論　　文

Student Engagement
in an Extensive Reading Activity:

the Effect of Book Selection Methods on Student's Interest, Enjoyment, and Success

沖　キャサリン

同志社女子大学・学芸学部・国際教養学科・助教（有期）

OKI Catherine

Department of International Studies, Faculty of Liberal Arts,
Doshisha Women's College of Liberal Arts, Assistant professor (contract)

Abstract

Extensive Reading (ER) has several recommended guidelines (e.g., Day & Bamford, 1997), including students choosing books by themselves, according to their level and interest. This study sought to understand how selection as individuals versus small groups could affect student engagement in ER activities, looking at the emotional engagement components of interest and enjoyment, and reported success at three stages of book selection, reading, and small group discussion. To explore whether there was an effect due to the way the books were selected, paired-sample t-tests were used on student self-report surveys, revealing significant differences when books were selected individually at the discussion stage for interest and enjoyment and at the reading stage for success.

While these results are inconclusive due to the small sample size, the study can serve as a model for future studies with a larger sample.

Extensive Reading (ER) was first coined by Palmer (in Bamford and Day, 1997), creating a dichotomy between Intensive Reading (IR), which entails detailed reading of texts above one's level, for the purpose of "complete and detailed understanding", and ER, which refers to reading large amounts of text with the intention of understanding its meaning generally, rather than comprehending each and every word or phrase. Graded Readers, foreign language texts simplified to meet learners' levels (Warring, 1997), are often used for ER, unlike in IR where shorter, complex text written in the foreign language are often used.

ER is also further defined according to its timing, with Robb and Kano (2013) making the distinction between reading in class as *replacement* for other activities or reading done outside of class in

additive ER. They argue that additive is the best way to expose students to increased amounts of text because time for reading in class is limited with teachers already having enough to cover, and they argue for the distinction to be made when reporting research results, though not all do.

Regardless, previous research has found that ER contributes to language learning by showing a rise in reading rates (e.g., Bell, 2001; Tanaka & Stapleton, 2007; Beglar et al., 2012), vocabulary (e.g., Lao & Krashen, 2000; Poulshock, 2010), reading comprehension (e.g., Tanaka & Stapleton, 2007; Yamashita, 2008), enjoyment (Doyle & Parrish, 2012), and motivation (e.g., Nation, 1997; Grabe, 2009). It is thus unsurprising that ER has been implemented in many English language programs across Japan and that implementation is its own area of ER study.

In successfully implementing ER programs, ten characteristics were found in ER programs, including that learners should choose books at a level appropriate to their current abilities and have the freedom to stop reading if the selected book is uninteresting or of an inappropriate level (Bamford & Day, 1997, Day & Bamford, 1998). It is thought that books must be chosen by students because they know "what they can and can't cope with," and for book selection self- or whole-class are thought viable options (The Extensive Reading Foundation Guide, n.d.). Bassett (2008) says, "Students are more motivated to read something if they have chosen it

themselves" (p. 18), and while Rosszell and Brown (2009) argue that individually chosen readers, group chosen readers, and class assigned readers have their own unique motivational benefits, under the individual reader approach they include that students are motivated because they can choose their own books. However, this study is a reflection based on the teachers' experiences of implementing ER, not an empirical look at ER.

A decade later, VanAmelsvoort (2017) compared implementation across two years at a private Japanese university, and believes student engagement in ER increased, as measured by the students reading more, when ER implementation included orienting students to the "requirements and expectations" (p. 100) needed to successfully participate in the ER program in 2016 compared to when they did not in 2015. However, the definition of engagement is left out and moreover, appears to be used interchangeably with motivation, a common occurrence that has led to efforts to distinguish these constructs in order to better understand research on them.

Skinner et al. (2009) see the *psychological processes* involved in learning as motivation but engagement as the *actions that result* from motivation, so in learning activities, engagement is how actively a student participates. It is the "energized result" of motivation (Wang & Degol, 2014, p. 138). Egbert (2003) frames engagement within the psychological state of flow and shows how task design supports student engagement, which

"leads to improved performance caused by repetition, motivation, exploration, satisfaction, more time on task, and willingness to risk" (p. 502) further leading to higher competence in language skills. Therefore, how ER tasks are implemented will likely affect student engagement, but the construct needs to be made clearer in order to do so.

Fredericks and McCloskey (2012) explain that engagement is a multidimensional construct which includes behavioral (task participation), emotional (negative or positive feelings about tasks), and cognitive (use of learning strategies and self-regulation) engagement, each made up of multiple components. For example, interest and enjoyment are components of emotional engagement. Additionally, engagement and student outcomes may also be seen as reciprocal in that greater engagement could lead to greater success and could thus lead to further engagement (Wang & Degol, 2014). Fredericks et al. (2004) explain that while enjoying and joining in may be the start of engagement, the result can be "commitment or investment and thus may be a key to diminish student apathy and enhancing learning" (p. 82). In terms of ER, it is only through students reading more that they will be exposed to vast amounts of texts associated with the gains that ER brings, so finding ways for emotional engagement and task success to increase in ER activities could facilitate students reading more as they move from simple participation in ER to a commitment to read.

Therefore, understanding what may heighten student levels of engagement by measuring their interest, enjoyment, and success in ER tasks, could provide for more evidence-based ER task implementation in order to increase chances that students will be exposed to more English by reading more. In light of the selection recommendations for implementing ER and in an effort to quantify the emotional components of interest and enjoyment while also looking at their reported success in these tasks, this study explored whether the method of selecting online graded readers, either done on their own or with a small group of peers, affected student engagement in an additive online ER activity.

Method

First, this study is an exploratory study seeking to gain better insight into this issue with a small, specific group of students and to identify where future studies on ER engagement in the Japanese L2 context could focus. Additionally, this design may serve as a model for similar lines of inquiry with larger samples that could be used to challenge or reinforce assumptions about how ER should be conducted and, more specifically, how books should be selected in ER programs at Japanese universities.

The twenty participants were all first-year students in their first semester, taught by the teacher-researcher in two classes in the Department of International Studies, though due to absences, three students were dropped from the study. At the start of the study in Spring, students' TOEFL iBT scores averaged 41 points but at the end of Fall semester averaged 53 points. Students had not

experienced ER previously, and this was the only class with an ER component.

Over three weeks, ER, the online graded reader website XReading (2021), and the study were introduced (Oki, 2018). Here, students established their graded reading levels with all falling at a level five or six. This was used to group them into three groups with three or four students each.

Next, students participated in six weekly cycles and each cycle consisted of an ER activity that started by selecting a book in class. Depending on the week, books were selected in small groups or as individuals. Table 1 shows the schedule for book selection method.

Table 1　*Book Selection Cycles*

Group-selection cycles	Individual-Selection cycles
Week 1	Week 3
Week 2	Week 5
Week 4	Week 6

In each cycle, after selecting a book, they read it for homework, and in the following class participated in discussions with their groups. A new discussion question was given to the students a week ahead to enable preparation and all discussions lasted ten minutes. In order to make the discussion questions relevant, the XReading assignments were set so students could only read works of fiction. For example, "How would you feel if something like this happened to you or someone you know?" could be used to discuss any fiction book where characters drive the story but would not necessarily work when discussing a piece of non-fiction, which may be about any number of topics unrelated to human experiences, like how a product is made.

After each stage of a cycle, book selection, reading, and discussion, a bilingual Google Forms survey was used to collect participants' self-reports to gauge their feelings about the engagement components of interest, enjoyment, and success toward the ER task. Students answered on a 6-point Likert scale with 6 being Strongly Agree and 1 being Strongly Disagree. Table 2 shows a selection of representative items from the questionnaire (upon request a full version can be made available).

Reported scores for individual

Table 2　*A Selection of Representative Items from the Questionnaire*

Stage	Selection	Reading	Discussion
Interest	I felt bored while looking for a good reader.	I felt bored while reading the story.	My group's discussion was interesting.
Enjoyment	I enjoyed the process of selecting this reader.	I enjoyed reading the story.	Overall, I enjoyed discussing the story.
Success	My preferences were an important part of the selection process.	I feel good that I could read a whole book in English.	I was an active participant in the discussion.

＊Note: scores for interest at the selection and reading stages were reverse-ordered for analysis

selection and group selection were combined into a composite score for each of the 3 stages, before being compared with paired sample *t*-tests. Alpha was set at .0167, based on the Bonferonni adjustment accounting for multiple comparisons. Cohen's *d* was calculated to provide an additional standardized measure for understanding the size of the effect of the two treatments (Cohen, 1988).

Results

The results of the analysis on students' reported engagement for the components of interest, enjoyment, and success in the two different book selection methods are reported below.

Interest

First, paired-sample *t*-tests were conducted to compare reported engagement through the component of student interest in the activity between individually- and group-selected books. Table 3 shows at the Selection Stage there was not a statistically significant difference in the means for individually selected (M=5.06, SD=.914) and group-selected (M=5.10, SD=1.07) conditions;

$t(16)=-.356$, $p=.727$. Likewise, at the Reading Stage a statistically significant difference in the means for individually selected (M=4.80, SD=.553) and group-selected (M=4.65, SD=.960) conditions was not found; $t(16)=.691$, $p=.500$. On the other hand, at the Discussion Stage, there appears a statistically significant difference in the means for individually selected (M=5.10, SD=.839) and group-selected (M=4.84, SD=.842) conditions; $t(16)=2.19$, $p=.043$. Additionally, at this stage there is a medium size effect ($d=.530$).

Enjoyment

Table 4 summarizes the results of the paired-sample *t*-tests conducted to compare reported engagement through the component of student enjoyment in the activity between individually- and group-selected books are presented. At the Selection Stage a statistically significant difference in the means for individually-selected (M=4.88, SD=1.11) and group-selected (M=4.90, SD=.052) conditions was not found; $t(16)=-.115$, $p=.910$. Likewise, at the Reading Stage for individually selected (M=4.75, SD=.618) and group-selected (M=4.55, SD=1.01)

Table 3 *Results of Interest Scores by Reader Selection Methods*

Interest	Individual		Group		*t*(16)	*p*	Cohen's *d*
	M	*SD*	*M*	*SD*			
Selection Stage	5.06	.914	5.10	1.07	-.356	.727	-.086
Reading Stage	4.80	.553	4.65	.960	.691	.500	.166
Discussion Stage	5.10	.836	4.84	.842	2.19	.043	.530

Table 4 *Results of Enjoyment Scores by Reader Selection Methods*

Enjoyment	Individual		Group		$t(16)$	p	Cohen's d
	M	SD	M	SD			
Selection Stage	4.88	1.11	4.90	.052	-.115	.910	-.029
Reading Stage	4.75	.618	4.55	1.01	.838	.415	.203
Discussion Stage	5.16	.825	4.86	.754	2.50	.023	.607

Table 5 *Results of Success Scores by Reader Selection Methods*

Success	Individual		Group		$t(16)$	p	Cohen's d
	M	SD	M	SD			
Selection Stage	5.20	.755	5.27	.795	.474	.642	.114
Reading Stage	5.02	.901	4.43	1.19	2.37	.031	.576
Discussion Stage	5.06	.818	4.90	.822	1.22	.240	.300

conditions there was not a statistically significant difference in the means; $t(16)=.838$, $p=.415$. On the other hand, at the Discussion Stage, a statistically significant difference appears in the means for individually selected (M=5.16, SD=.825) and group-selected (M=4.86, SD=.754) conditions; $t(16)=2.50$, $p=.023$. Additionally, at this stage there is a medium size effect ($d=.607$).

Success

　　Table 5 shows the findings concerning success. Paired-sample *t*-tests were conducted to compare students' reported success in the activity between individually- and group-selected books. At the Selection Stage there was not a statistically significant difference in the means for individually selected (M=5.20, SD=.755) and group-selected (M=5.27, SD=.792) conditions; $t(16)=-.474$, $p=.642$. However, at the Reading Stage there appears a statistically significant difference in the means for individually selected (M=5.02, SD=.901) and group-selected (M=4.43, SD=1.19) conditions; $t(16)=2.37$, $p=.031$. In addition, the effect size was medium ($d=.576$) On the other hand, at the Discussion Stage, there was no statistically significant difference in the means for individually selected (M=5.06, SD=.818) and group-selected (M=4.90, SD=.822) conditions; $t(16)=1.22$, $p=.240$.

　　The reported differences in interest, enjoyment, and success for all three stages were slight and showed no strong effects due to the changing conditions of how students selected the readers. Statistically significant differences were

shown in the discussion stage in favor of independently selected books over group-selected books for the components of interest and enjoyment, and again, at the reading stage for success, making it seem that individually selected books were at times more engaging than group-selected books.

Discussion

The majority of the mean differences between the two conditions and at various stages were so small that they were determined to be likely due to chance variation and not an effect due to the way the books were selected. While keeping in mind that these results do not show a strong effect towards either condition, the results supporting the individual-selection condition for interest and enjoyment in discussion and success at the reading stage are discussed, along with overall engagement.

Looking at the results about feelings of interest and enjoyment in discussing individually selected books showed statistically significant differences. Discussing books that their partners had not read means shared schema is lower, requiring everyone to communicate for understanding. Rosszell & Brown (2008) argue that individual selection of books "enables true exchanges of information" (p.10) as students overall want to learn about one another's books. Furthermore, they argue that enjoyment comes from being able to criticize their individual books and say that enjoyment may suffer when groups read the same book because opinions of the book may differ, and they may feel that criticism of the book

will be taken as a personal affront. Therefore, individual selection may increase interest and enjoyment for students during discussion.

Looking at the statistically significant difference in students' feelings of success after reading in favor of individually selected books, it might be that students feel more successful having read a whole book in English because they could choose a book better in line with their level and interest. As the guidelines for ER recommend, they could also stop reading and choose a different book in these weeks. Conversely, in group-selection cycles, they could not stop and change books; in order to participate in the weekly discussion, even if the reading level or interest level was inappropriate, they had to finish. However, in individual weeks, they could and in some cases did switch books. If this led to a better or smoother reading experience, it may have contributed to them feeling that they more successfully complete an English book and may lead to more reading.

Finally, throughout, overall engagement tended to be higher than an average of 3 out of 6 points on the Likert scale, and thus, these students seemed on average and across all three components to be more engaged in the activities of selecting, reading, and discussing graded readers than not. This indicates that ER could be worth continuing in the program, though consideration as to how it is implemented should be further explored.

Limitations and Future Research

Due to the exploratory nature of this

study, these results should not be seen as conclusive, but instead as an attempt to think more about the premise of how students should choose books when doing ER. Additionally, the small sample size (N=17) should caution against strong conclusions and widespread applications based on the results. Nonetheless, these findings can shed light on how a replication study with a larger sample could seek to understand more about how ER can be done and make evidence-based changes to its implementation. Thus, further research on a larger sample is needed.

Additionally, as this was conducted using a self-report survey method which was not anonymous and given by their teacher, it may be that students were not entirely honest in their reporting (Appleton et al., 2006). Future research could use means to make students' responses anonymous. Survey fatigue may have also set in as the cycles went on (Lavrakas, 2008) and in measuring emotional response, perhaps methods beyond self-report could be used. For example, observing psychophysiological reactions (like students' facial expressions or monitoring their pulse) or experience sampling methods (ESM) could serve as another approach to understanding affective components of engagement (Wang & Degol, 2014).

It is also possible that because non-fiction books were restricted, students who prefer non-fiction may have been less engaged. Likewise, students with weaker speaking skills may have felt burdened by the follow-up activity always being a discussion and might have

preferred a writing activity. Furthermore, Helgeson (2005) suggests that asking students to do the same post-reading activity week after week can lead to boredom and so, while discussion questions were changed weekly to keep the activity from feeling monotonous, it may be that discussion fatigue or boredom also set in. Therefore, changing the study to run over two semesters with a break between each cycle or changing the follow-up activity to encompass a variety of skills and so that all book genres could be included might mitigate these limitations.

Moreover, a qualitative element should be added to a future study. All discussion of these results was speculative, so asking students why they answered as they did would help to better understand the nuanced differences between their engagement in individual-versus group-selection. Having a better understanding of how students participate in each stage of the activity could further define these engagement components in the context of Japanese university students doing ER activities. For example, this could help teachers understand more about what students find interesting and enjoyable in discussions, be it expressing themselves, asking and/or answering questions, or any other reasons for reporting interest between the two conditions.

Conclusion

This study sought to understand more about student engagement in ER activities by conducting a quantitative study comparing individual versus group

book selection. To measure engagement, the emotional components of interest and enjoyment were measured along with student reports of task success, using student self-report surveys given at the end of the three stages of book selection, book reading, and book discussion. The results of the paired sample *t*-tests showed significant differences when books were selected individually for interest and enjoyment at the discussion stage and success at the reading stage. While this may indicate that individual selection leads to more engagement and a feeling of reading success, the small sample size cautions against making broad conclusions and instead the researcher would argue that a replication study with a larger sample size is necessary. That notwithstanding, students did report higher than average engagement, regardless of the selection method, throughout, possibly indicating that ER is generally an engaging activity.

References

Appleton, J. J., Christenson, S. L., Kim, D., & Reschly, A. L. (2006). Measuring cognitive and psychological engagement: Validation of the Student Engagement Instrument. *Journal of School Psychology, 44*, 427-445. doi:10.1016/j.jsp. 2006.04.002.

Bamford, J., & Day, R. R. (1997). Extensive Reading: What is it? Why Bother? *The Language Teacher, 21*(5). https://jalt-publications.org/old_tlt/files/97/may/extensive.html

Bassett, J. (2008). Stories: Feeding language learners for life. *The Language Teacher, 32*(7), 17-20.

Beglar, D., Hunt, A., & Kite, Y. (2012). The effect of pleasure reading on Japanese university EFL learners' reading rates. *Language Learning, 62*, 665-703.

Bell, T. (2001). Extensive reading: Speed and comprehension. *The Reading Matrix, 1*(1). http://www.readingmatrix.com/articles/bell/index.html

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. Hillsdale, NJ: L. Erlbaum Associates.

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

Doyle, H., & Parrish, M. (2012). Investigating students' ways to learn English outside of class: A researchers' narrative. *Studies in Self-Access Learning Journal, 3*(2), 196-203.

Egbert, J. (2003). The study of flow theory in the foreign language classroom. *The Modern Language Journal, 87*(iv), 499-518.

The Extensive Reading Foundation Guide. (n.d.) ［PDF］ https://erfoundation.org/guide/ERF_Guide.pdf

Fredericks, J., Blumenfeld, P., & Paris, A. (2004). School Engagement: Potential of the Concept, State of the Evidence. *Review of Educational Research, 74*(1), 59-109.

Fredricks J. A., & McColskey, W. (2012) The Measurement of Student Engagement: A Comparative Analysis of Various Methods and Student Self-report Instruments. In: Christenson S., Reschly A., Wylie C. (eds) *Handbook of Research on Student Engagement*. Springer. https://doi.org/10.1007/978-1-4614-2018-7_37

Grabe, W. (2009). *Reading in a second language: Moving from theory to practice*. Cambridge University Press.

Helgesen, M. (2005). Extensive reading reports －Different intelligences, different levels of processing. *Asian EFL Journal, 7*(3), 25-33.

Lao, C. Y., & Krashen, S. (2000). The impact of popular literature study on literacy development in EFL: More evidence for the power of reading. *System*, *28*, 261-270.

Lavrakas, P. J. (2008). *Encyclopedia of survey research methods*. Sage Publications, Inc. http://dx.doi.org/10.4135/9781412963947.n480

Nation, P. (1997). The language learning benefits of extensive reading. *The Language Teacher*, *21*(5). www.jalt-publications.org/tlt/files/97/may/benefits.html.

Oki, C. (2018). Setting up a quantitative study for a collaborative research project. *Extensive Reading in Japan*, *11*(1), 18-20.

Poulshock, J. (2010). Extensive graded reading in the liberal arts and sciences. *Reading in a Foreign Language*, *22*, 304-322.

Robb, T., & Kano, M. (2013). Effective extensive reading outside the classroom: A large-scale experiment. *Reading in a Foreign Language*, *25*(2), 234-47.

Rosszell, R., & Brown, D. (2009). Two contrasting approaches to extensive reading. In A. M. Stoke (Ed.), *JALT2008 conference proceedings*. JALT.

Skinner, E. A., Kindermann, T. A., Connell, J. P., & Wellborn, J. G. (2009). Engagement and disaffection as organizational constructs in the dynamics of motivational development. In K. Wentzel & A. Wigfield (Eds.), *Handbook of motivation at school* (pp. 223-245). Erlbaum.

Takase, A. (2002). Motivation to read English extensively. *Forum for Foreign Language Education*, *1*, 1-17.

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*)1, 115-126.

VanAmelsvoort, M. (2017). Extensive Reading Onboarding: Program Design for Icreasing Engagement. *Juntendo Journal of Global Studies*, *2*, 98-106.

Wang, M., & Degol, J. (2014). Staying Engaged; Knowledge and Research Needs in Student Engagement. *Child Development Perspectives* *8*(3), 137-143. https://doi.org/10.1111/cdep.12073

Waring, R. (1997). Graded and Extensive Reading -Questions and Answers. *The Language Teacher*, *21*(5). https://jalt-publications.org/old_tlt/files/97/may/waring.html

XReading. (2021). https://xreading.com/

Yamashita, J. (2008). Extensive reading and development of different aspects of L2 proficiency. *System*, *36*, 661-672.