

18. Cumulative Guided Passages as an Alternative to Extensive Reading: An Empirical Study in an EFL Context¹

Bakil Ali ALWALSS ²

APA: Alwalss, B. A. (2026). Cumulative Guided Passages as an Alternative to Extensive Reading: An Empirical Study in an EFL Context. *RumeliDE Dil ve Edebiyat Araştırmaları Dergisi*, (53), 319-336.
DOI: <https://doi.org/10.5281/zenodo.22243364>

Abstract

This empirical study investigated whether cumulative guided passages (CGPs) can enhance reading comprehension, vocabulary, and grammar development in the absence of formal extensive reading programs. CGPs represent a structured, text-based instructional resource that systematically recycles key vocabulary and grammatical structures across a sequence of related passages, offering a justified alternative to commercial Graded Readers in resource-constrained contexts. The study involved two groups: an experimental group (n = 15) provided with weekly guided passages, and a control group (n = 17) receiving no instruction. Four assessments were administered pre- and post-experiment to evaluate progress in reading and language components (grammar and vocabulary). Although the experimental group showed nearly double the gain of the control group in both reading and language measures, independent-samples t-tests revealed no statistically significant difference: reading ($M = 18.80$, $SD = 2.14$) compared to control ($M = 17.76$, $SD = 2.41$), $t(30) = 1.28$, $p = .106$. Therefore, the null hypothesis could not be rejected. Nevertheless, CGPs provided additional language input and exposure, suggesting pedagogical value despite non-significant statistical outcomes—particularly for institutions where Graded Readers are logistically or financially difficult to attain. The paper recommends extended implementation periods (minimum ten weeks) with increased passage frequency to allow declarative knowledge to become proceduralized and automated. Further research is recommended to investigate the influence of guided reading-for-writing (RFW) approaches on specific writing skills.

Keywords: Cumulative guided passages, extensive reading, intensive reading, language materials

¹ **Statement (Thesis / Paper):** It is declared that scientific and ethical principles were followed during the preparation process of this study and all the studies utilised are indicated in the bibliography.

Conflict of Interest: No conflict of interest is declared.

Funding: No external funding was used to support this research.

Copyright & Licence: The authors own the copyright of their work published in the journal and their work is published under the CC BY-NC 4.0 licence.

Ethics Statement: Ethical permission was granted by Jazan University University Ethics Commission with the decision dated 9th November and numbered REC-47/05/1662.

Source: It is declared that scientific and ethical principles were followed during the preparation of this study and all the studies used are stated in the bibliography.

Similarity Report: Received – Turnitin /Ithenticatge, Rate: 0 %

Ethics Complaint: editor@rumelide.com

Article Type: Research article, Article Registration Date: 07.05.2026-**Acceptance Date:** 02.09.2026-**Publication Date:** 03.09.2026; **DOI:** <https://doi.org/10.5281/zenodo.22243364>

Peer Review: Two External Referees / Double Blind

² Yüksek Lisans Öğrencisi, College of Arts & Languages, Seiyun University / MA Student, Seiyun Üniversitesi, Sanat ve Dil Fakültesi (Seiyun, Yemen), **e-mail:** balwalss@jazanu.edu.sa **ORCID ID:** <https://orcid.org/0009-0009-1499-4046>
ROR ID: <https://ror.org/01xp8dp03>

Geniş Kapsamlı Okumaya Alternatif Olarak Birikimli Rehberli Metinler: Yabancı Dil Olarak İngilizce Öğretimi Bağlamında Ampirik Bir Çalışma³

Öz

Bu ampirik çalışma, biçimlendirilmiş kapsamlı okuma programlarının bulunmadığı `durumlarda, kümülatif rehberli metinlerin (KRM) okuduğunu anlama, kelime bilgisi ve dil bilgisi gelişimini destekleyip destekleyemeyeceğini araştırmıştır. KRM'ler, ilgili metinler dizisi boyunca temel kelime ve dil bilgisi yapılarını sistematik olarak yeniden dolaşıma sokan, yapılandırılmış, metin tabanlı bir öğretim kaynağını temsil etmekte ve kaynak kısıtlı bağlamlarda ticari Dereceli Okuyuculara (Graded Readers) haklı bir alternatif sunmaktadır. Çalışma iki grubu içermektedir: haftalık rehberli metinlerle desteklenen deney grubu (n = 15) ve herhangi bir öğretim almayan kontrol grubu (n = 17). Okuma ve dil bileşenlerindeki (dil bilgisi ve kelime bilgisi) ilerlemeyi değerlendirmek için deney öncesinde ve sonrasında dört değerlendirme yapılmıştır. Deney grubu hem okuma hem de dil ölçümlerinde kontrol grubunun neredeyse iki katı kazanç göstermesine rağmen, bağımsız örneklem t-testleri istatistiksel olarak anlamlı bir fark ortaya koymamıştır: okuma (Ort. = 18.80, SS = 2.14) kontrol grubu (Ort. = 17.76, SS = 2.41) ile karşılaştırıldığında, $t(30) = 1.28, p = .106$. Bu nedenle, sıfır hipotezi reddedilememiştir. Bununla birlikte, KRM'ler ek dil girdisi ve maruziyeti sağlayarak, istatistiksel olarak anlamlı olmayan sonuçlara rağmen pedagojik değer önermektedir—özellikle Dereceli Okuyuculara lojistik veya mali açıdan ulaşmanın zor olduğu kurumlar için. Makale, bildirimsel bilginin süreçsel hale gelmesi ve otomatikleşmesi için daha uzun uygulama süreleri (en az on hafta) ve artırılmış metin sıklığı önermektedir. Rehberli okuma-yazma (reading-for-writing) yaklaşımlarının belirli yazma becerileri üzerindeki etkisini araştırmak için daha fazla araştırma yapılması tavsiye edilmektedir.

Anahtar kelimeler: Kümülatif rehberli metinler, kapsamlı okuma, yoğun okuma, dil materyalleri

³ **Beyan (Tez/ Bildiri):** Bu çalışmanın hazırlanma sürecinde bilimsel ve etik ilkelere uyulduğu ve yararlanılan tüm çalışmaların kaynakçada belirtildiği beyan olunur.
Çıkar Çatışması: Çıkar çatışması beyan edilmemiştir.
Finansman: Bu araştırmayı desteklemek için dış fon kullanılmamıştır.
Telif Hakkı & Lisans: Yazarlar dergide yayınlanan çalışmalarının telif hakkına sahiptirler ve çalışmalarını CC BY-NC 4.0 lisansı altında yayımlanmaktadır.
Etik İzni: Jazan Üniversitesi Etik Komisyonu tarafından 9.11.2025 tarihli, REC-47/05/1662 sayılı kararla etik izni verilmiştir.
Kaynak: Bu çalışmanın hazırlanma sürecinde bilimsel ve etik ilkelere uyulduğu ve yararlanılan tüm çalışmaların kaynakçada belirtildiği beyan olunur.
Benzerlik Raporu: Alındı – Turnitin, Oran: % 0
Etik Şikayeti: editor@rumelide.com
Makale Türü: Araştırma makalesi, **Makale Kayıt Tarihi:** 07.05.2026-**Kabul Tarihi:** 02.09.2026-**Yayın Tarihi:** 03.09.2026; **DOI:** <https://doi.org/10.5281/zenodo.22243364>
Hakem Değerlendirmesi: İki Dış Hakem / Çift Taraflı Körleme

Introduction

Reading is an active cognitive process of deriving meaning from written text, requiring the decoding of visual symbols into meaningful language while applying background knowledge to comprehend, analyse, and anticipate ideas. Reading constitutes a "complex psycholinguistic construct" (Koda, 2012, p. 308) involving three core operations: "(a) decoding (extracting linguistic information directly from print), (b) text-information building (integrating the extracted information into phrases, sentences, and paragraphs), and (c) reader-model construction (synthesizing the amalgamated text-information with prior knowledge)" (Koda, 2012, p. 308).

Building on foundational models, Duke and Cartwright (2021) argue that reading comprehension must extend beyond decoding and linguistic comprehension to include inference generation, comprehension monitoring, working memory, and attention control. Reading represents an active, multifaceted cognitive endeavour through which readers construct meaning via dynamic interaction among prior knowledge, textual information, and situational context.

A holistic perspective unfolds as follows: readers contribute distinct skills—word recognition and lexical knowledge—enabling cognitive focus on meaning (Willingham, 2006). Schema theory activates background knowledge encompassing world knowledge and text-type awareness (Carrell & Eisterhold, 1983). Readers' objectives influence their deployment of metacognitive strategies, affecting processing depth and reading speed (Kintsch & van Dijk, 1978). This interactive process yields a coherent mental model integrating textual information with reader knowledge (McNamara & Magliano, 2009). Proficient readers can summarize, explain, apply, and critique textual meaning. Language programs introduce reading courses to cultivate proficient readers and enhance grammar and vocabulary. This study investigates whether cumulative guided passages (CGPs) can improve reading skills and accelerate language development.

Three common terminologies

To contextualize CGPs within existing reading pedagogy, three established approaches require distinction. First, there is intensive reading which constitutes an analytical, detail-focused approach wherein learners engage with brief, level-appropriate texts to achieve deep comprehension and master language features (Nuttall, 1996). This approach is often contrasted with extensive reading (ER). ER represents a holistic, fluency-oriented method wherein learners engage with substantial quantities of linguistically accessible material, focusing on overall meaning rather than analyzing individual elements (Day, 2013; Hill, 2013). Finally, there is the uncommon term of extended reading, which refers to sustained engagement with longer, cohesive texts over time, building stamina and offering immersion in sustained context (Brown & Lee, 2015; Paran, 2003).

A fourth term

This study proposes cumulative guided passages (CGPs)—a reading format leaning towards intensive reading while offering greater intensity and topic variety. Cumulative guided passages (CGPs) are defined as a teacher-selected sequence of short, graded texts (500–1500 words) that systematically recycle key vocabulary and grammatical structures across a series of related passages. CGPs are designed to function as a scaffolded reading intervention that provides meaningful language input and reading practice in contexts where formal extensive reading programs or commercial Graded Readers are

unavailable. They serve primarily as a substitute for extensive reading, while also incorporating principles of intensive reading through embedded comprehension tasks and vocabulary recycling. CGPs function where ER programs are absent, serving as a structured alternative to commercial Graded Readers. The fundamental principle involves controlled, systematic language recycling: key vocabulary, grammatical structures, and concepts are progressively reintroduced throughout the sequence, facilitating natural acquisition without external levelling.

CGPs differ from Graded Readers in five respects: (1) incremental progression—difficulty increases weekly; (2) internal cohesion—unified themes enhance engagement; (3) embedded scaffolding—glossaries and comprehension checks promote autonomy; (4) pedagogical precision—specific lexical objectives provide teachers greater input control; and (5) ensured availability—institutions can compile passages where Graded Readers are unavailable.

Rationale and Significance

Despite theoretical support for ER, L2 teachers often underutilize it, as intensive reading aligns more directly with classroom instruction. Jeon and Day (2016) identified three factors undermining ER effectiveness: lack of curricular integration, inadequate time allocation, and absence of post-reading activities. Additionally, teachers' traditional roles and syllabus coverage limit time for ER (Renandya & Jacobs, 2002). Zarei (2017) demonstrated that task-based extensive reading significantly improved EFL learners' writing accuracy, supporting the rationale for CGPs.

Fields (2017) proposed a teacher-directed variation of extended reading wherein students independently read authentic materials outside class and complete weekly comprehension quizzes. This paper addresses the absence of Graded Readers as the cornerstone of ER. In resource-constrained contexts, logistical demands—multiple copies, varied titles, storage—prove prohibitive for large institutions (Nation & Waring, 2019).

In academic contexts, learners require independent reading for comprehension and learning (Koda & Yamashita, 2019). Given time constraints in academic contexts, such as universities, ER cannot meet varied text-type requirements for large student populations. This experiment introduces cumulative guided reading through two weekly Graded Passages (500–1500 words each), contrasting with the "book per week" model (Nation & Wang, 1999). The objective is to investigate feasibility of producing hybrid Graded Reader booklets compiled from short passages selected, and adapted, from diverse genres, providing language input that supports development, reading ability, and vocabulary acquisition (Lai, 1993). Grabe and Stoller (2009, 446) identify reading program objectives including vocabulary building, reading fluency, and effective strategy use.

Reading and the EFL Classroom

Intensive reading instruction comprises three stages: pre-reading, while-reading, and post-reading. Our cumulative guided reading experiment functions as a fourth stage—extending classroom learning through independent practice of previously taught strategies. Passages (500–1500 words) are graded linguistically appropriate or slightly above learners' levels, prioritizing high-frequency vocabulary as "an important and justifiable criterion" (Williams, 2006, p. 359).

Grabe (2002) identifies dilemmas relevant to this experiment: reading fluency, vocabulary development, engagement in ER, and appropriate strategy use. Graded passages address these

requirements. Multiple variables influence comprehension: purpose, text type, word density, and background knowledge (Cook, 2016). Effective comprehension requires motivated readers employing flexible strategies aligned with reading goals (Grabe & Stoller, 2019).

Language programs teach learners to skim, scan, comprehend detail, and make inferences (Grellet, 1981). Reading strategies—titles, pictures, text organization, contextual clues—render reading an active process (Carrell et al., 1988). Conscious strategies become automated through practice, becoming subconscious skills. Williams (2006, p. 363) distinguishes skill as "an acquired ability, which has been automatised, and operates subconsciously" from strategy as "a conscious procedure carried out in order to solve a perceived problem."

Extensive Reading (ER)

Grabe (2006) identifies ER as an implication requiring promotion by language teachers. Text factors, reading processes, and reader factors warrant consideration in reading programs (Anderson, 1999; Grabe & Stoller, 2014; Grellet, 1981). ER research consistently demonstrates vocabulary as the primary beneficiary. Suk (2017, p. 84) found that "the most noticeable gain from extensive reading was in vocabulary acquisition." Hafiz and Tudor (1989) demonstrated ER's positive influence on writing and reading skills. Mason and Krashen (1997) confirmed ER accelerates L2 proficiency. Day and Bamford (2002) advocate integrating ER into reading programs.

This study views ER as extended self-disciplined language practice—not merely for improving reading skills but for gaining knowledge and building confidence. Reading sub-skills taught in intensive classes become automatic through additional guided passage practice (cf. DeKeyser & Suzuki, 2025). Graded Readers provide required language practice reflecting ER characteristics (Hill, 2013). Grabe and Yamashita (2022) offer five explanations for ER's limited adoption: curricula prioritize discrete skills; ER demands considerable resources; emphasis on comprehension over speed; teachers may lack qualifications; and administrators prefer direct teaching.

The literature review culminates in two definitions highlighting the contrast between teacher-guided analysis (intensive) and independent fluency-building (extensive). CGPs occupy a middle ground, offering structured guidance within an extended reading framework. ER constitutes "independent reading of relatively long self-selected texts with minimal teacher intervention" (Williams, 2006, p. 361), involving "rapid reading of large quantities of material for general understanding" (Carrell & Carson, 1997, p. 50). Intensive reading involves purposeful interaction with short, challenging texts for profound comprehension (Nuttall, 1996).

Research Hypotheses

The proposal I present in the following paragraphs is quite different from the description of extensive reading defined by Carrell and Carson (1997) and Williams (2006). The cumulative guided practice involves the following features:

- Limited in text lengths (no more than two passages per week).
- Allocated time per week is defined, i.e. two or three hours.
- Materials are selected by the teacher not the student, normally slightly higher than their level.

- The selected level is generalized for the group of students (though their language proficiency levels may differ). Cumulative guided passages are designed for pre-intermediate and intermediate learners. Students who fall below these levels will find it challenging to grasp vocabulary and grammar, while those at higher levels can improve their language skills through self-directed study, drawing on their past experiences in language learning.
- The reading purpose is open to the student (for comprehension, word recognition, speed reading, vocabulary development, and so on).
- Reading strategies are provoked indirectly through the different comprehension questions following selected passages.
- The teacher plays the role of a language ‘supervisor’, and language ‘consultant’.
- Evaluation of the learners’ improvement is left to the teacher to do it through reading home assignments.

In essence, CGPs represent a teacher-facilitated, systematically scaffolded approach to extended reading—distinct from both purely intensive instruction and fully independent extensive reading. The assumption in this study is that the experiment group will score in the post-tests (reading and language) better than the control group (who receive no extra ‘supplementary’ training in reading). This statement is based on the belief that the students would get motivated for reading extra material because they would feel that they are:

- consolidating and improving their language,
- activating their passive vocabulary,
- adding more new words,
- forced to read for comprehension, and
- practising their reading skills and strategies.

The Null Hypothesis (H_0) states that extended reading (represented as CGPs) has no effect on the language development as it resembles in nature extra material given to language learners, or functions as a self-study activity for devoted learners. On the other hand, *the Alternative Hypothesis (H_a)* stipulates that extended reading (represented as CGPs) enables the learners to build a better reading strategy, expand their vocabulary, consolidate grammar, and eventually facilitates the quick development of language in terms of reading, grammar, and vocabulary.

Research Design

Two groups were involved in the empirical study: the experimental group and the control group. The two groups represent the independent variables, whereas their scores of reading tests, and language tests represent the dependent variables. Therefore, statistical analysis will focus on the latter to test the hypotheses stated above.

The present study should be regarded as an exploratory pilot investigation. Its primary aim was to examine whether cumulative guided passages show promise as a supplementary reading resource. The design, which compared an experimental group receiving weekly guided passages with a control group receiving no additional instruction, does not allow for isolating the specific effect of the cumulative design versus extra reading time. Future research with an active control group receiving another type of supplementary reading practice would provide a stronger test of the specific value of CGPs.

Material

In choosing the passages, we considered specific characteristics (including several of those mentioned in Gascoigne (2008, p. 78). These features included interest, familiarity, academic-oriented passages, linguistic difficulty ranges between the A2+ to B1 CEFR scale (Council of Europe, 2018), authentic texts to preserve text structure. Four passages were chosen from different learning published materials intended for the PET practices (Cambridge Preliminary English Test, currently known as Cambridge Preliminary B1), and four longer passages intended for FCE (currently known as Cambridge B2 First Certificate) at times without comprehension questions. Four other passages, ranging between 500 to 1500, were adopted from online blogs, such as food, travel, health, business, and news. These passages were modified to suit the intended levels of A2 and B1, which renders them semi-authentic.

Participants

The participants are English undergraduates in their third semester of studying English. They have spent their first two semesters to grasp the basics of English grammar, high-frequency vocabulary, rudimentary language skills in reading, writing, speaking, and listening. Their time is not fully devoted to English since they must pass other university requirements. All are males, ranging in age between 19 to 21 years old. The subjects were distributed as follows: 15 for the experimental group, and 17 for the control group. These two groups were originally two separate sections in the academic system, who enrolled in the same academic year. Originally, they were 21 and 22 respectively. Unfortunately, some participants withdrew from the training during the first three weeks.

Analysis

MS-Excel was used to perform all the reported analyses and their results (see Abbott, 2011). Given that we have two groups undergoing two distinct types of assessments (Reading & Language) at two separate times (pre- and post-assessment), this resulted in a total of eight testing occurrences, which were eventually consolidated into four sets of testing (Table 1).

Table 1 Test Distribution among the Two Groups

	Component	Points	Number of Subjects	
			Experimental	Control
Pre-Assessment	Language 1	75	15	17
	Reading 1	25	15	17
Post-Assessment	Language 2	75	15	17
	Reading 2	25	15	17
Total	Language 1 & Reading 1	100	15	17

Language 2 & Reading 2

100

15

17

To compare the means of the post-test results and to investigate the effect of cumulative guided passages on the improvement of reading comprehension, along with the advancement of grammar and vocabulary, the t-test is considered the appropriate analytical method in this scenario. The primary goal of the t-test is to evaluate the means of these two independent groups in order to examine the presence of a statistically significant difference (of influence) between them. (Wilcox, 2021). Nevertheless, a pertinent question that emerged was which type of t-test should be employed, given that there are two prevalent forms: the equal and unequal sample size t-tests. The former is referred to as the normal student's (pooled) t-test, while the latter is known as Welch's t-test, which is used to account for the variances in the two samples.

Prior to choosing and executing the necessary t-test, it is essential to evaluate the data for normality and the homogeneity of variances. To address these assumptions and determine the appropriate type of t-test, we performed the Shapiro-Wilk Test, which is considered one of the most effective tests for normality (Shapiro & Wilk, 1965). It is especially useful for identifying non-normality in smaller sample sizes, which is applicable in our experiment. The results are presented in Table 2.

Table 2 Shapiro-Wilk Test Results

	<i>Experimental Group</i>			<i>Control Group</i>		
	<i>Language Test 2</i>	<i>Reading Test 2</i>	<i>Total</i>	<i>Language Test 2</i>	<i>Reading Test 2</i>	<i>Total</i>
Shapiro-Wilk statistic W	0.981	0.899	0.971	0.960	0.968	0.949
Calculated Shapiro-Wilk p-value:	0.978	0.093	0.876	0.629	0.773	0.442
Critical value of W (5% significance level)	0.882	0.882	0.882	0.892	0.892	0.892

Consequently, the student t-test is considered the appropriate method for testing our hypotheses. The decision to deploy the student t-test for paired samples was based on the values of the Shapiro-Wilk statistic W, which were higher than the critical value of W at 5% significance level in all six instances. For more clarification, the Shapiro-Wilk test revealed for the totals of both groups that the data was normally distributed: $W(14) = 0.97$, $p = .88$, and $W(16) = .95$, $p = .44$, for the totals of the experimental and the control groups respectively.

Results

The five-week duration, while relatively brief, is not inherently problematic; indeed, detectable gains within such a timeframe would strengthen the case for CGP effectiveness. The results indicated a modest enhancement in both language acquisition and reading scores for both groups. An examination of the descriptive statistics reveals that both groups gained from the language training, as evidenced by the increased difference between the pre-test and post-test results, particularly in the averages, minimum, and maximum values. Equally, the sample variance and range decreased, suggesting a collective impact. It is commonly recognized that a higher variance signifies that data points are more dispersed, whereas a lower variance suggests that they are concentrated around the mean. These variances offer us deeper insights into the distribution of data within these samples. Table 3 provides a summary of the overall

scores.

The treatment group's total score increased by nearly 10 points (from 59.67 to 69.20), while the control group's gain was approximately 5 points (from 62.18 to 67.24). The treatment group post-language test ($M = 50.40$, $SD = 7.58$) had a higher average score than the pre-language test ($M = 43.93$, $SD = 8.75$). This improvement is also evident in the Reading Test. Moreover, the Reading Test indicated the same improvement ($M = 15.73$, $SD = 1.75$ and $M = 18.80$, $SD = 2.14$). The change of the experimental group's means was increased by slightly over three points. On the other hand, the control group post-language test ($M = 49.47$, $SD = 7.76$) had a slightly higher average score than the pre-language test ($M = 46.06$, $SD = 8.41$). Though the control group exhibited a similar progress to that of the experimental group, increase in their average scores in the reading (pre- and post-tests) was almost 50% less than that of the experimental group ($M = 16.12$, $SD = 2.45$ and $M = 17.76$, $SD = 2.41$). Interestingly, the range of both groups remained intact, which indicates more training is required to get a substantial gain.

Beyond statistical significance, the practical significance of the observed gains warrants attention. The experimental group's nearly double improvement compared to controls (10 points versus 5 points on the combined 100-point measure) represents a substantively meaningful difference in raw score terms. In many educational contexts, such gains over a five-week intervention would be considered educationally valuable, even if the small sample size precluded statistical significance. This pattern aligns with emerging calls in applied linguistics to complement null hypothesis significance testing with effect size reporting and consideration of practical significance (Plonsky & Oswald, 2014; Larson-Hall & Plonsky, 2015).

Table 3 Descriptive Statistics of Language and Reading

	Mean	STDEV	Sample Variance	Range	Minimum	Maximum
<i>Experimental Group (Count 15)</i>						
Language Test 1	43.93	8.75	76.64	32	27	59
Language Test 2	50.40	7.58	57.40	28	36	64
Reading Test 1	15.73	1.75	3.07	6	13	19
Reading Test 2	18.80	2.14	4.60	6	16	22
Test 1 Totals	59.67	10.17	103.38	35	42	77
Test 2 Totals	69.20	9.44	89.03	33	53	86
<i>Control Group (Count 17)</i>						
Language Test 1	46.06	8.41	70.81	29	30	59
Language Test 2	49.47	7.76	60.26	26	36	62
Reading Test 1	16.12	2.45	5.99	9	11	20
Reading Test 2	17.76	2.41	5.82	9	13	22
Test 1 Totals	62.18	10.69	114.28	37	42	79
Test 2 Totals	67.24	9.94	98.82	33	51	84

Gain-Score Analysis. To examine whether the experimental group improved more than the control group over time, gain scores (post-test minus pre-test) were calculated for both reading and language

measures. For reading, the experimental group showed a mean gain of 3.07 ($SD = 0.80$), while the control group showed a mean gain of 1.65 ($SD = 1.22$). An independent-samples t-test on the reading gain scores revealed a difference approaching statistical significance, $t(30) = 1.76$, $p = .088$ (two-tailed), with a mean difference of 1.42 (95% CI [-0.23, 3.07]). For language, the experimental group showed a mean gain of 6.47 ($SD = 2.03$), while the control group showed a mean gain of 3.41 ($SD = 3.08$), $t(30) = 1.10$, $p = .280$ (two-tailed), with a mean difference of 3.06 (95% CI [-2.58, 8.70]). These results suggest a directional trend favouring the experimental group in reading, though the differences did not reach conventional levels of statistical significance. All p-values are reported as two-tailed.

The findings indicated a modest enhancement in both language attainment and reading performance for both groups. Nevertheless, though the control group performed better than the experimental group in both pre-test scores of the language test and the reading test (46.06 and 16.12), they did not retain their ranking in the post-test scores. They were outperformed in the post-test results by 50.40 and 18.80. But, do these results carry a statistically significant difference? This is the focus of this paper to explore if such variations have a true significance. In other words, our hypotheses will be tested against these findings.

Overall, the improvement is evident. However, whether it is more effective or less successful compared to similar language programs and training is a separate matter. The pertinent question is: are these differences statistically significant and genuinely meaningful? To address this, we must conduct a t-test, but we still need to ascertain which type of t-test to apply: equal or unequal (Wilcox, 2021). Although these Q-Q plots suggest normality, leading to the use of an 'equal' t-test, it is crucial to further validate this, as equal and unequal t-tests are both forms of the independent two-sample t-test, differing in their assumptions regarding the variances (spread) of the two groups being compared. The equal variance t-test (pooled) assumes that the variances are similar and utilizes a combined estimate, whereas the unequal variance t-test (Welch's t-test) does not assume equal variances and adjusts for degrees of freedom, providing robustness when variances or sample sizes vary significantly (Wilcox, 2021).

A two-sample F-test for variances was performed to assess the variability in test scores between the experimental group and the control group based on the results of the post-tests: language, reading, and the totals for each set (i.e., language and reading) (refer to Table 4). For example, the totals from both tests indicated homogeneity of variances. Levene's test confirmed that the assumption of equal variances was met for the experimental group ($M = 69.20$, $SD = 9.44$) and the control group ($M = 67.24$, $SD = 9.94$), $F(14, 16) = 0.90$, $p = .41$.

The null hypothesis of the F-test posits that the variances are equal when the F-value exceeds the F-critical value at $p = 0.05$. Given that the p-value from the F-statistic is greater than the significance level ($\alpha = 0.05$) in all three cases (Table 4), it was concluded that the variances were not significantly different, thus a standard independent-samples t-test was employed as the suitable method to evaluate our hypotheses. Results (Table 5) show that the t-statistical value is higher than the significance level of alpha at 0.05, which clearly indicate that the null hypothesis is retained.

Table 4 F-Test of Two-Sample Variations

F-Test Two-Sample for Variances			
	<i>Language Test 2</i>	<i>Reading Test 2</i>	<i>Totals of Both Tests</i>
df	14, 16	14, 16	14, 16
F Value	0.953	0.791	0.901
P(F<=f) one-tail	0.468	0.333	0.426
F Critical One-Tail	0.409	0.409	0.409

Table 5 T-Test of Equal Variations

t-Test: Two-Sample Assuming Equal Variances			
	<i>Language Test 2</i>	<i>Reading Test 2</i>	<i>Totals of Both Tests</i>
df	30	30	30
t Statistical Value	0.342	1.276	0.571
P(T<=t) One-Tail	0.368	0.106	0.286
t Critical One-Tail	1.697	1.697	1.697
P(T<=t) two-tail	0.735	0.212	0.572
t Critical two-tail	2.042	2.042	2.042

Effect Sizes

In addition to null hypothesis significance testing, effect sizes (Cohen's *d*) were calculated to evaluate the magnitude of the observed differences between groups. For reading scores, the experimental group ($M = 18.80$, $SD = 2.14$) outperformed the control group ($M = 17.76$, $SD = 2.41$) with a small-to-moderate effect size ($d = 0.46$, 95% CI [-0.25, 1.16]). For language scores, the difference between the experimental group ($M = 50.40$, $SD = 7.58$) and the control group ($M = 49.47$, $SD = 7.76$) yielded a negligible effect size ($d = 0.12$, 95% CI [-0.58, 0.82]). For the combined total scores (language and reading together), the effect size was small ($d = 0.20$). These effect sizes suggest that the cumulative guided passages had a modest practical impact on reading comprehension specifically, though the study was underpowered to detect this effect as statistically significant. The negligible effect on language scores indicates that the five-week intervention period may have been insufficient to produce measurable gains in grammar and vocabulary. In sum, despite visible gains favouring the experimental group, the statistical tests confirm that these differences do not exceed what might occur by chance. Thus, the null hypothesis is retained.

Discussion

The null hypothesis—that cumulative reading passages would yield no meaningful effect beyond supplementary material—could not be rejected. The experimental group's modestly higher scores ($M = 18.80$ vs. 17.76) did not reach statistical significance, $t(30) = 1.28$, $p = .212$ (two-tailed). The short training period is a plausible explanation for the non-significant finding. This interpretation is supported by the consistent pattern of gains favouring the experimental group across both measures (Figures 2-4). With additional time—perhaps 10 weeks or more—and increased passage frequency, these trends might strengthen. Skill Acquisition Theory (DeKeyser & Suzuki, 2025) provides a useful lens: five weeks may be insufficient for declarative knowledge (strategies taught in class) to become proceduralized through independent practice. CGPs may require longer implementation to move learners along the skill-acquisition continuum.

The present findings align with recent meta-analytic evidence on the effectiveness of structured reading interventions. Taylor et al. (2022), in a comprehensive synthesis of 52 studies on reading fluency interventions, found that interventions lasting fewer than eight weeks yielded effect sizes substantially smaller than those implemented over longer periods. This pattern directly supports the interpretation that the five-week duration in the current study, while sufficient to show directional gains, may have been insufficient for achieving statistical significance. Similarly, Webb and Chang (2022), in their longitudinal study of vocabulary acquisition through graded reading, demonstrated that meaningful vocabulary gains required a minimum of 10-12 weeks of consistent exposure, reinforcing the recommendation for extended implementation periods.

While the results did not reach conventional levels of statistical significance, the pedagogical implications warrant consideration. In applied linguistics research, particularly in classroom-based studies with practical constraints, non-significant findings do not necessarily indicate absence of educational value (Larsen-Freeman, 2022). The experimental group's consistent outperformance across multiple measures—nearly double the gain of controls in reading scores—suggests that CGPs provided measurable benefits that may accumulate over longer implementation periods. This interpretation is consistent with calls within the field to move beyond binary significance testing and toward more nuanced evaluation of intervention effects (Norris, 2015).

Limitations

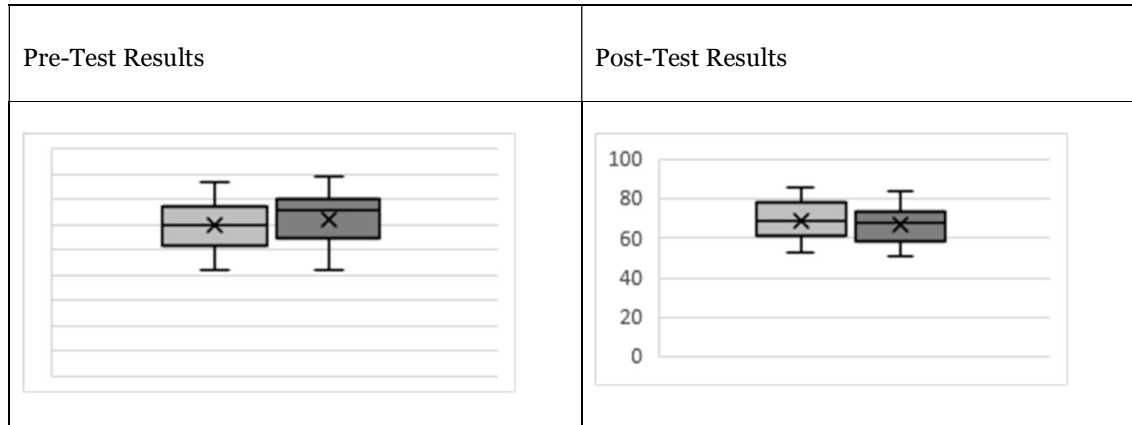
Several limitations of this study should be acknowledged. First, no a priori power analysis was conducted to determine the appropriate sample size. The sample ($n = 15$ for the experimental group, $n = 17$ for the control group) was determined by the availability of participants within two intact classroom sections rather than through statistical power calculations. Consequently, the non-significant findings may be attributable in part to insufficient statistical power, and future replication studies should employ larger sample sizes informed by power analysis (e.g., minimum 30–40 participants per group) to detect small-to-moderate effects. Second, the study duration was limited to five weeks. As Skill Acquisition Theory suggests, this period may be insufficient for declarative knowledge (reading strategies taught in class) to become proceduralized and automated through independent practice. Future research should extend the implementation period to a minimum of ten weeks with increased passage frequency to allow for more robust skill development. Third, the sample was exclusively male (aged 19–21), which limits the generalisability of the findings. While there is no theoretical reason to expect fundamentally different reading strategy development processes in female learners, the extent to which these findings apply to female populations remains an empirical question for future investigation. Fourth, effect sizes, though reported, yielded confidence intervals that crossed zero, indicating imprecision in the estimates—a reflection of the small sample size. Future studies with larger samples would produce more precise effect size estimates. Despite these limitations, the consistent pattern of gains favouring the experimental group across multiple measures suggests that cumulative guided passages hold pedagogical promise, particularly in resource-constrained contexts where extensive reading programs are not feasible.

The value of extended reading is well documented: it supports comprehension (Pichette, 2005), increases reading rate (Tanaka & Stapleton, 2007), and builds vocabulary (Horst, 2009). Jiang, Grabe, & Carrell, (2020), accordingly, recommend extensive reading as an essential curriculum component. For contexts where full ER programs are impractical, CGPs offer a viable alternative. Moreover, it is equally important to motivate learners to practise reading extra material outside the class to appreciate various genre structures, which would assist in developing writing skills (Grabe & Kaplan, 1996; Hirvela,

2017). Recent research by Zhang and Zhang (2023) on reading-writing connections in EFL contexts further underscores the potential of guided reading interventions to facilitate transfer effects across modalities.

Guidance through CGPs may foster language awareness, helping learners recognize their strengths and preferences—an outcome not captured by our post-test measures but potentially valuable for long-term development. This may lead the learner to apply and connect any acquired piece of declarative knowledge to different practical situations pertaining to other language skills. Though cumulative guided passages did not show a substantial statistical impact, they still made an observable change in the experimental group as compared to the control group. More guidance with additional practice may lead to greater gains, provided that these passages are well planned and organised. As Ellis (2023) notes in his recent synthesis of instructed second language acquisition, structured input interventions that may not yield immediate statistically significant effects often produce cumulative benefits that emerge over extended timeframes—a finding directly relevant to the current study.

Figure 2 Totals of the Two Groups' Performance (out of 100)



Note: The experimental group marked in grey colours and situated to the left, and the control group is in black colours to the right.

Figure 3 Language Performance Before and After the Training (out of 75)

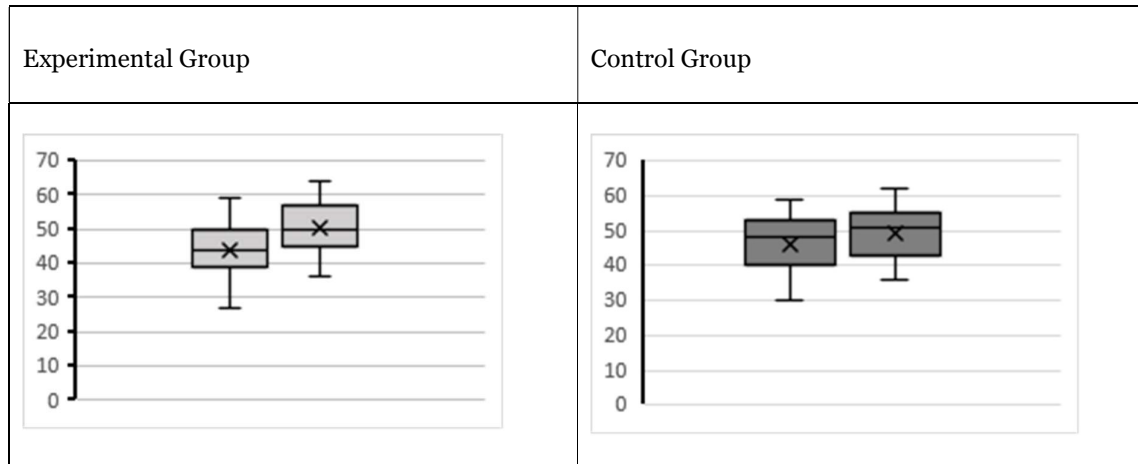
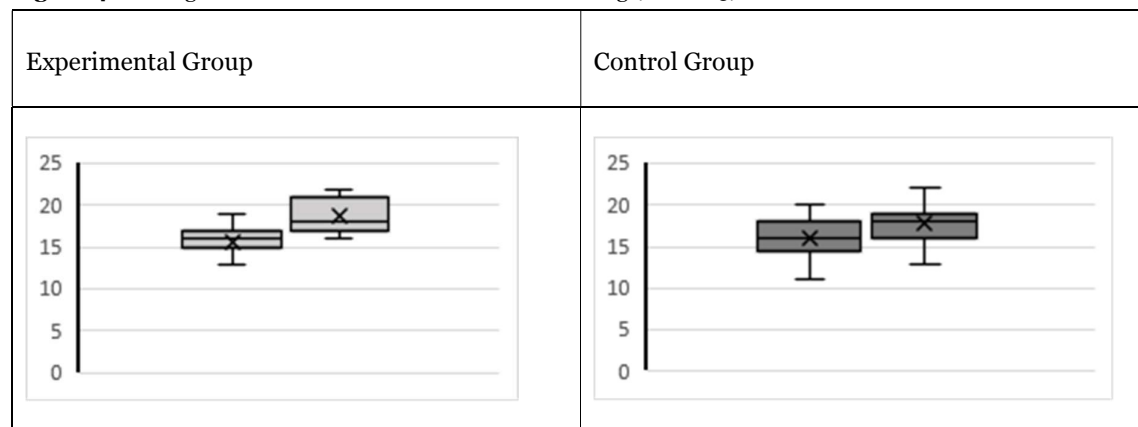


Figure 4 Reading Performance Before and After the Training (out of 25)

Recommendations

One limitation beyond this study's scope is the possible influence of students' L1 reading practices. Future research might examine how Arabic reading strategies (e.g., text type preferences, reading purposes, frequency) transfer to L2 reading (cf. Grabe & Stoller, 2019; Koda & Zehler, 2008). Finally, it is recommended that future studies investigate the reading-for-writing (RFW) approach by emphasizing text types that students will later need to produce (Hirvela, 2016). Such research could illuminate the transfer effects of guided reading on writing development—a connection well established in the literature (Grabe & Kaplan, 1996; Nation & Macalister, 2021).

Based on the pattern of gains observed and the theoretical framework of skill acquisition, the following specific recommendations for implementation are proposed. First, a minimum implementation period of ten weeks is recommended to allow sufficient time for declarative knowledge to become proceduralized. Second, the frequency of passages should be increased to three per week, with varying lengths (500-1200 words) to maintain engagement while providing adequate exposure. Third, pre- and post-test intervals should be designed to capture incremental gains at multiple time points (e.g., weeks 5, 10, and 15) to better understand the trajectory of skill development. These recommendations align with recent dosage studies in reading intervention research (Nation, 2022; Webb, 2021).

Several methodological enhancements would strengthen future investigations of CGPs. A larger sample size (minimum 30-40 participants per group) would provide greater statistical power to detect moderate effect sizes. Including a delayed post-test (e.g., 4-6 weeks after intervention completion) would help assess retention of gains and the durability of any observed effects. Additionally, incorporating qualitative measures—such as learner diaries, semi-structured interviews, or think-**Barkhuizen** protocols—would provide richer insights into how learners engage with CGPs and what strategies they employ during independent reading (Barkhuizen, 2022).

Beyond replication studies with longer implementation periods, future research should explore several additional dimensions. First, the effectiveness of CGPs with different proficiency levels (beginner, advanced) merits investigation, as scaffolding needs may vary across proficiency bands. Second, comparative studies examining CGPs against other ER alternatives—such as digital reading platforms, audiobooks, or teacher read-alouds—would help situate CGPs within the broader landscape of reading interventions. Third, research examining the motivational and affective dimensions of CGP

implementation, including learner attitudes, self-efficacy beliefs, and reading enjoyment, would complement the cognitive focus of the current study (Saito et al., 2023).

Conclusion

This study examined whether cumulative guided passages could serve as a practical alternative to extensive reading programs in contexts where graded readers are unavailable. The results, while not statistically significant, point to several conclusions. For instance, a syllabus, in contrast to a curriculum, is generally recognized as the outline of goals and objectives that guide the selection and assessment of content within a foreign language program. Consequently, an extra reading component, such as the cumulative reading passages, may be included in these syllabuses, enabling teachers to choose the reading materials they deem suitable for their class level (see Freddi, 2015). The modest gains suggest that longer and more intensive practice may yield better results. It was recommended that extending training to at least ten weeks with three varied passages weekly, allowing declarative knowledge to become proceduralised and automated.

Despite the non-significant statistical outcome, this study makes several contributions to the field of L2 reading pedagogy. First, it introduces and operationalizes the concept of cumulative guided passages as a structured, resource-sensitive alternative to commercial Graded Readers—a contribution of particular relevance to educational contexts where extensive reading programs are logistically or financially prohibitive. Second, the study provides empirical data on the short-term effects of CGPs, demonstrating directional gains that, while not statistically significant, suggest pedagogical promise. Third, the detailed description of CGP design principles (incremental progression, internal cohesion, embedded scaffolding, pedagogical precision, and ensured availability) offers a practical framework that teachers and curriculum developers can adapt to their local contexts.

Although cumulative guided reading did not show a substantial statistical effect here, it remains a promising alternative to extensive reading programs, particularly for building strategy, fluency, and vocabulary. With more structured implementation and encouragement of supplementary reading, such materials could significantly support overall language skill development. The consistent pattern of gains across measures suggests that with longer implementation and more intensive exposure, CGPs may yet prove pedagogically valuable—particularly in resource-constrained settings where extensive reading programs remain aspirational. As recent work on sustainable language teaching practices has emphasized (Macalister & Nation, 2020), locally adaptable, low-resource interventions like CGPs represent an important direction for the field.

In sum, the study examined whether cumulative guided reading passages enhance language development compared to conventional supplementary materials. While the experimental group achieved slightly higher average scores ($M = 18.80$) than the control group ($M = 17.76$), the difference was not statistically significant. Accordingly, the null hypothesis could not be rejected. However, the consistent pattern of gains across measures suggests that with longer implementation and more intensive exposure, CGPs may yet prove pedagogically valuable—particularly in resource-constrained settings where extensive reading programs remain aspirational.

References / Kaynakça

- Abbott, M. L. (2011). *Understanding educational statistics using Microsoft Excel and SPSS*. John Wiley & Sons.
- Anderson, N. J. (1999). *Exploring second language reading: Issues and strategies*. Heinle & Heinle.
- Barkhuizen, G. (2022). Qualitative research in language teaching and learning. In J. I. Lontas (Ed.), *The TESOL encyclopedia of English language teaching* (2nd ed., pp. 1–7). Wiley.
- Brown, H. D., & Lee, H. (2015). *Teaching by principles: An interactive approach to language pedagogy* (4th ed.). Pearson.
- Carrell, P. L., & Carson, J. G. (1997). Extensive and intensive reading in an EAP setting. *English for Specific Purposes*, 16(1), 47–60. [https://doi.org/10.1016/S0889-4906\(96\)00031-2](https://doi.org/10.1016/S0889-4906(96)00031-2)
- Carrell, P. L., & Eisterhold, J. C. (1983). Schema theory and ESL reading pedagogy. *TESOL Quarterly*, 17(4), 553–573. <https://doi.org/10.2307/3586613>
- Carrell, P. L., Devine, J., & Eskey, D.E. (Eds.). (1988). *Interactive approaches to second language reading*. Cambridge University Press.
- Cook, V. (2016). *Second language learning and language teaching* (5th ed.). Routledge.
- Council of Europe. (2018). *Common European framework of reference for languages: Learning, teaching, assessment. Companion volume with new descriptors*. <https://rm.coe.int/cefr-companion-volume-with-new-descriptors-2018/1680787989>
- Day, R. R. (2013). Graded Readers. In C. A. Chapelle (Ed.), *The encyclopedia of applied linguistics*, (pp. 1–4). Blackwell Publishing.
- Day, R. R., Bamford, J. (2002). Top ten principles for teaching extensive reading. *Reading in a Foreign Language*, 14(2), 136–141. <https://nflrc.hawaii.edu/rfl/item/61>
- DeKeyser, R. M., & Suzuki, Y. (2025). Skill acquisition theory. In B. VanPatten, G. D. Keating, & S. Wulff (Eds.), *Theories in second language acquisition: An introduction* (4th ed., pp. 157–182). Routledge.
- Duke, N. K., & Cartwright, K. B. (2021). The science of reading progresses: Communicating advances beyond the simple view of reading. *Reading Research Quarterly*, 56(S1), S25–S44. <https://doi.org/10.1002/rrq.411>
- Ellis, R. (2023). *Instructed second language acquisition*. Routledge.
- Fields, M. (2017). Extended reading: A teacher-directed alternative to extensive reading. *The Journal of Asia TEFL*, 14(1), 197–204. <http://dx.doi.org/10.18823/asiatefl.2017.14.1.15.197>
- Freddi, M. (2015). Text and corpus: Mixing paradigms in EAP syllabus and course design. In Thompson, P., & Diani, G. (Eds.), *English for academic purposes: Approaches and implications*. Cambridge Scholars Publishing.
- Gascoigne, C. (2008). Independent second language reading as an interdependent process. In S. Hurd & T. Lewis (Eds.), *Language learning strategies in independent settings* (pp. 67–83).: Multilingual Matters.
- Grabe, W. & Kaplan, R. B. (1996). *Theory and practice of writing: An applied linguistic perspective*. Longman,
- Grabe, W. & Stoller, F. L. (2009). Teaching the written foreign language. In K. Knapp & B. Seidlhofer (Eds.), *Handbook of foreign language communication and learning* (pp. 439–463). Walter de Gruyter.
- Grabe, W. & Stoller, F. L. (2014). Teaching reading for academic purposes. In M. Celce-Murcia, D. M. Brinton, & M. A. Snow (Eds.), *Teaching English as a second or foreign language* (4th ed.), (pp. 189–205). National Geographic Learning
- Grabe, W. & Stoller, F. L. (2019) *Teaching and researching reading* (3rd ed.). Routledge.

<https://doi.org/10.4324/9781315726274>

- Grabe, W. & Yamashita, J. (2022). *Reading in a second language: Moving from theory to practice* (2nd ed.) Cambridge University Press. <https://doi.org/10.1017/9781108878944>
- Grabe, W. (2002). Dilemmas for the development of second language reading abilities. In J. C. Richards & W. A. Renandya (Eds.), *Methodology in language teaching: An anthology of current practice* (pp. 276-286). Cambridge University Press.
- Grabe, W. (2006). Areas of research that influence L2 reading instruction. In E. Usó-Juan & A. Martínez-Flor (Eds.), *Current trends in the development and teaching of the four language skills* (pp. 279-301). Mouton de Gruyter.
- Grellet, F. (1981). *Developing reading skills: A practical guide to reading comprehension exercises*. Cambridge University Press.
- Hafiz, F. M. & Tudor, I. (1989). Extensive reading and the development of language skills. *ELT Journal*, 43(1), 4-13. <https://doi.org/10.1093/elt/43.1.4>.
- Hill, D. R. (2013). Graded readers. *ELT Journal*, 67(1), 85-125. <https://doi.org/10.1093/elt/ccs067>
- Hirvela, A. (2016). Academic reading into writing. In K. Hyland & P. Shaw (Eds.), *The Routledge handbook of English for academic purposes*, (pp. 127-138). Routledge.
- Hirvela, A. (2017). *Connecting reading & writing in second language writing instruction* (2nd ed.). University of Michigan Press. <https://doi.org/10.3998/mpub.8122864>
- Horst, M. (2009) Developing definitional vocabulary knowledge and lexical access speed through extensive reading. In Han, Z-H., Anderson, N. (Eds.), *Second language reading research and instruction: Crossing the boundaries* (pp. 40-64). University of Michigan Press.
- Jeon, E.-Y., & Day, R. R. (2016). The effectiveness of ER on reading proficiency: A meta-analysis. *Reading in a Foreign Language*, 28(2), 246-265. <https://doi.org/10.64152/10125/66901>
- Jiang, X. Grabe, W. & Carrell, P. L. (2020). Reading. In N. Schmitt & M. P. H. Rogers (Eds.), *An introduction to applied linguistics* (3rd ed.) (pp. 259-278). Routledge.
- Kintsch, W., & van Dijk, T. A. (1978). Toward a model of text comprehension and production. *Psychological Review*, 85(5), 363-394. <https://doi.org/10.1037/0033-295X.85.5.363>
- Koda, K. & Zehler, A. M. (Eds.). (2008). *Learning to read across languages: Cross-linguistic relationships in first-and second-language literacy development*. Routledge.
- Koda, K. (2012). Development of second language reading skills: Cross-linguistic perspectives. In S. M. Gass & A. Mackey (Eds.), *The Routledge handbook of second language acquisition* (pp. 303-318). Routledge.
- Koda, K., & Yamashita, J. (Eds.). (2019). *Reading to learn in a foreign language: An integrated approach to foreign language instruction and assessment*. Routledge.
- Lai, F. K. (1993). The effect of a summer reading course on reading and writing skills. *System*, 21(1), 87-100. [https://doi.org/10.1016/0346-251X\(93\)90009-6](https://doi.org/10.1016/0346-251X(93)90009-6)
- Larson-Hall, J., & Plonsky, L. (2015). Reporting and interpreting quantitative research findings. In L. Plonsky (Ed.), *Advancing quantitative methods in second language research* (pp. 219-245). Routledge.
- Macalister, J., & Nation, I. S. P. (2020). *Language curriculum design* (2nd ed.). Routledge.
- Mason, B. & Krashen, S. (1997). Extensive reading in English as a foreign language. *System*, 25(1), 91-102. [https://doi.org/10.1016/S0346-251X\(96\)00063-2](https://doi.org/10.1016/S0346-251X(96)00063-2)
- McNamara, D. S., & Magliano, J. P. (2009). Toward a comprehensive model of comprehension. In B. H. Ross (Ed.), *The Psychology of Learning and Motivation: Advances in research and theory* (Vol. 51, pp. 297-384). Academic Press.
- Nation, I. S. P. & Macalister, J. (2021). *Teaching ESL/EFL reading and writing* (2nd ed.). Routledge.

- Nation, I. S. P. & Wang, K. (1999). Graded readers and vocabulary. *Reading in a Foreign Language*, 12(2), 355-380.
- Nation, I. S. P. (2022). Dosage and timing in L2 reading development. *Reading in a Foreign Language*, 34(2), 215-232.
- Nation, I. S. P., & Waring, R. (2019). *Teaching extensive reading in another language*. Routledge. <https://doi.org/10.4324/9780367809256>
- Norris, J. M. (2015). Statistical significance testing in second language research: Basic problems and suggestions for reform. *Language Teaching*, 48(4), 487-504.
- Nuttall, C. (1996). *Teaching reading skills in a foreign language* (2nd ed.). Heinemann.
- Paran, A. (2003). Intensive reading and extensive reading. In J. C. Richards & W. A. Renandya (Eds.), *Methodology in language teaching: An anthology of current practice* (pp. 205-210). Cambridge University Press.
- Pichette, F. (2005) Time spent reading and reading comprehension in second language learning. *Canadian Modern Languages Review*, 62, 243-262. <https://doi.org/10.3138/cmlr.62.2.243>
- Plonsky, L., & Oswald, F. L. (2014). Meta-analysis in second language acquisition. *Language Learning*, 64(4), 879-910.
- Renandya, W. A. & Jacobs, G. M. (2002). Extensive reading: Why aren't we all doing it Barkhuizen? In J. C. Richards & W. A. Renandya (Eds.), *Methodology in language teaching: An anthology of current practice* (pp. 295-302). Cambridge University Press.
- Saito, K., Dewaele, J.-M., & Hanzawa, K. (2023). Motivation and reading engagement in EFL contexts. *Studies in Second Language Acquisition*, 45(2), 312-338.
- Shapiro, S. S., & Wilk, M. B. (1965), An analysis of variance test for normality (complete samples), *Biometrika* 52, 591-611. <https://doi.org/10.2307/2333709>
- 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>
- Tanaka, H., & Stapleton, P. (2007) Increasing reading input in Japanese high school EFL classrooms: An empirical study exploring the efficacy of extensive reading. *The Reading Matrix* 7(1), 115-126.
- Taylor, J., Roehrig, A. D., & Connor, C. M. (2022). Reading fluency interventions: A meta-analysis of the last decade. *Scientific Studies of Reading*, 26(3), 189-210.
- Webb, S. (2021). Vocabulary learning through extensive reading. *Applied Linguistics*, 42(5), 873-896.
- Webb, S., & Chang, A. C. S. (2022). Longitudinal vocabulary gains through graded reading. *TESOL Quarterly*, 56(1), 112-138.
- Wilcox, R. R. (2021). *Introduction to robust estimation and hypothesis testing* (5th ed.). Academic Press.
- Williams, E. (2006). Teaching reading: Individual and social perspectives. In Uso-Juan, E., & Martinez-Flor, A. (Eds.), *Current trends in the development and teaching of the four language skills* (Vol. 29) (pp. 355-380). Mouton de Gruyter.
- Willingham, D. T. (2006). How knowledge helps. *American Educator*, 30(1), 30-37. <https://www.aft.org/ae/spring2006/willingham>
- Zhang, X., & Zhang, R. (2023). Reading-writing connections in EFL: A longitudinal study. *Journal of Second Language Writing*, 59, 100-118.