Exploring the Effect of Extensive Reading on Young English Learners’ Second Language Writing Achievement in a Vietnamese Primary School Setting

Minh Châu Nguyễn
University of Finance-Marketing, Vietnam
Email: nmchau@ufm.edu.vn

John R. Baker
Ton Duc Thang University, Vietnam
Email: drjohnrbaker@tdtu.edu.vn

Received: 26 December 2021
Reviewed: 12 September 2022-1 December 2023
Accepted: 28 December 2023
Published: 31 December 2023

Abstract
Extensive reading (ER) provides young English language learners (YELLs) many benefits, including the sustainable development of writing (e.g., the reading-writing relationship). However, this area is under-researched. To address this, this quasi-experimental study investigated the effect of ER on YELLs’ second language writing achievement using an enhanced ER treatment. Pre- and post-Cambridge Primary English Test (PET) samples were collected and evaluated via the PET analytic rubric total band and subscale scores (Content, Communicative Achievement, Organization, Language). The results indicated that the EG outperformed the CG in total band score and three subscale bands (Content, Organization, Language), and the EG and CG performed comparably well in the fourth (Communicative Achievement). As prior research specific to ER’s effect on YELLs’ writing achievement in the Vietnam context is noticeably understudied, this study offers a starting point for future investigations.

Keywords: EFL; extensive reading; second language writing; young English language learners; Cambridge PET

Introduction
Early extensive reading (ER) literature explicated the belief that there is a positive reading-writing relationship between ER and writing in that extensive reading positively affects a person's ability to write: A well-read person simply has a much larger and richer set of images of what a text can look like (Flower et al., 1980). Krashen, for instance, in a series of arguments, posited that ER is a sustainable activity that results in panacea-like benefits for all sorts of L2 learning, to include writing (Chan & Krashen, 1992; Krashen, 1982a; 1982b; 1987; 1994), specifically arguing that ER is responsible for the acquisition of planned discourse and that students acquire the feel for the style of sophisticated writing via large amounts of reading for meaning.
Contrasts of reading with formal writing instruction have also been made, arguing that the former, rather than the latter, is responsible for good writing. Smith (1988), for instance, addressed the great amount of instructional time needed to make substantial writing gains, explaining that there is insufficient time to learn to write in formal educational settings. Instead, it must be that “we learn to write without formal effort” and “without knowing we are learning or what we learn” (p. 560), e.g., through reading. Krashen (1987), similarly, discounting the need for formal writing instruction, posited that neither formal instruction nor writing practice is responsible for good writing, but that ER “is responsible for our competence in writing style” (p. 14).

Following the increased support in L2 literature regarding the positive effect ER has on L2 writing, a limited number of empirical investigations of its effect on young English language learners’ (YELLs) L2 writing began to be undertaken in international contexts in the late 1980s (Ahmed & Ahmed, 1987; Elley & Mangubhai, 1983; Hafiz & Tudor, 1989; Iliyas et al., 2015; Im et al., 2010; Irvine, 2007; Sari, 2013; Tsang, 1996). However, such investigations were, and still are, noticeably understudied in the Vietnamese YELL context. To address this, this study is intended to explore this lacuna in the literature.

**Literature review**

Education in Vietnam has a long history (London, 2011), with primary attendance and graduation rates slowly rising from less than 3% in the 1800s through the 1980s (Wright, 2020), when Vietnam, in the 1990s, accepting that the education system needed to be adapted to serve the needs of the new market economy (Biddington & Biddington, 1997), introduced new education objectives (e.g., Vietnam’s Universal Primary Education Policy). Following this, enrollment grew to 98.6% by 2010 (Attfield & Vu, 2019), and curricula advanced to 35 weeks a year starting at grade 3. The National Foreign Languages Project 2008-2020 also provided contributory objectives, i.e., students at all educational levels should have a good command of a foreign language to compete in the global market (Freeman & Drean, 2017).

To meet the National Foreign Languages Project’s goals, learning objectives were set, a specific number of words to be learned (e.g., grade 5, 180–200 words) (Nguyen & Nguyen, 2017), and a standard set of materials suggested by the MOET became commonplace (Moon, 2005). However, these texts, which typically have short passages, have been, following contrasts of the sort of reading material (short passages) found in traditional EFL textbooks with that used in ER (longer texts, e.g., graded readers) (Krashen 1982a), have been criticized as inappropriate for primary level YELLs’ language acquisition (Moon, 2005).

Accepting that providing appropriate materials for YELLs at the primary level is of great importance to sustainably facilitate learning gains, teachers in Vietnam have been allowed some agency to select materials (DOET document 1563/HD-SGD & ĐT) (Nguyen et al., 2014). School supervisors, for instance, have been found to show more liberal outlooks, e.g., teachers should “be flexible and creative with their teaching methods provided that their decisions and choices work well with the students in the class” (Nguyen et al., 2014, p. 9).

Summarizing this more liberal atmosphere, Le et al. (2014) explained that there “are very few documents, publications, or explicit guidance on how to teach and that teachers are allowed to be flexible with their teaching methods provided that they convey the mandated pool of knowledge and cater to diverse students’ needs and capacities” (p. 9).

Accepting that students in Vietnam are, in what Day et al. (1998) described as an English-input impoverished environment, current EFL textbooks (which have short passages) need to be supplemented, and there is some freedom in the selection of materials, many teachers in Vietnam...
have turned to ER, albeit empirical investigations regarding ER’s effect on YELLs’ L2 writing achievement in this context are underexplored.

Empirical explorations of extensive reading and its effect on second language writing in Vietnam

Following Vietnam’s long historical trajectory of educational advances and reforms, ER similarly found its place in Vietnam’s educational context. In line with the pattern illustrated in international ER primary education literature (Ahmed & Ahmed, 1987; Elley & Mangubhai, 1983; Hafiz & Tudor, 1989; Iliyas et al., 2015; Im et al., 2010; Irvine, 2007; Sari, 2013; Tsang, 1996) research that demonstrated that ER is a sustainable activity which provides a variety of L2 benefits, to include writing, a limited number of empirical ER investigations in the Vietnamese context began to be published in 1999 (Anh, 2015; Chang & Renandya, 2017; Do, 2008; DoHuy et al., 2006; Duong et al., 2017; Hue, 2017; Mai, 2018; Nguyễn, 2018; Renandya et al., 1999; Thi, 2018; Tran, 2018; Vu & Nguyen, 2017). However, these studies (N = 13) illustrated a gap in the literature, as they went in a different direction than those undertaken in international settings. That is, while many of the international studies were done with YELLs, studies in Vietnam focused on older learners, and the majority of these (N = 10), although they showed positive gains in other L2 areas, did not address writing achievement (Anh, 2015; Chang & Renandya, 2017; Do, 2008; DoHuy et al., 2006; Duong et al., 2017; Nguyễn, 2018; Thi, 2018; Tran, 2018; Vu & Nguyen, 2017). The remaining three studies with older learners, however, explored the effect of ER on students’ L2 writing achievement, showing positive effects (Hue, 2017; Mai, 2018; Renandya et al., 1999)

Gap in the literature

Extensive reading has a long history, and its effects on L2 writing achievement with YELLs have received moderate attention in international literature. Additionally, Vietnam has reached a point that using ER to facilitate writing achievement can be adopted. However, the extant literature shows that ER has received limited attention in the Vietnamese context, where studies have explored its effect on L2 writing with adults. As such, this relationship with Vietnamese YELLs remains understudied. To address this gap, this study explored the effect of ER on L2 writing achievement with YELLs in a Vietnamese primary school setting. To investigate this, two research questions were posed:

RQ1. Is there a significant difference in total band score gains between the EG and CG as specified by the Cambridge PET Writing rubric?

RQ2. Is there a significant difference in band score gains between the EG and CG on the subscales of the Cambridge PET writing rubric (Content, Communicative Achievement, Organization, Language)?

Research method

To gather the data needed to address the two research questions, a quasi-experimental study was conducted in a Vietnamese primary school setting (Figure 1)
Following the research design (Figure 1), using convenience sampling, a control group (CG) and one experimental group (EG), were formed from two intact fifth-grade classes. Participation was explained to the children in class (Brown, 2000), and child and parent consent forms were subsequently sent home with the children and collected. Participation was voluntary, no coercion or incentives were employed (Creswell & Creswell (2018), and ethical approval was received by the university research committee. A total of 45 students consented to participate. Two students in the EG, however, did not complete the study, i.e., they transferred to different schools within the semester. Four additional students’ data in the EG were not utilized due to non-participation. In total, 39 students completed the study (CG, n = 11: 4 male, 7 female, aged 10-11 years; EG, n = 28; 14 male, 14 female, aged 10-11 years). These demographics illustrate that the CG and EG were similar in age and gender.

The study began with a pretest. Each group (CG and EG) was given a writing prompt from the Cambridge PET Writing Tests Practice Exam Book (Part 2-Story Narrative) and asked to write a short narrative using the same writing prompt:
Your English teacher has asked you to write a story. Your story must begin with this sentence: Jo looked at the map and decided to go left. Write your story.

Both tests were administered under identical testing conditions (i.e., PET writing test timing protocol was followed, 25 minutes), and PET Rubric and protocol were employed for scoring. Student responses were evaluated individually by two evaluators with five or more years of experience. In the event of a discrepancy, the response was scored by a similarly experienced third evaluator.

After the pretest, both groups were taught using the standard school curriculum: Using the EFL textbook, instruction involved oral practice (short spoken dialogues), listening, grammar, vocabulary, phonics, reading short passages, retelling dialogues in the form of short stories, and guided sentence-level writing. The EG, however, received an enhanced ER treatment (Mermelstein, 2015).

Drawing on the extant literature, an enhanced ER treatment was designed. This included several steps. The first step was a prereading assessment, Betts’ (1946) Five Finger Test, to determine students’ reading levels for the purposes of choosing texts (i.e., graded readers) for the class library. Texts were chosen to be below, at, and above the students' i levels (current level of competence) (Chan & Krashen, 1992; Day et al., 1998). Additional subjective features that contribute to readability were also considered (background knowledge, interest) (Gunning, 2003; Weaver, 2000).

The next step was to introduce the class library. Here, students were shown how to self-select books (Day, 2002 & Taguchi, 2004) using features of interest and the aforementioned 95 percent vocabulary comprehension technique (Betts 5 Finger Method) (Anderson, 1999)

Once the library had been introduced, the enhanced ER treatment began. This, drawing on the extant literature, included the following:

- Thirty minutes per class was allotted to SSR (sustained silent reading) (Day, 2002 & Taguchi, 2004; Waring & McLean, 2015)
- The teacher modeled SSR behavior by sitting and reading books from the class library during SSR periods (Day et al., 1998).
- The teacher encouraged students by meeting with them to talk about what they were reading and providing additional activities (e.g., storytelling) (Waring & Mclean, 2015; Renandya & Jacobs, 2002).
- Students were encouraged to read at least one book a week (Brumfit, 1997).
- Students kept a reading log of the number of books and pages read (Irvine, 2007; Sari, 2013).
- Students self-selected books to take home (Hafiz & Tudor, 1989)
- A poster with students’ names and progress was displayed in class to provide encouragement (Davis, 1995; Renandya et al., 1999).

At the end of one semester, a post-test was given to each group, a PET Part 2 narrative writing test identical to the pretest, i.e., the same writing prompt was given, and the writing samples were analyzed in the same fashion according to PET writing test scoring procedure.

To explore students’ writing achievement, pre- and post-performance of the CG on the overall band score and subscale band scores were compared using inferential statistics (Mann-Whitney U test).
Results

This quasi-experimental study investigated the effect of ER on the L2 writing achievement of YELLs in a Vietnamese primary school setting as indicated by an inferential statistical analysis (Mann-Whitney U test) of the students’ achievement on the Cambridge PET’s overall and subscale bands (Communicative Achievement, Organization, Language). The results for each area are presented in the following sections.

Findings for RQ1: total band performance

RQ1 investigated the following question: Is there a significant difference in total band score gains between the EG and CG as specified by the Cambridge PET writing rubric? It was found that the EG (Mdn = 3.5) outperformed the CG (Mdn = 2) by 1.5 bands. The Mann-Whitney U test indicated that this result approached significance (U = 108; z = -1.44; p = .079; r = .23) (Tables 1, 2).

<table>
<thead>
<tr>
<th></th>
<th>N</th>
<th>Mdn</th>
<th>Mean Rank</th>
<th>Sum of Ranks</th>
</tr>
</thead>
<tbody>
<tr>
<td>EG</td>
<td>28</td>
<td>3.5</td>
<td>21.64</td>
<td>606</td>
</tr>
<tr>
<td>CG</td>
<td>11</td>
<td>2</td>
<td>15.82</td>
<td>174</td>
</tr>
</tbody>
</table>

Findings for RQ2: subscale band performance

RQ2 investigated the following question: Is there a significant difference in band gains between the EG and CG on the subscales of the PET writing rubric (Content, Communicative Achievement, Organization, Language)? To explore the effect of ER on YELLs achievement in each subscale area, RQ2 was separated into four sub-questions: RQ2a, RQ2b, RQ2c, and RQ2d). To order the results, the following subsections address each.

Findings for RQ2a: the subscale band of content

RQ2a explored the following question: Is there a significant difference in band score gains between the EG and CG’s scores on the subscale of Content as specified by the Cambridge PET Writing rubric? It was found that the EG (Mdn = 1) outperformed the CG for the subscale band of content (Mdn = 0.75) by .25 bands. The Mann-Whitney U test indicated that this result approached significance (U = 103.5; z = -1.6; p = .06; r = .26) (Tables 3, 4).

<table>
<thead>
<tr>
<th></th>
<th>N</th>
<th>Mdn</th>
<th>Mean Rank</th>
<th>Sum of Ranks</th>
</tr>
</thead>
<tbody>
<tr>
<td>EG</td>
<td>28</td>
<td>1</td>
<td>21.8</td>
<td>610.5</td>
</tr>
<tr>
<td>CG</td>
<td>11</td>
<td>.75</td>
<td>15.41</td>
<td>169.5</td>
</tr>
</tbody>
</table>
Table 4. Findings for RQ2a: The Subscale of Content

<table>
<thead>
<tr>
<th></th>
<th>Values</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mann-Whitney U</td>
<td>103.5</td>
</tr>
<tr>
<td>Z</td>
<td>-1.6</td>
</tr>
<tr>
<td>Asymptotic Significance (1-tailed)</td>
<td>.06</td>
</tr>
<tr>
<td>Exact Significance (1-tailed)</td>
<td>.06</td>
</tr>
</tbody>
</table>

Findings for RQ2b: the subscale band of communicative achievement

RQ2b explored the following question: Is there a significant difference in band score gains between the EG and CG’s scores on the subscale of Communicative Achievement as specified by the PET writing rubric? It was found that the EG and CG performed comparably well in this area (Mdn = 1). However, the Man-Whitney U test indicated that this result was not significant (U = 152; z = .06; p = .475; r = .01) (Tables 5, 6).

Table 5. Findings for RQ2b: the subscale of communicative achievement

<table>
<thead>
<tr>
<th></th>
<th>N</th>
<th>Mdn</th>
<th>Mean Rank</th>
<th>Sum of Ranks</th>
</tr>
</thead>
<tbody>
<tr>
<td>EG</td>
<td>28</td>
<td>1</td>
<td>20.07</td>
<td>562</td>
</tr>
<tr>
<td>CG</td>
<td>11</td>
<td>1</td>
<td>19.82</td>
<td>218</td>
</tr>
</tbody>
</table>

Table 6. Findings for RQ2b: the subscale of communicative achievement

<table>
<thead>
<tr>
<th></th>
<th>Values</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mann-Whitney U</td>
<td>152</td>
</tr>
<tr>
<td>Z</td>
<td>-0.06</td>
</tr>
<tr>
<td>Asymptotic Significance (1-tailed)</td>
<td>.481</td>
</tr>
<tr>
<td>Exact Significance (1-tailed)</td>
<td>.475</td>
</tr>
</tbody>
</table>

Findings for RQ2c: the subscale band of organization

RQ2c investigated the following question: Is there a significant difference in band score gains between the EG and CG’s scores on the subscale of Organization as specified by the Cambridge PET writing rubric? It was found that the EG (Mdn = 1) outperformed the CG (Mdn 0) for the subscale band of content by one band. The Mann-Whitney test indicated that this result was significant (U = 93.5; z = -1.95; p = .031; r = .31) (Tables 7, 8).

Table 7. Findings for RQ2c: the subscale of organization

<table>
<thead>
<tr>
<th></th>
<th>Values</th>
</tr>
</thead>
<tbody>
<tr>
<td>EG</td>
<td>28</td>
</tr>
<tr>
<td>CG</td>
<td>11</td>
</tr>
</tbody>
</table>

Table 8. Findings for RQ2c: the subscale of organization

<table>
<thead>
<tr>
<th></th>
<th>Values</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mann-Whitney U</td>
<td>93.5</td>
</tr>
<tr>
<td>Z</td>
<td>-1.95</td>
</tr>
<tr>
<td>Asymptotic Significance (1-tailed)</td>
<td>.031</td>
</tr>
<tr>
<td>Exact Significance (1-tailed)</td>
<td>.026</td>
</tr>
</tbody>
</table>
Findings for RQ2d: the subscale band of language

RQ2d addressed the following question: Is there a significant difference in band score gains between the EG and CG’s scores on the subscale of Language as specified by the Cambridge PET writing rubric? It was found that the EG (Mdn = 1) outperformed the CG (Mdn = .25) for the subscale band of language by .75 bands. The Mann-Whitney U test indicated that this result was significant (U = 90.5; z = -2; p = .025; r = .32) (Tables 9, 10).

Table 9. Findings for RQ2d: The Subscale of Language

<table>
<thead>
<tr>
<th></th>
<th>N</th>
<th>Mean Rank</th>
<th>Sum of Ranks</th>
</tr>
</thead>
<tbody>
<tr>
<td>EG</td>
<td>28</td>
<td>1</td>
<td>22.27</td>
</tr>
<tr>
<td>CG</td>
<td>11</td>
<td>.25</td>
<td>14.23</td>
</tr>
</tbody>
</table>

Table 10. Findings for RQ2d: The Subscale of Language

<table>
<thead>
<tr>
<th></th>
<th>Values</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mann-Whitney U</td>
<td>90.5</td>
</tr>
<tr>
<td>Z</td>
<td>-2</td>
</tr>
<tr>
<td>Asymptotic Significance (1-tailed)</td>
<td>.025</td>
</tr>
<tr>
<td>Exact Significance (1-tailed)</td>
<td>.023</td>
</tr>
</tbody>
</table>

Summary of the results

The results indicated that the EG outperformed the CG in the area of total band scores by 1.5 bands, and that this approached significance (p = .079). The results further indicated that the EG outperformed the CG in three of the four subscale bands (Content, Organization, Language) by between .25 and one full band, the first of which approached significance (p = .6), and the latter two of which were significant (p = .31; p = .25). It was additionally found that the EG and CG performed comparably well (one band) in the fourth (Communicative Achievement), albeit this result was not significant (p = .475).

Discussion

This study explored the effect of ER on YELLs’ L2 writing achievement in a Vietnamese primary school setting. The first research question, RQ1, investigated whether there was a significant difference in total band score gains between the EG and CG as specified by the Cambridge PET writing rubric. The results indicated that the EG outperformed the CG in the area of total band scores by 1.5 band gains and that this approached significance. This result suggests that the ER treatment positively affected the EG’s writing performance. This supports global literature that has found that ER has a positive reading-writing relationship effect on L2 writing (Chan & Krashen, 1992; Krashen, 1982a; 1982b; 1987; 1994). This result is also in accordance with global literature which has specifically demonstrated that ER positively affects YELLs’ L2 writing achievement (Ahmed & Ahmed, 1987; Elley & Mangubhai, 1983; Hafiz & Tudor, 1989; Iliyas et al., 2015; Im et al., 2010; Irvine, 2007; Tsang, 1996), to include performance on the Cambridge PET writing test (Sari, 2013).

Addressing the second research question, RQ2 (Is there a significant difference in band score gains between the EG and CG on the subscales of the Cambridge PET writing rubric: Content, Communicative Achievement, Organization, Language?), it was found that the EG outperformed the CG in three of the four subscale bands (Content, Organization, Language) by between .25 and one full band, the first of which approached significance, and the latter two of
which were significant. These results also suggest that the ER treatment positively affected the EG’s writing performance. Additionally, these results are aligned with arguments that reading supports apprenticing writers in various ways relevant to good writing (Grabe, 2003; Krashen, 2004).

Examining the two areas that were found to be significant (Organization, Language), the significant result regarding organization supports literature that has found that ER facilitates L2 writers’ organization (Hyland, 2004; Jordan, 2004; Kaplan, 1996; Krashen, 1982b), as well as ER studies with YELLS (Irvine, 2007), in that a well-read person simply has a much better idea of what text can look like (Flower & Hayes, 1980). The results do not, however, support the limited amount of literature that has reported no significant gains in organization (Im, 2010; Tsang 1996). The results regarding language also support literature that has shown that ER facilitates L2 writing in the area of language (Elley & Mangubhai, 1983; Im, 2010; Irvine, 2007; Tsang, 1996).

Conclusion

Examining the results as a whole, this study extends global literature that has demonstrated that ER positively affects YELLS’ L2 writing achievement but has yet to demonstrate such gains in the Vietnamese context (Ahmed & Ahmed, 1987; Elley & Mangubhai, 1983; Hafiz & Tudor, 1989; Iliyas et al., 2015; Im et al., 2010; Irvine, 2007; Sari, 2013; Tsang, 1996). The results similarly extend literature which has shown that ER supports L2 writing achievement in the Vietnamese context (i.e., with older learners, adults) but has yet to demonstrate such gains with YELLS (Hue, 2017; Mai, 2018; Renandya et al., 1999).

Suggestions for future study

As prior research specific to ER and its effect on YELLS’ (primary students) writing achievement in the Vietnam region is noticeably absent, this study’s findings are beneficial to instructors at primary schools in Vietnam and potentially those in the larger Asian context and abroad. The results additionally further ER literature and thus may inform the broader research community. However, the study contains limitations in administration and scope that provide directions for further research.

Firstly, regarding administration, the study was initially intended to have roughly equal-sized groups (EG, CG). However, while approved by the school administration, participation in the investigation was voluntary. As such, all of the students in the EG, excited about the idea of ER, consented. However, a portion of the CG elected not to participate. As this study was done with preformed intact classes, the EG was larger than the CG. Hence, future studies are encouraged to explore groups that are more equitable in size. Secondly, the EG subjects were aware of the special nature of the ER program, a challenge relevant to ER studies, and this may have contributed to the positive gains (Elley & Mangubhai, 1983). As such, future studies are suggested where the CG and EG are selected from different schools. Thirdly, separate instructors taught the EG and CG groups. Although there was regular communication between the two instructors to ensure that the teaching of the standard curriculum was as equal as possible, having one researcher teach both classes could have helped to keep many variables constant (e.g., rapport with students, pedagogy). This is a potential area for further exploration. Fourthly, the location for ER used by the EG was less than optimum in terms of comfort (temperature, excessive outside noise, size, and furniture), i.e., the classroom lacked air conditioning, was near the playground, and was without comfortable spaces for the students to retire while reading. As such, the lack of comfort begs the question whether these conditions affected the results of the EG’s performance. Further research in this area
is needed. Additionally, the reading time available for ER was limited to 30 minutes, twice per week. Thus, the question arises as to whether additional time allotted to ER would have provided different results. A final area with regard to the writing prompt emerged. Although the extant literature illustrates that employing the same writing prompt for the pre and post-test is a suggested practice (Sari, 2014), several students questioned the face validity of being tested repeatedly with the same prompt. As such, this raises a question of validity that might be addressed in future explorations.

The matter of scope is also relevant. This study explored the effects of ER treatment on the L2 writing achievement of YELLS in Vietnam at the elementary level, and thus future investigations are suggested to explore whether the results are generalizable to Vietnam’s many other education settings (e.g., junior high schools, high schools, colleges, universities, and private language centers).

Declaration of conflicting interest
The authors report no conflicts of interest. Inquiries regarding this article may be directed to the corresponding author John R. Baker, at the Creative Language Center, Ton Duc Thang University, Ho Chi Minh City, Vietnam, via email (drjohnrbaker@tdtu.edu.vn).

Funding acknowledgements
The research received no external funding.

References


