

The Perceptions of Students and Teachers on the Practice of Assessment in Extensive Reading

Misook Kim*

Kim, Misook. (2019). The perceptions of students and teachers on the practice of assessment in extensive reading. *English Teaching*, 74(4), 179-203.

This paper examined the use of assessment techniques in extensive reading and students' and teachers' perceptions on the practice of extensive reading. A total of 232 university students who were placed at different proficiency levels of classes and 13 professors participated in the study by completing a questionnaire. In-depth interviews were carried out with a selected number of student and professor volunteers. Data were analyzed using *t*-tests to compare two different groups of students. Results revealed that writing short answers and assessments involving discussion with peers and professor were utilized more at the basic level, but students at the basic level reported discussion with peers and professor as the most effective assessment technique. In contrast, at the intermediate level, multiple-choice questions and T/F statements were used more as assessments, but students at the intermediate level reported that writing one's opinion on themes was the most effective technique. Both groups believed that reading graded readers is helpful in improving reading skills and vocabulary as well as understanding other cultures. Results further revealed that professors preferred to use writing short-answer questions and in-class discussion as assessments on graded readers; however, they responded that a writing task was the most adequate type of assessment. In addition, they reported that graded readers are helpful in improving all language skills and favored the use of graded readers as part of the language program.

Key words: extensive reading, graded readers, assessment, extrinsic motivation, affective domain

1. INTRODUCTION

The extensive reading (ER) approach, which has been widely accepted by researchers

* **Misook Kim**, Professor, Humanitas College, Kyung Hee University

and teachers, means reading a large quantity of easy reading materials by picking books that learners are interested in. Davis (1995) states that ER can provide “far more ‘good’ English than, unassisted, the English teacher could ever hope to do” (p. 329), and it can be more enjoyable than most English lessons. Empirical studies have demonstrated that ER contributes to successful language learning skills, including reading comprehension (Bell, 2001; Grabe, 2009; Mason & Krashen, 1997; Yamashita, 2008), vocabulary acquisition (Hurst, 2005; Nation, 2009; Pigada & Schmitt, 2006), speaking (Cho & Krashen, 1994), and writing (Ro & Park, 2016). Furthermore, literature has attested that ER has helped students develop positive attitudes toward reading and get motivated to read (Al-Homoud & Schmitt, 2009; Camiciottoli, 2001).

Day and Bamford (1998) put forward ten principles for teaching ER, and one of them is stated simply as “Reading is its own reward.” They discourage comprehension questions but recommend post-reading activities, such as writing about students’ favorite characters and writing about the best or worst book they have read to check what students understood. In the context of EFL, especially in a university setting in Korea, it can be difficult not to evaluate ER in a language program. To be more specific, making college students read voluntarily extra books outside the class in the language that they are learning is challenging. Students have a high load of reading requirements for classes they take, and unless books are a part of the course evaluation, they would not read voluntarily. In addition, often students lack the time to read for pleasure even in their mother tongue, and therefore, they should be given external reasons to read in L2. Mori (2015) argues that when EFL students are not interested in reading in an EFL class, teachers could create an environment to provide extrinsic motivation to help them read. One quick, easy method to make students read books for the course might be by giving some kind of evaluation relevant to a course grade as Mori, Gobel, Thepsiri and Pojanapunya (2010) suggest. Then, the students will read reluctantly in order to fulfill the course requirements and pass the course.

There is a substantial number of empirical studies on ER that have investigated the benefits of ER, but there are very few studies that have looked into assessment techniques in ER especially with respect to different proficiency levels of students when it has been commonly practiced (Stoeckel, Reagan & Hann, 2012). It is believed that proficiency level will affect student preferences of assessment techniques in ER. The present study thereby aims to examine the practice of ER, quantitatively and qualitatively and university students’ and teachers’ perceptions on assessment techniques in ER in a language program by looking into types of assessment used, perceived effectiveness of assessment techniques, and perceived impact on English skills.

2. LITERATURE REVIEW

2.1. Assessment in Extensive Reading

Researchers state that the assessment in ER may have undesirable consequences (e.g., Day & Bamford, 1989; Krashen, 2004; Richards, 1997), but it has been widely used, and some publishers even provide supplementary ER evaluation materials (i.e., quizzes). However, a study by Stoeckel et al. (2012) shows a case that assessment on ER did not influence students' reading attitudes. They investigated the effect of ER quizzes on attitudes of Japanese university students who were placed at different proficiency levels based on the scores on the Global Test of English Communication. Students in both control and experiment group, comprised of two groups each (a group of high-intermediate and intermediate levels of students and a group of low level of students) were required to read 10 graded readers (1 per week) for 10 weeks. Then the control group was required to write eight short responses and give two oral reports whereas the treatment group was requested to do the same and take a weekly quiz on graded readers. At the end of semester, all participants completed the reading attitudes survey that they had done in the beginning of the experiment. They found that there were no differences between the treatment group that was given short quizzes and the control group. That is, ER quizzes did not impact reading attitudes (anxiety, comfort, linguistic value, and practical value) of the treatment group, and students in the treatment group indeed gained intellectual benefits from the quizzes.

In addition, studies argue that students should be held accountable for reading outside classroom to bring about the benefits of ER that literature has attested. For example, van Amelsvoort (2016) acknowledges that reading should be required, and students need to be monitored to develop their intrinsic motivation for reading since students are not willing to read on their own when they are just encouraged to read. In addition, Robb and Kano (2013) who introduced M-Reader, a database of short, timed quizzes on graded readers (GR) to monitor students' reading also argue that students should be held accountable for reading outside classroom by giving an assessment. It is argued that for those students who are reluctant to read, external incentives should be given to them until they have the intrinsic motivation to read for their own (Chang & Renandya, 2017).

2.2. Teachers' Perceptions on Extensive Reading

A study by Chang and Renandya (2017) examined the perceptions of 119 Asian teachers who had been engaged in teaching L2 (116 teach English) on extensive reading and analyzed the responses of online questionnaires that the teachers from Taiwan, Vietnam, Korea, Indonesia and Japan had completed. The results showed the L2 teachers hold a

strong belief that ER is effective in improving students' overall language competence and are in favor of an extensive reading program in the curriculum. However, many participants reported that students' lack of motivation to read is the most challenging problem. Chang and Renandya hence argue that by administering quizzes to track what and how much students read, teachers can hold accountable for their work as Macalister (2010) notes, "If extensive reading is included as part of the teaching program, it may be linked to assessment as a means of ensuring its acceptability to various stakeholders" (p. 71).

For an extensive reading program to be successful, teachers' beliefs and knowledge play an important role. Byun (2011) examined the effects of extensive reading on secondary school English teachers' motivation and confidence in reading. Fourteen teachers participated in the study, and they were required to read 7-10 readers for 2 weeks and were given lectures about ER by teacher trainers. Throughout the in-service training course, the participants also performed various activities such as giving oral reports, book talks and drawing a picture and talking based on the books that they read. The findings show that at the initial stage of the experiment, the majority of the participants showed some resistance to reading and held a negative belief about the effect of ER; however, at the end of experiment, their attitudes have changed positively, and they reported that they gained confidence since they could choose books to read and were willing to implement an ER program in their classroom.

In addition, a study by Macalister (2010) examined the attitudes of teachers who were teaching university preparation courses on ER through questionnaires and phone interviews, and he found that there was a lack of familiarity with ER and a shared understanding about the nature of ER among teachers. On the other hand, Tien (2015), who conducted a study on the perceptions of 5,711 students and 36 teachers of an ER program by asking questions, claimed that although there was some resistance in the beginning stage of the implementation, both the attitudes and beliefs of students and teachers changed positively. Teachers reported that students can learn about different cultures and see things from different perspectives. The results of a study by Chung and Yang (2012) who investigated teachers' beliefs on ER using questionnaires and interviews are in line with this conclusion. The findings show that the prior belief of teachers (that ER would mainly benefit reading skills and the more competent students would enjoy reading more than lower level of students) changed after the experiment, and the researchers observed that ER could help students build up confidence, and not only more competent students would enjoy reading but also the lower level of students would enjoy reading. They also claim that teachers' positive attitudes toward ER play a significant role for the success of the program.

Studies from Stoeckel et al. (2012) and Robb and Kano (2013) showed that assessment in ER do not affect students' attitudes in reading. However, it seems difficult to make a

generalization due to a scarcity of empirical studies which have examined the impact of assessment in ER especially with different proficiency levels of students. Hence in the present study, perceptions of students from different proficiency levels and teachers will be examined focusing on the following research questions.

1. How do students with different levels of English perceive the assessment techniques used for extensive reading?
2. What is students' attitudes on the assessment techniques used for extensive reading and their perceived impact on their English skills through extensive reading?
3. What are professors' attitudes on the assessment techniques used for extensive reading and their perceived impact on students' English skills?

3. METHOD

3.1. Context of the Present Study

In an EFL setting like Korea at a tertiary level of education, a majority of four-year colleges require students to take only one or two English classes at most as a graduation requirement (Kim & Kim, 2017). In the university where the researcher works, there is only one English class offered, and it is mandatory for graduation. In order to make students to be exposed to more input, an ER approach is incorporated as part of the curriculum. Students are required to read a number of GRs as part of the class requirements regardless of the level of class they are enrolled, and they are assessed after reading.

This ER program includes a couple of differences from the traditional ER. In a traditional ER program, students would choose books that they are interested in reading, and they can read as many books as they want. Also there is a variety of materials on a wide range of topics available. However, in the present study, the number of GRs that students need to read is determined by the department in which the mandatory English courses are offered. They need to read four GRs throughout the 16-week semester. Another difference is that teachers who are assigned to teach the mandatory English courses get to select four books for their students from a list prepared by the department. In choosing four GRs, at least one must be a nonfiction book.

The English language program encompasses three levels: basic, intermediate, and advanced. Each level requires one main textbook and four GRs. Teachers are given a list of GRs, consisting of 45 readers with levels from 2 to 6 containing both fiction and nonfiction works. The basic level must use 2 readers each from levels 2 and 3; the intermediate must include 2 readers from level 3 and 2 from level 4; and the advanced level has to select 2

readers from level 5 and 2 from level 6. Assessment of reading GRs is part of the class requirement, and teachers are allowed to select their own preferred assessment of GRs.

3.2. Participants

A total of 232 students (95 males and 137 females) who were enrolled in English classes took part in the study. Among the 232 students, 126 students were from the basic classes, and 106 were from the intermediate classes. They were all freshmen from different departments. In this study, no data from the advanced classes are included due to a small number of classes offered (only 2 classes).

Thirteen professors (3 females and 10 males) who were teaching the required English courses participated in the study. Among the 13 professors, one held a PhD, two were in a PhD program, and the rest of them had a Master's degree in various fields of studies. The number of years that they have been employed by the university ranged from 7 to 18 years. They all had been teaching the required English courses at the time of data collection, using four GRs for each class.

3.3. Instruments

A set of questionnaire for students was developed by the researcher by adopting ideas that met the purpose of the present study based on the previous literature (Macalister, 2010; Stoeckel, Reagan & Hann, 2012). The questionnaire consisted of 25 questions: 1-3 concerned demographic information; 4-6 the use of devices to aid reading; 7-14 the topic of assessment; 15-21 the student's perceived improvement of English skills; 22-24 the affective learning domain; and 25 was an open-ended question about GRs. On a number of questions, the Likert scale, ranging from 1 'strongly disagree' to 5 'strongly agree' was used to collect responses. The reliability of the questionnaire was computed, and the Cronbach's alpha turned out to be .85, which shows a high degree of reliability among the items.

For professors, a questionnaire was also created, which was adopted from the student questionnaire, but it changed the perspectives to make them appropriate for professors. It consisted of 23 questions: 1-3 concerned demographic information; 4 the topic of monitoring reading; 5-7 the use of devices to aid reading; 8-14 the assessment types; 15-21 the perceived impact on students' English skills; and 22-23 were open-ended questions about GRs.

A semi-structured one-to-one interview (Appendix A) was conducted with students ($n = 9$, 6 from the basic and 3 from the intermediate group) and professors ($n = 5$) who volunteered. Interviews with students were conducted in Korean whereas interviews with

professors were carried out in English. The interviews were recorded with participants' consent, transcribed, and translated into English (in the case of student interviews) by the researcher.

3.4. Procedure

The mandatory English classes run for 16 weeks, meeting twice a week for 75 minutes each time. At week 14 of the semester, the researcher emailed 15 foreign faculty members who were teaching the required English courses, explaining the purpose of the proposed research and asking for their cooperation. The email included two web links for online questionnaires (one for students and one for professors). The professors were teaching three to four mandatory English classes, and they were asked to choose one of their classes by themselves for the experiment. Then they were asked to complete two tasks: first to send the web link for the student online questionnaire to one class of theirs, asking their students to complete the questionnaire. Second, they were requested to complete the professor questionnaire themselves (see Appendix B). A total of 232 student questionnaires were collected, and 13 questionnaires were returned from professors. A semi-structured interview was conducted with volunteer students and professors in the office of the researcher, taking about approximately 15 minutes for each participant.

3.5. Data Analysis

The responses on the student and professor questionnaires were analyzed by using SPSS version 23. Then the mean score and standard deviation for each item were computed in order to gain a general view of responses for each group. In the case of the student questionnaires, *t*-tests were performed in order to compare the mean scores of the two different levels of students.

4. RESULTS AND DISCUSSION

4.1. Results of Student Responses

Student responses to the use of devices while reading GRs are shown in Table 1. There were slight differences between in the mean scores of the basic and intermediate groups. The mean scores of both groups for the use of the dictionary were highest followed by the use of the Internet and translator apps. Upon close analysis, the mean score of the basic group for the use of the dictionary was slightly higher than that of the intermediate group ($M = 2.24$, $M = 2.07$). However, the mean difference between the two groups was not

significant ($t = 1.929, p = .055$) concerning the effect size. With respect to the use of translator apps, the mean score of the basic group ($M = 1.75$) was higher than that of the intermediate group ($M = 1.52$). The difference between the two groups ($t = 2.245, p = .026$) was significant at $p < .05$ level showing the effect size of .29, which indicates that the difference between the groups is medial. In regard to the use of the Internet, the mean score for the basic level was 1.91, whereas for the intermediate level it was 1.77, and there was no significant difference ($t = 1.596, p = .112$) between the two groups.

TABLE 1
Use of Devices While Reading

Variable	Level	$N = 232$	M	SD	t	p	d
Use of dictionary	Basic	126	2.238	0.650	1.929	.055	0.25
	Intermediate	106	2.066	0.707			
Use of translator apps	Basic	126	1.754	0.776	2.245*	.026	0.29
	Intermediate	106	1.528	0.745			
Use of the Internet	Basic	126	1.912	0.645	1.596	.112	0.20
	Intermediate	106	1.774	0.680			

* $p < .05$

These results imply that a significant number of students from both groups use a dictionary more than the other two devices when encountering unknown words in reading GRs. Conceivably, students do not feel comfortable facing even a single word that seems to break the flow of a story while reading, or they are accustomed to looking up words. Thus, they use a dictionary while reading GRs. Both groups also utilized the Internet to some extent. Since reading GRs is mostly done outside class as an assignment, it seems they may have used it to search for information on topics such as the authors and background information about the books as Student 4 from the interview reported.

I searched related characters and incidents of The African Queen and Mandela which I hadn't known before. (Student interviewed 4 (I))

Various types of assessments were used with respect to GRs in class as shown in Table 2. In general, assessments that had clear, definite answers, such as multiple-choice questions (MCQ), true or false statements (T/F), and writing short answers (WSA) were used more than assessments that were more complex and cognitively demanding, such as writing a summary (WS) and writing one's opinion on themes (WOT) in both levels. For the basic level, the mean score of WSA was highest ($M = 1.56, SD = .497$), followed by discussion with peers and professor (DPP) ($M = 1.52$), and the mean score of WOT was the lowest ($M = 1.21, SD = .412$). In contrast, the mean score of MCQ for the intermediate group showed the highest level ($M = 1.60, SD = .491$), followed by T/F statements ($M = 1.50$), and the mean

score of WS was the lowest ($M = 1.19$, $SD = .400$).

TABLE 2
Types of Assessment Used on GRs in Class

Variable	Level	$N = 232$	M	SD	t	p	d
Multiple-choice questions (MCQ)	Basic	126	1.500	.502	-1.584	.115	.20
	Intermediate	106	1.603	.491			
True or false statements (T/F)	Basic	126	1.484	.501	-.240	.811	.03
	Intermediate	106	1.500	.502			
Writing short answers (WSA)	Basic	126	1.563	.497	1.394	.165	.18
	Intermediate	106	1.471	.501			
Writing a summary (WS)	Basic	126	1.317	.467	2.067*	.040	.26
	Intermediate	106	1.198	.400			
Writing one's opinion on themes (WOT)	Basic	126	1.214	.412	-1.096*	.047	.27
	Intermediate	106	1.330	.473			
Discussion with peers and professor (DPP)	Basic	126	1.516	.501	3.197**	.002	.40
	Intermediate	106	1.311	.466			

* $p < .05$, ** $p < .01$

An interesting point to note is that the mean score of discussion with peers and professor (DPP) for the basic group ($M = 1.51$) was slightly higher than that of the intermediate group ($M = 1.31$). The difference in the mean scores between the two groups was significant at $p < .05$ ($t = 3.197$, $p = .002$) with the effect size of .40. This fact indicates that DPP was used more often in the basic classes than in the intermediate classes. It is to be expected that carrying out discussion with peers and professors would be more challenging for students at the basic level than the intermediate level. Since professors are given authority to use any assessment they wish to, they may prefer such assessment though it may be challenging to their students, thus disfavoring MCQ, T/F statements, and WSA.

Both student groups showed favorable responses to statements that seek the perceived effectiveness of different types of assessments on GRs with a slight difference in mean scores as shown in Table 3. The mean score for each assessment was higher than a score of 3.0, except for the mean score of WS ($M = 2.92$) for the intermediate group. For the basic group, the mean score of DPP was the highest ($M = 3.62$) and WS was the lowest ($M = 3.05$). On the other hand, for the intermediate group, the mean score of WOT was the highest ($M = 3.64$) with WSA being the lowest ($M = 2.92$).

These results imply that students at the basic level believe that DPP is a more effective assessment technique than WS. In contrast, the intermediate group believes that WOT is more effective with WS being the least effective. Writing a summary requires a good understanding of a book first in order to give an overview of a book to the reader, and this task is much more complex and challenging than discussing or writing one's opinion. Therefore, students might disapprove it. It seems that an assessment that is less burdensome on memory is preferred by students.

TABLE 3
Perceived Effectiveness of Each Assessment Type

Variable	Level	<i>N</i> = 232	<i>M</i>	<i>SD</i>	<i>t</i>	<i>p</i>	<i>d</i>
Multiple-choice questions	Basic	126	3.341	1.089	.535	.593	.07
	Intermediate	106	3.264	1.098			
True/false statements	Basic	126	3.373	1.033	.910	.364	.11
	Intermediate	106	3.245	1.102			
Writing short answers	Basic	126	3.389	1.050	2.355**	.019	.30
	Intermediate	106	3.047	1.158			
Writing a summary	Basic	126	3.056	1.189	.840	.402	.11
	Intermediate	106	2.925	1.176			
Writing one's opinion on themes	Basic	126	3.516	1.101	-.852	.395	.10
	Intermediate	106	3.642	1.140			
Discussion with peers and professor	Basic	126	3.620	1.057	1.503	.134	.19
	Intermediate	106	3.396	1.200			

** $p < .01$

In examining the mean differences between the two groups with respect to the perceived effectiveness of different types of assessments, there were no significant differences, except for WSA. The mean difference between the two groups was significant at $p < .05$ ($t = 2.35$, $p = .019$) with the effect size of .30. This fact indicates that the basic level of students preferred the WSA technique more than the intermediate group, even though they did not think it was the most effective assessment technique. In the interviews, similar responses from both groups were found as shown below.

I like T/F or MCQ because they give me less burden to read books. Reading the GRs is time-consuming and it itself is burdensome, so I prefer T/ F and MCQ. (Student interviewed 2 (I))

I like T/F. Writing a summary or opinion might be good, but since the professor will read my writing, I can get either a high or low point depending on how she interprets. (Student interviewed 8 (B))

One follow-up question in the questionnaire asked students to choose one assessment that they preferred and to provide a reason for their choice. Their responses were categorized by level and listed in the order of frequency (most-mentioned), excluding items receiving less than five occurrences in Table 4.

For the basic level, both writing short answers and discussion with peers and professor were preferred most, followed by writing one's opinion on themes, while for the intermediate level, only discussion with peers and professor was the most favorable assessment type, with writing short answers being least favorable. Student 134 (Basic) explained why he preferred discussion with peers and professor and writing one's opinion

on themes.

Reading GRs is not to study the content of the books. It is to improve reading skills. Thus asking detailed information on a test goes against the purpose of readers. To check whether or not students have read the book, T/F statements are okay, but asking detailed information will make us to memorize the books. So I think writing one's opinion on themes or discussing the books would be more appropriate. (Student 134)

TABLE 4
Student Preferred Assessment Type

In order of most mentioned	Basic	Intermediate
1	Writing short answers Discussion with peers and professor	Discussion with peers and professor
2	Writing one's opinion on themes	Multiple-choice questions
3	T/F statements	Writing one's opinion on themes
4	Multiple-choice questions	Writing a summary
5	Writing a summary	T/F statements
6		Writing short answers

On the other hand, Student 78 (Intermediate) showed a completely opposite view from Student 134.

Writing short answers is the most reasonable assessment on GRs. In the case of MCQ, they required memorization of specific details, so I don't think they were effective for reading skills. Also, asking students' opinion is not suitable because it doesn't check whether or not they read the book. Therefore, I prefer writing short answers that do not require very specific details. (Student 78)

It is interesting to observe that writing short answers and discussion with peers and professor were mostly used at the basic level while MCQ and T/F statements were used most at the intermediate level. However, a type of assessment that both groups prefer most is discussion with peers and professor. This fact seems to imply two points. First, students want opportunities to express their ideas and thoughts. Second, they prefer sharing their ideas with their peers and listening to their peers. This is likely to lessen their cognitive burden for the assessment. By listening to their peers, students can come up with ideas to share and participate in discussion and debate.

Table 5 below shows descriptive statistics of student responses concerning the perceived impact on English skills. Both groups displayed similar opinions; they perceived that

reading GRs is helpful in improving reading skills most ($M = 3.68$ for the basic, 3.64 for the intermediate), followed by learning other cultures ($M = 3.46, 3.48$) and vocabulary learning ($M = 3.28, 3.19$). On the other hand, they perceived reading GRs is not very helpful in improving grammar ($M = 2.78, 2.77$). Nonetheless, the mean scores of both groups concerning the statement “Reading GRs was helpful in improving my ability.” were 3.3 and 3.2 respectively.

TABLE 5
The Perceived Impact on Language Improvement

Variable GRs are helpful in ...	Level	$N = 232$	M	SD	t	p	d
Learning new words	Basic	126	3.286	.962	.670	.504	.08
	Intermediate	106	3.198	1.027			
Learning grammar	Basic	126	2.786	1.000	.092	.926	.01
	Intermediate	106	2.774	.988			
Improving reading skills	Basic	126	3.682	.854	.354	.724	.04
	Intermediate	106	3.642	.907			
Improving writing skills	Basic	126	2.944	.998	-.061	.951	.00
	Intermediate	106	2.952	1.081			
Improving speaking skills	Basic	126	2.746	1.087	-.841	.401	.01
	Intermediate	106	2.868	1.113			
Learning other cultures	Basic	126	3.468	1.025	-.095	.925	.01
	Intermediate	106	3.481	1.044			
Improving my ability	Basic	126	3.389	.937	.936	.350	.01
	Intermediate	106	3.273	.931			

In examining the mean differences between the basic and intermediate groups with respect to five skills, there were no significant differences, and the effect size for each variable was small as shown in the table. However, the interview with Student 9 (Basic) revealed a very different benefit that was not mentioned on the questionnaire.

GRs are not helpful in learning English. They just gave me opportunities to read masterpieces in a short version, and perhaps this will lead students to read the original books. (Student interviewed 9)

In sum, students believe reading GRs helps them to learn words and improve their reading skills; however, they did not hold the same beliefs regarding other skills, such as grammar, speaking, and writing though research has shown that the impact of ER on writing, grammar, and speaking is noticeable (Bell, 2001; Cho & Krashen, 1994; Ro & Park, 2016). Nevertheless, in general, they were positive about the improvement of their language ability.

In examining the affective domain with regard to reading GRs, both groups showed somewhat positive responses as shown in Table 6. Concerning the enjoyment of reading GRs, both groups held a slightly positive attitude with the mean scores of 3.17 for the basic and 3.18 for the intermediate group. With respect to an aspect of a sense of achievement through GRs, the mean scores of both groups were considerably high ($M = 3.48$ each). However, the mean scores of both groups for the statement, "I've gained confidence through reading GRs" were 2.97 for the basic group and 3.07 for the intermediate group, showing the lowest level. The fact that students felt a sense of achievement is an important point to note for language teachers because having a sense of achievement will push students to challenge themselves to read more and to relieve the burden of reading the next books.

TABLE 6
The Affective Domain on Reading GRs

Variable	Level	$N = 232$	M	SD	t	p	d
I enjoyed reading GRs.	Basic	126	3.175	.877	-.039	.969	.00
	Intermediate	106	3.180	.933			
I felt a sense of achievement after each book.	Basic	126	3.484	1.056	.090	.928	.01
	Intermediate	106	3.481	1.034			
I gained confidence through reading GRs.	Basic	126	2.976	.933	-.773	.440	.10
	Intermediate	106	3.075	1.020			

A study by Byun (2011) showed a similar result and emphasized the importance of confidence in language learning. She states that confidence will be a driving force for students to challenge something that they have been reluctant to do or been afraid of doing. Students 1 (Intermediate) and 9 (Basic) in the interview expressed how their feeling has changed.

When I first had to read the GRs, I felt burdensome because I had never read English books except textbooks. As I read the second and third book, my burden for reading decreased. (Student interviewed 1)

My reluctance to read English books has decreased. At first, I was a bit worried if I could finish the book by the deadline, but when I started reading, I could understand and it was easy to read. When I read the second one, I felt less burdensome. (Student interviewed 9)

One open-ended question on the questionnaire sought to disclose students' perspectives on the strengths and weaknesses of reading GRs. Although not all participants expressed their opinions, more strengths were presented than weaknesses. The top five recurrent

responses are displayed in Table 7 below.

TABLE 7
Strengths and Weaknesses of GRs

	Strengths		Weaknesses	
	Basic	Intermediate	Basic	Intermediate
1	Language improvement (reading, vocabulary, grammar)	Language improvement (reading, writing, vocabulary)	Books are expensive.	Some books are boring to read.
2	Stories are interesting.	GRs cover various topics.	GRs are not helpful for improving speaking.	I can't choose a book on my own.
3	Books are short.	GRs made me read English books.	Four books per semester are too much.	GRs are not helpful for speaking.
4	GRs are good to learn other cultures and knowledge.	GRs are easy to read.	Some books were boring.	They are difficult.
5	Books are easy to read.	GRs are stories-based.	GRs are difficult.	

Both groups reported that one big advantage of GRs lies in improvement of language skills, especially reading and vocabulary. Another advantage the basic level considered was that stories were interesting and short. For the intermediate level, GRs cover various topics, and the reading requirements gave them a chance to read English books. With respect to weaknesses, both groups showed similar ideas though with different frequencies. The basic level mentioned that books are expensive, and GRs are not helpful in improving speaking. In contrast, the intermediate level remarked that some GRs are boring, and they can't choose a book by themselves.

This result needs to be interpreted with caution. Unlike the advantages of GRs reported by both groups, the frequencies of weaknesses for each item were far fewer, and thus, it should not be understood as representative of both groups.

4.2. Results of Professor Responses

Results showed that professors use various ways to monitor students' reading of GRs. Five professors reported that they use a combination of quizzes, in-class discussions and writing assignments while three professors responded they give only quizzes. One professor stated that he monitors students by distributing questions in class, and another one mentioned that students' comprehension is checked through assignments. Both did not further specify the exact tasks they required students to perform.

Table 8 below shows the responses of professors' perceptions on the use of devices when students read GRs. The mean score for the use of the dictionary was considerably high ($M = 4.15$), followed by the use of the Internet ($M = 3.61$). The mean score for the use

of translation apps was lowest ($M = 2.62$). The results of the professors' responses were similar to those of the students. It is intriguing to note that the professors think it is okay for students to use a dictionary while reading GRs but not translation apps. However, this fact must be interpreted with caution. The interviews reported clearly what the professors meant with regard to the use of dictionary.

I don't want my students to use a dictionary immediately when they come across unknown words. They first need to use the context to figure the meanings. If they can't get them, then they should use a dictionary. (Professor interviewed 1)

I didn't give any instruction on the use of a dictionary. Hope they use a dictionary when they finish the whole page rather than one sentence after another. But I think they should use context to guess the meanings. (Professor interviewed 4)

TABLE 8

Use of Devices While Students Are Reading GRs

Variable	<i>N</i>	<i>M</i>	<i>SD</i>	Min.	Max.
It is okay to look up words while reading GRs.	13	4.15	.800	3	5
It is okay to use translation apps while reading GRs.	13	2.62	1.386	1	5
It is okay to use the Internet while reading GRs.	13	3.61	1.120	1	5

Both professors and students considered that looking up words while reading GRs is acceptable. However, many studies have claimed that guessing the meaning of words is a very effective strategy to teach (Nation, 2009). Perhaps professors need to explain the purpose of ER to students and not knowing a couple of words in a page would not be a huge problem in understanding the book as a whole.

Various types of assessments were used by the professors as shown in Table 9. The mean score for writing short answers (WSA) was highest ($M = 4.54$) followed by giving discussion questions (DQ) ($M = 3.31$). The mean scores for giving writing prompts to write opinions and writing a summary were relatively higher ($M = 3.15, 3.0$) than MCQ and T/F statements. Although students preferred to be given objective questions, such as MCQ and T/F on GRs, professors were in favor of only WSA and DQ. The results undoubtedly imply that professors seem to ensure that students read GRs by requiring the writing of short answers, an approach that is easier to create than other types of questions. In addition, it seems professors are in favor of in-class discuss questions assessment on GRs. Perhaps, they strive to give opportunities for students to express their ideas and thoughts through discussion.

TABLE 9
Types of Assessment Used

Variable	<i>N</i>	<i>M</i>	<i>SD</i>	Min.	Max.
Multiple-choice questions (MCQ)	13	2.53	2.025	1	5
True/false statements (T/F)	13	2.08	1.754	1	5
Writing a short answer questions (WSA)	13	4.54	1.198	1	5
Writing a summary (WS)	13	3.00	1.825	1	5
In-class discussion questions (DQ)	13	3.31	1.974	1	5
Writing prompts (WP)	13	3.15	1.908	1	5

Among the assessments, professors were asked to select and write what they believe constitutes the most adequate type of assessment. Their responses varied but coincided well with the results of the questionnaire. As shown in Table 10, some responses mentioned that a form of writing assessment such as writing a summary or a book report and writing opinions is most adequate to use for GRs.

TABLE 10
Most Adequate Type of Assessment

Profs	Response
Prof 2	Writing. It gives the students the opportunity to consider the book in more detail.
Prof 3	Multiple choice. If they read the book and did the questions in the back and were present for class discussion, they usually do well. If they didn't, they obviously do poorly.
Prof 7	I prefer discussion because I tend to emphasize conversation and pair-work while in class.
Prof 10	Writing/ discussion. Relevant to students' own lives, and helps with writing skills.
Prof 11	I think a combination is best, but if I had to choose one, I'd say writing a summary to show they've understood the gist of the material and then allow them to express their opinion on the material as an adjunct to the summary.

The responses of professors' perceived impact on student English skills showed positive sentiment in all language skills as shown in Table 11. The mean score for improvement of reading skills was highest ($M = 4.23$) followed by improvement of vocabulary ($M = 4.08$) and learning other cultures ($M = 3.92$). This fact is in line with the results of the students. In addition, the mean scores for improvement of writing and speaking, however, were slightly lower ($M = 3.70, 3.77$) than the other skills, yet they were higher than those of the students.

In general, professors responded, like students, that reading GRs is helpful in improving reading skills and vocabulary learning. In addition, they stated that it is helpful in understanding other cultures and improving grammar, speaking, and writing. Furthermore, they reported that reading GRs can aid students to gain confidence.

TABLE 11
Benefits of Reading GRs

Variable	<i>N</i>	<i>M</i>	<i>SD</i>	Min.	Max.
Reading GRs can enhance students' reading skills.	13	4.23	1.012	2	5
Reading GRs can enhance students' speaking skills.	13	3.77	.725	3	5
Reading GRs can improve the vocabulary power of students.	13	4.08	1.115	2	5
Reading GRs can improve students' writing skills.	13	3.70	1.109	2	5
Reading GRs can help learning grammar.	13	3.85	1.068	2	5
Reading GRs can help students learn other cultures.	13	3.92	.640	3	5
Students can gain confidence through reading GRs.	13	4.00	.816	2	5

In asking professors' perspectives on the strengths and weaknesses of the use of GRs, they expressed various opinions as presented in Table 12. It seems that one strength many professors consider is that GRs cause students to be exposed to different genres that students may not have experienced before in L2. Concerning weaknesses, there were varied responses from having a limited book list, students disliking reading, and GRs being not authentic. Notwithstanding the weaknesses, Professor 13 reported that they would keep using GRs even if they would become an option in the program, for they give a new experience to students. As Professor 1 said, "The students come alive with GRs in a way that they do not with the main textbook."

TABLE 12
Strengths and Weaknesses of GRs

Profs	Strengths	Weaknesses
Prof 3	I believe it gives students a chance to read interesting stories and other material in English that, if left to themselves, they may not bother to try reading.	I don't think there are any significant disadvantages unless one uses up a lot of class time making them read when that time could be used for other purposes.
Prof 5	GRs can help students fulfill the basic concept of ER for vocabulary attainment as well as the benefit to have other material than just the textbook to enrich the students' experience in class.	At times it may not be related well with the textbook's topic.
Prof 8	The main textbook can lack a variety of genres or rhetoric while GRs can be a chance for students to discover the joy of easy reading and be motivated to read the original novel.	The limited choice of GRs to choose from and they are modified. Since the readers are short and modified, the sentences are choppy and may miss the point of reading authentic texts.

5. CONCLUSIONS AND IMPLICATIONS

The present study strove to investigate the practice of ER and students' and teachers'

perceptions about the assessment techniques of reading GRs. The results reveal that the types of evaluation used by professors for the readers showed a difference between the two levels. For the basic level, writing short answers and discussion with peers and professor were more frequently utilized while at the intermediate level, multiple-choice questions and T/F statements were more frequently used. It can be conjectured that a certain type of assessment is more likely to be used, depending on the professors' teaching experience and philosophy. Professors of the intermediate level of class presumably wanted to test whether or not students have read a book by giving objective tests. In addition, they perhaps wanted a simple assessment that is less time-consuming to grade.

In asking about students' perceived effectiveness of each assessment type, the basic level students stated that discussion with peers and professor was the most effective type; on the other hand, the intermediate group believed that writing one's opinion on themes was the most effective technique. Although the two groups favored a different type of assessment, both approaches required the use of productive skills (speaking and writing). This fact seems to imply that students believe that an assessment that requires an expression of their ideas and thoughts orally or in writing is an effective assessment technique. Undoubtedly, students from both levels prefer discussion with peers and professor as an assessment for readers, for presumably it seems less challenging than the other assessments, and they believe that they can learn from their peers by listening to their thoughts as reported in the interviews.

Both levels of students believe that reading GRs is helpful in improving reading skills and vocabulary and understanding other cultures. Furthermore, they reported that they felt a sense of achievement after finishing each book and that they enjoyed the reading. This conclusion is supported by the strengths of GRs that the two groups reported. They stated that readers were interesting and helpful in improving language skills (mainly reading and vocabulary). In addition, the books were short and covered a variety of topics. They also liked the fact that they were given a chance to read something different, something that they had never read. This experience itself proved rewarding to students and gave more input to their learning.

A significant number of professors stated that using a dictionary was acceptable when students read GRs, and the professors were favorable toward the use of the Internet as well. In addition, the majority of professors responded that a form of writing is the most desirable type of assessment to help students enhance their critical thinking skills by analyzing, reflecting, and linking and expressing their thoughts. When asked about the benefits of reading GRs, they held a positive attitude, indicating that readers were helpful in improving all skills and in building confidence. In addition, additional strengths of readers were addressed by professors; they like readers because they give students a chance to read, and they offer various genres and a variety of topics. Although they voiced the

view that the selection of readers was too limited, and the number of books was excessive, they said that they would keep using readers even if they would become an option in the program, because the GRs helped students to explore different genres and topics that they had never experienced. Nevertheless, the results of the professors revealed that there was a lack of understanding about the nature of ER. According to the principles of ER, dictionaries should be rarely used when reading. However, many professors in the present study think it is okay to use dictionary while reading. It seems teacher education or training on the topic of ER seems imperative as teachers' prior learning experiences influence all aspects in teaching, such as procedures, materials, classroom interaction patterns and their roles in class (Borg, 2003; Ebsworth & Schweers, 1997).

The findings of the present study illuminate a number of points to take into consideration before implementing an ER approach in a language program in the future. As described earlier in the present study, the department where the researcher works unquestionably needs some modifications to facilitate a more successful program and to provide a greater satisfaction to both students and professors. In an EFL education setting like a tertiary level, it is often logistically difficult to implement the traditional ER program that the literature suggests (e.g., Day & Bamford, 1998) because schools may not have enough budget to buy over thousands GRs for students to read, and students take more than five classes per semester, and this means that they literally do not have time to spare to read for pleasure. Therefore, some compromises in the program seem necessary. For example, the school could provide a longer list of readers for professors to select from when it is not feasible for students to select books to read. Furthermore, when assessment of GRs is an integral part of the program, a variety of assessment techniques that include speaking and writing tasks would be desirable as students are eager to have opportunities to express their ideas and thoughts. Then students would engage in reading more and obtain a greater satisfaction from the books.

The present study examined the perceptions of students with different proficiency levels and teachers on the practice of assessment in ER. The findings of this study are in line with the studies by Stoeckel et al. (2012) and Chang and Renandya (2017). That is students from basic and intermediate levels hold positive views on ER, teachers are in favor of ER, and they are aware of the benefits of ER. However, not only the previous studies but also the present study did not investigate the effectiveness of different assessments using actual scores that students obtained. In a follow-up study, examining the scores of different types of assessments on GRs would seem like a good direction to take.

REFERENCES

- Al-Homoud, F., & Schmitt, N. (2009). Extensive reading in a challenging environment: A comparison of extensive and intensive reading approaches in Saudi Arabia. *Language Teaching Research, 13*(4), 383–401.
- Bell, T. (2001). Extensive reading: Speed and comprehension. *The Reading Matrix, 1*(1), 1-13.
- Borg, S. (2003). Teacher cognition in language teaching: A review of research on what language teachers think, know, believe, and do. *Language Teaching, 36*(2), 81-109.
- Byun, J-H. (2011). The effects of extensive reading on secondary school English teachers' motivation for and confidence in reading. *Journal of the Korea English Education Society, 10*(1), 119-144.
- Camicciotoli, B. C. (2001). Extensive reading in English: Habits and attitudes of a group of Italian university EFL students. *Journal of Research in Reading, 24*(2), 135–153.
- Chang, A., & Renandya, W. A. (2017). Current practice of extensive reading in Asia: Teachers' perceptions. *The Reading Matrix, 17*(1), 40-58.
- Chung, H. J., & Yang, J. W. (2012). Identifying teachers' beliefs in extensive reading. *Modern Studies in English Language and Literature, 54*(4), 253-276.
- Cho, K-S., & Krashen, S. D. (1994). Acquisition of vocabulary from the Sweet Valley Kids series: Adult ESL acquisition. *Journal of Reading, 37*(8), 662-667.
- Davis, C. (1995). Extensive reading: an expensive extravagance? *ELT Journal, 49*(4), 329-336.
- Day, R., & Bamford J. (1998). *Extensive reading in the second language classroom*. Cambridge, England: Cambridge University Press.
- Ebsworth, M., & Schweers, C. (1997). What researchers say and practitioners do: Perspectives on conscious grammar instruction in the ESL classroom. *Applied Language Learning, 8*(2), 237-260.
- Grabe, W. (2009). *Reading in a second language*. New York, NY: Cambridge University Press.
- Hurst, M. (2005). Learning L2 vocabulary through extensive reading: A measurement study. *Canadian Modern Language Review, 61*(3), 355–382.
- Kim, M., & Kim, Y. M. (2017). A study on the perceived needs of students, NESTs and NNESTs concerning GELP curricula development. *Journal of Learner-Centered Curriculum and Instruction, 17*(10), 197-218.
- Krashen, S. (2004). *The power of reading: Insights from the research*. Portsmouth, NH: Heinemann.
- Macalister, J. (2010). Investigating teacher attitudes to extensive reading practices in higher education: Why isn't everyone doing it? *RELC Journal, 41*(1), 59-75.

- Mason, B., & Krashen, S. (1997). Extensive reading in English as a foreign language. *System*, 25(1), 91-102.
- Mori, S. (2015). If you build it, they will come: From a "Field of Dreams" to a more realistic view of extensive reading in an EFL context. *Reading in a Foreign Language*, 27(1), 129-135.
- Mori, S., Gobel, P., Thepsiri, K., & Pojanapunya, P. (2010). Attributions for performance: A comparative study of Japanese and Thai university students. *JALT Journal*, 3(1), 5-28.
- Nation, P. (2009). *Teaching ESL/EFL reading and writing*. New York, NY: Routledge.
- Pigada, M., & Schmitt, N. (2006). Vocabulary acquisition from extensive reading: A case study. *Reading in a Foreign Language*, 18(1), 1-28.
- Richards, J. (1997). *From reader to reading teacher. Issues and strategies for second language classroom*. New York, NY: Cambridge University Press.
- Ro, E., & Park, J. (2016). Students' attitudes toward undertaking writing activities on extensive reading. *The Journal of Asia TEFL*, 13(2), 186-203.
- Robb, T., & Kano, M. (2013). Effective extensive reading outside the classroom: A large-scale experiment. *Reading in a Foreign Language*, 25(2), 234-247.
- Stoeckel, T., Reagan, N., & Hann, F. (2012). Extensive reading quizzes and reading attitudes. *TESOL Quarterly*, 46(1), 187-198.
- Tien, C-Y. (2015). A large scale study of extensive reading program for non-English majors: Factors and attitudes. *International Journal of Applied Linguistics and English Literature*, 4(4), 46-54.
- Yamashita, J. (2008). Extensive reading and development of different aspects of L2 proficiency. *System*, 36(4), 661-672.
- Van Amelsvoort, M. (2016). Extensive reading onboarding: Challenges and responses in an optional program. *Juntendo Journal of Global Studies*, 1, 95-106.

APPENDIX A

A. Interview questions for students

1. Did you like reading graded readers?
2. Were there any changes before and after reading graded readers?
3. Do you think reading graded readers is related to English learning?
4. What benefits do you think reading graded readers give to you?
5. What assessment type do you think is most adequate for graded readers? Why?

6. If reading graded readers became optional, do you think you would read books?
7. Do you know what extensive reading is? Can you explain?

B. Interview questions for instructors

1. What kind of monitoring did you use in class?
2. What do your students think about reading graded readers?
3. Do you think extensive reading has a positive effect on language acquisition?
4. Do you know what extensive reading is?
5. What kind of assessment do you think is most adequate for graded readers?
6. Do you think students read graded readers if they become optional in class?
7. If graded readers are optional, would you keep using them? Why or why not?

APPENDIX B

A. Student questionnaire

1. Gender Female () Male ()
2. Class level you enrolled Basic () Intermediate ()
3. Major: _____

Read each statement and check the box that best represents you.

	Statement	never	sometimes	always
4	When I read GRs, I used a dictionary (paper and online) to look up the meanings of words.			
5	When I read GRs, I used a translator app.			
6	When I read GRs, I used the Internet to look for information.			

7. Choose an assessment(s) used in your class. Choose ones that apply to you.
 - a. Multiple-choice questions
 - b. True and false statements
 - c. Writing short answers
 - d. Writing a summary
 - e. Writing one's opinion on themes
 - f. Discussion with peers and professor
 - g. Others: _____

Choose one that best expresses your thought.

1: strongly disagree 2: disagree 3: neutral 4: agree 5: strongly agree

	Statement	1	2	3	4	5
8	Multiple choice questions are an effective method to use on graded readers.					
9	T/F statements are an effective method to use on graded readers.					
10	Writing short answers are an effective method to use on graded readers.					
11	Writing a summary is an effective method to use on graded readers.					
12	Writing one's opinion on themes is an effective method to use on graded readers.					
13	Discussion with peers and professor is an effective method to use on graded readers.					

14. Which assessment do you prefer most? Explain.

Choose one that best expresses your thought.

1: strongly disagree 2: disagree 3: neutral 4: agree 5: strongly agree

	Statement	1	2	3	4	5
15	Reading GRs is helpful in learning new words.					
16	Reading GRs is helpful in learning grammar.					
17	Reading GRs is helpful in improving reading skills.					
18	Reading GRs is helpful in improving writing skills.					
19	Reading GRs is helpful in improving speaking skills.					
20	Reading GRs is helpful in learning other cultures.					
21	Reading GRs is helpful in improving my English ability.					
22	I enjoyed reading GRs.					
23	I felt a sense of achievement finishing each book.					
24	I gained confidence through reading GRs.					

25. Write the strengths and weakness of the use of GRs.

26. Would you be willing to have an interview with the researcher Yes / No

Email: _____

B. Professor questionnaire

1. Gender: Female () Male ()

2. Major in MA and/or PhD program: _____

3. Number of years of teaching at the current university: _____

4. How do you monitor students' reading (graded readers)?

Read statements and check the box that best represent you.

1: strongly disagree 2: disagree 3: neutral 4: agree 5: strongly agree

	Statement	1	2	3	4	5
5	It's okay to look up words students don't know while reading GRs.					
6	It's okay to use a translation app to understand while students read GRs.					
7	It's okay to use the Internet to look for information relevant to GRs while students read GRs.					

Read statements and check the box that best represent you.

1: never 2: seldom 3: sometimes 4: often 5: always

	Statement	1	2	3	4	5
8	I use multiple-choice questions to assess on GRs.					
9	I use T/F statements to assess on GRs.					
10	I use writing short answers to assess on GRs.					
11	I give writing a summary task to assess on GRs.					
12	I give discussion questions to discuss with their classmates to assess on GRs.					
13	I give writing prompts (themes, issues, problems conflict) relevant to GRs to assess.					

14. Which of the assessments do you think is most adequate? Why?

Read statements and check the box that best represent you.

1: strongly disagree 2: disagree 3: neutral 4: agree 5: strongly agree

	Statement	1	2	3	4	5
15	Reading GRs enhances students' reading skills.					
16	Reading GRs enhances students' speaking skills.					
17	Reading GRs enhances students' vocabulary skills.					
18	Reading GRs enhances students' writing skills.					
19	Reading GRs enhances student's grammar.					
20	Reading GRs helps students understand other cultures.					
21	Students can build up confidence through reading GRs.					

22. What are the advantages and disadvantages of GRs in class?

23. If GRs become optional, would you keep using them? Why or why not? Explain.

24. Would you be willing to have an interview with the researcher? Yes / No

Email: _____

Applicable levels: Secondary, tertiary

Misook Kim
Professor
Humanitas College
Kyung Hee University
26 Kyungheedaero, Dongdaemun-gu
Seoul, Korea
Email: michellems@khu.ac.kr

Received on October 28, 2019
Reviewed on November 21, 2019
Accepted on December 5, 2019