



An analysis of graded readers for compulsory secondary education in Spain: A mixed study*

Antonio Daniel Juan Rubio

(University of Granada)

Juan Rubio, Antonio Daniel. 2024. An analysis of graded readers for compulsory secondary education in Spain: A mixed study. *Linguistic Research* 41(Special Edition): 1-24. The use of graded readers, whether compulsory or optional, is a common practice in many secondary schools in Spain. Many English departments include these readers in their annual teaching programmes because their use promotes effective learning and improves students' language skills, both productive and receptive. The main objective of this paper was to examine the material that publishers included in their collections of graded readers for the subject of English as a foreign language in the 4th year of Compulsory Secondary Education (CSE). For this purpose, a mixed study was conducted following a concurrent design and using a 16-item questionnaire developed by the researcher and validated by several experts according to Lawshe's content validity ratio model. The sample analysed 173 graded readers belonging to nine publishers with such material for CSE by means of a parametric analysis of variance (ANOVA). The research hypothesis was that publishers do not include the same material in their graded readers' collections. The research results confirm that not all publishers include the same material in their collections. (University of Granada)

Keywords educational material, graded readers, mixed study, publishers, questionnaire, Secondary Education

1. Introduction

This article aims to develop a revision of the current offer of graded readers for the 4th year of Compulsory Secondary Education (CSE) by the main publishers, offering a descriptive analysis of their contents with the aim of identifying the resources offered

* We would like to thank the Language Research editor and two anonymous reviewers for their helpful comments on earlier versions of this manuscript.

by each collection to facilitate the comprehension process for students and the development of their communicative skills, and encourage their interest in reading and cultural enrichment.

For many years, the use of graded readers has been a common practice in English classes in the stage of CSE because, as their name suggests, these readings aim to adapt and adjust to the reader's language skills. Reading real texts is an impossible task, especially in the case of English learners with a very basic command of the language and a very limited vocabulary (De la Fuente McCauley 2019). Therefore, using graded readers is a good way to introduce learners to extensive texts, broadening their exposure to the language. In addition, the usual graded reader's distribution into short chapters, together with frequent illustrations and the use of simple language, facilitates their reading.

Simultaneously, the incorporation of new technologies cannot be forgotten into the field of education and their repercussions. Until a couple of decades ago, an English reading book was based on a physical book that had been adapted to the learner's expected level and, in some cases, included a glossary of vocabulary. Nowadays, graded readers are available in different formats and are complemented by interactive activities incorporated into each section. Given the wide range of materials available, the importance of their selection is clear. The aim is for the teacher to opt for resources that best suit the needs of the students, make their language learning process more enjoyable, and arise their interest so that they can take on a more active role in improving their language (González Otero 2016).

This new methodology adapted to the students and the diversity of activities helps us realise that the same resource can be worked on from different approaches, and that it must be adapted to the different levels that can be found in the same classroom (Alonso 2020). For this reason, it is essential to be familiar with the current range of graded readers available for learning English to be able to select materials that are best suited to the classroom and facilitate activities that encourage active participation.

Consequently, the general objective of this study was to determine the material included by publishers to work with graded readers in a 4th CSE classroom. For this purpose, nine collections, one from each publisher, and 173 representative graded readers were randomly selected for examination through a questionnaire measured on a numerical scale.

2. Theoretical framework

Even at the risk of omitting some important studies, a representative selection was made of the scientific production on the subject matter under study, as the theoretical framework in a mixed study tends to be broader by considering multiple approaches and perspectives to achieve an in-depth understanding of the study phenomenon. To do so, not only articles or book chapters were reviewed, but also master's dissertations to make it as varied as possible, since it is difficult to find studies dealing with this topic for Secondary Education in Spain.

The most recent article was published by Serrano Serrano in 2023, in which the author analysed the benefits of graded readers with scientific content on vocabulary acquisition in 96 Primary Education students in Spain. Serrano Serrano examined 39 graded readers over one academic year and concluded that “the findings indicate that the practice of extensive reading resulted in notable improvements in vocabulary acquisition during the first half of the school year; however, the advantages were less evident in the second half” (2023: 10). In her article, the author focused on aspects related to vocabulary, and omitted other types of content.

The second article was by Alghizzi and Elyas (2022), who analysed the effect that graded readers had on students' reading comprehension during the Covid-19 pandemic. In their longitudinal experimental quantitative research, the authors distributed a questionnaire to 130 students selected from a university in Saudi Arabia. In their study, they showed that the students' reading comprehension had improved, although not gradually: “these results imply that although the reading comprehension of the control group participants had temporarily increased significantly, such an increase was not gradual” (2022: 13). Unlike our study, the authors neither focused on the same educational stage, nor they examined the material for graded readers.

In the previous year, Puspitasary and AUFAR (2021) published an article in which they investigated four students' perceptions of graded readers through interviews. This study was similar to ours because it focused on the materials provided by graded readers, although in their case, they analysed students at a private university in Indonesia, and the sample was quite small. According to the authors, “the students perceived graded readers as materials that gave them benefits because they provided different levels, vocabulary, pages, and representative illustrations” (2021: 415). It should also be noted that they employed qualitative research with a phenomenological

design.

Another article published in 2021 was by Ateek, who examined the impact of graded readers on students' reading fluency and vocabulary knowledge. This study involved ten low-intermediate Jordanian students using a mixed method through testing and interviewing: "The findings of the study indicate that the impact of the extensive reading approach was positive on the learners' reading fluency and vocabulary knowledge" (2021: 126). The author examined 250 graded readers from three different series in terms of syntax, sentence length, complexity, and vocabulary: the Cambridge English Readers (Cambridge), Oxford Bookworms Series (Oxford), and Penguin Readers (Pearson).

A study published by Rivera Jurado and López Pérez (2020) brought English literature closer to that of primary and secondary school students through several graded readers of Edgar A. Poe's short stories. According to the authors, "in addition to contributing to improving students' linguistic competence, the graded readers also fulfil a pleasurable function since the student does not find the enjoyment impeded by observing the vocabulary, the structures or the progression" (2020: 225). In their article, they analysed Poe stories from five publishers reviewed in our study according to their subject matter: MacMillan, Pearson, Oxford, Vicens Vives, and Burlington.

The previous year three studies on graded readers were found, two articles, and one master's dissertation. Starting with the article published by Kara (2019), she studied the effects that graded readers had on students' reading comprehension skills. The author examined these effects in a group of 30 students from a Turkish university. In the author's opinion, "the amount of reading had a significant effect on students' reading comprehension scores. Students who read more books achieved more reading comprehension development in comparison to those who read less" (2019: 161). In this study, only university students were involved, and the graded readers' material was not examined. Furthermore, it focused specifically on Oxford Bookworms series from Oxford.

Serrano Serrano and Pellicer Sánchez (2019) published another article on the year in which they related extensive reading through graded readers and auditory input to improve students' reading fluency and comprehension. According to the authors, "our results suggested a significant relationship between the amount of time learners spent processing the text and comprehension scores" (2019: 17). The study focused on 36 primary school students who had to read the graded reader "Uncle Jack and

the Bakonzi tree” based on comprehension questions, but without analysing the content of the graded reader as in our study.

The master’s dissertation was presented by De la Fuente McCauley (2019) in which he examined the use of graded readers for teaching English. However, in his work, the author focused on a single reader, “Frankenstein”, for Bachillerato with two different versions by means of a quantitative analysis. The aim was to show which version was more useful for teaching English, the original or the adapted version: “There are two tendencies which can be easily identified: a heavy simplification process where the text loses much of its original richness and linguistic value; and a lighter one, in which the final product maintains some of the features that can be appreciated in the original text” (2019: 37). Similar to our study, De la Fuente McCauley analysed the elements of both graded readers based on three levels: grammatical structures, verb tenses, and sentence length (level 1), content and adaptation changes (level 2), and activities, glossary, and resources (level 3).

The most prolific year in terms of publications was 2017, with five selected articles. One of them was published by Chang and Millett (2017) in which they analysed the effects that three graded readers in the “Sherlock Holmes” series (Sherlock Holmes and the Dukun’s Son, Sherlock Holmes Short Stories, and The Last Sherlock Holmes Story) had on students’ reading comprehension, reading speed, and perceptions. The study involved 53 18-year-old Taiwanese high school students through pre and post-tests. Although the students were of a similar educational stage to ours, the study only examined two variables, reading speed, and reading comprehension, without looking at other elements as in our study: “Although there were some differences in text organisation, the students showed that they read faster with related texts than with unrelated ones. These results may have some important pedagogical implications in teaching beginner readers to develop L2 reading fluency” (2017: 15).

A second article, published by Madarova (2017), referred to the use of graded readers in Spain and how to motivate students to read in class. This study focused on the secondary school stage like ours, although it did not specify any specific grade, nor did it analyse the materials of the graded readers: “Even students who might enjoy reading in their native language might be discouraged from reading a book in English. Teachers try to remedy such situations by assigning a graded reader to help students to complement their reading” (2017: 86). In her opinion, one of the disadvantages of using graded readers was that, in general, they did not include

characters with which to empathise, and the language was complicated.

Martinez (2017) compared the use of graded readers to improvements in vocabulary and reading comprehension. In this article, the author analysed two adult classes with a total of 38 students at a community college by means of a pre-test and a post-test in a mixed study: “As the researcher, I was able to see the power of graded readers first hand. At least once a week, I would go to the site where the study was being conducted to bring graded readers. All of them appeared focussed on what they were reading” (2017: 55). Therefore, this study did not only examine the material of graded readers but also focused on measuring pupil improvement.

The fourth article was published by Holster et al. (2017), who investigated the difficulty of graded readers. After their analysis, the authors recommended to students that their main concern when choosing a graded reader should be its length, an aspect not considered in our study: “These findings suggest that the current practice of levelling graded readers by headword levels derived from native-speaker corpus is largely ineffective” (2017: 239). Thus, they analysed 300 graded readers by measuring the number of words, sentence length, and readability statistics, using the SPSS tool to calculate the correlation between different levels of difficulty.

Finally, the fifth and last article of the year was published by Albay (2017), in which he discussed the benefits to graded readers. According to the author, this allowed students to practice reading skills and, at the same time, provided an excellent opportunity to reinforce vocabulary: “Reading significantly improves language ability and reading comprehension of learners. Learners by means of graded readers experience how structures and words function in texts” (2017: 178). However, this short article did not emphasise any particular stage of education, nor did it discuss the elements of graded readers, thus limiting to a generic review of their benefits.

In 2016, an article with a theme similar to ours was found. Rodrigo’s article (2016) validated different levels of graded readers according to the publisher. After analysing 203 titles, the aim was to provide a uniform system with which to catalogue graded readers according to the number of headwords and level, but without analysing other elements: “In order to help students reach the reading skills needed to face the demand of upper-level courses, deal with authentic texts, and improve their language skills, graded readers have shown to be an optimal tool” (2016: 78).

Two other articles related to this topic were published in 2015. One of them was published by Gillis-Furutaka (2015) in which she reviewed the way graded readers

were classified according to their levels of difficulty. The author acknowledged that the most common classification method was based on the number of headwords, but according to her, other aspects should also be considered: “The findings suggest that authors and editors need to pay closer attention to the likely age range of the target readers, cultural issues, use of idiomatic and figurative language, literary devices, illustrations, and plot structure when determining the readability of grade readers” (2015: 15). As can be seen, some elements coincide with those of our study, but not with the educational stage, since qualitative interviews were conducted with 83 senior high school and university students.

The other article of the year, published by Ali and Saiden (2015), examined the materials of two graded readers with children with reading difficulties in primary school. Through a qualitative study, the authors concluded that “the findings showed that the children have increased skills in reading. There is a pedagogical implication for the use of alternative strategies in the teaching of reading for struggling readers” (2015: 2632).

In 2014, two articles of interest were selected. One of them, published by Tian (2014), explained how reading books can be used extensively to improve students’ writing. The author argued for the potential value of graded readers in enhancing creative writing because “the language control in guided readers is achieved through three methods: lexical control, structural control, and information control” (2014: 57). Analysing four graded readers, this study focused on syntactic complexity, lexical density, and discourse organisation in terms of linguistic difficulty. It also added that graded readers should contain topics relevant to students, as well as illustrations.

The article by Azmuddin et al. (2014) promoted the use of extensive reading by graded readers in Malaysian public universities. The authors measured the level of proficiency in reading comprehension and reading development of 125 students by examining 344 graded readers from four different publishers and series: MacMillan Readers (MacMillan), Cambridge English Readers (Cambridge), Oxford Bookworms Series (Oxford), and Penguin Readers (Pearson), as in our study: “This analysis suggest that extensive reading improved students’ performance in both progress tests and extensive reading tests. Students performed significantly better on the post-test than on the pre-test” (2014: 112).

In 2013, a master’s degree dissertation by Kredátusová (2013) analysed the benefits of extensive reading through graded readers for secondary school students. According

to the author, “graded readers are ideal source of material for lower-intermediate independent readers. If students find their appropriate level, they can make progress in reading quite smoothly without teachers’ instruction” (2013: 18). Unlike our study, this only examined reading books in terms of the number of headwords.

But certainly, one of the first authors to take an in-depth look at graded readers was Claridge (2012), who examined the production of graded readers by the major publishers of the time: “Cambridge Readers” (Cambridge), “Oxford Bookworms” (Oxford), “MacMillan Guided Readers” (MacMillan), and “Penguin Readers” (Pearson). According to the author, publishers did not take students’ opinions into account and based their output on the demands of teachers and booksellers. Claridge reviewed 703 graded readers in terms of subject matter, language level and design: “In examining graded readers for any one factor which might either encourage or discourage learners from reading, it is hard to pinpoint an area where the publishers as a group could be at fault. The four described here all agree that a good story, well-written for the level, is the ideal model for a graded reader” (2012: 117).

3. Methodology

The review of the theoretical framework has shown that there are not a wide range of studies that analyse the use of graded readers in Secondary Education in Spain. Therefore, this study aimed to clarify this situation, as it is currently an unresolved issue. In this section, the research design, objective and hypothesis, sample, instrument, and procedure for data collection and analysis are developed.

3.1 Methodological design

The design is nothing more than a strategy to be implemented to achieve research objectives. Therefore, Creswell (2021) defined mixed research as a methodology in which the researcher brings together quantitative (closed) and qualitative (open) data, integrates or combines the two, and draws inferences that provide a broader view than quantitative or qualitative data alone. A fundamental premise of this approach is that combining statistical trends (quantitative data) with personal experiences (qualitative data) provides a better understanding of the research problem than using

either data set alone.

Ortega Sánchez (2023) also argued that mixed research represents a natural complement to qualitative and quantitative research, recommending that the two methods be combined so that the strengths present in both are chosen and the weaknesses, which are fewer, allow the researcher to visualise the research problem more fully, are reduced.

In our research, within the mixed method, a concurrent design was applied according to the classification made by Hernández Sampieri et al. (2014), and the instrument used was a questionnaire. This design allows for the integration of multiple perspectives and approaches in the same research, which provides a more complete and enriching understanding of the phenomenon studied to simultaneously collect qualitative and quantitative data. The independent variable was the selection of nine publishers. The dependent variable was a list of the 16 items to be completed in the questionnaire. Ethical considerations are fundamental in any type of research, including mixed research. Therefore, some key ethical considerations in conducting this research are privacy and confidentiality, data protection, transparency, and honesty.

3.2 Objective, research question and hypothesis

The main objective of this study was to examine the material that publishers include in their graded readers' collections on the subject of English as a foreign language in the 4th year of Compulsory Secondary Education. To frame the research objective, it is important in mixed research to consider the subject of the study and the combination of approaches. Considering this objective, the following research question was posed: "Which publishers include the most material in their collections of graded readers?". The concurrent design allowed us to analyse the data to identify the materials included in the collections.

Our research hypothesis was that publishers would not include the same material in their graded readers' collections. The concurrent approach used allowed us to better explore the relationship between the two variables. Since a normality test of the instrument was subsequently conducted, as well as its subsequent validation, it was necessary to establish a null hypothesis and an alternative hypothesis to accept the

correct hypothesis. In this case, the null hypothesis stated that the data obtained had a normal distribution. In contrast, the alternative hypothesis specified that the data did not have a normal distribution.

3.3 Sample

According to Hernández Sampieri et al. (2014), population is defined as the set of all units of study. In mixed research, both population determination and sampling are essential to ensure the representativeness and validity of the results obtained. Population determination involves identifying the group of elements that are the object of study, based on the research objective and question posed in advance. The sampling process involved selection of an adequate sample that represented the study population.

Considering the concurrent mixed design used, mixed sampling was conducted, which involved selecting a representative sample through random sampling (a series from each publisher) and selecting specific cases through purposive sampling (the most complete series). To do so, Fischer and Navarro Vega's (1997) formula was followed for adequate sample selection. Applying such a formula for the proportion of a finite population ($n=271$), we needed to examine 160 graded readers. However, a representative sample of 173 readers was chosen, larger than needed, which gave the sample broad validity and representation.

Therefore, the most widely used graded readers' collections were selected for the 4th year of Compulsory Secondary Education in Spain based on the following criteria: grade (4th CSE), CEFR level (intermediate), and number of headwords (1,000-1,400). In addition, the number of graded readers offered by each publisher for the stage followed an uneven distribution, with Oxford being the most representative with 48 titles, and Express Publishing the least with just four.

Consequently, the selected sample included the following graded readers' collections ($n=173$): "Activity Readers" (Burlington, 19 titles), "Experience Readers" (Cambridge, six titles), "MacMillan Readers" (MacMillan, 30 titles), "Bookworms Readers" (Oxford, 48 titles), "Richmond Secondary Readers" (Richmond, nine titles), "Teen Graded Readers" (Eli Publishing, eight titles), "Penguin Active Readers" (Pearson, 17 titles), "Black Cat" (Vicens Vives, 32 titles), and "Illustrated Readers" (Express Publishing, four titles).

Analysis	Illustrations & photographs									
	Vocabulary									
	Final glossary									
	Phonetic transcription									
	Audiobook									
	Digital format									
	APP									
Supplementary Materials	Before-reading activities									
	While-reading activities									
	After-reading activities									
	Answer key									
	Downloadable activities									
	Teachers' materials									

Legend: 1- Burlington. 2- Cambridge. 3- Eli Graded Readers. 4- Express Publishing. 5- MacMillan. 6- Oxford. 7- Pearson. 8- Richmond. 9- Vicens Vives.

To prevent the possible negative effects of the questionnaire application, Cronbach's alpha coefficient was used to check the internal consistency of the questionnaire, which yielded a value of 0.93. This value means that the instrument was highly reliable; therefore, the questionnaire had high internal consistency and high reliability, and its measurements were stable and consistent. In the following table, the reliability of the questionnaire can be checked.

Table 2. Reliability of the questionnaire

Number items	Sum of item variance	Total instrument variance	Cronbach's alpha coefficient
16	2415.1851	19376.6666	0.93371

3.5 Data collection procedure

The data collection and analysis procedure consisted of three phases: exploratory, fieldwork, and information processing. The exploratory phase is preliminary research conducted to clarify the research procedure to constitute the fieldwork. In this phase, the objective of our research was delimited and developed a theoretical and methodological study. Furthermore, the instrument to use was chosen and the sample to be examined.

The second phase was based entirely on the available fieldwork. Here, we put

the theoretical construction elaborated in the previous phase into practice, with the necessary time for data collection. The second phase was conducted by completing a questionnaire using data from nine publishers. For this purpose, a template was created to enter the data obtained from each publisher, and then transferred it to an Excel spreadsheet.

Finally, the third phase concerned the evaluation, understanding, and interpretation of the empirical data, articulated with the theory on which our study was based. The last phase was conducted after obtaining all data. In this phase, a descriptive analysis of the variables was conducted, expressed in terms of frequency.

3.6 Data analysis

The data analysis was performed using Microsoft Excel, a statistical data processing software offered by Microsoft. The procedure followed was to prepare an Excel spreadsheet, enter the data from each publisher, and then statistically process the data. For our study, descriptive statistics to determine the means and standard deviations were used.

To verify which hypothesis was valid (null or alternative), a normality test of the instrument was conducted. For this purpose, the Kolmogorov-Smirnov normality test was used, which allowed us to measure the degree of agreement between the distribution of the data set and a specific theoretical distribution. After this test, the calculated Kolmogorov-Smirnov value (Ks c-value) was 0.1774, while the Kolmogorov-Smirnov table value (Ks t-value) which reflects the maximum permissible error, was 0.2734, with a p-value of 0.2538. Since the Ks c-value was lower than the Ks t-value and the p-value was greater than the accepted degree of significance (0.05), the null hypothesis had to be accepted because the data had a normal distribution. Consequently, a parametric test was used to analyse the results.

These data were corroborated by the Ryan-Joiner test. With a p-value of 0.3059, the calculated Ryan Joiner value (RJ c-value) was 1.0023, and the Ryan Joiner table value (RJ t-value) was 0.9104. In this case, since the RJ c-value was greater than the RJ t-value and the p-value was greater than the degree of significance, the null hypothesis was accepted. The standard deviation was 147.6439 for both tests.

Table 3. Kolmogorov-Smirnov and Ryan-Joiner tests

Value	t-value	p-value	Test
0.1774	0.2734	0.2538	Kolmogorov-Smirnov
1.0023	0.9104	0.3059	Ryan-Joiner

Second, to perform an inferential analysis and verify the differences between the nine publishers, we searched for kurtosis (K) and the coefficient of asymmetry (CA) between different items. These are valuable statistical measures that provide information regarding the shape of the data distribution and the presence of outliers. The following table shows the data obtained for each case for the entire instrument:

Table 4. Kurtosis and Coefficient of asymmetry

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
K	0.58	0.86	4	0.11	3.04	1.64	0	1.31	0.45	9	1.16	0.36	0.07	0.11	2.05	1.25
CA	1.04	1.36	2.12	0.93	1.84	1.46	0	0.51	1.31	3	1.28	1.08	0.91	0.93	4.43	1.3

Finally, as the p-value obtained in both normality tests was greater than 0.05, which is the accepted degree of probability, a parametric test was used to further analyse the results. Therefore, an inferential analysis was conducted by means of ANOVA analysis of variance, a statistical tool to compare the means of the nine publishers (173 graded readers) in terms of the 16 items scrutinised in the questionnaire.

4. Results

The results derived from the analysis of the questionnaire used for the research are presented, first analysing the overall questionnaire results. An inferential ANOVA analysis was conducted of the 16 items and the results of the collections of the nine publishers for each item. The analysis of variance test ANOVA is a statistical method used to determine whether the results of a test are significant in rejecting the null hypothesis or accepting the alternative hypothesis. The null hypothesis states that there are no differences between variables. On the other hand, the alternative hypothesis states that the means are different for each variable. In the following table, the results of the ANOVA analysis of variance can be checked according to the parametric test.

Table 5. ANOVA analysis of variance

<i>Item</i>	<i>Total</i>	<i>Sum</i>	<i>Mean</i>	<i>Variance</i>	<i>Standard error</i>	<i>Standard deviation</i>
1	9	94	10.4444	164,52778	4.2756	12.8268
2	9	112	12.4444	301,77778	5.7905	17.3717
3	9	12	1.3333	8	0.9428	2.8284
4	9	173	19.2222	218,69444	4.9294	14.7883
5	9	61	6.7777	122,44444	3.6884	11.0654
6	9	124	13.7777	256,69444	5.3405	16.0216
7	9	0	0	0	0	0
8	9	164	18.2222	162,94444	4.2549	12.7649
9	9	75	8.3333	126,5	3.7490	11.2472
10	9	16	1.7777	12,444444	0.8888	2.6666
11	9	124	15.6666	243,25	5.1988	15.5964
12	9	137	15.2222	202,94444	4.7486	14.2458
13	9	172	19.1111	222,11111	4.9677	14.9033
14	9	173	19.2222	218,69444	4.9294	14.7883
15	9	139	15.4444	223,52778	4.9836	14.9508
16	9	143	15.8888	237,86111	5.1409	15.4227

According to the data, since the f-value obtained (2.3031) was greater than the critical value (1.7451), the alternative hypothesis was accepted because there were significant differences among the different variables examined. Pearson's correlation analysis was also conducted to determine the t-student value. Since the t-student value obtained (1.0151) was lower than the critical value, there is sufficient statistical evidence to conclude that the correlation coefficient is different from 0.

For each dimension, not only the quantitative data were considered, but also the qualitative data regarding the material included in the collections by each publisher. The nine publishers examined were those who currently have graded readers for 4th CSE in Spain. Therefore, the results were analysed in terms of the material included.

Similarly, Tukey's method allows the comparison of each dimension to determine whether there are significant differences. If the calculated Q sub-alpha value (4.848) is less than the sample difference, there is a significant difference between the variables, which validates the alternative hypothesis, not all publishers include the same material in their collections of graded readers.

The first dimension focused on the context of the author and work and included two items: historical and cultural context (1), and description of the author and works (2). The Tukey honestly significant difference test (HSD) shows the following results

for this dimension, where differences can be seen in eight out of the nine collections for each item.

Table 6. HSD Tukey Dimension 1

	1	Sample Difference (SD)	2	Sample Difference (SD)
Mean	10.4444		12.4444	
Activity Readers (AR)	17	6.5556	0	12.4444
Experience Readers (ER)	2	8.4444	0	12.4444
Teen Readers (TR)	8	2.4444	8	4.4444
Illustrated Readers (IR)	4	6.4444	4	8.4444
MacMillan Readers (MR)	30	19.5556	20	7.5556
Bookworms Series (BS)	1	9.4444	48	35.5556
English Active Readers (EAR)	0	10.4444	0	12.4444
Richmond Readers (RR)	0	10.4444	0	12.4444
Black Cat (BC)	32	21.5556	32	19.5556

Legend: AR- Burlington. ER- Cambridge. TR-Eli Graded Readers. IR-Express Publishing. MR-MacMillan. BS-Oxford. EAR-Pearson. RR-Richmond. BC-Vicens Vives.

The second dimension referred to the analysis of the graded reader and included eight items. The first four items were the description of the main characters (3), illustrations and photographs (4), vocabulary (5), and the final glossary (6). In the following table, the results are checked for items 3-6 in the second dimension. We can see a certain disparity of results in the first four items of this second dimension, where three items stand out, with almost no differences between the publishers' collections.

Table 7. HSD Tukey Dimension 2 (3-6)

	3	SD	4	SD	5	SD	6	SD
Mean	1.3333		19.2222		6.7777		13.7777	
AR	0	1.3333	19	0.2222	0	6.7777	19	5.2230
ER	0	1.3333	6	13.2222	0	6.7777	6	7.7777
TR	8	6.6667	8	11.2222	8	1.2223	8	5.7777
IR	4	2.6667	4	15.2222	4	2.7777	4	9.7777
MR	0	1.3333	30	11.2222	0	6.7777	30	16.2223
BS	0	1.3333	48	29.2222	0	6.7777	48	34.2223
EAR	0	1.3333	17	2.2222	17	10.2223	0	13.7777
RR	0	1.3333	9	10.2222	0	6.7777	9	4.7777
BC	0	1.3333	32	13.2222	32	25.2223	0	13.7777

The remaining four items included the phonetic transcription of words (7),

audiobook (8), available in digital format (9), and available in an app (10). In the following table, the results can be checked for items 7-10 in the second dimension. The seventh and tenth items stood out in this part of the second dimension with hardly any significant differences in the collections.

Table 8. HSD Tukey Dimension 2 (7-10)

	7	SD	8	SD	9	SD	10	SD
Mean	0		18.2222		8.3333		1.7777	
AR	0	0	19	0.7778	0	8.3333	0	1.7777
ER	0	0	6	12.2222	0	8.3333	0	1.7777
TR	0	0	8	10.2222	8	0.3333	0	1.7777
IR	0	0	4	14.2222	4	4.3333	0	1.7777
MR	0	0	30	11.7778	24	15.6667	0	1.7777
BS	0	0	39	20.7778	30	21.6667	8	6.2223
EAR	0	0	17	1.2222	0	8.3333	0	1.7777
RR	0	0	9	9.2222	9	1.3333	0	1.7777
BC	0	0	32	13.7778	0	8.3333	0	1.7777

Finally, the third dimension focused on all the supplementary materials available in the graded readers' collections analysed to complement their content and guide students. This dimension included six different items, being the first three before-reading activities (11), while-reading activities (12), and after-reading activities (13). In the following table, the results are shown for items 11-13 in the third dimension. For these items, only three collections showed no significant differences.

Table 9. HSD Tukey Dimension 3 (11-13)

	11	SD	12	SD	13	SD
Mean	15.6666		15.2222		19.1111	
AR	19	3.3334	19	3.7778	19	0.1111
ER	4	11.6666	5	10.2222	5	14.1111
TR	8	7.6666	8	7.2222	8	11.1111
IR	4	11.6666	4	11.2222	4	15.1111
MR	0	15.6666	0	15.2222	30	10.8889
BS	48	32.3334	43	27.7778	48	28.8889
EAR	17	1.3334	17	1.7778	17	2.1110
RR	9	6.6666	9	6.2222	9	10.1111
BC	32	16.3334	32	16.7778	32	12.8889

The remaining three items included the answer key (14), downloadable activities

(15), and materials intended for teachers (16). This dimension obtained the most different results among the nine publishers for the six items examined. In the following table, the results are checked for items 14-16 in the third dimension. The results are very similar to the previous table.

Table 10. HSD Tukey Dimension 3 (14-16)

	14	SD	15	SD	16	SD
Mean	19.2222		15.4444		15.8888	
AR	19	0.2222	19	3.5556	19	3.1112
ER	6	13.2222	6	9.4444	6	9.8888
TR	8	11.2222	8	7.4444	8	7.8888
IR	4	15.2222	4	11.4444	4	11.8888
MR	30	10.7778	28	12.5556	0	15.8888
BS	48	28.7778	48	32.5556	48	32.1112
EAR	17	2.2222	17	1.5556	17	1.1112
RR	9	10.2222	9	6.4444	9	6.8888
BC	32	12.7778	0	0	32	16.1112

5. Discussion of results

After validating the 16 items of the questionnaire, following the indications of the five experts, and applying the content value ratio (CVR) of Lawshe's model based on the experts' assessments, and after filling in the data of the nine publishers, this section presents the analysis of the results obtained by means of the JASP statistical programme.

The findings of this study confirm our research hypothesis that not all publishers include the same material in their collections of graded readers. A total of 13 items in the questionnaire showed significant differences in the inferential analysis. Most of them reflected such significant differences, while only three cases did not, as shown in the figure.

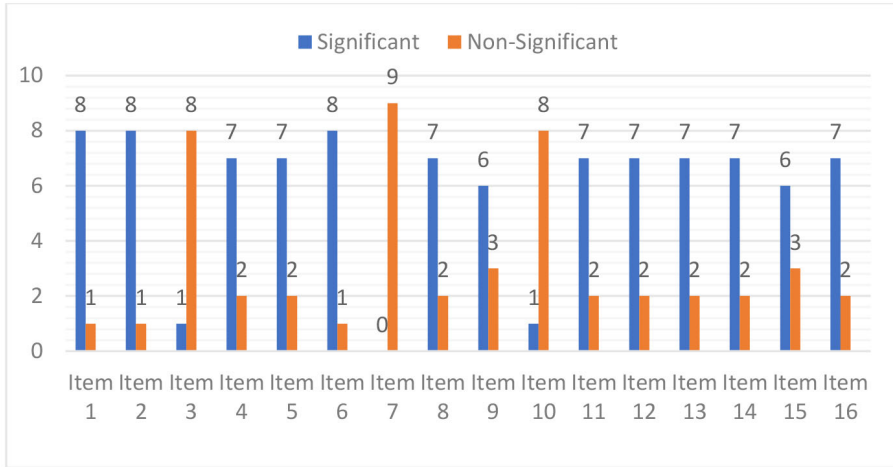


Figure 1. Significant vs. non-significant differences

Similarly, from a descriptive point of view, all items in the first and third dimensions showed significant differences among the nine collections. Only three cases of non-significant differences were found in the second dimension, the analysis of the graded reader. These cases referred to item 3 (the description of the main characters, Teen Readers), and item 10 (being available in an app, Bookworm Series) in which only one collection out of the nine examined did not show any difference in each item. The third case was item 7, the phonetic transcription of words, which was the only item in the questionnaire in which all collections coincided.

The other five items in the second dimension showed significant differences. For item 9, available in a digital format, three out of nine collections did not show any significant difference (Teen Readers, Illustrated Readers, and Richmond Readers). The Activity Readers and English Active Readers collections showed no differences in item 4 (illustrations and photographs), and 8 (audiobook), and the Teen Readers and Illustrated Readers collections in item 5 (vocabulary). Finally, the Richmond Readers collection was the only publisher that did not show differences in item 6, the final glossary.

In the case of the first dimension, context, and description of the author and works, both items showed the same distribution, with differences found in eight of the nine collections examined. The only exception was the collection of Teen Readers,

which did not show any significant difference.

The third dimension examined the supplementary materials found in the collections. This dimension had the most similar results, since significant differences were found in the six items analysed. The only difference was in item 15, downloadable activities, in which three out of the nine collections did not show any differences (Active Readers, English Active Readers, and Bookworm Series). The first two collections did not show any differences for the other five items either.

To further triangulate the validity and reliability of this study, our results were compared with those of previous studies, although none of them focused on our subject of study. The closest study to ours was that of Puspitsary and Aufar (2021) because they focused on materials provided by graded readers. The difference was the stage on which they focused since they examined students from a private university.

Madarova's (2017) article also focused on the use of graded readers in secondary education, although, in her case, she did not analyse the materials brought by these. Chang and Millet (2017) also focused on secondary education, but they simply examined one series in terms of reading comprehension and speed.

Several studies attempted to classify graded readers according to the number of headwords (Kredátusová 2013; Gillis-Furutaka 2015; Rodrigo 2016); length (Holster et al. 2017); subject matter, language level, syntax, and design (Claridge 2012; Tian 2014; Ateek 2021). On the other hand, Albay's article (2017) discussed the benefits of graded readers in terms of vocabulary acquisition, although none analysed the inclusion of supplementary materials.

Finally, some other studies also related to the use of graded readers in class, although they focused on different educational stages: Primary (Ali and Saiden 2015; Serrano Serrano and Pellicer Sánchez 2019; Rivera Jurado and López Pérez 2020; Serrano Serrano 2023), Bachillerato (De la Fuente McCauley 2019), university (Kara 2019; Azmuddin et al. 2021; Alghizzi and Elyas 2022), and adult students (Martínez 2017). None of these studies examined the materials offered by the different graded readers' collections.

6. Conclusions

Several conclusions can be drawn from the review of the theoretical framework.

First, most of the publications analysed stages other than Secondary Education, the stage our research is focused on, from Primary Education to university and adult students. Second, none of the studies examined focused specifically on the materials brought by the collections of the graded readers analysed or concentrated on nine different collections and publishers. This fact lends validity to our study because of its originality, among other aspects.

In our study, we established the objective of determining which materials publishers included in their graded readers' collections for the subject of English as a foreign language in the 4th year of CSE. To achieve this objective, a random sample of 173 graded readers was chosen from nine different collections, one from each publisher, including more readers than were required following Fischer and Navarro Vega's formula, which gave our sample broad validity and representation.

The answer to this objective came with the establishment of our research question by means of a concurrent design within the mixed method, which allowed us to bring together various perspectives and approaches. Our research hypothesis stated that publishers do not include the same materials in their graded readers' collections. The concurrent approach used allowed us to better explore the relationship between the variables.

The instrument used was a questionnaire with 16 items divided into three dimensions (context, analysis, and supplementary material) developed by the researcher after external validity by five experts, following Lawshe's content validity ratio model. The internal consistency of the questionnaire was also tested after obtaining Cronbach's alpha coefficient, which showed high consistency of the instrument with descriptive data.

Subsequent inferential analysis of the data using both Kolmogorov-Smirnov and Ryan-Joiner normality tests ensured the validity of the established hypothesis, which led us to conduct a parametric statistical test. The tool selected was the analysis of variance ANOVA, which yielded sufficient data on correlation, variance, standard deviation, and error to reject the null hypothesis and accept the alternative.

Ultimately, these analyses helped us demonstrate that our research hypothesis was certain, and the collections did not include the same material because there were significant differences in 13 of the 16 items scrutinised regarding the graded readers. The three collections that showed the highest number of significant differences were Experience Readers (Cambridge), MacMillan Readers (MacMillan), and Black Cat

(Vicens Vives) with 14 different items each. On the other hand, two collections had the lowest number of differences with six items: Activity Readers (Burlington), and English Active Readers (Pearson).

There are two possible limitations of this study. On the one hand, due to spatial limitations, the representative sample of graded readers was limited to 173 out of the 271 in total, choosing one collection from each publisher. Second, we limited the examination of graded readers to one specific year of CSE, without considering the rest of the stage. Therefore, a larger sample of graded readers should be further examined to corroborate our hypothesis, including more collections and levels.

References

- Albay, Mustafa. 2017. The benefits of graded reading. *International Journal of Social Sciences & Educational Studies* 3(4): 177-180.
- Alghizzi, Talal M. and Tariq Elyas. 2022. The effect of graded-reading websites/applications on EFL undergraduates' reading comprehension during the COVID-19 pandemic. *Electronics* 11(Special Edition): 1-22.
- Ali, Manisah M. and Muhammad N. Saiden. 2015. The use of graded reading materials for children with reading difficulties. *Procedia, Social and Behavioral Sciences* 191(1): 2631-2633.
- Alonso, Edgar L. 2020. *La comprensión lectora en la enseñanza del inglés en Secundaria*. Master's Thesis. University of La Laguna.
- Ateek, Mohammed. 2021. Extensive reading in an EFL classroom: Impact and learners' perceptions. *Eurasian Journal of Applied Linguistics* 7(1):109-131.
- Azmuddin, Ruhil A., Zuraina Ali, Ezihaslinda Ngah, Liyana M. Tamili, and Nooradelena M. Rusmil. 2014. Extensive reading using graded readers. *International Journal of Research in Social Sciences* 3(8): 109-113.
- Chang, Anna C. and Sonia Millett. 2017. Narrow reading: Effects on EFL learners' reading speed, comprehension, and perceptions. *Reading in a Foreign Language* 29(1): 1-19.
- Claridge, Gillian. 2012. Graded readers: How the publishers made the grade. *Reading in a Foreign Language* 24(1): 106-119.
- Creswell, John W. 2021. *A concise introduction to mixed methods research*. Thousand Oaks, CA: Sage.
- De la Fuente McCauley, Juan I. 2019. *Graded readers as a tool for teaching EFL: A comparative analysis on Frankenstein*. Master's Thesis. University of Valladolid.
- Fischer, Laura and Alma Navarro Vega. 1997. *Introducción a la investigación de mercados*. Ciudad de México: McGraw-Hill.

- García Córdoba, Fernando. 2015. *El cuestionario: Recomendaciones metodológicas para el diseño de cuestionarios*. Ciudad de México: Editorial Limusa.
- Gillis-Furutaka, Amanda. 2015. Graded reader readability: Some overlooked aspects. *Journal of Extensive Reading* 3(1): 1-19.
- González Otero, Rebeca. 2016. *Innovative resources based on ICTs and authentic materials to improve EFL students' communicative needs*. Dublin: Research Publishing.
- Hernández Sampieri, Roberto, Carlos Fernández Collado, and Pilar Baptista Lucio. 2014. *Metodología de la investigación*. Ciudad de México: McGraw-Hill.
- Holster, Trevor A., Jack W. Lake, and William R. Pellowe. 2017. Measuring and predicting graded reader difficulty. *Reading as a Foreign Language* 29(2): 218-244.
- Kara, Bayram. 2019. The effect of graded readers on reading comprehension skill of EFL students. *Language Teaching and Educational Research* 2(2): 160-172.
- Kredátusová, Maria. 2013. *The benefits of extensive reading in EFL*. Master's Thesis. University of Masaryk.
- Madarova, Slavka. 2017. Reflection about increasing students' motivation for reading in EFL classes through use of authentic materials that support creative production of language. *Encuentro* 26(1): 83-93.
- Martinez, Liza E. 2017. Adult Hispanic ESL students and graded readers. *The CATESOL Journal* 29(2): 41-59.
- Ortega Sánchez, Delfín. 2023. *¿Cómo investigar en didáctica de las ciencias sociales? Fundamentos metodológicos, técnicas e instrumentos de investigación*. Barcelona: Octaedro.
- Puspitasari, Evi and Almas AUFAR. 2021. A phenomenological exploration of EFL students' experiences with graded readers. *Advances in Social Science, Education and Humanities Research* 518(1): 412-418.
- Rivera Jurado, Paula and Magdalena López Pérez. 2020. Acercar la literatura inglesa en Primaria y Secundaria desde los cuentos de Edgar Allan Poe. *Aula de Encuentro* 20(1): 221-238.
- Rodrigo, Victoria. 2016. Graded readers: Validating reading levels across publishers. *Hispania* 99(1): 66-86.
- Serrano Serrano, Raquel. 2023. Extensive reading and science vocabulary learning in L2: Comparing reading-only and reading-while-listening. *Education Sciences* 13(493): 1-15.
- Serrano Serrano, Raquel and Ana Pellicer Sánchez. 2019. Young L2 learners' online processing of information in a graded reader during reading-only and reading-while-listening conditions: A study of eye-movements. *Applied Linguistics Review* 13(1): 1-36.
- Tian, Meng. 2014. Making full use of graded readers to facilitate EFL creative writing teaching and learning. *ABAC Journal* 34(2): 56-63.

Antonio Daniel Juan Rubio

Assistant Professor

24 Antonio Daniel Juan Rubio

Language and Literature Didactics Department
University of Granada
C/ Profesor Vicente Callao
Granada 18011, Spain
E-mail: daniel.juan@ugr.es

Received: 2024. 02. 07.

Revised: 2024. 08. 13.

Accepted: 2024. 08. 13.