

Revelations from Three Consecutive Studies on Extensive Reading

Sy-ying Lee

National Taipei University, Taiwan, ROC
lwen@mail.ntpu.edu.tw

Abstract ■ This paper presents three consecutive studies on the effect of extensive reading on the development of reading and vocabulary for Taiwanese university non-English majors. Each study used a different approach, with subsequent studies adjusting the methodology in response to the results of the previous year. These results confirm other findings, using different subjects in other countries, that (1) extensive reading can be integrated into an EFL curriculum, termed in-class sustained silent reading, at the university level; (2) extensive reading is at least as effective and efficient as traditional instruction in acquiring English as a foreign language and is more effective than traditional instruction when the treatment duration is longer; (3) book access and self-selection of reading are two keys to the success of a sustained silent reading program.

Keywords ■ [please supply 6-8 keywords].

Introduction

Extensive Reading (ER) is not a new idea in English-as-native-language contexts. It has been well supported by a number of different researchers who have related their results to well-established theoretical frameworks. The consensus among researchers in this area is that we acquire language by understanding messages, by being exposed to a large quantity of print that is comprehensible and compelling, by joining a literacy club in which all the members are learning on a ‘social and collaborative basis’, with no risk of being evaluated and excluded (Krashen 1982, 1985a, 1985b, 2004; Smith 1983, 1988; Vygotsky 1978).

For first language development and the development of literacy, Smith hypothesizes that ‘Learning about language is not the primary aim, but



rather the by-product of some other activities... Language is learned for its uses at the time' (Smith 1988: 7). There is a great deal of evidence supporting this view: studies consistently show that children who grow up in a print-rich environment display superior competence in several different aspects of language and literacy, including syntax, vocabulary, spelling, and knowledge of history, culture, literature, and practical information (Cipielewski and Stanovich 1990, 1992; Chomsky 1972; Goodman and Goodman 1982; Nagy, Herman, and Anderson 1985; Nagy, Anderson, and Herman 1987; Shu, Anderson, and Zhang 1995; Stanovich and Cunningham 1992; West, Stanovich and Mitchell 1993).

The impact of reading exposure on one's first language development has also been found in Mandarin using a survey study. In Lee (1995, 1996), 200 subjects from three senior high schools representing three different levels of academic achievement took part in the survey, including students from the best senior high school to those from a mediocre high school in Taipei. In this study, home environment (indicated by parental education, parents' reading behavior, parents' view toward reading, and number of books owned in the home) significantly predicted the subjects' free reading behavior, which in turn was the only significant variable among others (leisure writing and writing apprehension) that predicted the subjects' Chinese writing performance for a nationwide entrance examination.

Clearly, we cannot assume that what works for the child in the first language situation will also work for the second or foreign language acquirer, but so far there is good evidence that second language acquirers also profit from free voluntary reading. Those who report doing more reading in their second language outside of school do better in writing (Janopolous 1986; Lee 2005a), have greater grammatical competence (Lee, Krashen, and Gribbons 1996), and on the TOFEL examination (Constantino, Lee, Cho and Krashen 1997; Gradman and Hanania 1991).

There is also consistence evidence that pedagogical approaches based on the comprehension of messages through interesting reading materials, adapted to the ESL and EFL context, are successful. Research done in ESL/EFL situation consistently shows that a curriculum incorporating extensive reading is a teaching practice at least as effective as a curriculum based on skill-building. More often than not, the ER curriculum works better. Extensive reading, with no direct instruction on formal aspects of language, has been shown to be very effective for children, teenagers, and college students in acquiring a second or foreign language (Cho and

Krashen 1994; Cho 1995; Cho and Kim 2004; Elley and Mangubhai 1983; Elley 1980, 1989, 1991; Hafiz and Tudor 1990; Lai 1993; Lao and Krashen 2000; Mason and Krashen 1997; Mason 2003; McQuillan 1994; Tsang 1996; Tudor and Hafiz 1989; Young 2001). More important, extensive reading is very pleasant to do and is a lifelong approach for language acquisition and intellectual growth (Krashen 2004).

Taiwan has been a productive laboratory for the study of extensive reading in school, or ‘sustained silent reading’ (SSR), investigating the potential of extensive reading as a tool for language acquisition and literacy development. Results thus far have shown that students participating in sustained silent reading make gains that are equivalent to or better than gains made by comparison students in classes not including SSR, in reading and vocabulary (Cheng 2003; Hsu and Lee 2005; Lee 2005b, 2005c, 2005d; Liu 2005; Sheu 2004; Sims 1996; Yuan and Nash 1992), writing (Lee and Hsu 2005), grammar (Sheu 2004), and attitudes toward reading (Lee 1998; Sheu 2004). In addition to the positive results favoring readers, researchers also suggest that an extensive reading class can be ‘less labor intensive for the teacher’ (Yuan and Nash 1992). It allows teachers also to enjoy some good quality reading as a member in the same literacy club they are inviting their students to join.

The duration of the studies listed above was either one semester or one academic year. Studies with the same duration, however, sometimes produce different results. Hence the inquiry: What makes an SSR program better than another? What are the conditions underlying successful SSR? Our ultimate goal is to make it possible for foreign language students to have at least some of the advantages second language students have, to be able to profit from extensive reading.

In order to determine why some studies produced better results than others, the author conducted three studies in three consecutive years, a one-semester study and two one-year studies, involving three different approaches to extensive reading, with each approach an improvement over the one preceding it.

The three studies were done at the same university. Experimental classes were taught by the same teacher (the researcher), and compared with control groups following the regular first-year university English curriculum. Students participating in these studies were not taking other classes using English as a medium of instruction and had little exposure to English outside of school. Moreover, because classes are taught by different teachers with different styles and approaches, more than one

comparison group was included in the studies. This was done to increase ecological validity, to make it more likely that the findings can be generalized beyond the confines of the studies reported here.

Study I:

Pure SSR in One Semester (12 Weeks, Spring 2001)

This study examined the impact of extensive reading under less-than-optimal conditions: Students read for only 12 weeks, had access to a limited amount of reading (215 graded readers), were asked to write summaries of what they read, and their in-class reading took place only once a week. In addition, it is likely that the students were not serious about English class. The study took place in the second half of a year-long course; the first semester was devoted to viewing films with Chinese subtitles. The results of the pretests shown in Table 1 revealed not only their significantly lower proficiency level before treatment, they might also reflect their low motivation in learning English. For obvious reasons, a new instructor was brought in for the second semester, this researcher. These conditions were not set up on purpose: they were a result of practical constraints. Nevertheless, the situation offered an opportunity to see how robust extensive reading is, and to determine if it is worthwhile to utilize an extensive reading approach when the situation is not optimal.

Consistent with common practice in sustained silent reading, students were not tested on the content of what they read. The ER class also included class discussion on language acquisition theory, including a presentation of the research evidence showing the efficacy of reading as a means of developing competence in a second or foreign language. This was done to give students an orientation to the sustained silent reading approach, and to give them confidence that self-selected reading would indeed positively impact their language acquisition.

Two comparison groups were used. Comparison group 1 used a textbook and did traditional reading comprehension and writing exercises. In comparison group 2, outside reading was encouraged, but no record of the reading was kept. In addition, the instructor of group 2 devoted about 70 to 80% of class time to explaining vocabulary students encountered in the assigned text as well as related words. Neither comparison class did grammatical analysis or form-focused exercises. Both included culture, role-plays, discussions, presentations, and direct teaching of reading strategies and vocabulary.

The Nation Vocabulary Test (Nation 1990), and a cloze test constructed by Mason (Mason and Krashen 1997), consisting of approximately 1600 words with every tenth word deleted, were used as pre- and posttests.

Because of pretest differences, posttest scores were adjusted by ANCOVA. The reading group did slightly better than the first comparison group on the vocabulary test, but the difference was not significant ($p = .32$). Comparison 2 did better than the reading group, and the difference fell just short of statistical significance ($p = .07$). Inspection of the raw data (Table 1) shows that the comparison 2 group's advantage was due to performance on the part of the test that contained less frequent words, those at the 5000 word level. This may be due to the fact that subjects in the experimental group read mostly graded readers written at the 2000 and 3000 word level. It can be postulated that given more time for the reading group, allowing the subjects to move on to higher level materials, they would very possibly show improvement on the less frequency words.

Table 1. *Vocabulary Test Results*

	2000 word level			3000 word level			5000 word level			Total gain	Adjusted total
	pre	post	gain	pre	post	gain	pre	post	gain		
Exp. (S.D.)	41.4 (9.9)	45.2 (8.1)	3.8	32.8 (12.2)	35.6 (11.2)	2.8	22.2 (10.1)	25.3 (9.8)	3.1	9.7	116.2
Com. 1 (S.D.)	47.6 (8.5)	49.1 (7.8)	1.5	41.5 (12.7)	42.1 (12.0)	0.6	31.0 (12.0)	33.2 (12.4)	2.2	4.3	114.4
Com. 2 (S.D.)	49.2 (4.7)	50.7 (3.6)	1.5	42.0 (8.5)	45.1 (7.1)	3.1	29.3 (11.0)	34.4 (8.8)	5.2	9.8	120.1

Note. The vocabulary level tests were taken from Nation (1990)

[Typesetter: suggest all cols of numbers are decimal point aligned]

On the cloze test, the experimental and comparison 2 groups made clear gains, but comparison group 1 did not, as shown in Table 2. The reading group made larger gains than comparison group 1 ($p < .05$). Comparison 2 did better than the readers, but the difference was not significant.

The results of this study are consistent with previous reports of the efficacy of using graded readers (Mason and Krashen 1997), and with the desirability of sharing language acquisition and reading theory with students (Lee 1998). It was also shown that students of English as a foreign language can improve without producing language, without form-focused activities, and without being tested on what they read.

Table 2. *Cloze Test Results* [keep with title]

	<i>Pretest</i>	<i>Posttest</i>	<i>Gain</i>	<i>Adjusted mean</i>
	<i>M (SD)</i>	<i>M (SD)</i>		
Exp (n = 65)	37.9 (12.0)	44.0 (11.5)	5.6	46.4
Com 1 (n = 38)	44.0 (13.0)	45.0 (12.6)	1.0	43.3
Com 2 (n = 38)	46.3 (11.4)	51.3 (10.4)	5.0	47.9

By Figure 1, which provides a visual representation of all the numerical data, it can be observed that the Experimental group made most gain on the 2000 word level test and the cloze test. Control Group 2 made most gains on the 5000 word level test as well as the cloze test; while Control Group 1 made gains that were more reserved.

The limitations of Study I were obvious: the experimental group was significantly inferior to their two comparison groups on all measures in the pretests; and the two comparison groups might have reached the ‘ceiling effect’ on the 2000 VLT, for which the full score was 54. These flaws were repaired in the study in the following academic year by lengthening the treatment duration, providing more books, and having groups with equal performance on all pretest measures.

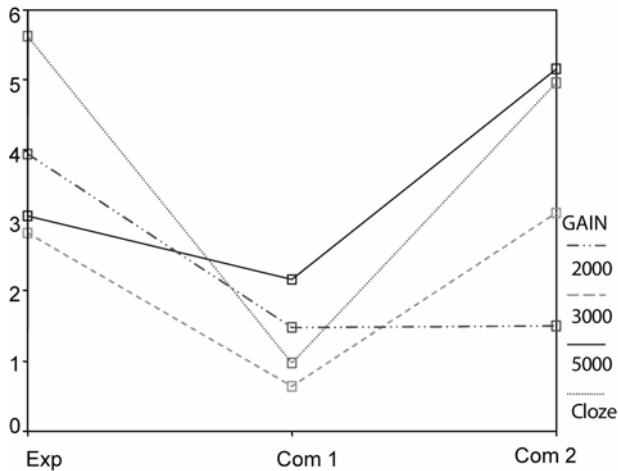


Figure 1. *A Summary of the Gains on Four Tests*

Study II:

SSR Using Assigned Reading (A Full Academic Year SSR, Fall 2002–

*Spring 2003, with the Second Semester Conducted as an Assigned
Reading Literature Class)*

In this study, some of the flaws of Study I were repaired. All groups had equivalent levels of English competence at the start of the treatment and the treatment lasted one year. In this study, three comparison classes (Com 1 = 40; Com 2 = 45; Com 3 = 54), randomly selected from the 26 Freshman English classes at the same university and one experimental class (N = 67) taught by the researcher were involved.

In this study, the vocabulary measure developed by Schmitt (2000) replaced the Nation measure used in Study I. Schmitt's measure covers 2000, 3000, 5000, 10,000 and academic level words. According to Nation (2001), Schmitt's measure was 'a major improvement' over the original measure because lower frequency words and academic level words are also assessed.

The results of a Multivariate Analysis of Variance indicated that the three comparison groups were not significantly different on the pretests, so they were combined into one group. Table 3 presents the descriptive pretest data for the experimental group and the three comparison groups. An Independent t test showed that the experimental group (n = 67) and the combined comparison group (n = 139) were not significantly different on all pretests (for total vocabulary, $F = 1.84$, $p = .145$; for the cloze, $F = 1.65$, $p = .18$).

Table 3. *A Mean Comparison Process Comparing the Pretest Scores of Each Group*

<i>GROUP</i>		<i>2000</i>	<i>3000</i>	<i>5000</i>	<i>10000</i>	<i>Academic words</i>	<i>Total vocab</i>	<i>Cloze</i>
Com 1	Mean	26.3	20.8	16.3	3.6	19.2	86.2	44.0
	S. D	3.5	5.9	5.6	4.6	5.2	19.5	10.4
Com 2	Mean	26.9	22.5	18.4	4.8	21.1	93.8	47.9
	S. D	3.2	4.8	5.4	4.0	4.8	18.6	9.6
Com 3	Mean	27.5	22.6	17.5	5.2	21.2	94.0	48.4
	S. D	3.3	6.2	6.7	3.9	6.3	22.1	10.0
Exp.	Mean	26.3	20.9	17.0	3.7	19.8	87.7	47.1
	S. D	4.5	6.5	6.7	4.0	6.2	22.5	10.0

Note. The vocabulary level tests were taken from Schmitt (2000)

The comparison groups had textbook-oriented instruction, reading, analyzing and discussing texts, students did presentations based on issues related to the assigned readings, and there was direct instruction on language skills and learning strategies covered in each chapter of the text. There were also regular quizzes and examinations.

During the first semester, the experimental group did self-selected reading of graded readers. Students chose from 570 graded readers varying in difficulty from 300 headwords to 3300 headwords. Students devoted half of the once weekly three-hour class to reading, 20 minutes to checking in and out books from the classroom collection, and the rest of the class time to shared reading, giving short presentations or interacting with classmates. Students were required to record what they read (titles, pages, time spent on reading) and write short reflections on what they read in either English or Chinese. These reading logs were handed in each week. Grades were based on participation and students' logs (time spent reading, pages read, and reflections on reading).

During the second semester, students were required to read five texts: *Stuart Little*, *Charlotte's Web*, *The Trumpet of the Swan*, *The Little Prince*, and *Tuesdays with Morrie*. In addition, students were required to choose another two books from a list of suggested readings. The list consisted of books related to current popular films, such as *The Bridges of Madison County*, *Bridget Jones' Diary* and books from the *Harry Potter* series, as well as teachers' suggestions.

Table 4 presents pre- and posttest scores for the vocabulary test. The effect of the in-class SSR treatment was determined by examining differences between gain scores (Table 5). At each level of the vocabulary test, the experimental group made better gains. Because multiple t-tests were used, the alpha level, the level of significance necessary to achieve statistical significance, was adjusted using the Bonferroni procedure (Rosenthal and Rosnow 1984). Using the adjusted alpha of .008 (.05/6), the experimental group significantly outperformed the comparison group on the combined vocabulary test, and on the 10,000 and 3000 word level test. It is not surprising that there was no difference between groups on the 2000 word level test. Chen (1999) found that an average university student (first year non-English major) in Taiwan possesses approximately 3000 to 3500 English words. Even though there was no significant difference between the groups on the academic level test, these results are, like the others, a challenge to traditional instruction, because the experimental group performed as well as the instructional group without direct instruction and examinations. Gains on the cloze test for the experimental group and the comparison group were nearly identical (Table 6).

Table 4. *Vocabulary Test Results on the Pretests and Posttests*

<i>Comparison</i>	<i>Experimental</i>
-------------------	---------------------

	<i>Pretest means (sd)</i>	<i>Posttest means (sd)</i>	<i>Pretest means (sd)</i>	<i>Posttest means (sd)</i>
2000	27.0 (3.3)	27.6 (2.4)	26.3 (4.5)	27.9 (2.3)
3000	22.1 (5.7)	23.5 (4.9)	20.9 (6.5)	24.0 (4.7)
5000	17.5 (6.0)	19.4 (5.6)	17.0 (6.7)	20.5 (5.5)
10000	4.6 (4.1)	6.0 (4.6)	3.7 (4.0)	7.3 (4.3)
ACADEMIC	20.6 (5.5)	22.4 (5.5)	19.7 (6.2)	22.4 (5.4)
TOTAL	91.7 (20.3)	99 (18.7)	87.7 (22.5)	102.1 (17.7)

Note. Each level of the vocabulary test contained 30 items, taken from Schmitt (2000)

Table 5. *Gain Scores for the Vocabulary Test*

	<i>DIFF pre/post</i>			
	<i>COMP</i>	<i>EXP</i>	<i>t</i>	<i>p</i>
2000	0.6	1.6	2.01	0.046
3000	1.4	3.1	2.72	0.007*
5000	1.9	3.5	2.55	0.012
10000	1.4	3.6	3.84	0.00013*
ACADEMIC	1.8	2.7	0.55	0.583
TOTAL	7.3	14.4	4.35	0.000014*

Table 6. *Cloze Test Results*

	<i>PRE</i>	<i>POST</i>	<i>DIFF</i>
Comp	46.9 (10.1)	51.8 (9.8)	4.9
Exp	47.1 (10.0)	52.1 (8.3)	5.0

This study reports a modest victory for the group that did self-selected reading followed by assigned reading over traditional instruction in vocabulary growth, and a tie in reading comprehension, as measured by the cloze test. Previous studies using the same cloze test show some experimental (self-selected reading) and comparison groups making five point gains on this test in just one semester (Study I, using university students; Hsu and Lee 2005, using junior college students). In this study, the average gain was five points over two semesters. Thus, neither group made impressive gains on the cloze test.

A likely candidate for the unimpressive results on the cloze tests is the kind of books that were assigned. The list consisted of books that teachers felt were interesting; teachers' views, however, may not be the same as students' views (Ujji and Krashen 2002). In fact, some students remarked

that they merely flipped through the pages of the assigned books, with little comprehension, and several students considered E.B. White's books too childish. With such a lack of enthusiasm about the reading, in fact, one wonders how the students made as much progress as they did.

It appears to be the case that for reading to do a reader any good, to result in language and literacy development, it needs to be more than comprehensible. It needs to be interesting, or even compelling. An interesting hypothesis is that the reader needs to be 'lost in the book' (Nell 1988). Sometimes assigned reading is comprehensible and compelling, and results in real gains, sometimes it does not (Krashen 2004: 51-52). There are good reasons to assign reading in language arts or foreign language programs, for the purpose of discussion and to ensure exposure to certain crucial readings, but we need to be sure that what was assigned is really right for the students. Subjects in McQuillan (1994) preferred assigned reading over self-selected reading. The assigned texts, however had been popular with students in previous studies. In the present study, the materials were selected based on their popularity among native speakers and because they were adaptations of blockbuster movies. They had not been tried before with other groups of students. Figures 2 and 3 present visual representations of the gain differences between groups.

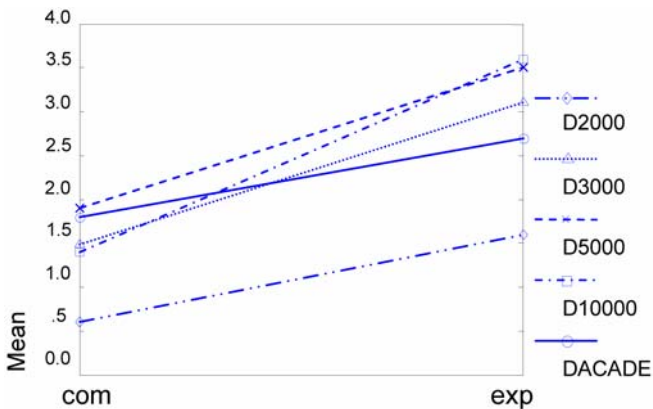


Figure 2. *Vocabulary Gains Made by the Experimental Group and the Three Comparison Groups Combined*

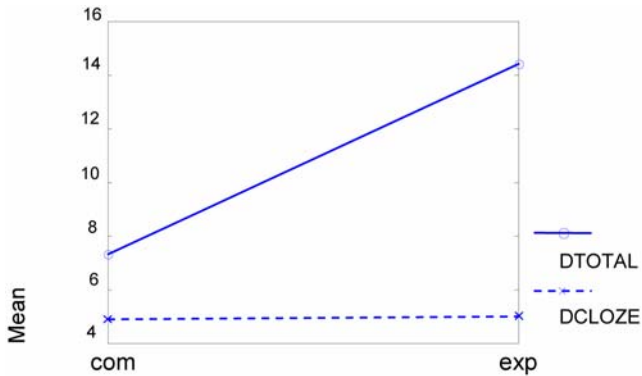


Figure 3. Gains on the Cloze Test and of All the Vocabulary Tests
 [T/S: I assume lines will be in black. Ensure picture is clear]

Study III:

SSR Allowing Full Self-selected Reading (Fall 2003–Spring 2004)

The purpose of this study was to compare another version of SSR—a full year of self-selected reading—to assigned reading. In this study, the combined comparison group used in Study II was used as the basis of comparisons, since the measures used and the duration of instruction were the same. Subjects in the experimental group (Exp 2, $N = 41$) did self-selected reading, choosing from a collection of about 1200 titles for one academic year. Experimental group 1 was the assigned reading group from the previous study, presented here again to facilitate comparison. Tables 7 and 8 present the descriptive data for all groups. Table 9 indicates the gain scores made by the three groups on all measures. MANOVA was performed on pretests and showed no difference among groups when the study began. The same statistical procedure was then employed to determine if there were significant differences among groups on the gain scores for all measures (Table 10).

Table 7. Means and Standard Deviations on Pretests

Groups		2000	3000	5000	10000	ACADE	TOTAL	CLOZE
Com	Mean	27.0	22.1	17.5	4.6	20.6	92.0	46.9
	SD	3.3	5.7	6.0	4.2	5.6	20.4	10.1
Exp 1	Mean	26.3	20.9	17.0	3.7	19.75	87.7	47.1
	SD	5.0	6.5	6.7	4.0	6.2	22.49	10.0
Exp 2	Mean	26.4	20.6	17.3	3.8	19.7	87.8	44.4
	SD	3.8	5.7	5.6	3.2	6.2	20.0	8.2

Note. N for Com group = 139; Exp1 = 67; Exp 2 = 41

Table 8. *Means and Standard Deviations on Posttests*

<i>Groups</i>		<i>2000</i>	<i>3000</i>	<i>5000</i>	<i>10000</i>	<i>ACADE</i>	<i>TOTAL</i>	<i>CLOZE</i>
Com	Mean	27.6	23.5	19.4	6.0	22.4	98.9	51.8
	SD	2.4	4.9	5.7	4.6	5.6	18.7	9.8
Exp 1	Mean	27.9	24.0	20.5	7.3	22.4	102.1	52.1
	SD	2.3	4.7	5.5	4.3	5.4	17.7	8.3
Exp 2	Mean	27.8	24.8	21.6	8.1	22.6	104.8	58.9
	SD	2.3	4.3	4.7	3.1	4.1	14.7	7.9

Table 9. *Means and Standard Deviations on Gains*

<i>Groups</i>		<i>D2000</i>	<i>D3000</i>	<i>D5000</i>	<i>D10000</i>	<i>DACADE</i>	<i>DTOTAL</i>	<i>DCLOZE</i>
Com	Mean	.66	1.5	2.0	1.4	1.8	7.2	4.6
	SD	2.5	4.1	4.4	3.8	4.8	11.3	6.7
Exp 1	Mean	1.6	3.1	3.5	3.6	2.6	14.4	5.0
	SD	4.0	3.8	3.6	4.1	4.8	10.8	7.6
Exp 2	Mean	1.3	4.2	4.4	4.3	2.9	17.0	14.6
	SD	3.1	4.5	3.2	3.3	4.1	11.1	7.4

Table 10. *Multiple Comparisons with Scheffe Post Hoc Test on Gain Scores*

	<i>(I) GROUP</i>	<i>(J) GROUP</i>	<i>Mean Difference (I-J)</i>	<i>Std. Error</i>	<i>Sig.</i>
D2000	Com	Exp1	-.8879	.4705	.17
	Com	Exp2	-.5878	.5689	.59
	Exp1	Exp2	.3000	.6222	.89
D3000	Com	Exp1	-1.6830	.6185	.03*
	Com	Exp2	-2.7324	.7478	.00*
	Exp1	Exp2	-1.0494	.8180	.44
D5000	Com	Exp1	-1.6613	.5873	.02*
	Com	Exp2	-2.3291	.7102	.01*
	Exp1	Exp2	-.6678	.7768	.69
D10000	Com	Exp1	-2.1937	.5708	.00*
	Com	Exp2	-2.7911	.6902	.00*
	Exp1	Exp2	-.5974	.7549	.73
DACADE	Com	Exp1	-.8335	.7082	.50
	Com	Exp2	-.7194	.8564	.70
	Exp1	Exp2	.1140	.9367	.99
DTOTAL	Com	Exp1	-7.2593	1.6283	.00*
	Com	Exp2	-9.1598	1.9689	.00*
	Exp1	Exp2	-1.9005	2.1536	.68
DCLOZE	Com	Exp1	-.3687	1.0817	.94
	Com	Exp2	-9.9542	1.3080	.00*
	Exp1	Exp2	-9.5855	1.4307	.00*

Note. The mean difference is significant at the .05 level.

Results showed that Experimental groups 1 and 2 (assigned SSR and self-selected SSR) significantly outperformed the combined comparison group on the 3000, 5000, 10,000 vocabulary level tests, but not on the 2000 and

academic levels. As previously mentioned, the average Taiwanese university student may possess up to 3000 to 3500 vocabulary words in English; the 2000 vocabulary level test may therefore not be able to reflect their growth at this level. And again, there was no significant difference among groups in the gains on academic level words, although both experimental groups made slightly better gains than comparisons in raw scores. As for the cloze test, the self-selected reading group (Exp 2) performed significantly better than the comparison groups combined and also outperformed the experimental group that did assigned reading in Study II (Exp 1). This result confirms the importance of fully respecting students' free selection of materials, a practice that gives them the responsibility for their own learning and enhances their motivation to read. Figures 4 and 5 present visual representations of the gain score differences among the experimental and comparison groups on the vocabulary and cloze tests.

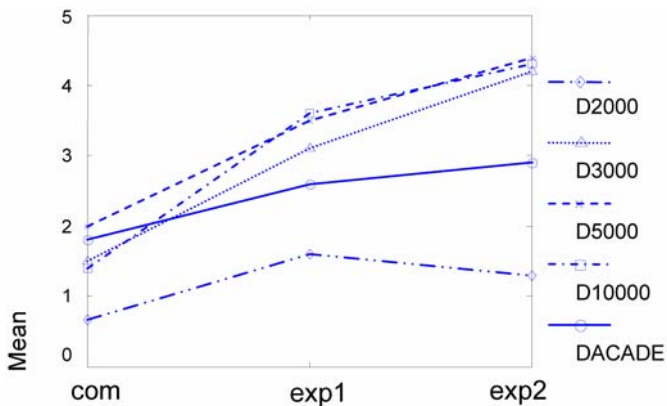


Figure 4. *Vocabulary Gains Made by the Experimental Group and the Three Comparison Groups Combined*

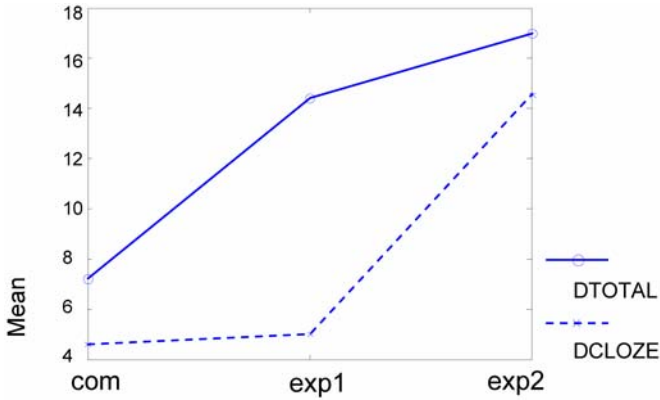


Figure 5. *Gains on the Cloze Test and of the Total Gain on the Vocabulary Tests*

Students' logs confirmed the benefits of self-selection. While students in Experimental Group 1 found E.B. White boring, the books Experimental Group 2 students chose were 'so interesting and fun to read'. Students discussed their reading with each other, and recommended books to their classmates. Books written by Judy Blume, Louis Sachar, and books from the Sweet Valley series were especially popular, and were rarely back on the shelves until the end of the year.

Discussion

This series of studies intended to contribute to providing a more solid basis for the design of SSR programs for EFL students. The results provide responses to some of the concerns and doubts some might have about integrating extensive reading into the curriculum. Further, the three consecutive studies also addressed some issues pertaining to the conduct of the ER research; in addition to the use of multi-group comparisons, the potential novelty of the treatment (e.g. the Hawthorne effect) and possible confounds from other input sources are discussed.

First of all, the longer the duration of the study, the better the result (Krashen 2004), but even one semester of extensive reading is typically at least as effective and efficient as one semester of formal instruction (Lee 2005b; Hsu and Lee 2005). In two one-semester studies done with junior high students in Taiwan, Sheu (2004) obtained positive results for most measures in favor of the reading groups. If Sheu's study and Study I presented in this article are the result of the Hawthorne effect, one of the

most often heard criticisms of extensive reading, then the results of two year-length studies (Study II and Study III) might serve to answer this claim. In these two studies, the experimental groups outperformed their counterparts taught by three experienced university professors. Liu (2005), who conducted a full year study at the same university using students from the school of humanistic studies, also obtained positive results favoring the reading group.

Another concern is that in some studies, students have other sources of input in English (Horst 2005). This is of course true in second language, as contrasted with foreign language studies. This was not the case here. All subjects were freshmen who were not majoring in English. Most were majoring in law and commerce, and always complained about their full load of schoolwork. The gains obtained by the experimental groups in these studies were most likely a result of their extensive reading experience during the treatment semesters.

Perhaps the main criticism of extensive reading is that it ‘takes time’, and we need to accelerate students’ acquisition of English through more systematic instruction on the underlying subskills and learning strategies, and require students to do more output practice to ‘consolidate’ their learning. The results reported here, however, show that reading alone is at least as efficient as regular instruction and for long-term studies is actually more efficient. Studies intending to test a ‘balanced approach’ that included strategy training or output practice in addition to extensive reading showed that reading alone is equally efficient.

A study done in Taiwan investigated whether the teaching of word-guessing strategy enhanced reading comprehension (Cheng 2003). Subjects were asked to do ER after school but practice word-guessing strategies using the assigned textbooks in class. Results showed that the use of word-guessing strategies had a significant impact only on tests of word recognition; it did not, however, result in better performance on a test of reading comprehension and the measure on the ability of applying the word-guessing strategies. The study and use of word-guessing strategies did not, in other words, add to the power of reading.

Mason’s (2003) three-semester study on Japanese college students clearly showed that more writing practice with correction and revision did not make the subjects better readers and more accurate writers in English as a foreign language. Her subjects all did extensive reading in English. Those who only wrote brief reading reflections in their first language, Japanese, gained as much as another group that wrote summaries in

English, and also gained as much as another group that also wrote summaries in English and had their errors corrected. Those who read and wrote summaries in Japanese were far more efficient, gaining significantly more per hour devoted to English.

There is a concern that reading graded readers will not influence the acquisition of less frequent words. The vocabulary measures used in Studies II and III included lower frequency words (10,000 words) and academic words, in addition to words at other levels (Schmitt 2000). Both reading groups outperformed comparisons at the 10,000 level and did just as well on the academic level subtests. These results present a clear challenge to those who believe in the necessity or superiority of instruction on low frequency words and academic words; the experimental groups did not follow any textbooks and were not given any drill assignments and tests of any kind except for the pre- and posttests at the beginning and end of the experiments.

A final concern is that the simple syntax and reduced vocabulary of graded readers will limit language development. Student reading logs, however, revealed that although students began with graded readers, they gradually selected more and more authentic reading on their own. According to the logs, the use of graded readers during the first semester was helpful, building both competence and confidence in reading English.

It is highly likely that the success of Study III was due to the fact that students were able to select their own reading and had improved access to books (from 215 books in Study I to 1000 books in Study III), including graded readers and authentic materials that were both interesting and comprehensible. In Study II, reading materials were assigned, and were chosen based on their popularity among native speakers. But they were either not compelling enough or too difficult for these students. It may be the case that depriving students of the right to choose their own reading also deprives them of the pleasure of reading.

The studies described here confirm that SSR programs result in an impressive improvement in both vocabulary and reading comprehension. Other studies have found that SSR has a 'transfer effect', impacting other language skills as well, such as listening (Elley and Mangubhai 1983) and writing (Elley 1991; Hafiz and Tudor 1990; Lai 1993; Lee and Hsu 2005; Tsang 1996; Mason 2003; Tudor and Hafiz 1989).

The demonstration of the success of SSR does not imply that all classroom activities other than reading itself are forbidden. Those that make texts more comprehensible and that increase student interest in

books should be quite helpful. These include read-alouds (Trelease 2001 [2006 in refs]), book talks, and share simple but effective strategies (e.g. skip unfamiliar words, infer word meanings from context, and feel free to change books when they are too difficult or uninteresting), explaining unknown aspects of language when necessary and of course demonstrating an enthusiasm for reading. Finally, self-selected reading does not exclude the study of literature; in fact, those who have done a great deal of voluntary reading will be far better prepared to appreciate texts that teachers select and assign for their literary value.

NOTE

[No marker for this in the text. Where should it be linked to?]

Inspection of the student logs revealed that some of the readers began to show more interest in language per se as the course proceeded, noticing interesting phrasing, vocabulary, and even grammar. It is unclear whether this kind of noticing contributes to language competence but it is an indication of progress: it suggests that the students' level of comprehension was high, that they were not struggling to understand the text, and could therefore notice interesting aspects of language form.

REFERENCES

- Chen, H.-J.
1999 'How Many Words Do They Know? Assessing Taiwanese College EFL Students' Receptive and Productive vocabularies', *Proceedings of the 16th Conference on English Teaching and Learning*: 83-79.
- Cheng, C.-K.
2003 'Extensive Reading, Word Guessing Strategies, and Incidental Vocabulary Acquisition', *Selected Papers from the Twelfth International Symposium on English Teaching*: 188-98.
- Cho, K.-S.
1995 'From Sweet Valley Kids to Harlequins in One Year', *California English* 1(1): 18-19.
- Cho, K.-S., and H.-J. Kim
2004 'Recreational Reading in English as a Foreign Language in Korea: Positive Effects of a 16-Week Program', *Knowledge Quest*: 36-38.
- Cho, K.-S., and S. Krashen
1994 'Acquisition of Vocabulary from the Sweet Valley High Kids Series: Adult ESL Acquisition', *Journal of Reading* 37: 662-67.
- Chomsky, C.
1972 'Strategies in Language Development and Reading Exposure', *Harvard Educational Review* 42: 1-33.
- Cipielewski, J., and K. Stanovich
1990 'Assessing Print Exposure and Orthographic Processing Skill in Children: A

Revelations from Three Consecutive Studies on Extensive Reading

- Quick Measure of Reading Experience', *Journal of Educational Psychology* 32: 733-40.
- 1992 'Predicting Growth in Reading Ability from Children's Exposure to Print', *Journal of Experimental Child Psychology* 54: 74-89.
- Constantino, R., S.-Y. Lee., K.-S. Cho and S. Krashen
1997 'Free Voluntary Reading as a Predictor of TOEFL Scores', *Applied Language Learning* 8: 111-18.
- Elley, W.
1980 'A Comparison of Content-interest and Structuralist Reading Programs in Nine Primary Schools', *New Zealand Journal of Educational Studies* 15.1: 39-53.
1989 'Vocabulary Acquisition from Listening to Stories', *Reading Research Quarterly* 24(2): 174-87.
1991 'Acquiring Literacy in a Second Language: The Effects of Book-based Programs', *Language Learning* 41: 375-411.
- Elley, W., and E. Mangubhai
1983 'The Impact of Reading on Second Language Learning', *Reading Research Quarterly* 19: 53-67.
- Goodman, K., and Y. Goodman
1980 'Spelling Ability of a Self-taught Reader', in F. Gollasch (ed.), *Language and Literacy: The Selected Writings of Kenneth S. Goodman 2* (London: Routledge): 135-42.
- Gradman, H.L., and E. Hanania
1991 'Language Learning Background Factors and ESL Proficiency', *The Modern Language Journal* 75: 39-51.
- Hafiz, F., and I. Tudor
1990 'Graded Readers as an Input Medium in L2 Learning', *System* 18(1): 31-42.
- Horst, M.
2005 'Learning L2 Vocabulary through Extensive Reading: A Measurement Study', *The Canadian Modern Language Review* 00 [number?] (March): 355-82.
- Hsu, Y.-Y., and S.-Y. Lee
2005 'Does Extensive Reading also Benefit Junior College Students in Vocabulary Acquisition and Reading Ability?', *The Proceedings of the 22nd International Conference in English Teaching and Learning*: 116-27.
- Janopoulos, M.
1986 'The Relationship between Pleasure Reading and Second Language Writing Proficiency', *TESOL Quarterly* 20: 763-68.
- Krashen, S.
1982 *Principles and Practice in Second Language Acquisition* (New York: Prentice Hall).
1985a *The Input Hypothesis: Issues and Implications* (Beverly Hills: Laredo).
1985b *Inquiries and Insights* (Menlo Park, CA: Alemany Press).
2004 *The Power of Reading* (Portsmouth, NH: Heinemann, 2nd edn).
- Lai, F.-K.
1993 'The Effect of a Summer Reading Course on Reading and Writing Skill', *System* 21: 87-100.
- Lao, C.-Y., and S. Krashen

- 2000 'The Impact of Popular Literature Study on Literacy Development in EFL: More Evidence for the Power of Reading', *System* 28: 261-70.
- Lee, S.-Y.
 2006 'A One-Year Study of SSR: University Level EFL Students in Taiwan', *International Journal of Foreign Language Teaching* 2(1): 6-8. <http://www.tprstories.com/ijflt/IJFLTSummer05.pdf>
- 1995 'A Cross-cultural Validation of The Reading Hypothesis: The Relationship of Pleasure Reading to Writing Proficiency and Academic Achievement among Taiwanese Senior High School Students' (Doctoral dissertation, University of Southern California, Los Angeles, CA).
- 1996 'The Relationship of Free Voluntary Reading to Writing Proficiency and Academic Achievement among Taiwanese Senior High School Students', *The Proceedings of the Fifth International Symposium on English Teaching*: 119-26.
- 1998 'Effects of Introducing Free Reading and Language Acquisition Theory on Students' Attitudes toward the English Class', *Studies in English Language and Literature* 4: 21-28.
- 2005a 'Facilitating and Inhibiting Factors on EFL Writing: A Model Testing with SEM', *Language Learning* 55(2): 335-74.
- 2005b 'How Robust Is In-class Sustained Silent Reading?', *Studies in English Language and Literature* 15: 65-76.
- 2005c 'The Robustness of Extensive Reading: Evidence from Two studies', *The International Journal of Foreign Language Teaching* 1(3): 13-19. Available at <http://www.tprstories.com/ijflt/IJFLTSummer05.pdf>
- 2005d 'Sustained Silent Reading Using Assigned Reading: Is Comprehensible Input Enough?', *The International Journal of Foreign Language Teaching* 1(4): 10-13. Available at <http://www.tprstories.com/ijflt/IJFLTfall05.pdf>
- Lee, Y. O., S. Krashen and B. Gribbons
 1996 'The Effect of Reading on the Acquisition of English Relative Clauses', *ITL: Review of Applied Linguistics* 113/114: 263-73.
- Lee, S.-Y., and Y.-Y. Hsu
 2005 'Does Extensive Reading also Effect Writing Improvement among Junior College Students?', Paper presented at *The 14th International Symposium on English Teaching*, Chien Tan, November, 2005, Taiwan.
- Liu, C.-K.
 2005 'Self-selected Reading Effects Significant Gains in Vocabulary Size and Reading Comprehension', Paper presented at *The 14th International Symposium on English Teaching*, Chien Tan, November, 2005, Taiwan.
- Mason, B.
 2003 'Sufficiency of Extensive Reading on the Development of Grammatical Accuracy' (Doctoral dissertation, Temple University, Tokyo).
- Mason, B., and S. Krashen
 1997 'Extensive Reading in English as a Foreign Language', *System* 25: 91-102.
- McQuillan, J.
 1994 'Reading versus Grammar: What Students Think Is Pleasurable for Language Acquisition', *Applied Language Learning* 5: 95-100.
- Nagy, W., P. Herman and R. Anderson

Revelations from Three Consecutive Studies on Extensive Reading

- 1985 'Learning Words from Context', *Reading Research Quarterly* 20: 233-53.
- Nagy, W., R. Anderson, and P. Herman
1987 'Learning Word Meanings from Context during Normal Reading', *American Educational Research Journal* 24: 237-70.
- Nation, P.
1990 *Teaching and Learning Vocabulary* (New York: Newbury House Publishers).
2001 *Learning Vocabulary in Another Language* (Cambridge: Cambridge University Press).
- Nell, V.
1988 *Lost in a Book* (New Haven: Yale University Press).
- Rosenthal, R., and R. Rosnow
1984 *Essentials of Behavioral Research: Methods and Data Analysis* (New York: McGraw-Hill).
- Schmitt, N.
2000 *Vocabulary in Language Teaching* (Cambridge: Cambridge University Press).
- Sheu, P.-H.
2004 'The Effect of Extensive Reading on Learners' Reading Ability Development', *Journal of National Taipei Teachers College* 17(2): 213-28.
- Shu, H., R.-C. Anderson, and H. Zhang
1995 'Incidental Learning of Word Meanings while Reading: A Chinese and American Cross-cultural Study', *Reading Research Quarterly* 30(1): 76-95.
- Sims, J.
1996 'A New Perspective: Extensive Reading for Pleasure', *The Proceedings of the Fifth International Symposium on English Teaching*: 137-144.
- Smith, F.
1983 *Joining the Literacy Club* (Portsmouth, NH: Heinemann).
1988 *Essays into Literacy* (Portsmouth, NH: Heinemann).
- Stanovich, K., and A. Cunningham
1992 'Studying the Consequences of Literacy within a Literate Society: The Cognitive Correlates of Print Exposure', *Memory and Cognition* 20(1): 51-68.
- Trelease, J.
2006 *The Read-Aloud Handbook* (New York: Penguin Books).
- Tsang, W.-K.
1996 'Comparing the Effect of Reading and Writing on Writing Performance', *Applied Linguistics* 17: 210-33.
- Tudor, I., and F. Hafiz
1989 'Extensive Reading as a Means of Input to L2 Learning', *Journal of Research in Reading* 12(2): 164-78.
- Ujiie, J., and S. Krashen
2002 'Home Run Books and Reading Enjoyment', *Knowledge Quest* 31(1): 36-37.
- Young, A.
2001 'Reading and the Non-academic Learner: A Mystery Solved', *System* 29(4): 51-466. [is this right?]
- Yuan, Y.P., and T. Nash
1992 'Reading Subskills and Quantity Reading', Selected Papers from *The Eighth Conference on English Teaching and Learning in the Republic of China*: 291-304.

Vygotsky, L.S.

1978 *Mind in Society* (Cambridge, MA: Harvard University Press).

West, R., K. Stanovich and H. Mitchell

1993 'Reading in the Real World and its Correlates', *Reading Research Quarterly*
28: 35-50.