Reading in a Foreign Language April 2021, Volume 33, No. 1
ISSN 1539-0578 pp. 78–102

Text Difficulty in Extensive Reading: Reading Comprehension and Reading Motivation

Ya-Han Yang
National Taiwan Normal University
Taiwan

Hsi-Chin Chu
National Taiwan Normal University
Taiwan

Wen-Ta Tseng
National Taiwan University of Science and Technology
Taiwan

Abstract

This study investigates the effects of the text difficulty of extensive reading materials on the reading comprehension and reading motivation of English as a foreign language (EFL) vocational high school students in Taiwan. Two experimental groups were assigned, on an individual basis, to read graded readers at either one level below ('*i*-1') or one level beyond ('*i*+1') their current level, while a control group followed their regular curriculum. The results showed that after treatment, the '*i*-1' group improved their overall comprehension and the subset of literal comprehension. They also outperformed the '*i*+1' group on the same measures. For reading motivation, the '*i*+1' group's overall motivation was promoted. Both groups enhanced their reading engagement, while only the '*i*-1' group inhibited reading avoidance. Moreover, the '*i*+1' group outperformed the '*i*-1' group in the perception of self-efficacy. Overall, the '*i*-1' level yielded better effects on reading comprehension; the '*i*+1' level, on reading motivation.

Keywords*:* extensive reading, foreign language reading, text difficulty, reading comprehension, reading motivation

Recently English as a second language (ESL) and English as a foreign language (EFL) graded readers designed for students' extensive reading have become popular choices for learners at beginning and intermediate levels. They are used in concert with or to supplement traditional textbooks that control vocabulary for a particular instructional level. A large body of research has thus assessed their utility in diverse types of learning, allowing meta-analysis studies to synthesize the overall and moderating effects of their use (Jeon & Day, 2016; Nakanishi, 2015). What these analyses have not examined, however, is how text difficulty might alter the extent of

readers' cognitive and affective development when engaging in extensive reading. This paper attempts to locate an optimal level of text difficulty in extensive reading (ER) in terms of two theoretical propositions: (a) the Input Hypothesis (proposed by Krashen, 1994a) and (b) the automaticity principle (proposed by Day & Bamford, 1998, and Samuels, 1994). These theories point to two contrastive orientations: (a) the Input Hypothesis regarding the process of acquiring language at a level beyond the student's current vocabulary capacity, and (b) the effects of reading below one's vocabulary level.

To explore the issue of text difficulty in ER, the present study varied the text difficulties by matching learners' levels with two ER text difficulty levels–one level below and one level above the students' current learner level–to assess students' development in reading comprehension and in reading motivation. The study was conducted in a summer ER program involving two groups of vocational high school students in Taiwan. A common practice in vocational high schools in Taiwan is to offer courses in the morning during the summer session that either review or preview lessons for the regular semester. Since no constraints on material selection and teaching approaches are imposed during the session, an alternative, devoting the entire class time to an ER program, was thus possible. Participants in the program were administered a vocabulary size test (Wan-a-rom, 2010) which yielded a range of vocabulary sizes between 500 and 2,600 words, based on which they were assigned to either the '*i*-1' or '*i*+1' group. Students in each group read self-selected, end-glossed EFL graded readers at a level designated by the researchers following the Graded Reader Level Scales (GRLS), devised to cross-refer books from various publishers. A third group, serving as a control group, received traditional textbook instruction centering on close reading and linguistic analysis. A three-group pretest-posttest experimental-control design was employed; the Test of English as Foreign Language (TOEFL) Junior Test was used to measure reading comprehension at two levels, literal and inferential, and a validated motivation scale was developed to assess reading motivation in three dimensions.

Literature Review

Theoretical Background

The Input Hypothesis (Krashen,1985, 1989, 1994a) proposed that language learners acquire language competence by processing comprehensible input with the lexicon and syntactic structures a little beyond the learner's current level of competence, that is, the '*i*+1' level. Comprehension of such text moves the learners' current competence in vocabulary, grammar, reading, and writing from the level of '*i*' to the next level of '*i*+1' (Krashen, 2004). Such attainment may occur because the small proportion of problematic parts, the extra 1 beyond the '*i*' level, is resolved by resorting to cues from the extended text context and by activating learners' world knowledge. In other words, this theory presumes that the readers' language knowledge will increase with repetition or elaboration of the ER content, and consequently a reader's language difficulty is more likely to be resolved.

On the other hand, Day and Bamford (1998) focused on the benefits of ER input in promoting reading automaticity and hypothesized that ER input at the '*i*-1' level, a level below the readers' current level, may allow readers a speedier sighting of and more frequent exposure to words in a

text (pp. 16–17). In a similar vein, Samuels' (1994) Automaticity Training proposed that reading easy, interesting, and meaningful materials may free students' attention from the decoding of words to the comprehension of texts (pp. 833–834). In other words, reading materials at the '*i*-1' level may be important for automaticity training. By reading '*i*-1' materials, processing speed may be facilitated. As a result, the contexts for target words are expanded, which promotes access to a rich memory reservoir for meaning mapping, and thus consolidates the acquisition of language.

Besides the influence on language-learning development, the difficulty level of reading texts may also influence readers' affective response. According to Krashen's (1994b) Affective Filter Hypothesis, one condition for comprehensible input is that the affective filters have to be low to allow for effective text processing. Since difficult or uninteresting texts may create a negative effect on a student's attitude or anxiety level, they could inhibit reader efficacy, thereby raising the affective filters for language acquisition. Conversely, easy and interesting texts might instill motivation, confidence, and self-assurance in a reader, which can lower his or her affective filter for language acquisition. Consequently, ER materials may be interesting for the readers, yet their difficulty level matters if their affective filters inhibit their reading comprehension.

Considered together, Krashen's and Samuels' theories, along with Day and Bamford's hypothesis, might reflect one or both of a foreign language reader's cognitive and affective attainment when engaged in ER at a particular difficulty level. To explore this option, the current study design explored two levels of text difficulty, '*i*+1' and '*i*-1', to identify whether these theoretical options do in fact match in practice. That possibility may be important for ER practice when learners and their teachers decide on their student level of language acquisition vis-à-vis the assessed FL difficulty level of the books they read.

Effects of Extensive Reading

There have been many studies on the effects of ER on the cognitive, affective, and behavioral dimensions of EFL/ESL learning. Cognitively, ER was reported to improve the participants' general reading ability and reading comprehension (Aka, 2019; Kargar, 2012; Suk, 2016; Yamashita, 2008), enhance incidental vocabulary acquisition (Boutorwick, Macalister, & Elgort, 2019; Suk, 2016), promote writing skills (Mermelstein, 2015; Park, 2016; Sakurai, 2017), and boost oral proficiency and speaking abilities (Cho & Krashen, 1994). Regarding the effects on affect, ER was found to facilitate students' motivation to read (Burgh-Hirabe & Feryok, 2013; Judge, 2011; Takase, 2007) and foster positive attitudes toward reading (Al-Homoud & Schmidt, 2009; Mikami, 2017; Tabata-Sandom, 2017; Yamashita, 2013). Also, ER has been shown to have an impact on reading behavior. Students experienced a reading flow (Kirchhoff, 2013), developed a reading habit (Asraf & Ahmad, 2003), changed their reading behavior (Rodrigo et al., 2014), and increased their reading rate (Beglar & Hunt, 2014; Huffman, 2014; McLean & Rouault, 2017; Suk, 2016).

Two recent meta-analysis studies on ER amassed the ER effects on the aspects of reading rate, reading comprehension, and vocabulary. Jeon and Day (2016) performed a meta-analysis of EFL/ESL ER studies published between 1980 and 2014. Two types of comparison were identified based on the study design: *experimental vs. control comparisons* (51 samples) and *pre-*

to posttest comparisons (20 samples). Of relevance to the present study is the former type, which yielded a small to medium overall effect size (d = .57), with the moderators of publication years, setting, age, input mode, and ER form impinging on the main effect. Especially noted is that treatment length and focus skill do not moderate the effect. Although focus skill does not moderate the main effect, one skill, reading comprehension, showed a medium effect size (d = .54), with the confidence interval not including zero. Moreover, age did moderate the main effect. Adults were found to benefit twice as much as adolescents from ER, which was attributed to adolescents' smaller cognitive reservoir of life experiences for extensive reading than adults. Besides, the form of ER was found to show varied effects: ER as a part of the curriculum, defined as a non-credit addition to an existing course, yielded the best effect, followed by ER as an extracurricular activity, while ER as a part of a reading course and ER as an independent course produced effects at the lower end. Notably, the two superior forms of ER practice, a non-credit addition to a curriculum and an extracurricular activity, do not entail the formal assessment that put students' grade at stake and consequently are less likely to hinder young learners' motivation to engage in ER. The findings shed light on the potential of summer programs as a site for ER practice because they pose no pressure of grading as a non-credit addition to regular semester courses.

Another meta-analysis study investigating the ER effect on the skills of reading speed, reading comprehension, grammar, and vocabulary was carried out by Nakanishi (2015), based on 34 ER studies, mostly EFL, published between 1989 and 2012. Similar effects to that of Jeon and Day (2016) were yielded for Group Contrasts, that is, *experimental vs. control comparisons* (22 samples) and Pre-post Contrasts (21 samples), with a medium overall ER effect size for group contrast (d =.46). Moderator analysis for group contrast on age level, length of instruction, areas of interest, test use, and test reliability was followed. Of relevance to this study is the area of interest, with reading comprehension indicating a medium effect size (d = .63), and the confidence interval not including zero. Moreover, measures of reading comprehension were found to be the most frequently used test type. Hence, we propose that the measure of reading comprehension be further classified into its two sub-components, literal and inferential comprehension, to account for the contemporary portrait of text comprehension, such as the textbases and situation models proposed by Kintsch (1998). Such division may enable the understanding of how readers of varied ER text difficulty distribute their mental resources disparately in their text processing. This study, therefore, measured reading comprehension at two levels.

Effects of Text Difficulty in ER Materials

There is, to date, a dearth of studies inspecting the effect of ER text difficulty. In a qualitative study, Wan-a-rom (2012) assigned 80 Thai high school students to read books at their current level based on their scores in a vocabulary size test. Students reported in their reading journals and in the interview, administered at the closure of the 6-week ER program, that their comprehension had improved. They were able to read fluently at a level of controlled difficulty, showed a more positive attitude, and achieve higher reading motivation. Another study (Lai, 1993) was conducted in Hong Kong as a summer reading program. Two hundred secondary school students (grades 7–9) were assigned to three level-groups and chose graded readers and short passages to read. Results showed that the two higher proficiency groups had significant

gains in reading comprehension and reading speed, while students at the lower level did not, since they read texts beyond their level. The results seemed to support the benefits of reading at students' current levels in terms of enhancing their comprehension, and the disadvantage of reading beyond one's capacity. While attempting to understand the issues of ER text difficulty, these two studies nevertheless did not directly compare the effect of varying difficulty levels, those beyond and below readers' current capacity, as did the following two studies.

Chiang (2016) investigated the effects of varying difficulty levels of ER materials on the reading comprehension and reading attitude of 54 freshman non-English majors who took a Freshman English course at a university in Taiwan. Students in 14 subgroups were assigned to two treatment groups, reading eight graded readers of '*i*–1' level and of '*i*+1' level respectively for two semesters. With a reference to the university placement test, the entry-level was set at Level 4 as classified by the Oxford Bookworm series. Hence, the '*i*-1' group read books at Levels 3 and 4, while the '*i*+1' group read books at Levels 5 and 6. Pre- and posttest comparisons yielded significant improvement in the attitudes of the '*i*-1' group, and in the English proficiency and its subset of reading comprehension of both groups. Nevertheless, between-group comparisons indicated no difference in any of the measures. Hence, ER text difficulty impacted neither reader attitude nor reading comprehension. Several factors may obscure the text-difficulty effects. For one, the ER treatment was embedded in a two-semester course, making it difficult to tease out its effect from those of other course components. Secondly, the graded readers were assigned by the teacher, leaving little room for choice in content geared toward the learners' interests. Finally, no information was given regarding how ER was implemented in the regular classroom. Without information and control of these factors, the study results may be inconclusive.

Another study inspecting the effect of varying text difficulty in ER was carried out by Bahmani and Farvardin (2017) involving 50 Iranian EFL college students with elementary-level proficiency in two intact classes. In addition to reading comprehension, the affective variable of anxiety was measured using the Foreign Language Reading Anxiety Scale (FLRA) (Saito et al., 1999). Based on the Headway Placement Test, A2 was located as the entry-level, which corresponded to the Level 2 of the Oxford Bookworm series. During the 4-month treatment, the class assigned as the '*i*-1' group self-selected two books at the Starter level and two at Level 1, and another class assigned as the '*i*+1' group chose two books at Level 3 and Level 4 as part of the class curriculum. The reading comprehension results corresponded to Chiang's (2016) findings in that no between-group difference was found, albeit significant within-group progress. However, the measure of reading anxiety indicated an interactive effect of text difficulty and time: the '*i*-1' group decreased while the '*i*+1' group increased their anxiety level after treatment.

In these two studies on ER text difficulty, the affective variables of attitude and anxiety were examined. A third variable, reading motivation, may warrant a probe since its positive effects on university students have been attested both quantitatively (Takase, 2007) and qualitatively (Burgh-Hirabe & Feryok, 2013; Judge, 2011). How ER text difficulty affects vocational high school students' motivational change is worth exploration. Most importantly, the learner-level assignment in these two previous studies was based on either random group assignment or convenience samples as intact classes, masking variation in individual ability level. In addition, these studies covered a broad range of difficulty levels, two levels above and below, which may have inflated the effects. To maximize the text difficulty effect, a study design employing a more

fine-tuned, individual-based assignment of graded readers in a narrower range around the '*i*' level is needed. Finally, without a control group, the significant findings yielded from within-group contrasts alone may weaken the basis for interpretation (Nakanishi, 2015). This study thus utilized a three-group experimental-control design to inspect the respective effects as well as the comparative effects of ER at one level higher and one level lower than the current level of vocational high school EFL students in Taiwan, who were assigned to read graded readers on an individual basis, on two components of reading comprehension and on three dimensions of reading motivation.

Research Questions

Four research questions were thus posed for this study:

1. What are the respective effects of the '*i*-1' and '*i*+1' ER materials on EFL vocational high school students' reading comprehension?

2. What are the comparative effects of the '*i*-1' and '*i*+1' ER materials on EFL vocational high school students' reading comprehension?

3. What are the respective effects of the '*i*-1' and '*i*+1' ER materials on EFL vocational high school students' reading motivation?

4. What are the comparative effects of the '*i*-1' and '*i*+1' ER materials on EFL vocational high school students' reading motivation?

Methods

Context, Participants, and Group Assignment

This experiment was conducted in the summer session, as a program to promote pleasure reading. The target vocational high school has supported extensive reading for English courses with funds to purchase graded readers. Although it is not credited in the student's semester grades, most of the teachers on site recommended ER as an extracurricular activity. Yet, very few students made use of this resource in the library due to the heavy course load they were to carry during the regular semester. Therefore, this project was implemented in the 6-week summer session, during which no constraints were imposed on the joint curriculum. Students and teachers thus had free class time and autonomy in their ER practice.

A total of 120 second graders, ranging in age from 17 to 19, from three classes at a vocational high school in Taiwan (equivalent to 11th graders in the regular high school) participated in this study. At the time of this experiment, they had been through formal English education for nine years since the third grade. Their English course was taught by the first author in the previous semester year, which includes fall and spring sessions, and in the summer session. Based on the mean English grades in the previous semester, the mid-range class (1 female, 40 male) from the Department of Power Mechanical Engineering was assigned as the control group. The other two

classes, one from the Department of Electronics and the other from the Department of Horticulture, altogether 79 students, were combined as a cohort and took the vocabulary size test (VST) validated by Wan-a-rom (2010). The VST is a 50-item test, with each item representing 100-word families. The cohort produced a range of five to 26 correct responses; hence their vocabulary size was between 500 and 2,600 words. Based on their VST scores, an S-shape sampling procedure was followed to allocate 39 students (10 female) to the '*i*-1' group, and 40 (15 female) to the '*i*+1' group. The comparability in reading proficiency across the three groups was derived by an Analysis of Variance (ANOVA) on the pretest scores of the TOEFL Junior Test of reading comprehension, with no difference found, $F = .50$, $p > .05$, $\eta_p^2 = .01$.

Graded Reader Assignment

To match the experimental students' vocabulary capacity with the difficulty level of the graded readers, the Graded Readers Level Scale (GRLS), developed by the Edinburgh Project on Extensive Reading to cross-refer book series based on the number of headwords, was employed (Waring, 2015). The experimental students' VST scores ranged from Level 5 (401–600 words) to Level 13 (2,401–2,700 words) on GRLS (Appendix A). For the assignment of graded readers, one lower level, Level 4 (301–400 words), was added for the lowest level of the '*i*-1' group, Level 5. For the highest level of the '*i*+1' group, no additional level was assigned because the single highest VST scorer (2,600 words) happened to be assigned to the '*i*-1' group, reading books at GRLS Level 12 (13-1). Altogether ten GRLS levels (4–13) were allocated for ER reading.

Two graded-reader collections, Macmillan and Penguin, were examined regarding their vocabulary size and book level against the ten GRLS levels, based on which 141 Macmillan book and 15 Penguin books were selected for this study. Appendix A displays the correspondence of GRLS levels to Macmillan classification (from 1 *starter* to 6 *advanced*) and Penguin classification (from 2 *elementary* to 5 *upper-intermediate*) as well as the matching of students in the two groups to graded readers of a level below or above. The assignment ensured that no two students of the same GRLS level in the two different groups read the same level book designated by the publisher, except for students of GRLS Level 5 in both groups reading Penguin Readers *2 elementary* and of Level 8 reading *3 pre-intermediate*. The impact of this elapse might be small because for each of these two Penguin book levels only three books were provided.

Treatment Procedure

During the 6-week summer session, students met twice weekly, each time for two class periods. The pre- and posttests were administered in the first and last weeks. In the 4 weeks in between, the experimental students went through four cycles of ER treatment, each following the same procedure covering one graded reader. The control group, on the other hand, was instructed on one textbook lesson-unit each week, with a focus on vocabulary and grammar learning, and on intensive, guided reading.

For the two experimental groups, a class library of 156 books, some of which had two or three copies, totaling 320 items, was set up. They were placed into six boxes, labeled from 1 to 6, corresponding with the levels of the Macmillan Book Collection. Each book item came with a

comprehension check sheet including four multiple-choice and two true-false questions developed by the English teachers at the research site, as a joint, school-wide effort to promote extensive reading.

The treatment activities were designed with reference to those suggested by Bamford and Day (2004). The weekly treatment cycle began in the first meeting with students selecting a book at the designated level and filling out a Library Index Card. They then engaged in silent reading for 60 minutes before they filled out a reading log, followed by book talks and teacher comments for the next 40 minutes. In the second meeting, the same procedure was administered, except that the reading log writing was replaced by filling out a comprehension check, upon book return. The few students who could not finish reading in class continued to read the rest of the book at home.

Pretest and Posttest Materials

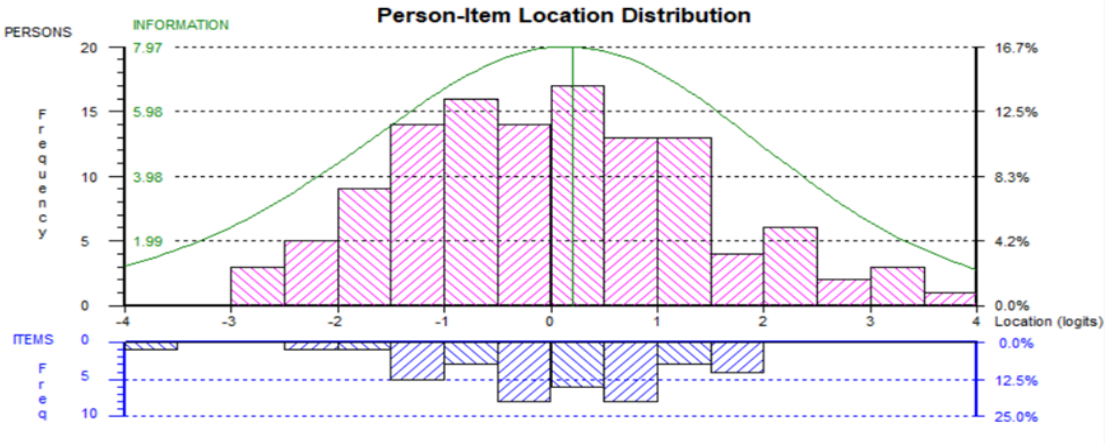
Reading comprehension test. The reading comprehension section of two TOEFL Junior Practice Test sets was used, with permission from the English Testing Service, for the pre- and posttests in a split-block design. Each reading comprehension test set includes six passages of varied text types and text length, each followed by four to 11 multiple-choice items, totaling 42 items. According to Educational Testing Service, the test measures middle school or lower-level high school students' language proficiency for reading, tapping their ability to understand main ideas, identify details, and make inferences about meaning, attitude, or point of view.

To check whether there was a match between the overall average of test difficulty and an average person's ability, we ran RUMM2030 (Andrich et al., 2010) on the pretest raw scores, and the results indicated an 'EXCELLENT' fit. Specifically, the mean person ability location, -0.021, was only marginally smaller than the mean difficulty location, 0.000, suggesting that the TOFEL Junior test was neither too difficult nor too easy for the ability level of the participants, and the test also achieved very high reliability (PSI value = 0.91, Cronbach's Alpha = 0.93).

The plot in Figure 1 indicates that the range of item difficulty locations (-4.0 and +2.0 logits) could be intersected mostly by the range of participants' ability locations (-3.0 and +4.0 logits). Importantly, the information curve covered a very wide range (-4.0 and +4.0 logits), indicating that the test was informative across all the various ability levels of the participants. Taken together, the test used was well-targeted to the proficiency levels of the participants.

Figure 1.

Person-item location distribution



To assess students' performance in terms of two levels of reading comprehension, altogether the 84 question items were further coded into two types, literal and inferential, by two English teachers at the research site. Inter-coder reliability of .96. was yielded; differences were resolved via discussion, yielding 25 literal questions and 17 inferential questions for one test set, and 31 literal questions and 11 inferential questions for the other.

Reading Motivation Questionnaire. The Reading Motivation Questionnaire (Appendix B) was composed of 11 items, of which nine were compiled in light of the treatment features with a reference to two motivation questionnaires, Mori's (2002) Motivation of Reading Questionnaire and Wigfield and Guthrie's (1995) Motivation for Reading Questionnaire. The other two items were designed by the researchers, one on *ease of management of difficult text* (item 2) and another on *reading speed* (item 3). A 4-point rather than a 5-point Likert scale from *strongly agree* (4) to *strongly disagree* (1) was adopted, as suggested by Nunnally and Bernstein (1967) to avoid eliciting neutral opinions. A reliability check on the pretest ratings, with three negatively stated items on *reading avoidance* reversed in value, showed a Cronbach's alpha value of .85, suggesting good internal consistency for the scale. Moreover, none of the item-total correlations were lower than .30.

Data Collection and Data Analysis

In the pretest, students took the VST for 30 minutes before they responded to the reading motivation questionnaire for 10 minutes. After a break, they worked on the TOEFL Junior Test for 70 minutes. The same procedure was followed in the posttest, except for the VST work.

For the data analysis, percentages of correct responses were calculated for the three measures of reading comprehension: overall comprehension and its two subsets, literal and inferential comprehension. Three rounds of Analysis of Covariance (ANCOVA), with group (3) as an independent variable and the pretest as the covariate, were performed on the three posttest measures of reading comprehension, respectively. In the case of significant difference, pairwise

comparisons between groups ensued and the Bonferroni *p*-value was corrected by setting the significance levels at .017 (.05/3) and .003 (.01/3). The same statistical procedure was also applied to the analysis of the Reading Motivation Questionnaire on overall reading motivation and its three components.

Results

Effects on Reading Comprehension

Tests of homogeneity of effect slopes on *overall comprehension* and its two subsets, *literal and inferential comprehension*, indicated no interaction between the pre- and posttest mean scores. Also, a strong relation was yielded between the pre- and posttests across the three measures of comprehension, with η_p^2 = .70, .39, and .58 for *overall*, *literal*, and *inferential* comprehension respectively (Appendix C).

After adjusting for the pretest scores, there was a significant difference among the three groups on two comprehension measures: *total comprehension*, F (2, 116) = 9.11, $p < .001$, η_p^2 = .14, and *literal comprehension*, F (2, 116) = 12.50, $p < .001$, η_p^2 = .18, both with a moderate effect size, but not on *inferential comprehension*, F (2, 116) = 1.57, $p > .05$, with a small effect size, η_p^2 = .03. (Appendix C). Pairwise comparisons were then followed on the *overall comprehension* and *literal comprehension* scores.

RQ1: *What are the respective effects of the 'i-1' and 'i+1' ER materials on EFL vocational high school students' reading comprehension?*

To answer this research question, two pairs of comparisons, between each experimental group and the control group, were performed. As shown in Tables 1 and 2, for *overall comprehension*, the '*i*-1' group (M = 55.91[a], SE = 1.67) outperformed the control group (M = 46.14[a], SE = 1.63) with a significant adjusted mean difference of 9.76[a] (SE = 2.34, 95% CI [5.13, 14.39]), $p < .003$, d = .93, while no significant difference between the '*i*+1' group and the control group was found. Hence, the '*i*-1' ER materials enhanced overall comprehension, whereas the '*i*+1' materials did not.

Table 1.

Adjusted Posttest Mean Scores for Reading Comprehension

	'*i*-1' (n = 39)	'*i*+1' (n =40)	Control (n =41)
Overall	55.91[a] (1.67)	49.14[a] (1.65)	46.14[a] (1.63)
Literal	61.25[a] (2.63)	47.79[a] (2.61)	43.60[a] (2.56)
Inferential	53.31[a] (2.36)	54.09[a] (2.33)	48.72[a] (2.30)

Note. Standard errors are in parentheses. [a]: mean score adjusted by pretest scores.

For the subset of *literal comprehension*, as revealed in Tables 1 and 2, the '*i*-1' group again

produced a significantly higher adjusted mean (M = 61.25[a], SE = 2.63) than the control group (M = 43.60[a], SE = 2.56), by a substantial adjusted mean difference of 17.65[a] (SE = 3.66, 95% CI [10.39, 24.90]), $p < .003$, d = 1.08; however, no significant difference was found between the '*i*+1' group and the control group. Therefore, ER at the '*i*-1' level boosted literal comprehension whereas ER at the '*i*+1' level did not.

Table 2.

Adjusted Mean Posttest Differences and Pairwise Comparisons on Reading Comprehension

	'*i*-1' vs. control	'*i*+1' vs. control	'*i*-1' vs. '*i*+1'
Overall	9.76 [a] (2.34)**	2.99 [a] (2.33)	6.78 [a] (2.35)*
	[5.13, 14.39]	[-1.63, 7.60]	[2.12, 11.43]
	d = .93	d = .29	d = .65
Literal	17.65 [a] (3.66)**	4.19 [a] (3.67)	13.46 [a] (3.74)**
	[10.39 , 24.90]	[-3.07, 11.45]	[6.06, 20.86]
	d = 1.08	d = .26	d = .82
Inferential	4.58 [a] (3.30)	5.37 [a] (3.28)	-.78 [a] (3.32)
	[-1.96, 11.12]	[-1.12 , 11.85]	[-7.36 , 5.79]
	d = .31	d = .36	d = .05

Note. Standard errors are in parentheses; confidence intervals in brackets. Bonferroni correction of p: * $p < .017$, ** $p < .003$.

RQ2: *What are the comparative effects of the 'i-1' and 'i+1' ER materials on EFL vocational high school students' reading comprehension?*

To answer research question 2, comparisons between the '*i*-1' group and the '*i*+1' group on overall comprehension and its subset of literal comprehension scores were performed. As shown in Tables 1 and 2, for *overall comprehension*, the '*i*-1' group (M = 55.91[a], SE = 1.67) significantly surpassed the '*i*+1' group (M = 49.14[a], SE = 1.65) with an adjusted mean difference of 6.78[a] (SE = 2.35, 95% CI [2.12, 11.43]), $p < .017$, d = .65. Thus, '*i*-1' graded readers are more effective than '*i*+1' readers in facilitating overall comprehension.

For the subset of *literal comprehension*, the '*i*-1' group (M = 61.25[a], SE = 2.63) again outperformed the '*i*+1' group (M = 47.79[a], SE = 2.61) by a significant adjusted mean difference of 13.46[a] (SE = 3.74, 95% CI [6.06, 20.86]), $p < .003$, d = .82 (Tables 1 and 2). As such, reading graded readers at the '*i*-1' level is superior to reading those at the '*i*+1' level for enhancing literal comprehension.

For *inferential comprehension*, an inspection of the mean scores show that the *inferential comprehension* produced by the '*i*+1' group (M = 54.09[a], SE = 2.33) was slightly higher than that by the *i*-1 group (M = 53.31[a], SE = 2.36) (Table 1), a contrast with the findings on *overall* and *literal comprehension*, for which the '*i*+1' group's performance significantly lagged behind the '*i*-1' group.

Effects on Reading Motivation

A factor analysis on the mean rating of the 11 reading-motivation questionnaire items yielded three component factors in the rotated pattern matrix: *self-efficacy*, *reading engagement*, and *reading avoidance*, each accounting for 43.62%, 16.53%, and 9.42% of the variance respectively (for the factor loadings, see Appendix B).

Tests of homogeneity of effect slopes on *overall reading motivation* and its three factors showed no significant interaction between the pretest and the posttest mean ratings. ANCOVA analyses on the four motivation indices also showed a high relation between the covariates, the pretest, and the posttest, with large effect sizes for overall reading motivation and its three factors (η_p^2 = .40, .36, .15, and .24 for *overall reading motivation, self-efficacy*, *reading engagement*, and *reading avoidance* respectively; see Appendix C).

After adjusting for the pretest ratings, significant differences among the three groups were found on all four indices of reading motivation: *overall reading motivation*, F (2, 116) = 4.34, $p < .05$, η_p^2 = .07; *self-efficacy*, F (2,116) = 6.52, $p < .01$, η_p^2 = .10; *reading engagement*, F (2,116) = 4.32, $p < .05$, η_p^2 = .07, and *reading avoidance*, F (2,116) = 3.08, $p < .05$, η_p^2 = .05 (Appendix C). Subsequent pairwise comparisons were therefore administered on the four measures.

RQ3: *What are the respective effects of the 'i-1' and 'i+1' ER materials on EFL vocational high school students' reading motivation?*

To answer research question 3, two pairs of comparisons, between each experimental group and the control group, were performed on the overall reading motivation and its three component factors. As shown in Tables 3 and 4, for *overall reading motivation*, the '*i*+1' group produced a significantly higher adjusted posttest mean rating (M = 2.58[a], SE = .06) than that of the control group (M = 2.32[a], SE= .06), with a significant adjusted mean difference of .26[a] (SE = 09, 95% CI [.09, .44]), $p < .017$, d = .66. On the other hand, the '*i*-1' group (M = 2.43[a], SE = .06) did not surpass the control group (M = 2.32[a], SE = .06). Thus, the '*i*+1' materials did boost overall reading motivation while the '*i*-1' materials did not.

Table 3.

Adjusted Posttest Mean Ratings for Reading Motivation

	i-1 (*n* = 39)	*i*+1 (*n* =40)	Control (*n* =41)
Overall	2.43[a] (.06)	2.58[a] (.06)	2.32[a] (.06)
Self-efficacy	1.74[a] (.08)	2.19[a] (.08)	1.96[a] (.08)
Reading engagement	2.90[a] (.09)	2.91[a] (.09)	2.58[a] (.09)
Reading avoidance	2.77[a] (.09)	2.62[a] (.09)	2.47[a] (.09)

Note. Standard errors are in parentheses. [a]: mean score adjusted by pretest scores.
For the factor of *self-efficacy*, no significant difference was yielded between the '*i*+1' group (M = 2.19[a], SE = .08) and the control group (M =1.96[a], SE = .08) or between the '*i*-1' group (M = 1.74[a], SE =.08) and the control group (Tables 3 and 4). A special note here is that while the '*i*+1'

group's mean adjusted rating was higher than that of the control group, the '*i*-1' group's was lower, with an adjusted mean lower than 2, indicating a less-than-satisfactory perception of self-efficacy.

For the factor of *reading engagement*, both the '*i*-1' group (M = 2.90[a], SE = .09) and the '*i*+1' group (M = 2.91[a], SE = .09) had a significantly higher rating than the control group (M = 2.58[a], SE = .09) (Table 3), with a gap of .32[a] (SE = .13, 95% CI [.07, .58)]) between the '*i*-1' group and the control group, $p < .017$, d = .55, and .33[a] (SE =.13, 95% CI [.08, .59]) between the '*i*+1' group and the control group, $p < .017$, d = .57 (Table 4). Hence, ER at both text difficulty levels engaged readers in reading to a similar degree.

For the factor of *reading avoidance*, the '*i*-1' group's rating (M = 2.77[a], SE =.09) significantly exceeded that of the control group (M = 2.47[a], SE =.09), with an adjusted mean difference of .31 (SE = .12, 95% CI [.06, .55]), $p < .017$, d = .56; meanwhile, the '*i*+1' group (M = 2.62[a], SE = .09) did not vary significantly from the control group, as shown in Tables 3 and 4. Therefore, the '*i*-1' ER materials inhibited reading avoidance, while the '*i*+1' materials did not.

Table 4.

Adjusted Mean Posttest Differences and Pairwise Comparisons of Reading Motivation

	i-1 vs. control	*i*+1 vs. control	*i*-1 vs. *i*+1
Overall	.11 [a] (.09) [-.06, .29] *d* = .28	.26 [a] (09)* [.09, .44] *d* = .66	-.15 [a] (.09) [-.33, .03] *d* = .37
Self-efficacy	-.22 [a] (.12) [-.45, .004] d = .43	.22 [a] (.12) [-.01, .50] *d* = .43	-.45 [a] ** (.12) [-.69, -.20] *d* = .85
Reading engagement	.32[a] (.13)* [.07, .58] *d* = .55	.33[a] (.13)* [.08, .59] *d* = .57	-.01[a] (.13) [-.27, .25] *d* = .02
Reading avoidance	.31[a] (.12)* [.06, .55] *d* = .56	.15 [a] (.12) [-.0 9, .40] *d* = .28	.15 [a] (.13) [-.09, .40] *d* = .28

Note. Standard errors are in parentheses; confidence intervals in brackets. Bonferroni correction of p: * $p < .017$, ** $p < .003$.

RQ4: *What are the comparative effects of the 'i-1' and 'i+1' ER materials on EFL high school students' reading motivation?*

To answer this research question, comparisons between the '*i*-1' group and the '*i*+1' group were performed on *overall motivation* and its three component factors. Only the factor of *self-efficacy* showed the effect of text difficulty. Despite a moderate adjusted mean of 2.19[a] (SE = .08) yielded (Table 3), the '*i*+1' group surpassed '*i*-1' group (M = 1.74[a], SE = .08) in their rating of self-efficacy. The significant adjusted mean difference of -.45[a] (SE =.12, 95% CI [-.69, -.20]) between

groups, $p < .003$, $d = .85$ (Table 4) indicated that the '*i*+1' graded readers instigated a significantly higher level of self-efficacy than the '*i*-1' graded readers. No difference was found between the two experimental groups on the other three indices, *overall reading motivation*, *reading engagement*, and *reading avoidance*.

Discussion

Effects of Easy and Difficult ER Texts on Reading Comprehension

The response to RQ1 is that the '*i*-1' ER materials promoted EFL vocational high school students' overall reading comprehension and its subset of literal comprehension, but the '*i*+1' ER materials did not. The effect sizes of overall reading comprehension, $d = .93$, and of literal comprehension, $d = 1.08$ (Table 2), for the '*i*-1' group in this study were much larger than the medium effect sizes of reading comprehension derived from experimental- vs. control-group contrasts in the two meta-analysis studies ($d = .54$ in Jeon & Day, 2016, & $d = .63$ in Nakanishi, 2015), for which text difficulty was not examined. The result thus supports the claim by Day and Bamford (2002) that reading easy materials facilitates the understanding of texts and bears out the automaticity theory proposed by Day and Bamford (1998) and Samuels (1994) that by reading easier texts students' attention to decoding may be released and better comprehension achieved. Besides the cognitive benefits, the ease in textual input may also trigger readers' affective response, that is, the lowering of their affective filter (Krashen, 1994b), which could also contribute to better comprehension.

On the other hand, the result that the '*i*+1' graded readers did not enhance overall comprehension and its subsets of literal and inferential comprehension, with respective effect sizes ($d = .29$ and $d = .26$, Table 2) much smaller than those synthesized in previous studies (Jeon & Day, 2016; Nakanishi, 2015), may not support the postulate of the Input Hypothesis (Krashen, 1994a) that input at one level beyond learners' capacity may promote acquisition. The result indicated that '*i*+1' ER materials could be too challenging for pleasure reading, which aims for general understanding rather than for content or language retention. The additional level of difficulty may command efforts in language processing for L2 learners, producing an extraneous cognitive load – as termed in cognitive load theory – that takes up capacity for intrinsic processing for meaning (Sweller, 2010). Still, even with the extraneous load posed in treatment reading materials, the '*i*+1' group's adjusted posttest mean was higher than that of the control group, albeit with an insignificant difference (Tables 1 and 2). At least additional language processing did not put the '*i*+1' group at a disadvantage in terms of reading proficiency development.

Despite significant improvements yielded by the '*i*-1' group in overall and literal comprehension, neither experimental group had their inferential comprehension enhanced after treatment. This finding may be interpreted as follows. First, the multiple-choice measure, contrived by the tester, may not adequately gauge the actual inference construed by readers, who tapped their knowledge and experiences to fill the textual gaps. Second, understanding at the inferential level may command deeper-level processing that takes time, perhaps longer than four weeks, to foster. Finally, it may take texts with a higher difficulty level than the levels covered in this study, in which sophisticated and elaborated ideas are embedded in complex language, to prompt

inferences during reading.

Effects of ER Text Difficulty on Reading Comprehension

The findings in response to RQ 2 showed that the '*i*-1' ER materials generated better overall and literal comprehension than the '*i*+1' materials. The present result echoes Wan-a-rom's (2012) findings on students' reports that matching learners' vocabulary size with text difficulty in ER promoted reading fluency and comprehension, and corroborates Lai's (1993) findings that showed the positive ER effects on high proficiency learners and null effects on low proficiency learners. However, the present finding did not correspond to the insignificant between-group effects found in the two previous studies examining ER text difficulty on university learners (Bahmani & Farvardin, 2017; Chiang, 2016). ER text difficulty did make a difference in the reading comprehension of vocational high school EFL learners in this study. Although adolescent learners may not be affected by ER as much as adult, university learners, as revealed in the meta-analysis by Jeon and Day (2016), they may be prone to the influence of text difficulty in ER materials, as reflected in the effect size of text difficulty, d = .65 (Table 2), generated by the two experimental groups of vocational high school learners in this study.

The varied effect sizes of text difficulty between two dimensions of comprehension, that is, d = .82 for literal comprehension and d = .05 for inferential comprehension (Table 2), revealed that inferential comprehension was not affected by linguistic difficulty unlike literal comprehension. In light of Kintsch's (1998) Construction-Integration model, the '*i*-1' materials, with lower text difficulty, may be able to drive the processing for textbase construction to a fuller extent, leading to higher literal comprehension than the '*i*+1' texts; on the other hand, the '*i*+1' materials, although constrained by language difficulty, could at least instigate the same level of processing for the establishment of text situations as the '*i*-1' texts, resulting in a similar level of inferential understanding.

Since reading comprehension involves the simultaneous activation of multiple components (Grabe, 2009) within a limited working memory span, the distribution of the processing levels matters. As put by Verhoeven (2011, p. 672), automated word recognition, in this case in reading '*i*-1' texts, frees mental resources for closer consideration of the meaning of a text – the exercise of processing at all levels. It is therefore conjectured that strengthening and stabilizing declarative memory of known lexical units (Nation, 2009, as cited in Suk, 2016, p. 74) may not be the sole factor of reading proficiency improvement for the '*i*-1' group; rather, the tacit acquisition of procedural knowledge addressing the connective acts during text processing may be the other boosting factor. Between the automatic word recognition and the meaning-making of the text, the '*i*-1' group may be bestowed greater capacity than the '*i*+1' group for procedural connections at varied processing levels: linking words for syntactic chunks, networking semantic units within and between episodes, and connecting propositions for causal-temporal structure, a characteristic of narrative texts for which readers are equipped with readily available schema (Yoshida, 2012). Hence, the '*i*-1' group, with capacity free from controlled processing of unknown words that taxed the '*i*+1' group, may be able to process the texts to a fuller extent than the '*i*+1' group, thus attaining higher reading proficiency.

Two features of the present ER program may have also contributed to the text difficulty effects.

First, the ER practice in a summer program may endorse a full, intensive engagement in the class time, allowing students to read four books in 4 weeks without the interruption of regular English lessons. By contrast, a long-term program (e.g., Chiang, 2016) embedding the ER components into the general English curriculum (e.g., Bahmani & Farvardin, 2017) may extend the reading of the same number of books to several months, weakening the linking of the storyline and thus lessening the fun of reading. Second, the classroom activities adopted in the present program may bolster students in managing pace and sustaining self-regulation–a skill that supports ER practice, especially for teenagers.

Several endeavors in the present study design may have also played a role in enhancing the effects. Matching each individual's vocabulary size with a level of graded readers and varying ER text difficulty by one level, instead of two, below or above may more precisely reflect the *i*-1 and *i*+1 precepts in the relevant theories (Day & Bamford, 1998; Krashen, 1994a). Also, by measuring reading comprehension in two dimensions, a more sophisticated variation in reading comprehension impacted by ER text difficulty was elucidated, as shown in the distinctive effects on literal and inferential comprehension.

Effects of Easy and Difficult ER Texts on Reading Motivation

In response to RQ 3, three significant findings were yielded: first, the '*i*+1' ER texts promoted overall reading motivation; second, both the '*i*-1' and the '*i*+1' ER materials enhanced reading engagement; third, the '*i*-1' ER materials inhibited reading avoidance. The positive ER effects on overall reading motivation generated by the '*i*+1' group may correspond to those obtained in previous studies involving university students using quantitative measures (Takase, 2007) or qualitative measures (Burgh-Hirabe & Feryok, 2013; Judge, 2011). However, this finding is not aligned with the prescription of the Affective Filter Hypothesis (Krashen, 1994b) that more difficult texts may raise the affective filter leading to low motivation, negative attitude, and low self-esteem. It is likely the '*i*+1' ER materials may contain inherent elements that suppress the raise of the affective filter: challenging language that entails more elaborate and thus more engaging content, for instance. These elements could be more robust than language ease in motivating students to read.

Secondly, both experimental groups' boost in the component of reading engagement supported the Pleasure Hypothesis (Krashen, 2004) in that self-selected interesting materials and meaning-based classroom activities promoted engagement in reading. The graded readers, whether easy or difficult, all came with an extensive, coherent storyline that promoted a deeper-level engagement than the short traditional textbook passages read by the control group. Above all, the autonomy and agency in book choice to meet personalized interest, meaning-making, and self-pacing in ER may have also enhanced the reading engagement (Lake & Holster, 2014) of the two experimental groups.

Finally, the '*i*-1' group's inhibition of reading avoidance after treatment supported Day and Bamford's (1998) assertions that ER materials within the readers' competence produced less reading avoidance. Supposing that the reduction of reading avoidance is a consequence of attitude change, the present finding may echo Chiang's (2016) finding that the '*i*-1' materials

promoted language learning attitude and may resonate with the desire to keep reading extensively expressed by individuals in Chiang's '*i*-1' group at the closure of the study. The ease and success in comprehending texts by the '*i*-1' group may help them to form an expectation of accomplishment, and thus they were less likely to dodge it in their future reading, as explained by the *expectancy-value* model (Feather, 1982).

Effects of ER Text Difficulty on Reading Motivation

Only one reading motivation component, self-efficacy, significantly distinguished the two ER groups after treatment, as a response to RQ4. The '*i*+1' group reported a higher level of self-efficacy than the '*i*-1' group after treatment. This effect may support Deci and Flaste's (1996) and Andreassen and Bråten's (2010) proposition that effort-tapping tasks, as in reading '*i*+1' texts, may demand higher-level processing and foster in learners a sense of confidence and satisfaction. By contrast, when tasks require little effort, as in reading '*i*-1' texts, self-efficacy may suffer.

The finding may not be in line with the prediction of the Affective Filter Hypothesis (Krashen, 1994b) regarding task difficulty, and thus may not support Bahmani and Farvardin's (2017) finding that the '*i*+1' level of ER reading boosted anxiety while the '*i*-1' level reduced anxiety. In Bahmani and Farvardin's (2017) study, the difficult texts were designated as two levels higher than the learner's current capacity. Perhaps one level higher, rather than two, as assigned in this study, may generate anxiety within a healthy range that can promote self-efficacy.

One anomaly was found in the higher rating of self-efficacy by the '*i*+1'group and the lower rating by the '*i*-1' group than the control group (Table 3). Specifically, the '*i*-1' group produced a less-than-satisfactory rating, lower than 2, in their perception of self-efficacy (Table 3). Feelings of satisfaction were not engendered due to the ease of input in the '*i*-1' level of reading, which could be so unchallenging that the perception of self-efficacy was stymied.

Taken the significant motivational effects together, the findings could also be interpreted in terms of Schraw and Lehman's (2001) model of feeling-related interest and value-related interest. The '*i*-1' level of reading may instigate feelings of pleasure and enjoyment, hence inhibiting future avoidance in reading, while the more difficult '*i*+1' level of reading may arouse a mastery orientation in the face of challenge, leading to valued goals and persistence with the activity (Schiefele, 1991, as cited in Springer et al., 2017, p. 522). Since feeling-related and value-related interest function independently (Springer et al., 2017), both levels of ER input may have a potential for continued pursuit in ER reading. In other words, repeated arousal of pleasure from the '*i*-1' level of reading and the value and efficacy fostered through the '*i*+1' level of reading could be two distinct routes for the development of ER motivation.

Pedagogical Implications

The present finding implies that an ER summer program, despite its intensity and short span, can yield positive effects on reading comprehension and reading motivation as can long-term

programs. Since the curriculum space for ER practice is often limited (Grabe, 2009, p. 312; Hedgcock & Ferries, 2009, p. 218) yet volumes of independent reading at school may contribute to students' reading ability (Taylor et al., 1990, as cited in Ash & Baumann, 2017, p. 384), ER practice as a summer program could be one solution. It has the advantages of the top two effective ER forms identified by Jeon and Day (2016), namely ER as *part of a curriculum* – a non-credit addition – and ER as *an extracurricular activity*, the former featuring class time regulated for reading and teacher-supported activities, and the latter students' autonomy and freedom from high-stake tests. With less course load competing for the ER practice, the summer program allowed students to read at least one graded book per week, a minimal effort for the benefits of extensive reading to be realized (Day & Bamford, 2002). Therefore, to create an initial positive (Day & Bamford, 1998) and breakthrough experience such as that reported by the ESL student, Yuko, in Spack (2004), a curriculum space for ER practice during the summer sessions can be created. Brief as it may be, the summer ER program could pave a new avenue for students in English reading, where they could, hopefully, stride on autonomously to worlds of words–especially on the Internet to which they have frequent access.

To develop the course, initially, the goal can be set for attaining pleasure, hence reading books at the '*i*-1' level. After reading avoidance is lessened, a shift to the '*i*+1' level for value building for self-efficacy may be launched with supports of strategy instruction to help students deal with challenging texts (Pinnell, 1985, as cited in Hoffman, 2017, p. 66). To lessen the cognitive efforts squandered on word recognition, vocabulary strategies can be exercised; to maximize the mental capacity for reading, global strategies, especially background knowledge activation (Shih et al., 2018), can be introduced.

In teaching practice, teachers can assist students in matching their lexical capacity with appropriate difficulty levels of ER readers (Mermelstein, 2015). The designation, '*i*-1' or '*i*+1', can be based on students' needs, instructional resources, and curriculum goals. Importantly, students need to be convinced of the value of ER (Hedgcock & Ferries, 2009, p. 220). With the cognitive and affective values of '*i*-1' and '*i*+1' ER texts elucidated, metacognitive awareness of ER reading can be developed, and ER reading can be sustained after the class is over. Moreover, the ER activities such as reading log, discussion, and comprehension check should include prompts that facilitate the integration of local and global ideas and the anchor of textual information with personal experience and knowledge.

Conclusion

To conclude, this study revealed that reading ER materials at a level below one's capacity might promote overall and literal comprehension and reduce reading avoidance while reading ER materials at a level above could enhance overall reading motivation and promote self-efficacy. The results also demonstrated that both difficulty levels facilitated reading engagement. By matching levels of graded readers with learners' vocabulary capacity on an individual basis, and by having learners complete book reading consistently and regularly, the ER practice implemented in the summer session could have an impact. Two factors may, however, limit the scope of generalization of the present findings. First, since the effects were found with EFL students at a vocational high school in Taiwan, future studies on the ER text difficulty should

involve learners of regular high school to gain a fuller understanding of the issue of ER text difficulty. Second, male participants outnumbered female participants in the current investigation because of the focus of this particular vocational high school. Future research may consider an equal gender balance of participants to gain further insights into ER text difficulty.

Acknowledgments

We would like to thank Professor Janet Swaffar and the anonymous reviewers for their valuable comments and suggestions.

References

Aka, N. (2019). Reading performance of Japanese high school learners following a one-year extensive reading program. *Reading in a Foreign Language*, *31*, 1–18.

Al-Homoud, F., & Schmidt, N. (2009). Extensive reading in a challenging environment: A comparison of extensive and intensive reading approaches in Saudi Arabia. *Language Teaching Research*, *13*, 383–401. https://doi.org/10.1177/1362168809341508

Andreassen, R., & Bråten, I. (2010). Examining the prediction of reading comprehension on different multiple-choice tests. *Journal of Research in Reading*, *33*, 263–283. https://doi.org/10.1111/j.1467-9817.2009.01413.x

Andrich, D., Sheridan, B., & Luo, G. (2010). *RUMM2030*. RUMM Laboratory. Retrieved from http://www.rummlab.com.au/

Ash, G. E., & Baumann, J. F. (2017). Vocabulary and reading comprehension: The nexus of meaning. In S. E. Israel (Ed.), *Handbook of research on reading comprehension* (2nd ed.) (pp. 377–409). The Guildford Press.

Asraf, R. M., & Ahmad, I. S. (2003). Promoting English language development and the reading habit among students in rural schools through the Guided Extensive Reading program. *Reading in a Foreign Language*, *15*, 83–102.

Beglar, D., & Hunt, A. (2014). Pleasure reading and reading rate gains. *Reading in a Foreign Language*, *26*, 29–48.

Bahmani, R., & Farvardin, M. T. (2017). Effects of different text difficulty levels on EFL learners' foreign language reading anxiety and reading comprehension. *Reading in a Foreign Language*, *29*, 185–202.

Bamford, J., & Day, R. R. (2004). *Extensive reading activities for teaching language*. Cambridge University Press.

Boutorwick, TJ, Macalister, J., & Elgort, I. (2019). Two approaches to extensive reading and their effects on L2 vocabulary development. *Reading in a Foreign Language*, *31*, 150–172.

Burgh-Hirabe, R., & Feryok, A. (2013). A model of motivation for extensive reading in Japanese as a foreign language. *Reading in a Foreign Language*, *25*, 72–93.

Chiang, M-H. (2016). Effects of varying text difficulty levels on second language (L2) reading attitudes and reading comprehension. *Journal of Research in Reading*, *39*, 448–468. https://doi.org/10.1111/1467-9817.12049

Cho, K.-S., & Krashen, S. D. (1994). Acquisition of vocabulary from the Sweet Valley Kids series: Adult ESL acquisition. *Journal of Reading*, *37*, 662–667. doi: 10.2307/20172388

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

Day, R. R., & Bamford, J. (2002). Top ten principles for teaching extensive reading. *Reading in a Foreign Language*, *14*, 136–141.

Deci, E. L., & Flaste, R. (1996). *Why we do what we do: Understanding self-motivation*. Penguin Books.

Feather, N. T. (1982). Expectancy-value approaches: Present status and future directions. In N. T. Feather (Ed.) *Expectations and actions: Expectancy-value models in psychology* (pp. 395–420). Lawrence Erlbaum.

Hedgcock, J. S., & Ferris, D. R. (2009). *Teaching readers of English students: Students, texts and contexts.* Routledge.

Hoffman, J. V. (2017). Comprehension is not simple: Considering the persisting dangers in the simple view of reading comprehension. In S. E. Israel (Ed.), *Handbook of research on reading comprehension* (2nd ed., pp. 57–69). The Guildford Press.

Huffman, J. (2014). Reading rate gains during a one-semester extensive reading course. *Reading in a Foreign Language*, *26*(2), 17–33.

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

Judge, P. B. (2011). Driven to read: Enthusiastic readers in a Japanese high school's extensive reading program. *Reading in a Foreign Languag*e, *23*, 161–186.

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

Kargar, A. A. (2012). The efficiency of extensive reading project (ERP) in an Iranian EFL context. *Theory and Practice in Language Studies*, *2*, 165–169. doi:10.4304/tpls.2.1

Kintsch, W. (1998). *Comprehension: A paradigm for cognition*. Cambridge University Press.

Kirchhoff, C. (2013). L2 extensive reading and flow: Clarifying the relationship. *Reading in a Foreign Language*, *25*, 192–212.

Krashen, S. (1989). We acquire vocabulary and spelling by reading: Additional evidence for the Input Hypothesis. *The Modern Language Journal*, *73*, 440–464. https://doi.org/10.1111/j.1540-4781.1989.tb05325.x

Krashen, S. (1994a). The input hypothesis and its rivals. In N. C. Ellis (Ed.), *Implicit and explicit learning of languages* (pp. 45–77). Academic Press.

Krashen, S. (1994b). The pleasure hypothesis. *Georgetown University round table on languages and linguistics, 1994: Educational linguistics, cross-cultural communication, and global interdependence* (pp.299–302). Georgetown University Press.

Krashen, S. (2004). *The Power of reading: Insights from the research*. Heinemann.

Krashen, S. D. (1985). *The input hypothesis.* Longman.

Lai, F.-K. (1993). The effect of a summer reading course on reading and writing skills. *System*, *21*, 87-100. https://doi.org/10.1016/0346-251X(93)90009-6

Lake, J., & Holster, T. (2014). Developing autonomous self-regulated readers in an extensive reading program. *Studies in Self-Access Learning Journal*, *5*, 394–403. https://sisaljournal.org/archives/dec14/lake_holster/

McLean, S., & Rouault, G. (2017). The effectiveness and efficiency of extensive reading at developing reading rates. *System*, *70*, 92–106. https://doi.org/10.1016/j.system.2017.09.003

Mermelstein, A. D. (2015). Improving EFL learners' writing through enhanced extensive reading. *Reading in a Foreign Language*, *27*, 182–198.

Mikami, A. (2017). Students' attitudes toward extensive reading in the Japanese EFL context. *TESOL Journal*, *8*, 471–488. https://doi.org/10.1002/tesj.283

Mori, S. (2002). Redefining motivation to read in a foreign language. *Reading in a Foreign Language*, *14*, 91–110. https://doi.org/10.1002/tesq.157

Nation, I. S. P. (2009). *Teaching ESL/EFL reading and writing*. Routledge.

Nunnally, J. C., & Bernstein, I. H. (1967). *Psychometric theory* (3rd ed.). McGraw-Hill.

Park, J. (2016). Integrating reading and writing through extensive reading. *ELT Journal*, *70*, 287–295. https://doi.org/10.1093/elt/ccv049

Pinnell, G. S. (1985). Helping teachers help children at risk: Insights from the Reading Recovery Program. *Peabody Journal of Education, 62*(3), 70–85. https://doi.org/10.1080/01619568509538485

Rodrigo, V., Greenberg, D., & Segal, D. (2014). Changes in reading habits by low literate adults through extensive reading. *Reading in a Foreign Language*, *26*(1), 73–91.

Saito, Y., Horwitz, E. K., & Garza, T. J. (1999). Foreign language reading anxiety. *The Modern Language Journal*, *83*, 202–218. https://doi.org/10.1111/0026-7902.00016

Sakurai, N. (2017). The relationship between the amount of extensive reading and the writing performance. *The Reading Matrix*, *17*(2), 142–164.

Samuels, S. J. (1994). Toward a theory of automatic information processing in reading, revisited. In R. B. Ruddell, M. R. Ruddell, & H. Singer (Eds.), *Theoretical models and processes of reading* (4th ed.) (pp. 816-837). International Reading Association.

Schiefele, U. (1991). Interest, learning, and motivation. *Educational Psychologist*, *26*, 299–323. https://doi.org/10.1080/00461520.1991.9653136

Schraw, G., & Lehman, S. (2001). Situational interest: A review of the literature and directions for future research. *Educational Psychology Review*, *13*, 23–52. https://doi.org/10.1023/A:1009004801455

Shih, Y.-C., Chern, C.-L., & Reynolds, B. L. (2018). Bringing extensive reading and reading strategies into the Taiwanese college classroom. *Reading in a Foreign Language*, *30*, 130–151.

Spack, R. (2004). The acquisition of academic literacy in a second language: A longitudinal case study, updated. In V. Zamel, & R. Spack (Eds.), *Crossing the curriculum: Multilingual learners* (pp. 19–46). Lawrence Erlbaum.

Springer, S. E, Dole, J. A., & Hacker, D. J. (2017). The role of interest in reading comprehension. In S. E. Israel (Ed.), *Handbook of research on reading comprehension* (pp.519–542). The Guilford Press.

Suk, N. (2016). The effects of extensive reading on reading comprehension, reading rate, and vocabulary acquisition. *Reading Research Quarterly*, *52*, 73–89. https://doi.org/10.1002/rrq.152

Sweller, J. (2010). Element interactivity and intrinsic, extraneous, and germane cognitive load. *Educational Psychology Review*, *22*, 123–138. https://doi.org/10.1007/s10648-010-9128-5

Tabata-Sandom, M. (2017). L2 Japanese learners' responses to translation, speed reading, and 'pleasure reading' as a form of extensive reading. *Reading in a Foreign Language*, *29*, 113–132.

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

Taylor, B. M., Frye, B. J., & Maruyama, G. M. (1990). Time spent reading and reading growth. *American Educational Research Journal*, *27*, 351–362. https://doi.org/10.3102/00028312027002351

Verhoeven, L. (2011). Second language reading acquisition. In M. L. Kamil, P. D. Pearson, E. Birr Moje, & P. P. Afflerbach (Eds.), *Handbook of reading research, Volume IV* (pp. 661–683). Routledge.

Wan-a-rom, U. (2010). Graded reading vocabulary-size tests for placement in extensive reading. *The Reading Matrix*, *10*, 15–44.

Wan-a-rom, U. (2012). The effects of control for ability level on EFL reading of graded readers. *English Language Teaching*, *5*, 49–61. http://dx.doi.org/10.5539/elt.v5n1p49

Waring, R. (2015, January 15). *The extensive reading foundation graded reader level scale: Chart*. Extensive Reading Foundation. Retrieved from http://www.robwaring.org/er/scale/ERF_levels.htm

Wigfield, A., & Guthrie, J. T. (1995). *Dimensions of children's motivations for reading: An initial study* (Reading Report No. 34). National Reading Research Center.

Yamashita, J. (2008). Extensive reading and development of different aspects of L2 proficiency. *System*, *36*, 661–672. https://doi.org/10.1016/j.system.2008.04.003

Yamashita, J. (2013). Effects of extensive reading on reading attitudes in a foreign language. *Reading in a Foreign Language*, *25*, 248–263.

Yoshida, M. (2012). The interplay of processing task, text type, and proficiency in L2 reading. *Reading in a Foreign Language*, *24*, 1–29.

Appendices

Appendix A

GRLS Level Assignment and Group Assignment to Corresponding Graded Book Levels

GRLS level assignment		Group assignment		Graded book level
Vocabulary size	GRLS level	i-1 Assignment	*i*+1 Assignment	Graded readers (Macmillan) (*Penguin)
301-400	4			1 starter *2 elementary
401-600	5	↗	↘	1 starter *2 elementary
601-800	6	↗	↘	2 beginner *2 elementary
801-1000	7	↗	↘	2 beginner *3 pre-intermediate
1001-1250	8	↗	↘	3 elementary *3 pre-intermediate
1251-1500	9	↗	↘	4 pre-intermediate *3 pre-intermediate
1501-1800	10	↗	↘	4 pre-intermediate * 4 intermediate
1801-2100	11	↗	↘	5 Intermediate * 4 intermediate
2101-2400	12	↗	↘	6 Advanced *5 upper-intermediate
2401-2700	13	↗		6 Advanced *5 upper-intermediate

Note. ↗ indicates assignment to one level below; ↘ assignment to one level beyond.
*: Penguin Graded Readers.

Appendix B

The 3-Factor Solution of the Reading Motivation Questionnaire

Factor	Item	Factor Loadings		
		1	2	3
Self-efficacy *Cronbach's alpha = .86* *% of variance explained: 43.62*	3. I know I can read fast and fluently.	.899		
	2. I know I can deal with difficult texts.	.857		
	4. I enjoy the challenge of difficult reading passages.	.788		
	1. I am good at reading English.	.706		
Reading engagement *Cronbach's alpha = .83* *% of variance explained: 16.53*	10. I make pictures in my mind when I read.		.818	
	9. I get immersed in interesting stories even if they are written in English.		.789	
	5. I like it when the questions in the books made me think.		.751	
	11. I tend to get deeply engaged when I read in English.		.730	
Reading avoidance *Cronbach's alpha = .71* *% of variance explained: 9.42*	7. Even though the content is simple, I do not have the desire to read in English.			.528
	6. Long and difficult passages put me off.			.831
	8. It's a pain to read in English.			.694

Note. 1= Self-efficacy, 2 = Reading engagement, 3 = Reading avoidance.

Appendix C

ANCOVA Summary Results on the Posttest for Reading Comprehension and for Reading Motivation

	Pretest (Covariate)		Group	
	F	η_p^2	*F*	η_p^2
Reading Comprehension				
Total	274.24***	.70	9.11***	.14
Literal	74.44***	.39	12.50***	.18
Inferential	162.99***	.58	1.57	.03
Reading Motivation				
Overall	76.98***	.40	4.34*	.07
Self-efficacy	64.91***	.36	6.52**	.10
Reading Engagement	20.70***	.15	4.32*	.07
Reading Avoidance	36.54***	.24	3.08*	.05

Note. Degree of freedom for Pretest: (1, 116); for Group: (2, 116). * $p < .05$, ** $p < .01$, *** $p < .001$.

About the Authors

Yang, Ya-Han received her M.A. in TEFL from National Taiwan Normal University. She teaches English at The Affiliated Tao-Yuan Agricultural & Industrial Senior High School of National Taipei University of Technology.
E-mail: icemorita@hotmail.com

Chu, Hsi-Chin, the corresponding author of this article, is an Associate Professor in the Department of English, National Taiwan Normal University. Her research interests include extensive reading, visual supports of reading, and cultural issues in written communication.
E-mail: hcchu@ntnu.edu.tw

Tseng, Wen-Ta is an Associate Professor in the Department of Applied Foreign Languages, National Taiwan University of Science & Technology. His current research interests include individual differences in language learning and psychometric analysis of language learning data.
E-mail: wenta.tseng@gmail.com