

## Contemporary Methods of Indonesian Research in Improving Students Writing Skills Through Extensive Reading Method

Muhammad Aldino Putra Difa<sup>1</sup>, Nur Mukminatien<sup>2</sup>, Francisca Maria Ivone<sup>3</sup>

<sup>1,2,3</sup>Universitas Negeri Malang, Indonesia

<sup>1</sup>[muhammad.aldino.2402218@students.um.ac.id](mailto:muhammad.aldino.2402218@students.um.ac.id), <sup>2</sup>[nur.mukminatien.fs@um.ac.id](mailto:nur.mukminatien.fs@um.ac.id),

<sup>3</sup>[francisca.maria.fs@um.ac.id](mailto:francisca.maria.fs@um.ac.id)

**Abstract.** Extensive Reading (ER) has been widely recognized as an effective approach for improving English as Foreign Language (EFL) students' writing skills. However, variations in ER implementation and contextual factors have produced differing findings, particularly in Indonesian EFL settings. This study aims to examine contemporary Indonesian research on the use of ER in enhancing students' writing skills and to identify the outcomes, differences, and challenges across studies. The study employed a Secondary Data Analysis design using the Preferred Reporting Items for Systematic Reviews and Meta-Analysis (PRISMA) framework to collect and analyze relevant studies. Five experimental studies conducted in Indonesian contexts were selected and compared based on ER implementation, writing assessment subscales, descriptive statistics, and effect sizes. The findings showed that ER generally contributed positively to students' writing performance, with treatment groups outperforming control groups in most studies. The greatest improvements were found in content and language use, indicating that ER supports idea development and grammatical competence through continuous exposure to meaningful reading materials. Effect size analysis also revealed medium to large impacts in several studies, confirming the pedagogical effectiveness of ER in Indonesian EFL classrooms. Nevertheless, one study reported non-significant results due to limited reading resources, low reading engagement, and insufficient literacy environments. Overall, ER is a promising approach for improving writing proficiency among Indonesian EFL students, although successful implementation requires adequate resources, sustained reading habits, and supportive instructional practices.

**Keyword:** Extensive Reading; Writing Skills; Indonesian EFL Learners; Writing Performance; Secondary Data Analysis

Received: 10-02-2026

Revised: 05-03-2026

Accepted: 30-03-2026

Corresponding Author: Muhammad Aldino Putra Difa; [muhammad.aldino.2402218@students.um.ac.id](mailto:muhammad.aldino.2402218@students.um.ac.id)



This work is licensed under a [Creative Commons Attribution-NonCommercial 4.0 International License](https://creativecommons.org/licenses/by-nc/4.0/)

## INTRODUCTION

Extensive reading (ER hereinafter) has been one of the widest advocated terms used as the umbrella encompassing plethora of approaches and mediums of pedagogy, particularly in developing writing skills (Mermelstein, 2015; Beglar & Hunt, 2014; Day & Bamford, 2002; Waring, 2006; Yamashita, 2013; Day, et al., 1998; McLean & Rouault, 2017; Taye & Tehsome, 2025; Hafiz and Tudor, 1989; Richards & Platt, 2003). As such, varying interpretation have been acknowledged by researchers alike. For example, as cited from Warring & McLean (2015), Carrell and Carson's (1997) definition of ER "generally involves rapid reading of large quantities of material or longer readings." This interpretation focusses on favouring sheer volume (e.g. a whole book or two) to measure students' general understanding and knowledge of said input. In the later interpretation of *The Top Ten Principles for Teaching ER* (as outlined in Day and Bamford (1998, 2002)); ER's definition has solidified and involved several more principles such as "the process of reading should be interesting, easy, fast, silent, pleasurable, individual and self-selected, and involve reading as much as possible while guided by a teacher as a role model" (Warring & McLean, 2015). Essentially, these four elements encompass the extensive reading description in practice based on later studies by Warring & McLean (2015) and Park & Ro (2015):

- a. Fluent comprehension
- b. High reading speed
- c. Reading large amounts of text
- d. Focus on meaning of text

However, even with said principles and elements; admittedly, researchers still cannot uniformly align said criterion with the myriads of situational aspects and conditions. In this case, the fine line of determining when a reading program will be called an ER program is cloudy to say the least. Warring & McLean (2015) summarized the questions researchers might have when exploring about ER with the four elements in mind:

*"Exactly what reading speed needs to be attained to call it 'fluent'? Is this a constant for all readers and texts or does it vary, and if so, how? What do we mean by 'fluent comprehension'? Will 90% comprehension, or 98%, suffice?"* (p.162)

As the result of this discussion, the majority of studies have varying degree of what can they call as ER program, especially with each specific-readily media output available (e.g. source and type of literature/book) in each research with topics regarding the use of ER in improving students' writing skill. This is, by no means, made research papers including ER to be incomparable or unjustifiable by result. This means that the study of human language knowledge and comprehension is just a *complex structure* individually (Brown & Hagoort, 2000; Marslen-Wilson & Tyler, 2007; Zwitserlood, 2018). Fortunately, researchers have unanimously agreed that; in order to perceive students' comprehension progress by ER, a metric measurement has to be included in the form of a *summative assessment*. This usually refers to some pre-tests and post-tests that students have to undergone to measure their improvements from an ER program (no matter the variance of ER utilized by the study). Some research papers (see Kargar Behbahani et al., 2025; Sakurai, 2017; Mermelstein, 2015; Lee & Hsu, 2009; Renandya & Jacobs, 2002) have been studied to refer to particular *subscales*, the design established by Jacobs, Zinkgraf, Wormuth, Hartfield, and Hughey (1981), which is commonly used to measure this summative type. This subscale includes: 1) Content, 2) Organization, 3) Vocabulary, 4) Language use, and 5) Mechanics (spelling). These subscales are told to be sufficient in Jacobs et al. (1981), however later studies proven that another subscale are needed to be added and modified in the design. This addition is said to be *fluency* and is argued in both Lee & Hsu (2009) and Mermelstein (2015) to be largely influential as it is interconnected with vocabulary learned in the program. Therefore, six of these

subscales are deemed as pivotal in learning students' improvements over the ER program in such studies going further.

Recent Indonesian research has also seen a surge in using ER approach to improve students writing skills. Some research even stated the use of ER is highly successful and important to implement (see Anandari & Iswandari, 2019; Wijaya, 2021) while others have identified issues regarding the implementation, specifically in Indonesia (see Renandya, et al., 2021). Although supplies of these evidences clearly show how the implementation of ER in Indonesia, the present study suggest an overview of ER results in Indonesia. Assuming that the current world trend on ER is the same with Indonesian research trend on writing skills has, this study's objective is to collect data from studies and to compare the results of ER implementations in Indonesia with other countries and give an overview of possible commonly shared issues, as well as to identify whether there are the specific issues in Indonesia.

## METHOD

Over several decades, classroom studies dealing with ER have included varying experimental research approaches in discerning control group (CG) and treatment/experiment groups (TG)' effect to a certain program. As experimental design studies are ample in supplies regarding ER in developing writing skills; not to mention the crystal transparency of their empirical evidence/ result data, this study will employ Secondary Data Analysis method by comparing results in numerical and graphical data with the description of data explanation (justifying each research unique scenario/ demographic). Due to the limitation of the study, this study will only explore the results of *pre-test* and *post-tests* from studies dealing with ER in students' writing skills from Indonesia and will also explore the distinction of these Indonesian studies (e.g. the type of ER intervention, program duration, and limitations. Thus, the present study wishes to make a comparison in: **a) The affecting systematic differences between each studied research and b) The Comparison of overall results deviation between the Indonesian studies.**

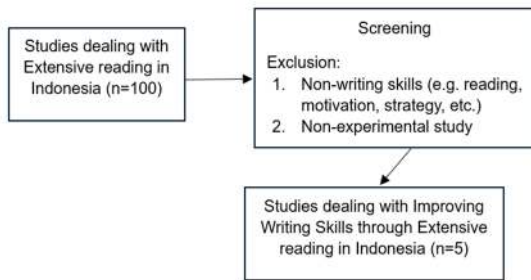
In order to justify the findings and to compare the results evenly, the present study utilize the modified measurement design from Jacobs et al. (1981) and Lee & Hsu (2009)'s subscales. These subscales used to measure the results are:

1. Content : The substance, ideas, and information expressed within a text.
2. Organization : The logical, coherent, and structural arrangement of the text.
3. Vocabulary : The range and precision of words used to facilitate a message.
4. Language use : The accuracy, appropriateness, and effectiveness of language to convey meaning.
5. Mechanics/ spelling : The technical accuracy and conventions used to ensure a piece of writing to make it clear.
6. Fluency : The speed, ease, and smoothness of the text produced.

On the note side, not all studies will have direct result with the same subscales included. As such, studies without one of these categories will still be included with said category omitted (still visible with no score, thus; ineligible from the scale comparison).

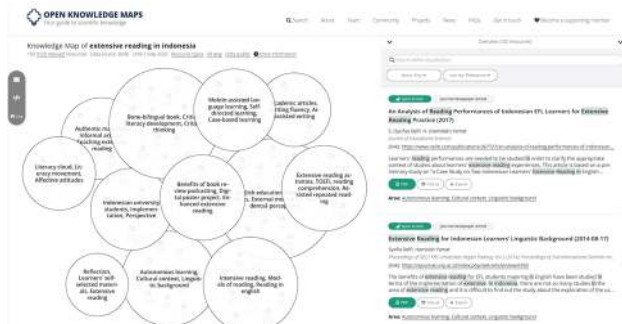
To collect relevant studies of ER in Indonesia, the Preferred Reporting Items for Systematic Reviews and Meta-Analysis (PRISMA) framework was selected as the most appropriate research design. This approach is widely supported in the literature (see Permana & Rohmah, 2024; Rohmah et al., 2024; Peng et al., 2022) for its structured method of analyzing articles as primary data sources. To ensure a rigorous process, several key components were emphasized: (1) clearly defining the research objectives, (2) following systematic procedures for article identification, (3) applying explicit strategies to reduce potential bias during selection and evaluation, (4) ensuring transparency throughout the methodology, and (5) conducting a thorough

and systematic analysis (Meline, 2006; Evans & Benefield, 2001). Only studies published in peer-reviewed journals or conference proceedings were considered eligible for inclusion.



**Figure 1** PRISMA flowchart of ER in developing Indonesian students’ writing skills

The following figure illustrates the data collection procedures, conducted in 30<sup>th</sup> April 2026. The articles are archived in the Open Knowledge Maps (<https://openknowledgemaps.org>). The use of this Open Knowledge Maps as a tool and database has seen a surge of popularity in studies regarding collective research analysis (Permana & Rohmah, 2024; Rohmah et al., 2024; Kraker et al., 2016), thus the present study deemed the tool as a sufficient and reliable source. The first step of the data collection was inputting ‘Extensive Reading in Indonesia’ keywords into the database resulting in 100 (one hundred) articles (see Figure 2). Of the 100 articles collected, 88 (eighty-eight) articles were excluded based on the first exclusion criteria due to its irrelevance in writing skill context (dealing about reading skills, digital skills, reading motivation, etc.). The second exclusion criteria deal with relevance on experimental design research, only managed to yield 5 concrete recent studies that are suitable for the data of the present study. These studies consist of the study from: Atun, et al., (2025); Nugrahini & Rakhmawati, (2022); Linuwih, (2021); Fitriansyah & Miftah, (2020); and Umar, et al., (2013). All of which, use descriptive writing as their measurement in their pre-tests and post-tests.



**Figure 2** The mapping based on the query from Open Knowledge Maps

After accomplishing the data collection procedure, the data analysis procedure will take place by comparing the mean score results of participants’ pre-tests and post-tests from the four said studies into a table which follows the adapted design from Jacobs et al. (1981) and Lee & Hsu (2009) criteria of performance in extensive reading program (described further in **Table 1**). Thus, the present study wishes to make a comparison in: a) The affecting social and cultural background differences between each studied research and b) The Comparison of overall results deviation between the Indonesian studies.

## RESULTS AND DISCUSSION

### The Affecting Systematic Differences

The present study identified several systematic differences across the five reviewed studies, particularly in terms of the ER methods applied, materials used, and duration of the ER programs. One notable difference appears in the study by Atun et al. (2025), which implemented ER through a skimming strategy. According to Maxwell (1969), cited in Diaz & Laguado (2013, p. 138), skimming refers to obtaining the main idea or gist of a text quickly while developing learners' inferencing skills through systematic practice. Ria (2016) further argues that skimming enables students to engage with texts according to their interests and needs. Although the implementation of skimming within ER remains debatable, the study still reported positive improvements in students' writing skills.

In terms of ER implementation, all reviewed studies conducted the treatment between the pre-test and post-test phases. However, the duration and materials varied. For instance, Linuwih (2021) implemented a six-week ER program in which students selected authentic books from predetermined sources. Meanwhile, Nugrahini & Rakhmawati (2022) used graded readers, specifically the Interchange series by Jack C. Richards, as the primary reading material. Despite using different materials, both studies reported positive outcomes from the ER programs.

Several studies also highlighted students' preferences regarding reading materials. Lien (2010) found that participants generally preferred authentic books over graded readers because authentic texts offered more engaging plots and concepts for adult EFL learners. Similarly, Adhitya & Novita (2021) reported that learners enjoyed self-selecting authentic reading materials and felt less pressured during the reading process. On the other hand, Nuttall (1982) emphasized that graded readers remain beneficial because they provide appropriate language levels, manageable text length, and accessible themes for learners at different proficiency levels.

### CG and TG Results in Indonesian Study

The present study identified and analysed the writing subscales used across the selected studies (see Table 1). The results were derived from participants' essay writing scores based on each study's respective scoring rubric. Although not all studies employed the complete framework proposed by Jacobs et al. (1981) and Lee & Hsu (2009), all major subscales were represented within the reviewed studies. Among these, the Content subscale appeared consistently across all studies, indicating its central role in assessing English writing performance. Previous studies have similarly positioned content as the primary criterion in writing assessment rubrics (Jacobs et al., 1981; Saddler & Andrade, 2004; Minnich et al., 2018).

**Table 1** The subscales studied within each research

| Subscale            | Atun, et al., | Nugrahini & Rakhmawati, | Linuwih, | Fitriansyah & Miftah, | Umar, et al., (2013) |
|---------------------|---------------|-------------------------|----------|-----------------------|----------------------|
| Content             | ✓             | ✓                       | ✓        | ✓                     | ✓                    |
| Organization        | X             | ✓                       | ✓        | ✓                     | ✓                    |
| Vocabulary          | ✓             | ✓                       | ✓        | X                     | ✓                    |
| Language use        | X             | ✓                       | ✓        | ✓                     | ✓                    |
| Mechanics/ spelling | ✓             | X                       | ✓        | X                     | ✓                    |
| Fluency             | X             | X                       | X        | ✓                     | X                    |

Variation was found in the inclusion of other subscales. The Organization subscale was absent in one study but present in the remaining four, while the Vocabulary subscale appeared in most studies with only one omission. The Language Use subscale was also included in the majority of studies, reflecting its importance in assessing grammatical accuracy and sentence construction. In contrast, the Mechanics/Spelling subscale was excluded in two studies, suggesting that some researchers viewed it as a secondary criterion. The Fluency subscale was the least represented, appearing in only one study. This finding indicates that fluency is not commonly prioritized in writing assessment rubrics related to ER studies in Indonesia.

The present study also compared the pre-test and post-test numerical results across the reviewed studies (see Table 2). Overall, all studies reported improvements in both control groups (CG) and treatment groups (TG), although the degree of improvement varied. In Atun et al. (2025), the CG increased from 49.9 to 60.3, while the TG improved more substantially from 55.2 to 72.8. Nugrahini & Rakhmawati (2022) similarly found higher gains in the TG (10.04 to 13.09) compared to the CG (9.95 to 11.71). In Linuwih (2021), both groups improved, with the TG showing greater progress (72.16 to 81.76) than the CG (71.41 to 75.97).

**Table 2** The descriptive statistic (mean score) procured within the Indonesia studies

| Study                          | CG pre-test | CG post-test | TG pre-test | TG post-test | Significance    |
|--------------------------------|-------------|--------------|-------------|--------------|-----------------|
| Atun, et al., (2025)           | 49.9        | 60.3         | 55.2        | 72.8         | Significant     |
| Nugrahini & Rakhmawati, (2022) | 9.95        | 11.71        | 10.04       | 13.09        | Significant     |
| Linuwih, (2021)                | 71.41       | 75.97        | 72.16       | 81.76        | Significant     |
| Fitriansyah & Miftah, (2020)   | 37.53       | 36.23        | 36.93       | 37.06        | Not Significant |
| Umar, et al., (2013)           | 74.5        | 76.30        | 70.5        | 81.20        | Significant     |

However, Fitriansyah & Miftah (2020) reported different findings. The CG slightly decreased from 37.53 to 36.23, while the TG showed only minimal improvement from 36.93 to 37.06, resulting in a non-significant outcome. In Umar et al. (2013), both groups improved, though the TG (70.5 to 81.20) outperformed the CG (74.5 to 76.30). Findings from Nugrahini & Rakhmawati (2022) revealed that students achieved particularly strong post-test gains in the Content and Language Use subscales. Other subscales across the reviewed studies also demonstrated improvement, although to a lesser extent.

The effect size analysis further demonstrated varying levels of treatment effectiveness across studies (see Table 3). Atun et al. (2025) produced a very large effect size  $d=6.04$ , indicating a substantial difference between treatment and control groups. Similarly, Linuwih (2021) reported a very large effect size  $d=2.24$ . Nugrahini & Rakhmawati (2022), meanwhile, demonstrated a medium effect size  $d=0.58$ . In contrast, standardized effect sizes could not be calculated for Fitriansyah & Miftah (2020) and Umar et al. (2013) due to the absence of standard deviation data.

**Table 3** Comparison to each studied research using effect sizes

| Study              | N (CG/TG) | CG Gain | TG Gain | Cohen's d | Effect Size |
|--------------------|-----------|---------|---------|-----------|-------------|
| Atun et al. (2025) | 25 / 25   | 7.40    | 10.60   | 1.04      | Large       |

|                               |         |       |       |      |        |
|-------------------------------|---------|-------|-------|------|--------|
| Nugrahini & Rakhmawati (2022) | 24 / 35 | 1.76  | 3.05  | 0.58 | Medium |
| Linuwih (2021)                | 34 / 34 | 4.56  | 9.60  | 1.24 | Large  |
| Fitriansyah & Miftah (2020)   | 32 / 32 | -1.30 | 0.13  | —    | Null   |
| Umar et al. (2013)            | 30 / 30 | 1.80  | 10.70 | —    | Null   |

## CONCLUSION

This study displayed that Extensive Reading (ER) in Indonesian research generally has a positive contribution toward improving Indonesian EFL students' writing skills. Across the reviewed experimental studies, treatment groups consistently achieved higher gains than control groups, indicating that ER can effectively support writing development in various educational contexts. The findings further reveal that ER particularly strengthens students' content development and language use, suggesting that continuous exposure to meaningful reading materials enhances learners' ability to generate ideas and apply grammatical structures more effectively in writing. In addition, the effect size analysis confirms that several ER interventions produced medium to large impacts on students' writing performance, highlighting the pedagogical value of ER in EFL instruction. Nevertheless, the findings also indicate that the effectiveness of ER is influenced by contextual factors such as access to reading materials, students' reading habits, motivation, and the type of ER implementation employed. The non-significant findings in one study emphasize that ER alone may not guarantee improvement when supporting literacy environments are limited. Therefore, successful ER implementation requires adequate reading resources, consistent reading engagement, and carefully designed instructional support. Overall, this study reinforces the importance of ER as a promising approach for enhancing writing proficiency in Indonesian EFL contexts while also highlighting the need for further research involving broader participant demographics, longer intervention periods, and more standardized assessment criteria to obtain more comprehensive findings

## REFERENCE

- Adhitya, N., & Novita, J. (2021). Students' responses towards self-selected reading in extensive reading class. *JELA (Journal of English Language Teaching, Literature and Applied Linguistics)*, 3(2), 1–7. <https://doi.org/10.37742/jela.v3i2.53>
- Anandari, C. L., & Iswandari, Y. A. (2019). Extensive reading in Indonesian schools: A successful story. *TEFLIN Journal*, 30(2). <https://doi.org/10.15639/teflinjournal.v30i2/137-152>
- Atun, M., Adam, S., Siahaan, S., Marpaung, J., & Subrata, D. (2025). Improving reading and writing literacy skills through extensive reading in Tanjung Gundap. *Anglo-Saxon: Jurnal Ilmiah Program Studi Pendidikan Bahasa Inggris*, 16(1), 1–12. <https://doi.org/10.33373/as.v16i1.7563>
- Bell, T. (2001). Extensive reading: Speed and comprehension. *The Reading Matrix*, 1(1).
- Brown, C. M., & Hagoort, P. (2000). On the electrophysiology of language comprehension: Implications for the human language system. In *Architectures and mechanisms for language processing* (pp. 213–237). Cambridge University Press.
- Diaz, S., & Laguado, J. C. (2013). Improving reading skills through skimming and scanning techniques at a public school: Action research. *Opening Writing Doors Journal*, 10(1), 133–150.
- Epting, L. K., Gallena, E. M., Hicks, S. A., Palmer, E. N., & Weisberg, T. (2013). Read and think before you write: Prewriting time and level of print exposure as factors in writing and revision. *Journal of Writing Research*, 4(3), 239–259. <https://doi.org/10.17239/jowr-2013.04.03.1>

- Fitriansyah, N., & Miftah, M. Z. (2020). Positive connection of extensive reading and writing fluency in EFL learning. *LET: Linguistics, Literature and English Teaching Journal*, 10(2), 44–70. <https://doi.org/10.18592/let.v10i2.4137>
- Gavelek, J., & Bresnahan, P. (2014). Ways of meaning making: Sociocultural perspectives on reading comprehension. In *Handbook of research on reading comprehension* (pp. 164–200). Routledge.
- Jacobs, H. L., Zinkgraf, S. A., Wormuth, D. R., Hartfield, V. F., & Hughey, J. B. (1981). *Testing ESL composition: A practical approach*. Newbury House.
- Kargar Behbahani, H., Razmjoo, S. A., & YarAhmadi, M. (2025). Extensive reading and L2 vocabulary development: Exploring L2 learning experience through the L2 motivational self-system. *Iranian Journal of Applied Language Studies*, 17(2), 135–152.
- Lee, S. Y., & Hsu, Y. Y. (2009). Determining the crucial characteristics of extensive reading programs: The impact of extensive reading on EFL writing. *The International Journal of Foreign Language Teaching*, 5(1), 12–20.
- Lien, H. Y. (2010). EFL college learners' perceptions of self-selected materials for extensive reading. *English Teacher*, 39(1), 194–204.
- Linuwih, E. R. (2021). The effectiveness of extensive reading in improving EFL academic writing. *Journal of English Language Teaching and Linguistics*, 6(1), 167–177.
- Marslen-Wilson, W. D., & Tyler, L. K. (2007). Morphology, language and the brain: The decompositional substrate for language comprehension. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 362(1481), 823–836.
- Minnich, M., Kirkpatrick, A. J., Goodman, J. T., Whittaker, A., Stanton Chapple, H., Schoening, A. M., & Khanna, M. M. (2018). Writing across the curriculum: Reliability testing of a standardized rubric. *Journal of Nursing Education*, 57(6), 366–370.
- Mol, S. E., & Bus, A. G. (2011). To read or not to read: A meta-analysis of print exposure from infancy to early adulthood. *Psychological Bulletin*, 137(2), 267–296. <https://doi.org/10.1037/a0021890>
- Nugrahini, Y., & Rakhmawati, I. (2022). The effect of extensive reading on EFL learners' writing performance. *Jurnal Eduscience*, 9(2), 515–531.
- Permana, P. G. A. M., & Rohmah, Z. (2024). Contemporary translangual English language policy and practice in Indonesia. *Cogent Education*, 11(1), Article 2404059. <https://doi.org/10.1080/2331186X.2024>
- Poorstoti, S., & Asl, H. D. (2016). Effect of extensive reading on grammatical accuracy and fluency of EFL learners' writing performance. *Journal of Applied Linguistics and Language Research*, 3(2), 188–201.
- Renandya, W. A., & Jacobs, G. M. (2002). Extensive reading: Why aren't we all doing it? In J. C. Richards & W. A. Renandya (Eds.), *Methodology in language teaching: An anthology of current practice* (pp. 295–302). Cambridge University Press.
- Renandya, W. A., Hidayati, M., & Ivone, F. M. (2021). Extensive reading: Top ten implementation issues. *JACET Journal*, 65, 11–21.
- Ria, T. N. (2016). The use of skimming to teach extensive reading. *Jurnal Egaliter*, 1(1).
- Saddler, B., & Andrade, H. (2004). The writing rubric. *Educational Leadership*, 62(2), 48–52.
- Sakurai, N. (2017). The relationship between the amount of extensive reading and the writing performance. *The Reading Matrix: An International Online Journal*, 17(2), 142–164.
- Suk, N. (2017). The effects of extensive reading on reading comprehension, reading rate, and vocabulary acquisition. *Reading Research Quarterly*, 52(1), 73–89. <https://doi.org/10.1002/rrq.152>
- Umar, S., Seken, K., & Artini, L. P. (2013). The effect of extensive reading and learning style on academic writing competency of semester five students of STKIP Hamzanwadi Selong. *Jurnal Pendidikan Bahasa Inggris Indonesia*, 1.

- Wijaya, K. (2021). The important role of extensive reading strategy in Indonesian EFL learning contexts. *Acuity: Journal of English Language Pedagogy, Literature and Culture*, 6(2), 107–117.
- Zwitserslood, P. (2018). Processing and representation of morphological complexity in native language comprehension and production. In *The construction of words: Advances in construction morphology* (pp. 583–602).