-concept and job performance. *Human Performance, 11*, 167–187. doi:10.1207/s15327043hup1102&3\_4
- Komiyama, R. (2013). Factors underlying second language reading motivation of adult EAP students. *Reading in a Foreign Language, 25*, 149–169.
- Krashen, S. (2015). The end of motivation. *New Routes, 55*, 34–35.
- Krashen, S. D. (2011). *Free voluntary reading* (1st ed). Santa Barbara, CA: Libraries Unlimited.
- Linacre, J. M. (2007). *A user's guide to WINSTEPS-MINISTEP: Rasch-model computer programs*. Chicago, IL: winsteps.com.
- MacIntyre, P. D. (1995). How does anxiety affect second language learning? A reply to Sparks and Ganshow. *The Modern Language Journal, 79*, 90–99. doi:10.2307/329395
- Matsui, T., & Noro, T. (2010). The effects of 10-minute sustained silent reading on junior high school EFL learners' reading fluency and motivation. *ARELE: Annual Review of English Language Education in Japan, 21*, 71–80. doi:10.20581/arele.21.0\_71
- McCracken, R. A. (1971). Initiating sustained silent reading. *Journal of Reading, 14*, 521–583.
- McLean, S. (2014). Evaluation of the cognitive and affective advantages of the Foundations Reading Library series. *Journal of Extensive Reading, 2*, 1–14.
- Mori, S. (2002). *The relationship between motivation and the amount of out-of-class reading* (Doctoral dissertation). Retrieved from ProQuest Dissertations and Theses database. (UMI No. 3040345).
- MReader [Computer software]. Available at <<http://mreader.org/>>.
- Nation, P. (2013). *What should every ESL teacher know?* Seoul, Korea: Compass Publishing.
- Nation, I. S. P., & Wang, K. (1999). Graded readers and vocabulary. *Reading in a Foreign Language, 12*, 355–380.
- Nishizawa, H., Yoshioka, T., & Fukada, M. (2010). The impact of a 4-year extensive reading program. In A. M. Stoke (Ed.), *JALT2009 Conference Proceedings* (pp. 632–640). Tokyo, Japan: JALT.
- Nuttall, C. (2005). *Teaching reading skills in a foreign language* (2nd ed.). Oxford: Heinemann Educational Books.
- Oettingen, G., & Zosuls, K. M. (2006). Self-efficacy of adolescents across cultures. In F. Pajares & T. Urdan (Eds.), *Self-efficacy beliefs of adolescents* (Vol. 5, pp. 245–265). Greenwich, CT: Information Age.
- Pilgreen, J. L. (2000). *The SSR handbook: How to organize and manage a sustained silent reading program* (1st ed.). Portsmouth, NH: Heinemann.
- Plonsky, L., & Oswald, F. L. (2014). How big is “big”? Interpreting effect sizes in L2 research. *Language Learning, 64*, 878–912. doi:10.1111/lang.12079
- Rosenthal, J. A. (1996). Qualitative descriptors of strength of association and effect size. *Journal of Social Service Research, 21*, 37–59. doi:10.1300/J079v21n04\_02

- Saito, Y., & Samimy, K. (1996). Foreign language anxiety and language performance: A study of learner anxiety in beginning, intermediate, and advanced level college students of Japanese. *Foreign Language Annals*, 29, 239–251.
- Schmidt, R., Boraie, D., & Kassabgy, O. (1996). Foreign language motivation: Internal structure and external connections. In R. Oxford (Ed.), *Language learning motivation: Pathways to the new century* (Technical Report #11, pp. 9–70). Honolulu: University of Hawai'i, Second Language Teaching & Curriculum Center. doi:10.1111/j.1944-9720.1996.tb02330.x
- Siyanova-Chanturia, A., & Webb, S. (2016). Teaching vocabulary in the EFL context. In W. A. Renandya, & P. H. Widodo (Eds.), *English language teaching today: Linking theory and practice* (pp. 227–239). (n.d.): Springer International Publishing.
- Takase, A. (2007). Japanese high school students' motivation for extensive reading. *Reading in a Foreign Language*, 19, 1–18.
- Takase, A., & Otsuki, K. (2012). New challenges to motivate remedial EFL students to read extensively. *Journal of Applied Language Studies*, 6(2), 75–94.
- Tremblay, P. F., & Gardner, R. C. (1995). Expanding the motivation construct in language learning. *The Modern Language Journal*, 79, 505–518. doi:10.1111/j.15404781.1995.tb05451.x
- Wang, J. H., & Guthrie, J. T. (2004). Modeling the effects of intrinsic motivation, extrinsic motivation, amount of reading, and past reading achievement on text comprehension between U.S. and Chinese students. *Reading Research Quarterly*, 39, 162–186. doi:10.1598/RRQ.39.2.2
- Wood, R. E., & Bandura, A. (1989). Impact of conceptions of ability on self-regulatory mechanisms and complex decision-making. *Journal of Personality and Social Psychology*, 56, 407–415. doi:10.1037/0022-3514.56.3.407

### About the Authors

Stuart McLean is an instructor at Osaka Jogakuin University. He holds an M.S.Ed. (TESOL), a Ph.D. in Forensic Medicine, and a P.G.C.E. He is currently an Applied Linguistics doctoral student at Kansai University. He has published in *Reading in a Foreign Language*, *Vocabulary Learning and Instruction*, *Language Teaching Research*, *TESOL Quarterly*, *System*, *Applied Linguistics*, and *Language Assessment Quarterly*. E-mail: stuart93@me.com

Joseph Poulshock is a professor of English linguistics in the Faculty of Economics at Senshu University. He holds an M.A. in TESOL from Biola University and a Ph.D. in Linguistics from the University of Edinburgh. He has published in *Reading in a Foreign Language*, *Zygon: The Journal of Religion and Science*, and *The Journal of Critical Realism*. E-mail: josephpoulshock@mac.com