A University-based
Extensive Reading Program Online

Kevin Ryan

Abstract

Extensive Reading (ER), or reading a large amount of easily understood materials for pleasure, has been used as an effective tool for acquiring all aspects of a second language for at least 20 years. One of the world leaders in applying this tool to university programs is Japan. This study looks at the adaptation of paper-based materials to an online environment, with attendant benefits and drawbacks in a university-based ER program. It intends to extend research on Xreading, an online software program that delivers a library of graded readers to students. Using data from this program, a clear picture of student use of ER is shown. No significant correlations with standard measures of language proficiency (GTEC or TOIEC) were found. Reasons for this lack of significance are discussed.

Introduction

Extensive Reading (ER) has been repeatedly shown to be an effective way of increasing language proficiency (Mason & Krashen, 1997; Krashen & Mason, 2015; Nakanishi, 2015; Jeon & Day, 2016; Day, 2015), including benefits to all four skills, both abroad and here in Japan (Susser & Robb, 1990; Ramonda, 2017). To clarify, according to Waring & McLean (2015), there are four main components of an ER program, and an additional 11 that are optional (See Table 1).

As ER was developing, students were first asked to read relatively simple books designed for children or teens until research showed that adult content with relatively few headwords (discrete vocabulary items) proved more appealing. Publishers have developed thousands of graded readers of this type, adapting and controlling for language difficulty while maintaining the informational content of the original book, or by simply writing content at a certain level of linguistic difficulty. Programs would provide students these books through a delivery system that might include a library or direct classroom distribution with the teacher selecting possible readers for students and allowing them to make the final choice.

Monitoring of student progress began with direct observation of students as they read in class. In some cases, students were asked to write a short report of the content of the material when they had finished reading. In other cases, students could take a short comprehension quiz to prove they had read and understood the material.

Online resources to support ER came into the mainstream with Moodle Reader, a set of quizzes for ER graded readers delivered in the Learning Management System (LMS) Moodle. Japan has always been a leader in the use of ER at universities, and this increased with the

Table 1. The core versus variable dimensions of ER

Core elements (necessary to retain the label 'ER')

      Fluent, sustained comprehension of text as meaning-focused input

      Large volume of material

      Reading over extended periods of time

      Texts are longer, requiring comprehension at the discourse level

Variable elements of an ER program (There are, no doubt, other variables: the below serve as examples.)

      ER is conducted in class or at home, or a combination thereof

      ER is required, or optional

      The reading is enjoyable, for pleasure, or not

      The reading is monitored (self declared, by the teacher), or not

      The reading is assessed, or not

      The presence or absence of follow-up activities (comprehension or language focus)

      The teacher reads or doesn't read with students in the classroom

      Graded or non-graded materials (provided they can be comprehended fluently)

      Longer or shorter texts

      The degree of freedom to select texts

      Requiring students to start with the simplest material available

(Waring & McLean, 2015, p.165)

development of Moodle Reader by Thomas (Tom) Robb at Kyoto Sangyo University leading the way (Robb & Kano, 2013). Noting the restrictions of using the quizzes within an LMS, Robb developed a web-based version of the approximately 3,000 quizzes for graded readers published by Oxford, Cambridge and many others. This new version MReader is now undergoing a revision of the bank of questions, using data from previous students and Rasch Analyses to improve the reliability and validity of each question (personal communication).

In 2015, Xreading was developed by Paul Goldberg as an online library of graded readers licensed from the publishers. This allowed students to access ER materials without having to go to the library. In addition, this allowed for teachers and administrators to both guide students through assignments and monitor progress on a much closer level. Xreading measures books and assignments in units of words instead of pages, as one example.

Milliner (2017) monitored ER use with Xreading for students at a Japanese university as part of a regular English program. He was unable to find a significant correlation between ER use and TOEIC scores. Milliner did not attempt to use any other data than "Words Read" to look at student ER use.

The calculation for the relationship between changes in student's overall TOEIC and TOEIC reading score and the volume of words read resulted in no relationship being observed ($r=-0.17815$, $n=19$, $p=.46559$). Similarly, no relationship was observed between reading volume and TOEIC reading score changes ($r=0.07567$, $n=19$, $p=.75817$). This result led the

author to conclude that students' ER may not have influenced the increase in students' TOEIC scores. (Milliner, 2017, p.55)

Research Focus

This study uses Milliner (2017) as a point of departure. With only 19 participants, significance was difficult for his study to achieve. Replicating his study with a larger N-size, more measures, and looking at more data provided by Xreading were the primary motivations to undertake this research.

Research questions:

1. Does Extensive Reading delivered online through Xreading software to students in a university English program correlate with increased TOEIC scores?

2. Are there any other factors that may enhance or detract from correlation with a higher TOEIC score?

These other factors may include reading speed, number of books started but not finished, level of the graded readers chosen, or scores on comprehension quizzes.

Participants

Subjects attend a prominent women's university in Tokyo as majors in the English department. The Extensive Reading (ER) program includes all first-year students (2018: n＝232) as part of a required reading class with a suggested 20% of the grade allotted to ER.

Students were introduced to Xreading software in the third week of their first semester. Thereafter weekly assignments were made until the end of the second semester. The author managed the Xreading software implementation. Weekly updates of how student goals were met (assignments completed) were distributed to faculty conducting Reading class.

During the Spring semester (April-July 2018) students were assigned a word count each week, depending on their level. The top 3 groups (Band 1) had a minimum of 6,000 weekly words to read, the medium groups (Band 2) read 5,000 words per week, while the lower groups (Band 3) were assigned 4,000 words per week. Settings for Xreading allowed a minimum of 60% (3/5 questions) on a simple comprehension quiz. Students were required to read at least 90% of a book before the quiz was available. A maximum read rate of 300 words/minute was set to prevent students from flipping through the book and attempting the quiz without reading (Allen-Tamai et al., 2018; Tagane et al., 2018). If students met all three criteria (read 90%, at less than 300 wpm, and get 60% on the quiz), their book was added to their Words Read total each week. During the 14 weeks available in the semester reading, students were required to meet their Assignment goals at least 12 of those weeks to get credit for the ER portion of their Reading class grade.

At the end of the Spring semester, students were given a choice to complete 16 more weeks of Xreading assignments during the 24 weeks available (8 weeks during the summer, 16 weeks

during Fall second semester). We used a weekly goal because after a monthly goal in the Fall semester of the previous year lead to a general loss of interest and a surge of reading at the end of each month.

Moreover, to maintain habitual reading and increase participation during summer break, students were given the option of a bonus doubling of their scores during the summer if they logged in and used Xreading at least 40 out of the 56 days (8 weeks). Fifty-four students availed themselves of this opportunity and thus did not have to complete any assignments during the second semester (October-January).

Measurement

The Extensive Reading software Xreading is a web-based application available to programs in Japan and a few other countries, depending on licensing arrangements with publishers. During the academic 2018 year Xreading served a library of approximately 1,000 graded readers at levels 1-14 to students on any internet-connected device. While Xreading serves content to the student, it also monitors important factors for feedback and research. Student data includes: 1) Books Added (checked out), 2) Books Completed (added to), 3) Words Read, 4) Read Time and Read Speed, and 5) Quiz data. Individual book data is available within the program to individual students, at a class level to instructors, and at a program level. Data is also available to be downloaded (.csv format).

Students were given a GTEC English test for placement during orientation in April 2018. The test was used to track students into 9 groups. Data was collected from 3 subsequent TOEIC Institutional Program (IP) tests; in May and July 2018, and January of 2019. Based on the July TOEIC test, students were re-tracked into 9 groups for the Fall semester beginning in October 2018.

Results

Over the academic year (April-Feb) students (n＝232) read 34,814,292 words in 22,362 books, taking 6,172 hours at an average reading speed of 113 words per minute. Students took 20,509 quizzes and passed 18,131 of them with an average score of 85.63%.

Table 2 shows a sample of 6 books with the data that Xreading reports on each book. Most of these items are self-explanatory. Two items that may be unclear are 1) Audio: some of the books come with audio of the text (some human, some computer), 2) Levels: Xreading has a 14-level scale for difficulty. See the Xreading website to see how this correlates with other scales.

With this data, we are able to get a picture of the student body and the program. During the introductory session, students were advised to start at Level 4, but were given tools to self-assess as to the difficulty so they could choose appropriately. In Table 2 we can see one student finished a Level 2 book (*The New Teacher*) in under 2 minutes, while another took 66 minutes to finish *Frankenstein*, presumably because it was much longer (7,001 words) and a higher level (8).

Table 2. Sample of 6 graded readers and their data provided by Xreading software

Books	*The New Teacher*	*Gulliver's Travels*	*Love Online*	*Alice's Adventures in Wonderland*	*Dan and the Missing Dogs*	*Frankenstein*
Publisher	MPI	Compass	Cengage	Oxford	Helbling	Macmillan
No. of Words	215	1241	2965	6408	6063	7001
Audio (h:m)	3	13		105	59	52
Xreading Level	2	7	4	6	5	8
Words Read	215	1241	2965	3688	6063	7001
Percent Read	100	100	100	57.55	100	100
Read Time (h:m:s)	0:01:43	0:11:33	0:21:36	0:32:41	1:01:35	1:06:19
Read Speed (Word/Min)	125.2	107.4	137.3	112.8	98.5	105.6
Listen Time (h:m:s)	0:00:00	0:00:00	0:00:00	0:03:04	0:40:03	0:56:20
Quiz Grade	60	100	80	100	80	80
Book Rating	2	3	2	3	3	2
Date (Check Out)	2018-08-12 22:28:29 JST	2018-08-12 22:11:50 JST	2018-08-12 21:16:07 JST	2018-06-06 15:04:43 JST	2018-05-31 12:03:58 JST	2018-05-26 15:50:24 JST
Date (Return)	2018-08-12 22:31:42 JST	2018-08-12 22:27:13 JST	2018-08-12 22:08:07 JST	2018-06-20 16:50:36 JST	2018-06-02 23:23:29 JST	2018-05-26 17:48:53 JST

Below in Table 3, we can see the breakdown of the level choices the students made throughout the year.

Table 3. Distribution of graded reader levels selected by students. Level 1 is the most basic

Book Level	13	12	11	10	9	8	7	6	5	4	3	2	1
Books Read	79	39	146	114	394	176	947	1567	5531	3315	4851	3156	2045
Percent	0.4	0.2	0.7	0.5	1.8	0.8	4.2	7	24.7	14.8	21.7	14.1	9.1

At the beginning of the study, students were tracked into 9 groups, and given slightly different goals for the semester (as noted above). When we look at the averages of Words Read and Books Completed for each of the 3 Bands (Upper, Medium, Lower) we discover that even though Band 3 were required to read slightly fewer words each week in the spring, the totals are relatively consistent across Bands. We also learn that Band 3 (Lower) worked harder on ER with more words read and a book level higher than Band 2 (See Table 4 below).

Table 4. Comparison of 3 Groups (Bands) on book difficulty, books read, and words read
Band Comparison (Upper, Medium, Lower)

	Avg. Book Level			Books Completed			Words Read		
	1	2	3	1	2	3	1	2	3
Valid N	76	78	76	78	78	76	76	78	76
Mean	4.463	3.850	3.903	69.19	88.95	107.7	152563	146798	154858
Std. Deviation	1.133	0.9877	1.001	47.51	42.71	58.79	70343	44450	59278
Skewness	0.365	1.201	−0.01	0.55	0.676	1.103	0.0397	−0.443	0.565
Kurtosis	−0.02	3.54	−0.64	−0.47	0.82	1.57	0.65	0.16	1.21
Minimum	2.100	1.900	2.000	0.000	9.000	4.000	6239	3014	2408
Maximum	7.500	8.200	6.100	200.0	226.0	299.0	374712	252287	342363

The only troubling measure is the Average Book Level for Band 2. The Kurtosis of 3.54 would indicate that almost all of Band 2 maintained the same level book throughout the year, and did not choose books of a higher level later in the year.

Milliner (2017) was unable to show any correlation between Xreading use (number of words read) and TOEIC scores, either for Total score or for Reading score. This reflects a similar experience with O'Neill using regular graded readers in 2012, and who also was unable to find any correlation with either overall or reading TOIEC scores.

Even with a larger sample size, this was also the case for the group of students in this study. As shown in Table 5, no overall correlation was found between any group or band, and any of

Table 5. Correlation of reading quantity (Words Read) and standardized test results. Change in TOEIC score was January (final) – May (initial) and did not include July.
ER reading quantity and Standardized Tests

		Spearman's rho	p
Words Read	- GTEC April	0.010	0.875
Words Read	- TOEIC Reading May	0.032	0.633
Words Read	- TOEIC Total May	0.018	0.786
Words Read	- TOEIC Reading July	−0.005	0.943
Words Read	- TOEIC Total July	−0.011	0.868
Words Read	- TOEIC Reading January	0.150	0.028
Words Read	- TOEIC Total January	0.099	0.148
Words Read	- Change TOEIC Reading	0.106	0.123
Words Read	- Change TOEIC Total	0.089	0.197

the standardized measures. Correlations for reading time, reading speed, level selection, reading time, or quiz grades also failed to achieve significance.

Even if we look at the group of motivated self-selected students (N＝54/232) who studied during the summer at least 40 out of the 56 days, no significant correlation can be made to any standard measure.

Discussion

Krashen & Mason (2015, p.8) found a significant correlation between TOEIC scores and ER for 6 adults reading large amounts of self-selected materials. Their conclusion, "Our results, and those of Nation, suggest that a second language reader can move from a low TOEIC score (250) to a very high score (950) with three years of self-selected reading, averaging about an hour a day." While this may seem amazing, we need to take into account that these were self-motivated adults reading large amounts of materials over periods of at least a year.

It may be overoptimistic to look for an influence by Extensive Reading on overall scores in a language program that includes at least 20 contact hours a week over 30 of the 40 weeks studied. While Xreading affords a wonderful matrix of measurements, trying to correlate reading to a standard test like TOEIC would be unlikely to show any significance unless a true experimental situation could be implemented. Significance in ER studies are often achieved with vocabulary measures or other specifically related tests (Horst, 2005).

An area of concern is the nature of TOEIC, a test of English related to business, as it relates to the general nature of ER. The administration of the IP tests may have been too frequent, leading to testing fatigue among students. This is notable in the July administration, done at the end of the semester on a very warm day, saw a drop in the t-test of reading scores compared with the other two TOEIC administrations in May and January.

Another observation is that as part of a Reading class, with a grade assigned, this is not a true ER administration. As Stoller (2015, p.157) notes, "This phenomenon suggests that we should establish one goal for the class initially, but students should know that upon reaching that goal, the teacher will work with them individually to set more ambitious goals." Students rarely went above and beyond the minimums set for words read. It was treated as another homework assignment.

Moving forward, this study indicates that the quantity of ER was lacking. Many other programs in Japan and abroad require much more reading. Koby (2016, p.396) noted, "80% of overall student assessment is based on reading volume; the remaining 20% is based on in-class activities." With this system, a passing grade of 50% is assigned to a volume of 415,000 words over four semesters, and over 1.1 million words for a 90% grade. Milliner's (2017) subjects read approximately 50% more than the subjects of this study (250,000 words to 150,000).

Xreading has afforded students a tremendous amount of freedom. More than 95% of students chose to read on their smartphones (192/216 from a final questionnaire). The clear nature of the

assignments also helped (Van Amelsvoort et al., 2018). The university has a large library of paper graded readers, but when given the option in 2016 to use paper books by checking them out at the library, only three students availed themselves of the option. Nonetheless, Xreading materials tend to focus more on the lower levels, and as such did not offer the best experience for our most accomplished students. This reflects a general trend in graded readers. "One possibility is that the wide-spread emphasis on creating interest in ER for lower-level students may have resulted in an accidental neglect of higher-level students." (Karlin & Romanko, 2010, p.195)

A plan to reinstate the paper-based ER monitored with MReader next year for the upper level groups and require a mix is currently being considered. Running two systems at the same time has its drawbacks, though, as noted by Koby (2017), there is often confusion about the different quiz systems.

But by far most important is the interaction with teachers (Ro, 2016). Facilitating teachers with information and feedback on student progress is not enough (Susser & Robb, 1990; Milliner & Cote, 2015). "Students can be reluctant to believe in the benefits of extensive reading, preferring instead to study directly for a test. The focus on reading for pleasure and the lack of follow up quizzes can lead students from high stakes testing cultures to disregard the benefits." (Quinn, 2018, p.121). Imparting the idea of ER as a pleasurable activity, developing reading as a habit means working with students, using class time on a regular basis for Sustained Silent Reading (SSR), and interacting with students to discuss the books they are reading, are all essential to the success of the program.

Colophon

Data collected using Xreading software. CSV files from Xreading concatenated into one spreadsheet using Google Sheets. Statistics were done using JASP.

References

Allen-Tamai M., Bruce, J., Bulach, J., Brooke, S., Butler, B., Dias, J., & Strong, G. (2018). Researching Extensive Reading and an online library. *Extensive Reading World Congress Proceedings* 4, pp.294–302.

Day, R. (2015). Extending Extensive Reading. *Reading in a Foreign Language* 27(2), pp.294–301.

Hagley, E. (2017). Extensive graded reading with engineering students: Effects and outcomes. *Reading in a Foreign Language* 29(2), pp.203–217.

Horst, M. (2005). Learning L2 vocabulary through extensive reading: A measurement study. *The Canadian Modern Language Review ∕ La revue canadienne des langues vivantes* 61(3), pp.355–382. Retrieved February 22, 2019 from Project MUSE database.

Jeon, E. & Day, R. (2016). The effectiveness of ER on reading proficiency: A meta-analysis. *Reading in a Foreign Language* 28(2), pp.246–265.

Karlin, O. & Romanko, R. (2010). Examining multiple variables within a single ER setting. *The Reading Matrix* 10(2), pp.181–204.

Kjeldgaard, M. (2015). Combining skills-based instruction and Extensive Reading: Using online resources

to support a reading curriculum. 「言語と文化」(Language and Culture), 32, pp.47–58.

Koby, C. (2016). The anatomy of an Extensive Reading syllabus. In P. Clements, A. Krause, & H. Brown (Eds.), *Transformation in language education.* Tokyo: JALT, pp.392–398.

Koby, C. (2017). Using a learning management system to enhance an Extensive Reading program. In K. Borthwick, L. Bradley & S. Thouësny (Eds.), *CALL in a climate of change: adapting to turbulent global conditions – short papers from EUROCALL 2017*, pp.189–193. Research-publishing.net. (https://doi.org/ 10.14705/rpnet.2017.eurocall2017.711)

Krashen, S. & Mason, B. (2015). Can second language acquirers reach high levels of proficiency through self-selected reading? An attempt to confirm Nation's (2014) results. *The International Journal of Foreign Language Teaching* 10(2), pp.1–9.

Macalister, J. (2015). Guidelines or commandments? Reconsidering core principles in Extensive Reading. *Reading in a Foreign Language* 27(1), pp.122–128.

Mason, B. & Krashen, S. (1997). Extensive Reading in English as a foreign language. *System* 25(1), pp.91–102.

Milliner, B. (2017). One year of Extensive Reading on smartphones: A report. *JALT CALL Journal* 13(1), pp.49–58.

Milliner, B. & Cote, T. (2015). One year of extensive reading on mobile devices: Engagement and impressions. In *Critical CALL – Proceedings of the 2015 EUROCALL Conference*, Padova, Italy. Research-publishing-net. pp.404–410.

Nakanishi, T. (2015). A meta-analysis of extensive reading research. *TESOL Quarterly* 49(1), pp.6–37.

O'Neill, B. (2012). Investigating the effects of extensive reading on TOEIC® reading section scores. *Extensive Reading World Congress Proceedings* 1, pp.30–33.

Quinn, K. (2018). Using evaluation to motivate students in an Extensive Reading program. *Journal of ELT Research* 3(2), pp. 120–141. (DOI: 10.22236/JER_Vol3Issue2)

Ramonda, K. (2017). The effects of implementing online extensive reading in the English classroom. *Tokyo University of Science Research Bulletin* 49, pp.291–305.

Ro, E. (2016). Exploring teachers' practices and students' perceptions of the extensive reading approach in EAP reading classes. *Journal of English for Academic Purposes* 22, pp.32–41.

Robb, T. & Kano, M. (2013). Effective extensive reading outside the classroom: A large-scale experiment. *Reading in a Foreign Language* 25(2), pp.234–247.

Stoller, F. (2015). Viewing Extensive Reading from different vantage points. *Reading in a Foreign Langauge* 27(1), pp. 152–159.

Susser, B. & Robb, T. (1990). EFL extensive reading instruction: Research and procedure. *JALT Journal* 12(2), pp. 161–185.

Tagane, Y., Naganuma, N. & Dougherety, P. (2018). Academic dishonesty in Extensive Reading programs: Stories and strategies from student interviews. *The Language Teacher* 42(1), pp.9–12.

Van Amelsvoort, M., Isozaki, A. & Yoshino, Y. (2018). Program features to nudge students to read more. *Extensive Reading World Congress Proceedings* 4, pp.55–63.

Waring, R. & McLean, S. (2015). Exploration of core and variable dimensions of extensive reading research and pedagogy. *Reading in a Foreign Language* 27(1), pp.160–167.

（ケビン・ライアン　英語コミュニケーション学科）